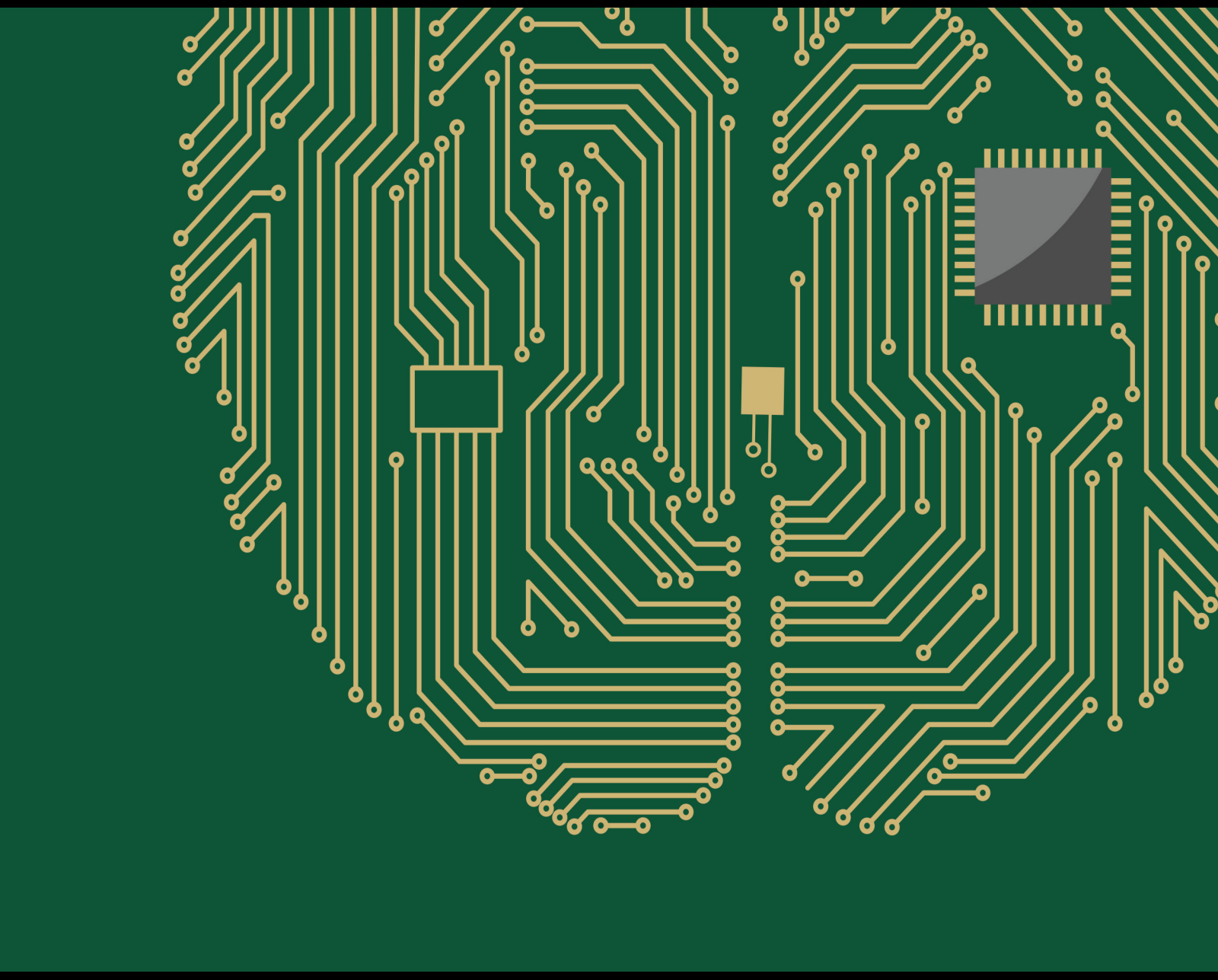
Computational Intelligence and Neuroscience

Swarm Intelligence and Neural Network Schemes for Biomedical Data Evaluation

Lead Guest Editor: Navid Razmjooy
Guest Editors: V. Rajinikanth and Suresh Manic



Swarm Intelligence and Neural Network Schemes for Biomedical Data Evaluation

Computational Intelligence and Neuroscience

Swarm Intelligence and Neural Network Schemes for Biomedical Data Evaluation

Lead Guest Editor: Navid Razmjooy
Guest Editors: V. Rajinikanth and Suresh Manic

Contents

Contents

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 6904217, 8 pages
https://doi.org/10.1155/2021/6904217

Research Article

Macular Hole Detection Using a New Hybrid Method: Using Multilevel Thresholding and Derivation on Optical Coherence Tomographic Images

Sahand Shahalinejad[1] **and Reza Seifi Majdar**[2]

[1]*Department of Electrical and Computer Engineering, Urmia Higher Education Institute, Urmia, Iran*
[2]*Department of Electrical Engineering, Ardabil Branch, Islamic Azad University, Ardabil, Iran*

Correspondence should be addressed to Sahand Shahalinejad; sahandshahali73@gmail.com

Received 15 July 2021; Revised 8 October 2021; Accepted 24 November 2021; Published 22 December 2021

Academic Editor: V. Rajinikanth

Optical coherence tomography (OCT) is a noninvasive imaging test. OCT imaging is analogous to ultrasound imaging, except that it uses light instead of sound. In this type of image, microscopic quality intratissue images are provided. In addition, fast and direct imaging of tissue morphology and reproducibility of results are the advantages of this imaging. Macular holes are a common eye disease that leads to visual impairment. The macular perforation is a rupture in the central part of the retina that, if left untreated, can lead to vision loss. A novel method for detecting macular holes using OCT images based on multilevel thresholding and derivation is proposed in this paper. This is a multistep method, which consists of segmentation, feature extraction, and feature selection. A combination of thresholding and derivation is used to diagnose the macular hole. After feature extraction, the features with useful information are selected and finally the output image of the macular hole is obtained. An open-access data set of 200 images with the size of 224×224 pixels from Sankara Nethralaya (SN) Eye Hospital, Chennai, India, is used in the experiments. Experimental results show better-diagnosing results than some recent diagnosing methods.

1. Introduction

Optical coherence tomography (OCT) is a fundamentally new type of optical imaging modality. OCT performs high-resolution, cross-sectional tomographic imaging of the internal microstructure in materials and biologic systems by measuring backscattered or backreflected light. OCT images are two-dimensional data sets that represent the optical backscattering in a cross-sectional plane through the tissue. Image resolutions of 1 to 15 μm can achieve one to two orders of magnitude higher than conventional ultrasound. Imaging can be performed in situ and in real-time [1]. The unique features of this technology enable a broad range of research studies and clinical applications [2]. Mendes and Abrah 2019 provide an overview of OCT technology, its background, and its potential biomedical and clinical applications. OCT, imaging the internal cross-sectional microstructure of tissues using measurements of optical backscattering or backreflection, was first demonstrated in 1991. OCT imaging was performed in vitro in the human retina and atherosclerotic plaque as examples of imaging in transparent, weakly scattering media and nontransparent, highly scattering media. OCT was initially applied for imaging in the eye and to date. OCT has had the largest clinical impact in ophthalmology. The first in vivo tomograms of the human optic disc and macula were demonstrated in 1993 [3]. OCT enables the noncontact, noninvasive imaging of the anterior eye as well as imaging of morphologic features of the human retina including the fovea and optic disc. Numerous clinical studies have been performed by many groups in the last several years. More recently, advances in OCT technology have made it possible to image nontransparent tissues, thus enabling OCT to be applied in a wide range of medical specialties. Imaging depth is limited by optical attenuation from tissue scattering and absorption. However, imaging up to 2 to 3 mm deep can be

achieved in most tissues. This is the same scale as that typically imaged by conventional biopsy and histology. Although imaging depths are not as deep as with ultrasound, the resolution of OCT is more than 10 to 100 times finer than the standard clinical ultrasound [4]. OCT has been applied in vitro to image arterial pathology and can differentiate plaque morphologies. Imaging studies have also been performed in vitro to investigate applications in dermatology, gastroenterology, urology, gynecology, surgery, neurosurgery, and rheumatology. OCT has also been applied in vivo to image developing biologic specimens for applications in developmental biology. OCT is of interest because it allows repeated imaging of developing morphology without the need to sacrifice specimens. Numerous developments in OCT technology have also been made. High-speed real-time OCT imaging has been demonstrated with acquisition rates of several frames per second. High-resolution and ultrahigh-resolution OCT imaging has been demonstrated using novel laser light sources, and axial resolutions as high as 1 μm have been achieved. Cellular level OCT imaging has recently been demonstrated in developmental biology specimens. CT has been interfaced with catheters, endoscopes, and laparoscopes which permit internal body imaging. Catheter and endoscope OCT imaging of the gastrointestinal, pulmonary, and urinary tracts as well as arterial imaging has been demonstrated in vivo in an animal model. In many cases, the disease is examined only superficially (without considering the complete characteristics involved in the disease). Many research groups are currently performing preliminary clinical studies.

The macular located in the middle of the retina is where most of the cone cells accumulate. A small depression in the middle of the macular called FOA has the largest cone cell. The macular is responsible for central vision, hue vision, and accurate detail recognition. The cylindrical cells are located in the peripheral part of the retina (around the retina) and allow night vision and the movement of objects around. The partial or complete absence of the macular sensory membrane is called macular perforation. The macular hole may be of an unknown and age-related cause (age-related macular hole) or maybe caused by trauma to the eye. The age-related type of the disease is most prevalent in older women in the seventh decade of life. The OCT images of normal and abnormal macular are shown in Figure 1.

2. Related Works

In connection with macular pathologies, we face some harms such as macular edema, age-related macular degeneration (AMD), macular edema, central serous retinopathy (CSR), and macular hole (MH) [5–7]. Macular holes lead to low vision and blindness, which can lead to retinal holes due to overuse of fovea [8]. The disease is more likely to occur in people over 40 years of age [9].

Early diagnosis of the disease will help a lot in its treatment because if the disease progresses little, it can be treated with the help of medicine or surgery. Therefore, knowing the characteristics of the hole such as shape, size, width, diameter, and length can be very effective [10–12]. OCT image is an effective tool to diagnose the condition of an MH [13–15].

The OCT is an applicable tool to diagnose the retina pathologies, which can visualize the 3D shape and structure of the retina without physical contact [16–18]. The high-resolution OCT images of the retina usually have speckle noise and shadows caused by retinal blood vessels and other abnormalities on retina, which are challenges in the segmentation process [19–26].

Because the disease can cause irreparable damage if misdiagnosed by algorithms, manual segmentation has always been a priority, but our algorithm ignores the use of this method with a very low error coefficient to detect the macular hole [27–32] Even in some cases, the correct diagnosis of a disease depends on several studies. Nevertheless, the need for fully automatic diagnostic methods for retinal pathology is now felt.

One of the important applications of OCT images is to detect MH (Figure 1). Therefore, the researchers want to propose full automated, novel, and trustfully methods for MH diagnosis and segmentation [33, 34]. New research studies in this area seek to cover each other to reduce each other's shortcomings and to make them acceptable methods by experts. The valuable information and features provided by OCT images can help researchers develop more valuable segmentation and automated diagnostic techniques to help patients lead better lives [35, 36].

In [37], a fully automatic method to identify the main layers of the retinal that delimits the retinal area is proposed. Therefore, an active contour-based model is used to segment the main retinal boundaries. This method uses the horizontal placement information of these layers and their relative location on the images to restrict the search space. A new OCT-based method to investigate epiretinal membrane (ERM) pathology in human eyes is proposed by Gonz'alez-L'opez et al. [38]. The new approach assesses automatically the ERM thickness and the space between the epiretinal membrane (ERM) and the retina surface. The adjusted mean arc length (AMAL) for segmenting OCT images for macular pathology is being used for segmenting OCT images for macular pathology [39]. This metric is used for automated OCT segmentation. In recent years, a segmentation method based on feature extraction using SFTA-based histological analysis has been introduced by Keller et al. [40]. In this research, a graph-based segmentation is used to find the layers of retina. The 3D level set segmentation approach based on the level set method can used to accurate segmentation of the MH [41]. This method is fully automatic and shows robust results in a variety of conditions. A novel layer guided convolutional neural network (LGCNN) is proposed by Nasrulloh et al. [42] to identify the common types of macular pathologies and normal retina. In this method, the retinal layer segmentation maps are generated by an efficient segmentation network, which can determine two lesion-related retinal layers associated with the meaningful retinal lesions. Then, two subnetworks in LGCNN integrate the information of these layers. Consequently, by focusing on the significant lesion-related layer regions, LGCNN can effectively modify OCT classification. Huang et al. [43] proposed a multi-instance multilabel-based lesions recognition method to detect and recognize simultaneously multiple lesions. The proposed method consists of the

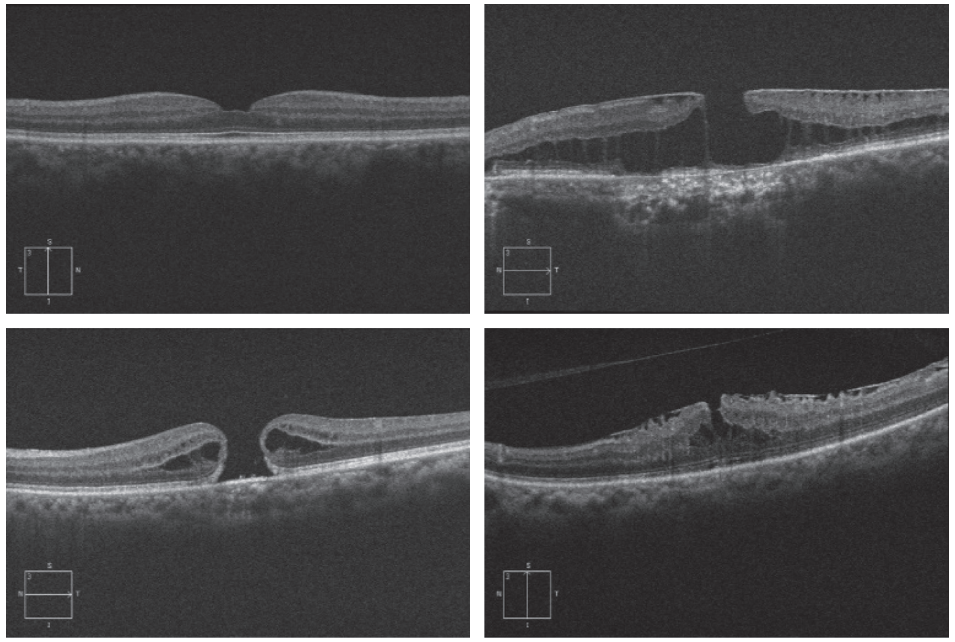


Figure 1: Samples macular images.

segmentation based on the different lesions and feature extraction and constructs multilabel detectors and recognizes the region of interests. A unique joint model that combines multiple information is proposed for retinal segmenting using OCT images [44]. A multimodal multiresolution graph-based method is proposed for internal limiting membrane segmentation within OCT images [45]. A hybrid method using multilevel thresholding and derivation on optical coherence tomographic images was proposed for automated detection [46, 47].

In this paper, a new combination is proposed to obtain more information for macular hole diagnostic. This is a multistep method, which consists of segmentation, feature extraction, and feature selection. A combination of thresholding and derivation is used to diagnose the macular hole. After feature extraction, the features with useful information are selected and finally the output image of the macular hole is obtained. The main contributions of the proposed method are (1) high sensitivity in the various OCT images, (2) better accuracy and reliability than the conventional methods, and (3) short processing time.

The remaining parts of this paper are organized as follows: In Section 3, the mentioned proposed method is introduced. Experiments and results can be found in Section 4. Comparison with some recent methods is demonstrated in Section 5, and the conclusion is given in Section 6.

3. Proposed Methods

The block diagram of the proposed method is shown in Figure 2. The proposed method consists of multiple steps. The preprocessing step is usually used before the main image analysis and data extraction, and its purpose is to obtain an accurate image that removes annoying data from the image. In medical imaging, major disorders are

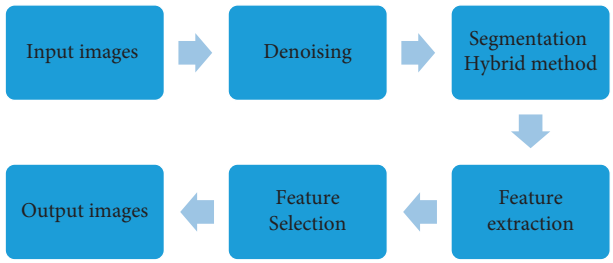


Figure 2: Macular hole detection block diagram.

observed, including noise due to high-frequency reception, different brightness in the field, and problems due to distant orientation. For this reason, preprocessors are applied to all images taken from a device. For this reason, these processes are usually device dependent and must be fast and efficient. Photogrammetric methods are used when the spatial or luminous properties of the noise are known. In this paper, we used an adaptive filter to remove noise from OCT images. In the next step, the proposed algorithm is implemented on the images, and images have been qualitatively improved; therefore, their characteristics must be determined and extracted. Most image data may be subdivided into enclosed areas, dots, and lines. To identify objects, they must be able to distinguish them from the context. It is usually best to convert the gray spectrum image to a binary image (black and white). Techniques such as image splitting and edge recognition work best on binary images but are sometimes applied to images in the gray or color spectrum. The purpose of feature extraction is to extract features that are directly related to the output. After this step, the extracted features with basic information are selected and finally the output image is obtained.

In the proposed method block diagram (Figure 2), a combination of thresholding and derivation is used for diagnosis. This method has been implemented with special emphasis on FOA depression and image features in this area, and finally a suitable segmentation has been used to distinguish this disease from other diseases and the rate of disease progression. Edge as the location of changes in lighting levels, the range of these changes should also be considered to decide on the presence of the edge and its exact location. In this case, if the edges of an image are exposed, the location of all the prominent and opaque objects in the image is determined. As a result, the use of an accurate edge detector directly helps to increase the feature recognition rate and the ability to accurately segment the image. The vector $f(x, y)$ represents the maximum rate of change of brightness.

$$f(x) = \frac{I_r - I_\ell}{2}\left(\operatorname{erf}\left(\frac{x}{\sqrt{2}\,\sigma}\right) + 2\right) + I_\ell. \tag{1}$$

Outset of edge (I_1) and end of edge (I_2) are defined in the following equation:

$$\begin{aligned} L(x) &= -1 \cdot I(x-1) + 0 \cdot I(x) + 1 \cdot I(x+1), \\ L_x &= \begin{bmatrix} +1.2 & 0 & -1.2 \end{bmatrix} L \text{ and } L_y = \begin{bmatrix} +1.2 \\ 0 \\ -1.2 \end{bmatrix} L, \end{aligned} \tag{2}$$

$$L_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} L \text{ and } L_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} L, \tag{3}$$

$$|\nabla L| = \sqrt{L_x^2 + L_y^2}, \theta = \arg\tan 2\left(L_y, L_x\right). \tag{4}$$

On the other hand, when the multilevel thresholding method is applied, it converts the gray level into a binary code. The key point in this method is to select the threshold value (or threshold values for the case where several levels are desired). Most images include objects with a uniform brightness level on a background with different brightness levels. For such images, brightness is a distinguishing feature that can separate the object from the background. Another way to choose the value of the threshold of light is to place the threshold value equal to the minimum point of the histogram between the two peaks.

4. Experiments and Results

4.1. Data Set. In this paper, we use the open-access database collected by Lakshminarayanan et al. 2019 [48]. This database includes more than 500 high-resolution OCT images of different pathological conditions. A raster scan protocol with a 2 mm scan length was used to obtain these images. The size of these images is 512×1024 pixels and took using a Cirrus HD-OCT machine (Carl Zeiss Meditec, Inc., Dublin, CA) at Sankara Nethralaya (SN) Eye Hospital, Chennai, India. In each volumetric scan, a fovea-centered image was selected by an experienced clinical optometrist (MKP). The axial resolution was 5 μm, and the transverse resolution was 15 μm (in tissue). The images were then resized to a 224×224 pixel size. The pathological conditions were determined by clinicians, and the labeling process was done by retinal clinical experts at SN hospital. Progressive stages and macular degeneration are given. This wide range of stages, less severe, medium severe, and more severe stages, in each disease would be ideal for the researchers; therefore, they can test their proposed method in various scenarios. MH OCT images are the main category of the database, and we used several images of MH cases.

4.2. Initialization. The raw data were used separately, in the mentioned method using MATLAB 2019a software to identify macular hole; in the first stage, OCT images were collected as input, with background uniformity and improved image resolution, under preprocessing. In the second stage, identification was performed using the segmentation method and the desired points were examined from the image. In the third and final step, a new hybrid method is implemented on the images. To implement the simulated results, Corei7 CPU, GeForce 7900GTX graphics processing unit, and 16 GB of memory were used.

4.3. Experiments. In this part, the implementation results of the proposed method on OCT images provided by Lakshminarayanan et al. [48] are presented. The image derivation is an important step in many image-processing algorithms. The simplest case may involve applying one algorithm. More complex cases include more accurate margin detection algorithms or applying several separable models, given that time processing is also important. Using two methods of thresholding and derivation, we introduced a new and more efficient method than other methods (Figure 3). In the recognition of surfaces, the most important features can be extracted from the surfaces (including corners and lines). The output of the process is a set of parts whose community includes the whole image or a set of lines extracted from the image. Each pixel in each section is similar in which it has specific properties such as hue, brightness, or texture. Adjacent sections are considered different from each other according to the mentioned features. By recognizing the pixels in the image, considering the existence of valuable and important information in the image, a segmentation algorithm was used in which each pixel is assigned a label so that pixels with the same label have similar properties These features must have properties so that with a set of these features, each image can be described uniquely in order to identify the identity of an image from the patterns of that image. Mean sensitivity and mean accuracy for 12 sample OCT images are shown in Table 1.

The accuracy is computed by comparing the automated results with the ophthalmologist's diagnosis opinions. The proposed algorithm helps the ophthalmologist to educate the patient about the progression of the disease. This

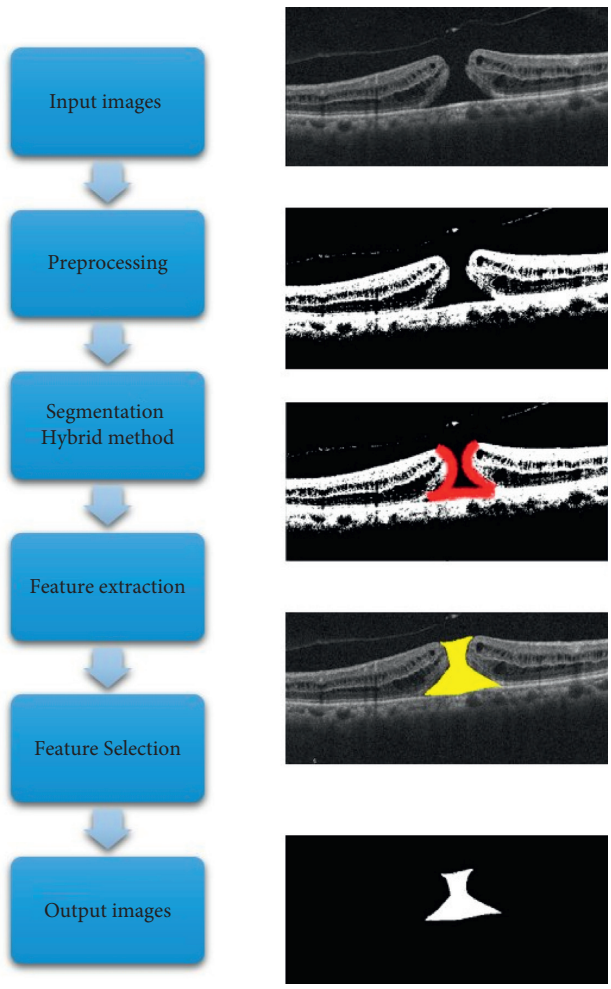


Figure 3: Macular hole detection by the proposed method.

Table 1: Results by applying the proposed algorithm to the OCT images.

Experiment	Mean sensitivity (%)	Mean accuracy (%)
Image 1	87.56	97.25
Image 2	86.65	98.76
Image 3	88.87	96.98
Image 4	89.20	97.75
Image 5	87.12	98.15
Image 6	89.32	96.99
Image 7	88.59	96.50
Image 8	86.75	94.20
Image 9	89.99	98.81
Image 10	88.22	97.98
Image 11	87.70	96.93
Image 12	86.13	98.82

algorithm can aid the ophthalmologist by analysing the huge number of samples in a short time and presenting only the samples exhibiting features of diseases. An automated pre-exclusion of normal cases might help to improve the program's efficiency. In future work, we consider to improve the accuracy in the detection of macular hole. The results of applying the proposed algorithm are clearly shown on the images at each stage of processing.

5. Comparison

In order to compare the results of our proposed method with some recent segmentation methods, we compare it with SVM, KNN, Navie Bayes, Decision Tree, MS-LGDF, and CMF [49]. SVM, KNN, Navie Bayes, and Decision Tree are four traditional segmentation methods; MS-LGDF is a segmentation method based on the local Gaussian distribution fitting (LGDF) energy functional which can collect various macular hole measurements. The segmentation results such as accuracy, sensitivity, Jaccard index, and DSC are shown and compared in Table 2.

According to the comparison, it can be concluded that the proposed method is much better than the other methods. Image segmentation methods are of two main categories: border-based logic and area-based logic, and each of them it is divided into several techniques [50]. The output of a process is a set of sections whose assembly comprises the entire image or a set of lines extracted from the image [49]. Each pixel in each section is similar in which it has specific properties such as color, brightness, or texture. Adjacent sections are considered different from each other according to the mentioned features [51, 52].

Table 2: Comparison between the proposed method and the other methods.

Image	Accuracy	Sensitivity	Jaccard index	DSC
SVM	80.5	79.2	65.3	78.2
KNN	75.7	72.4	63.8	73.1
Navie Bayes	78.4	78.2	64.2	74.2
Decision Tree	69.6	67.5	62.2	66.3
MS-LGDF	96.2	84.3	75.1	83.2
CMF	94.4	67.5	63.2	75.4
The proposed method	**97.5**	**88.3**	**78.5**	**85.3**

Table 3: Average run time comparison between the proposed method and some recent segmentation methods.

Methods	SVM	KNN	Navie Bayes	Decision Tree	MS-LGDF	CMF	Proposed method
Average run time (s)	68	72	59	70	89	86	54

In traditional methods, due to the complexity of the surface of the algorithms used, the accuracy was always low, the sensitivity was at the lowest level, and the long data processing time was clearly visible [52]. Image fragmentation has been done in areas such as computer vision and image processing, and due to its wide and wide application, it has suitable research fields. Despite the complexity of the algorithm, fragmentation has high accuracy and sensitivity and on the other hand fewer data processing time than traditional methods. The success of this research can be seen in other areas as well such as medicine. Remote and image retrieval is crucial, which helps to save, maintain, and protect human life [53].

The run times of the proposed combination method and other compared methods are tabulated in Table 3. The experiments were conducted on a computer with an Intel Corei7 CPU and 16 GB RAM running Windows 7 64Bit operating system. In addition, MATLAB version 2019a is used to obtain the computational times. As can be observed, the run time of the proposed method is better than that of the competition methods.

6. Conclusions

In this paper, we proposed an automatic and powerful method to segment OCT images to detect macular holes. This is a multistep method, which consists of segmentation, feature extraction, and feature selection. A combination of thresholding and derivation was used to diagnose the macular hole. After feature extraction, the features with useful information were selected and finally the output image of the macular hole was obtained. The proposed method was evaluated on an open-access data set from Sankara Nethralaya (SN) Eye Hospital, Chennai, India. Comparison with some state-of-the-art MH segmentation methods reveals the robustness of the proposed method. In the future, we want to develop this method for 3D segmentation and measurements.

Data Availability

All data used in the database are valid.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] C. Mendes and R. Abrahão, "Automatic segmentation of macular holes in optical coherence tomography images: a review," *Journal of Artificial Intelligence and Systems*, vol. 1, pp. 163–185, 2019.

[2] S. M. Caprani, S. Donati, L. Bartalena et al., "Macular hole surgery: the healing process of outer retinal layers to visual acuity recovery," *European Journal of Ophthalmology*, vol. 27, no. 2, pp. 235–239, 2017.

[3] G. R. Wilkins, O. M. Houghton, and A. L. Oldenburg, "Automated segmentation of intraretinal cystoid fluid in optical coherence tomography," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 1109–1114, 2012.

[4] V. Rajinikanth, S. C. Satapathy, S. L. Fernandes, and S. Nachiappa, "Entropy based segmentation of tumor from brain MR images–A study with teaching learning based optimization," *Pattern Recognition Letters*, vol. 94, pp. 87–94, 2016.

[5] N. Slokom, H. Trabelsi, and I. Zghal, "Segmentation of cyctoids macular edema in optical cohenrence tomography," in *Proceedings of the 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 303–306, Monastir, Tunisia, March 2016.

[6] D. C. Hood, S. C. Anderson, M. Wall, and R. H. Kardon, "Structure versus function in glaucoma: an application of a linear model," *Investigative Opthalmology & Visual Science*, vol. 48, no. 8, pp. 3662–3668, 2007.

[7] H. W. Van Dijk, F. D. Verbraak, P. H. B. Kok et al., "Decreased retinal ganglion cell layer thickness in patients with type 1 diabetes," *Investigative Opthalmology & Visual Science*, vol. 51, no. 7, pp. 3660–3665, 2010.

[8] H. Faghihi, F. Ghassemi, K. G. Falavarjani, G. Saeedi Anari, M. Safizadeh, and K. Shahraki, "Spontaneous closure of traumatic macular holes," *Canadian Journal of Ophthalmology*, vol. 49, no. 4, pp. 395–398, 2014.

[9] S. S. Thapa, R. Thapa, I. Paudyal et al., "Prevalence and pattern of vitreo-retinal diseases in Nepal: the Bhaktapur glaucoma study," *BMC Ophthalmology*, vol. 13, no. 1, pp. 9–1, 2013.

[10] J. Chhablani, M. Khodani, A. Hussein et al., "Role of macular hole angle in macular hole closure," *British Journal of Ophthalmology*, vol. 99, no. 12, pp. 1634–1638, 2015.

[11] H. A. Madi, I. Masri, and D. H. Steel, "Optimal management of idiopathic macular holes," *Clinical Ophthalmology (Auckland, N.Z.)*, vol. 10, pp. 97–116, 2016.

[12] D. H. W. Steel, C. Parkes, V. T. Papastavrou et al., "Predicting macular hole closure with ocriplasmin based on spectral domain optical coherence tomography," *Eye*, vol. 30, no. 5, pp. 740–745, 2016.

[13] E. Philippakis, M. Legrand, M. El Sanharawi, A. Erginay, A. Couturier, and R. Tadayoni, "Measurement of full-thickness macular hole size using en face optical coherence tomography," *Eye*, vol. 32, no. 3, pp. 590–596, 2018.

[14] T. Fabritius, S. Makita, M. Miura, R. Myllylä, and Y. Yasuno, "Automated segmentation of the macula by optical coherence tomography," *Optics Express*, vol. 17, no. 18, pp. 15659–15669, 2009.

[15] J. d. Moura, J. Novo, S. Penas, M. Ortega, J. Silva, and A. M. Mendonça, "Automatic characterization of the serous retinal detachment associated with the subretinal fluid presence in optical coherence tomography images," *Procedia Computer Science*, vol. 126, pp. 244–253, 2018.

[16] X. Sui, Y. Zheng, B. Wei et al., "Choroid segmentation from optical coherence tomography with graph-edge weights learned from deep convolutional neural networks," *Neurocomputing*, vol. 237, pp. 332–341, 2017.

[17] G. Gong, H. Zhang, and M. Yao, "Speckle noise reduction algorithm with total variation regularization in optical coherence tomography," *Optics Express*, vol. 23, no. 19, pp. 24699–24712, 2015.

[18] C. Tang, X. Zheng, and L. Cao, "OCT despeckling via weighted nuclear norm constrained non-local low-rank representation," *Laser Physics Letters*, vol. 14, no. 10, p. 106001, 2017.

[19] J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *Journal of Biomedical Optics*, vol. 4, no. 1, pp. 95–106, 1999.

[20] N. M. Grzywacz, J. de Juan, C. Ferrone et al., "Statistics of optical coherence tomography data from human retina," *IEEE Transactions on Medical Imaging*, vol. 29, no. 6, pp. 1224–1237, 2010.

[21] A. A. Lindenmaier, L. Conroy, G. Farhat, R. S. DaCosta, C. Flueraru, and I. A. Vitkin, "Texture analysis of optical coherence tomography speckle for characterizing biological tissues in vivo," *Optics Letters*, vol. 38, no. 8, pp. 1280–1282, 2013.

[22] J. Duan, C. Tench, I. Gottlob, F. Proudlock, and L. Bai, "Automated segmentation of retinal layers from optical coherence tomography images using geodesic distance," *Pattern Recognition*, vol. 72, pp. 158–175, 2017.

[23] H. Chen, S. Fu, H. Wang, Y. Li, and F. Wang, "Speckle reduction based on fractional-order filtering and boosted singular value shrinkage for optical coherence tomography image," *Biomedical Signal Processing and Control*, vol. 52, pp. 281–292, 2019.

[24] H. Lv, S. Fu, C. Zhang, and L. Zhai, "Speckle noise reduction for optical coherence tomography based on adaptive 2D dictionary," *Laser Physics Letters*, vol. 15, no. 5, Article ID 055401, 2018.

[25] S. Sadda, Z. Wu, A. Walsh et al., "Errors in retinal thickness measurements obtained by optical coherence tomography," *Ophthalmology*, vol. 113, no. 2, pp. 285–293, 2006.

[26] I. Krebs, P. Haas, F. Zeiler, and S. Binder, "Optical coherence tomography: limits of the retinal-mapping program in age-related macular degeneration," *British Journal of Ophthalmology*, vol. 92, no. 7, pp. 933–935, 2008.

[27] K. Hu, B. Shen, Y. Zhang, C. Cao, F. Xiao, and X. Gao, "Automatic segmentation of retinal layer boundaries in OCT images using multiscale convolutional neural network and graph search," *Neurocomputing*, vol. 365, pp. 302–313, 2019.

[28] G. N. Girish, A. R. Kothari, and J. Rajan, "Automated segmentation of intra-retinal cysts from optical coherence tomography scans using marker controlled watershed transform," in *Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1292–1295, IEEE, Florida USA, August 2016.

[29] C. Wang, Y. X. Wang, and Y. Li, "Automatic choroidal layer segmentation using markov random field and level set method," *IEEE journal of biomedical and health informatics*, vol. 21, no. 6, pp. 1694–1702, 2017.

[30] D. C. Fern´andez, H. M. Salinas, and C. A. Puliafito, "Automated detection of retinal layer structures on optical coherence tomography images," *Optics Express*, vol. 13, no. 25, pp. 10200–10216, 2005.

[31] K. A. Vermeer, J. Van der Schoot, H. G. Lemij, and J. F. De Boer, "Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images," *Biomedical Optics Express*, vol. 2, no. 6, pp. 1743–1756, 2011.

[32] N. George and C. V. Jiji, "Two stage contour evolution for automatic segmentation of choroid and cornea in OCT images," *Biocybernetics and Biomedical Engineering*, vol. 36, 2019.

[33] M. Berg, "IT-supported skill-mix change and standardisation in integrated eyecare: lessons from two screening projects in The Netherlands," *International Journal of Integrated Care*, vol. 7, p. e15, 2007.

[34] A. G. Marrugo, M. S. Mill´an, C. Gabriel, S. Gabarda, M. Sorel, and F. Sroubek, "Image analysis in modern ophthalmology: from acquisition to computer assisted diagnosis and telemedicine," in *Optics, Photonics, and Digital Technologies for Multimedia Applications II*vol. 8436, p. 84360C, International Society for Optics and Photonics, Bellingham, Washington, USA, 2012.

[35] N. Petrick, B. Sahiner, S. G. Armato et al., "Evaluation of computer-aided detection and diagnosis systems," *Medical Physics*, vol. 40, no. 8, Article ID 087001, 2013.

[36] A. Ramzan, M. Usman Akram, A. Shaukat, S. G. Khawaja, U. Ullah Yasin, and W. Haider Butt, "Automated glaucoma detection using retinal layers' segmentation and optic cup-to-disc ratio in optical coherence tomography images," *IET Image Processing*, vol. 13, no. 3, pp. 409–420, 2018.

[37] B. S. Gerendas, J.-S. Kroisamer, W. Buehl et al., "Correlation between morphological characteristics in spectral-domain-optical coherence tomography, different functional tests and a patient's subjective handicap in acute central serous chorioretinopathy," *Acta Ophthalmologica*, vol. 96, no. 7, pp. e776–e782, 2018.

[38] A. Gonz´alez-L´opez, J. de Moura, J. Novo, M. Ortega, and M. G. Penedo, "Robust segmentation of retinal layers in optical coherence tomography images based on a multistage

active contour model," *Heliyon*, vol. 5, Article ID e01271, 2 pages, 2019.

[39] A. Stankiewicz, T. Marciniak, D. Adam, M. Stopa, P. Rakowicz, and E. Marciniak, "Novel full-automatic approach for segmentation of epiretinal membrane from 3D OCT images," in *Proceedings of the 2017 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, pp. 100–105, IEEE, Poznan, Poland, September 2017.

[40] B. Keller, D. Cunefare, D. S. Grewal, T. H. Mahmoud, J. A. Izatt, and S. Farsiu, "Length-adaptive graph search for automatic segmentation of pathological features in optical coherence tomography images," *Journal of Biomedical Optics*, vol. 21, no. 7, Article ID 076015, 2016.

[41] S. C. Athira, R. M. Roy, and R. P. Aneesh, "Computerized detection of macular edema using OCT images based on fractal texture analysis," in *Proceedings of the 2018 International CET Conference on Control, Communication, and Computing (IC4)*, pp. 326–330, IEEE, Thiruvananthapuram, India, July 2018.

[42] A. V. Nasrulloh, C. G. Willcocks, P. T. G. Jackson et al., "Multi-scale segmentation and SurfaceFitting for measuring 3-D macular holes," *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 580–589, 2017.

[43] L. Huang, X. He, L. Fang, H. Rabbani, and X. Chen, "Automatic classification of retinal optical coherence tomography images with layer guided convolutional neural network," *IEEE Signal Processing Letters*, vol. 26, no. 7, pp. 1026–1030, 2019.

[44] L. Fang, L. Yang, S. Li et al., "Automatic detection and recognition of multiple macular lesions in retinal optical coherence tomography images with multi-instance multilabel learning," *Journal of Biomedical Optics*, vol. 22, no. 6, Article ID 066014, 2017.

[45] A. El Tanboly, M. Ismail, A. Switala et al., "A novel automatic segmentation of healthy and diseased retinal layers from OCT scans," in *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*, pp. 116–120, IEEE, Phoenix, AZ, USA, September 2016.

[46] M. S. Miri, V. A. Robles, M. D. Abràmoff, Y. H. Kwon, and M. K. Garvin, "Incorporation of gradient vector flow field in a multimodal graphtheoretic approach for segmenting the internal limiting membrane from glaucomatous optic nerve head-centered SD-OCT volumes," *Computerized Medical Imaging and Graphics*, vol. 55, pp. 87–94, 2017.

[47] D. H. W. Steel, L. Downey, K. Greiner et al., "The design and validation of an optical coherence tomography-based classification system for focal vitreomacular traction," *Eye*, vol. 30, no. 2, pp. 314–325, 2016.

[48] H. Hamdy Ghoraba, M. Leila, H. Ghoraba, M. A. Heikal, and H. Mansour, "Optical coherence tomography morphological features following modified internal limiting membrane surgical technique in traumatic macular holes," *Clinical Ophthalmology*, vol. 13, pp. 1963–1972, 2019.

[49] Y. Y. Liu, M. Chen, H. Ishikawa, G. Wollstein, J. Schuman, and J. M. Rehg, "Automated macular pathology diagnosis in retinal OCT images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding," *Medical Image Analysis*, vol. 15, no. 5, pp. 748–759, 2016.

[50] N. S. M. Raja, V. Rajinikanth, and K. Latha, "Otsu based optimal multilevel image thresholding using firefly algorithm," *Modelling and Simulation in Engineering*, vol. 2014, Article ID 794574, 2014.

[51] J. Aust, "Optimal multilevel image thresholding: an analysis with PSO and BFO algorithms," *Basic and Applied Science*, vol. 8, pp. 443–454, 2014.

[52] V. S. Lakshmi, S. G. Tebby, D. Shriranjani, and V. Rajinikanth, "Chaotic cuckoo search and Kapur/Tsallis approach in segmentation of t.cruzi from blood smear images," *International Journal of Computer Science and Information Security*, vol. 14, pp. 51–56, 2016.

[53] M. Szkulmowski, I. Gorczynska, D. Szlag, M. Sylwestrzak, A. Kowalczyk, and M. Wojtkowski, "Efficient reduction of speckle noise in optical coherence tomography," *Optics Express*, vol. 20, no. 2, pp. 1337–1359, 2012.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 7265644, 26 pages
https://doi.org/10.1155/2021/7265644

Review Article

The Progress of Medical Image Semantic Segmentation Methods for Application in COVID-19 Detection

Amin Valizadeh [1] **and Morteza Shariatee** [2]

[1]*Department of Mechanical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran*
[2]*Department of Mechanical Engineering, Iowa State University, Ames, IA, USA*

Correspondence should be addressed to Amin Valizadeh; amin.valizadeh@mail.um.ac.ir

Received 14 April 2021; Accepted 18 October 2021; Published 22 November 2021

Academic Editor: Suresh Manic

Image medical semantic segmentation has been employed in various areas, including medical imaging, computer vision, and intelligent transportation. In this study, the method of semantic segmenting images is split into two sections: the method of the deep neural network and previous traditional method. The traditional method and the published dataset for segmentation are reviewed in the first step. The presented aspects, including all-convolution network, sampling methods, FCN connector with CRF methods, extended convolutional neural network methods, improvements in network structure, pyramid methods, multistage and multifeature methods, supervised methods, semiregulatory methods, and nonregulatory methods, are then thoroughly explored in current methods based on the deep neural network. Finally, a general conclusion on the use of developed advances based on deep neural network concepts in semantic segmentation is presented.

1. Introduction

Semantic segmentation of medical images is also known as pixel-level classification. The task is to cluster the parts of the image side by side, which belong to a class of similar objects [1]. The other two key functions of the image are to classify the image's surface and define it. Image classification ensures that each image is exchanged as an equal group of images of similar groups, and monitoring also refers to the object's location and recognition. For predicting pixel level, image segmentation can be used as it categorizes each pixel. Furthermore, there is a task that identifies and separates joints called sample segmentation [2, 3]. Medical image semantic segmentation has a variety of applications, such as road sign detection [4], colon crypt segmentation [5], land-use classification, and land surface classification [6]. It is also widely used in medicine, such as brain and tumor detection [7] and discovering and tracking medical devices in surgery [8]. Numerous applications of segmentation in medicine are listed in some studies [8]. Scene resolution is of great significance in advanced driver assistance systems (ADAS) or car driving areas and depends extensively on semantic image segmentation ([9–11]). Research has developed a deep-learning (DL-) based system for assessing disease. This system automatically scans the location of the disease and measures the shape, size, and percentage of the disease on the CT image of people who have COVID-19 disease. In this study, a strategy (HITL) was proposed for the repeated production of training samples. This method is for radiologists to evaluate the results of DL segmentation, make changes, and frequently add more tutorials to update the model. As a result, they speed up the algorithm's development cycle [12]. As recent findings show, before choosing to use chest CT, a significant number of imaging studies need to be checked for patient diagnosis or patient screening. Artificial intelligence technology, particularly DL analysis tools, could potentially be created to support radiologists in triage, quantification, and data analysis. Artificial intelligence solutions can analyze several cases to determine if a chest CT scan shows lung abnormalities. If the software significantly increases the risk of developing the disease, the case will be reviewed by a radiologist or a physician for

further treatment/quarantine. Such systems or their modifications, after validation and testing, can be a key factor in the diagnosis and control of patients with the virus [13]. The pandemic of COVID-19 appears to have negative impacts on world health and well-being. An important method in combating COVID-19 is efficient screening in infected patients, and one of the most important screening methods is radiological test utilizing chest radiography. Preliminary studies have shown that patients infected with COVID-19 have problems with chest radiographs. COVID-Net is implementing a deep concealer neural network architecture to diagnose COVID-19 patients implying chest X-ray (CXR) images motivated by the academic community's open-source endeavors. This information is accessible to the public [14]. In the context of e-healthcare, Zhang et al. showed a privacy-preserving optimization of the clinical pathway query method (PPO-CPQ) [15]. Ala et al. have used a metaheuristic algorithm and optimized an appointment scheduling issue for healthcare systems depending on the quality of fairness service [16]. Also, Xu et al., to simulate pathogenesis diagnosis, proposed a computer technique called network differentiation [17]. Segmentation accuracy has greatly improved since the reemergence of the deep neural network. In general, traditional methods are called the methods that came before the deep neural network. The following parts of this convention are followed in this study and standard segmentation techniques are briefly analyzed in this article, and, most significantly, this development builds on the recent progress of adopting and organizing a deep neural network from different aspects. Furthermore, the image segmentation measurement and assessment databases are checked. The rest of this study is organized as follows: In the dataset and assessment criteria, Section 2 explores the semantic segmentation of the image. A description of traditional methods is given in Section 3. Section 4 outlines recent developments in detail. Finally, a description of the work performed is given with conclusions in Section 5.

2. Datasets and Evaluation Metrics

2.1. Datasets. Many general datasets are currently connected to image segmentation, such as PASCAL VOC, MS COC, ADE20K, and, in the field of the autonomous driving area, Cityscapes [11] and KITTI ([9, 10]). The challenge of visual object classes, or VOCs, consists of two components [15]: (1) image and annotation datasets that are available to the public and (2) annual workshops and competitions that are held online on some websites and sometimes in person. The main challenges have been dealt with since 2005. By 2012, the challenge included 20 classes. Educational and validation data contained 11530 images containing 27450 annotated objects with areas of interest and 6929 segmented images. Also, in image segmentation, datasets have been extensively utilized. The Microsoft COCO dataset [2] contains images of 91 objects, where a 4-year-old person can quickly identify with 2.5 million labeled samples in 328,000 images. Authors also introduced the dataset with a detailed statistical analysis compared to PASCAL data [15], ImageNet data [18], and SUN data [19]. An analysis of 50 chest X-ray images of 25 positive COVID-19 cases was confirmed because of the lack of an available COVID-19 dataset. Seven distinct architectures from neural network models are used in COVIDX-Net. Each deep neural network model can analyze the number of X-ray images to identify the patient's state as negative or positive COVID-19 [20]. The authors in [20] collected images from 5 different sources to test this idea and generate a dataset of 170 X-ray images and 361 CT images of COVID-19. Two explanations exist for the use of photographs from these sources. First, to help radiologists diagnose COVID-19 worldwide, it is important to design an advanced tool. Second, for the scientific community and the general public, photographs of these sources are openly available. Also, the images used in it will be publicly available in a GitHub repository [21]. With 150 classes of objects and materials, ADE20K data [17] is another scene analysis criterion. ADE20K data contains the object segmentation mask and component segmentation mask, unlike other datasets. There are also several pictures of parts of the head (like the mouth, eyes, and nose). In the training suite, there are precisely 20210 images, and there are 2,000 images in the validation suite and 3,000 images in the experimental suite [17]. Some of these images are depicted in Figure 1.

The Cityscapes dataset is a criterion that focuses on understanding the meaning of urban street scenes [11]. The collection contains 30 groups obtained from 50 towns in 5,000 fine annotated pictures. The selection period, which includes spring, summer, and autumn, is also several months. Figure 2 displays one of the images of this data in the annotations.

The KITTI dataset [9], another autonomous driving dataset recorded by driving on highways and in rural areas around Karlsruhe, is another example of semantic image data. On average, a maximum of 15 cars and 30 pedestrians can be seen in each image. Zhou et al. proposed a model for evaluating the clarity of screen content and natural scene images while blind [10]. Lv et al. proposed a deep-learning-based fine-grained visual computation [11]. Liu et al. have investigated the Style and Characters Inpainting Based on CGAN [12]. Road detection, stereo reconstruction, light current visual measurement, 3D object detection, and 3D tracking are the principal functions of this dataset (http://www.cvlibs.net/datasets/kitti/). One use of image segmentation is in automated vehicles. The system uses augmented reality to describe the amount of automation and its dependability to increase the system's confidence and reliability. In addition to the above databases, there are many others such as SUN, the Visual Database of Shadow Detection or Texture Segmentation (https://zenodo.org/record/59019#.WWHm3oSGNeM), Berkeley segmentation dataset [22], and LabelMe dataset [23] whose complete information can be found (http://homepages.inf.ed.ac.uk/rbf/CVonline/Imagedbase.htm). There are various imaging models in the field of medical data, most of which have been applications of DL methods on MRI, mammography, or CT scan imaging data. However, different areas of the body have this data or even other imaging samples. However, the main focus of this research is on three datasets, which are from the brain and

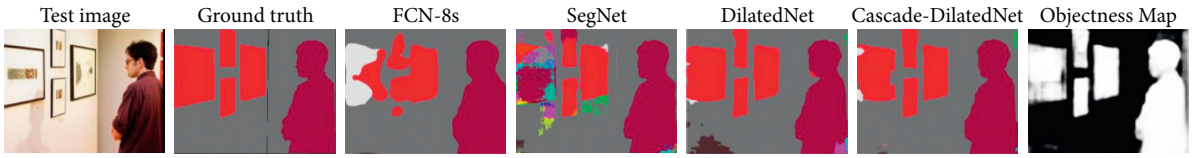


Figure 1: Examples of ADE20K data images. From left to right and from top to bottom, the first segmentation of object masks is seen. The second to fifth elements of photo segmentation are linked to the object's portions (e.g., body parts, glass parts, and photo board parts). In the sixth segment, parts of the head are displayed (such as the eyes, mouth, and nose) [17].

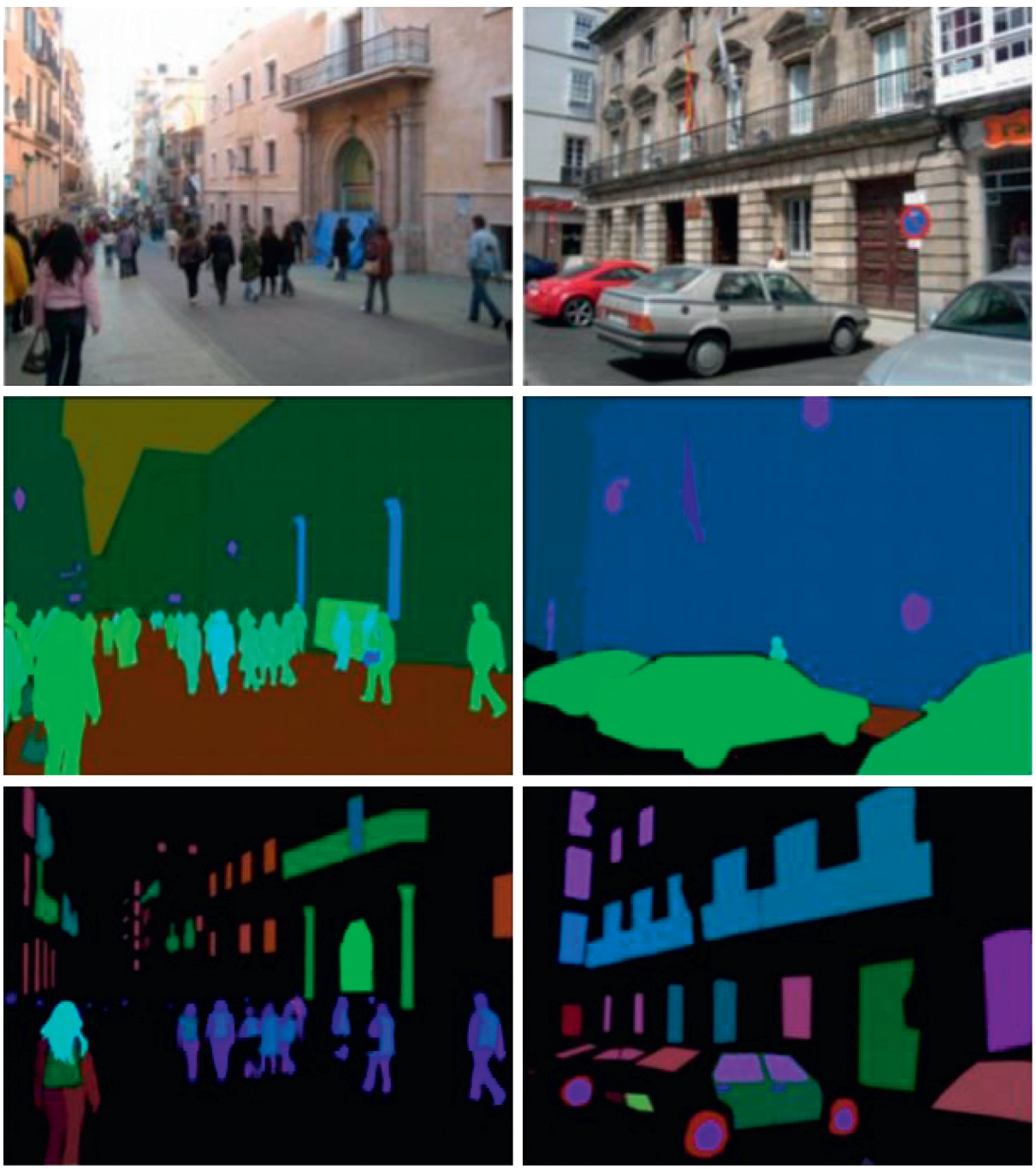


Figure 2: One of the images in the Cityscapes database [17].

chest areas. One of the most important brain MRI datasets working on diagnosing tumors, Alzheimer's, and MS is BraTS dataset (https://ieee-dataport.org/competitions/brats-miccai-brain-tumor-dataset). From the download path of this dataset, there will be four folders: T1, T2, Flair, and T1Ce, respectively. Each section has 155 sections, MRI image sections are weak, and there are 155 sections for a dataset and 210 sections for a high-grade Glioma dataset. At the same time, 75 sections are in another type of Glioma. Hence, the number is 285 cases. This dataset has versions between 2012 and 2019, and an example of the images of this dataset is in form (3). There is also another original dataset called TCGA-GBM, which has higher quality 3D images than BraTS (https://portal.gdc.cancer.gov/projects/TCGA-GBM) (see Figure 3).

There are also valid datasets in the field of mammography images used to diagnose breast tumors, one of the most important of which is a dataset called MIAS (https://www.mammoimage.org/databases).

Also, on the same website, other datasets called DDSM, AMDI, and IRMA are used to diagnose cancerous tumors in different shapes and sizes as benign and malignant. An example of these images is shown in Figure 4.

2.2. Evaluation Metrics. For image segmentation and scene analysis, standard performance assessment metrics include pixel resolution P_{acc}, middle resolution M_{acc}, region intersection upon the union M_{IU}, and connection area sharing frequency weight FW_{IU}. It is presumed that n_{ij} describes the number of class i pixels that are supposed to belong to class j, where there are various groups n_d and it is assumed that $t_i = \sum_J n_{ij}$ represents the number of pixels in class i. All of these relationships are written in the four following formulas [26]:

$$P_{\text{acc}} = \frac{\sum_i n_{ii}}{\sum_i t_i}, \tag{1}$$

$$M_{\text{acc}} = \frac{1}{n_{cl}} \sum_i \frac{n_{ij}}{t_i}, \tag{2}$$

$$M_{\text{IU}} = \frac{1}{n_{cl}} \sum_i \frac{n_{ij}}{t_i + \sum_j n_{ji} - n_{ii}}, \tag{3}$$

$$\text{FW}_{\text{IU}} = \frac{1}{\sum_k t_k} \sum_i \frac{t_i n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}}. \tag{4}$$

There are also other evaluation metrics for the segmentation of medical images, which are popularly used in scientific societies. These include accuracy, sensitivity, specificity, recall, ROC curve, and Area under Curve (AUC) rate.

3. Traditional Methods

Before presenting deep neural networks, features and classification methods are applied to the most important topics [19]. A feature is a piece of data that is applied to solving computational tasks in machine vision and image processing. That is the same context of machine learning precision and the identification of patterns [27]. For semantic segmentation in images, a number of features are used, such as pixel color, histogram of oriented gradients (HOG) ([28]), scale-invariant feature transform (SIFT) [29], local binary pattern (LBP) [30], SURF method [31], Harris Corner Detection [32], method of Shi-Tomasi [33], subpixel corner method [34], SUSAN edge method [35], Features from Accelerated Segment Test (FAST) [36], FAST-ER method [37], AGAST method [38], AGAST multiscale detection method [39], the bag-of-visual-words (BOV) [40], the Poselets method [41], the Textons method [42], and many other methods. Approaches to image semantic segmentation do not include supervised or unsupervised cases [43]. In particular, thresholding, which is commonly used in gray surface images, is a simple method. In the medical industry, optimization, classification, and diagnosis are very common, using imaging equipment ([44–46]). In general, in this regard, thresholding methods are very efficient. K-means clustering means an unsupervised clustering process. The K-means algorithm specifies that the number of clusters must be defined in advance. The points of k are initially randomly positioned in the property space. Additionally, each datum is allocated to the closest points. The gravity points are then successively transferred to the middle of the cluster. This process proceeds until the stop criterion is met [47]. The problem of segmentation can be used as an energy model, resulting from a compression method [48]. Intuitively, edges are an important part of segmentation, and there is also much research on edge recognition ([49–53]). Also, edge-based approaches and regional growth methods are other branches. Support vector machines (SVMs) are binary classifiers that are well studied and are employed in many jobs. Inseparable linear problems can be solved with the slack variables, too [54, 55]. The core approach was also used for integral tasks by the mapping of dimensional broader features. A Markov random field (MRF) is a set of randomized variables with an indirect diagram of a Markov attribute. Also, the Markov stochastic grid is a directionless graphical model (http://host.robots.ox.ac.uk/pascal/VOC/voc2010/results/index.html; [56]).

In general, a case study has been made between the methods of segmentation in images in the field of semantic segmentation, which is shown in Table 1. Classic studies on the detection of cancerous tumors from MRI images have also been presented. In [57], the method of Brownian motion of water molecules to produce contrast has been done. Also, in [58] an improved edge detection method for segmentation is presented. The watershed method and the hierarchical clustering algorithm have also been studied in [59]. Also, in [60], anisotropic diffusion based on segmentation and pattern based on group classification based on support vector machine and segmentation with FCM has been done. The application of a genetic algorithm and discrete wavelet transform thresholding method is presented in [59]. Qiao et al. have presented a local wavelet acoustic pattern and an MLP optimized by a modified Whale

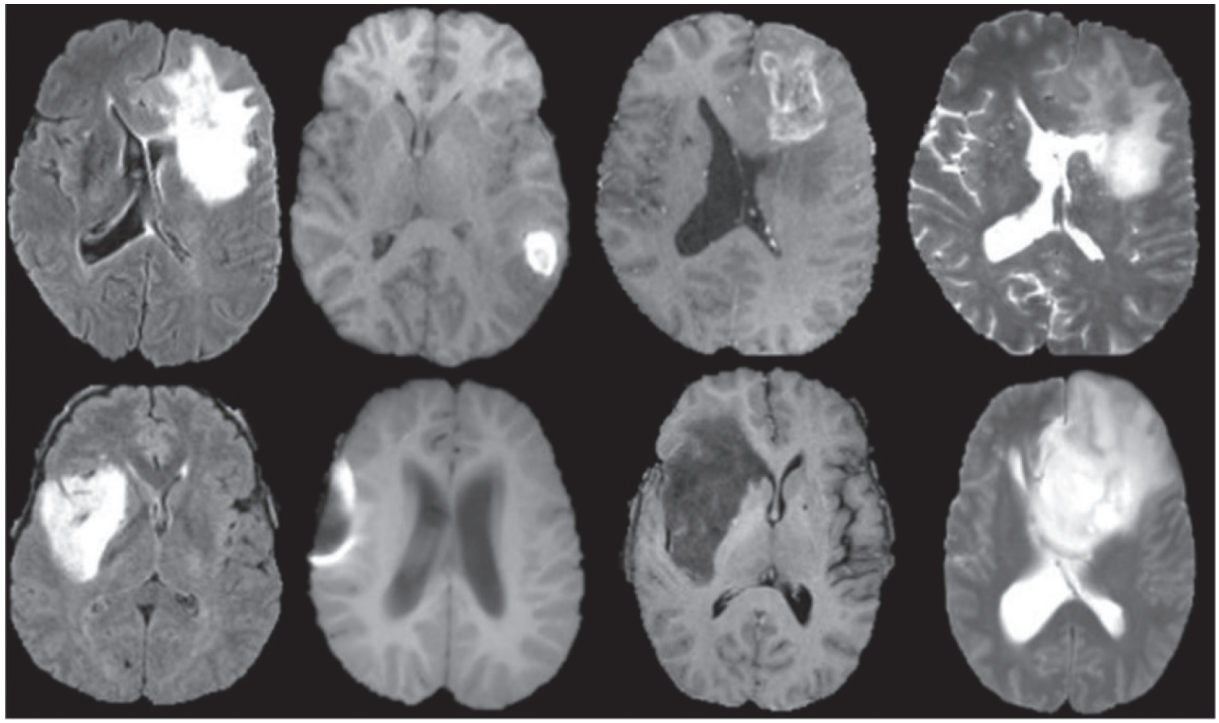


Figure 3: Example of BraTS dataset [24].

Optimization algorithm for classification of underwater objects [59].

Also, [65] presents a combined approach called ant colony optimization (ACO) algorithm and genetic algorithm. In [66], the chaotic firefly algorithm based on the FCM algorithm has been performed. The application of the optimal Particle Swarm Optimization (PSO) in [67] has also been studied, and in [68] a bat optimization algorithm for segmenting MRI images for different purposes is presented. In general, a case comparison between the existing methods in the segmentation field in MRI images has been done, which is shown in Table 2.

In the field of classical methods used to segment mammographic images, we can refer to the research [69] that has used the segmentation method of regional growth with a cellular neural network with a specific threshold. The use of a Back-Propagation (BP) neural network has also been considered in [70]. Applying the new Naïve Bayesian method in [71] has also been considered in this field. In [72], the regression-based evolutionary methods are used to diagnose breast cancer tumors to estimate and predict the remaining life based on the size of the tumor. In [73], the classification or diagnosis of breast cancer in mammographic images combined with wavelet analysis and genetic algorithm is presented. Xu et al. provided a method for identifying, classifying, and predicting nucleic acid-binding proteins [74].

Also, [75] presents a semisupervised adaptive algorithm named GrowCut for the segmentation of tumors of interest areas or ROI of mammographic images based on the amendment of the law of automatic evolution. In general, a case comparison has been made between the methods available for segmentation in mammographic images, which is shown in Table 3.

4. Recent Deep Neural Network Methods in Segmentation

4.1. Artificial Neural Network (ANN). Biological neurons are inspired by the artificial neural network (ANN). An artificial neuron is an essential element of an artificial neural network. Each artificial neuron has only inputs that weigh. Neurons issue a scale following a transfer function or activation function. An instance of a neural model is shown in Figure 5.

Based on artificial neurons, the accumulation of different neurons is automatically Autoencoder [81], Restricted Boltzmann Machine (RBM) [82], Recurrent Neural Network (RNN) or convolutional neural network (CNN) [83], Long Short-Term Memory (LSTM) [84], and other models. The basic architecture is shown in Figure 6.

A shared weight architecture, influenced by biological mechanisms, is used by the convolutional neural network (CNN) [83]. The connection pattern between neurons mimics the development of the visual cortex of the animal. Acceptance is another essential term, indicating that individual cortical neurons can respond to stimuli only in a small region of the visual field. They also have immutable or complex spatial properties dependent on the architecture with shared weight and spatial characteristics. Due to this excellent structure, the convolutional neural network has gained significance which caused image classification, segmentation, and detection. In the following section, recent developments using convolutional neural networks in the semantic segmentation of the image will be presented.

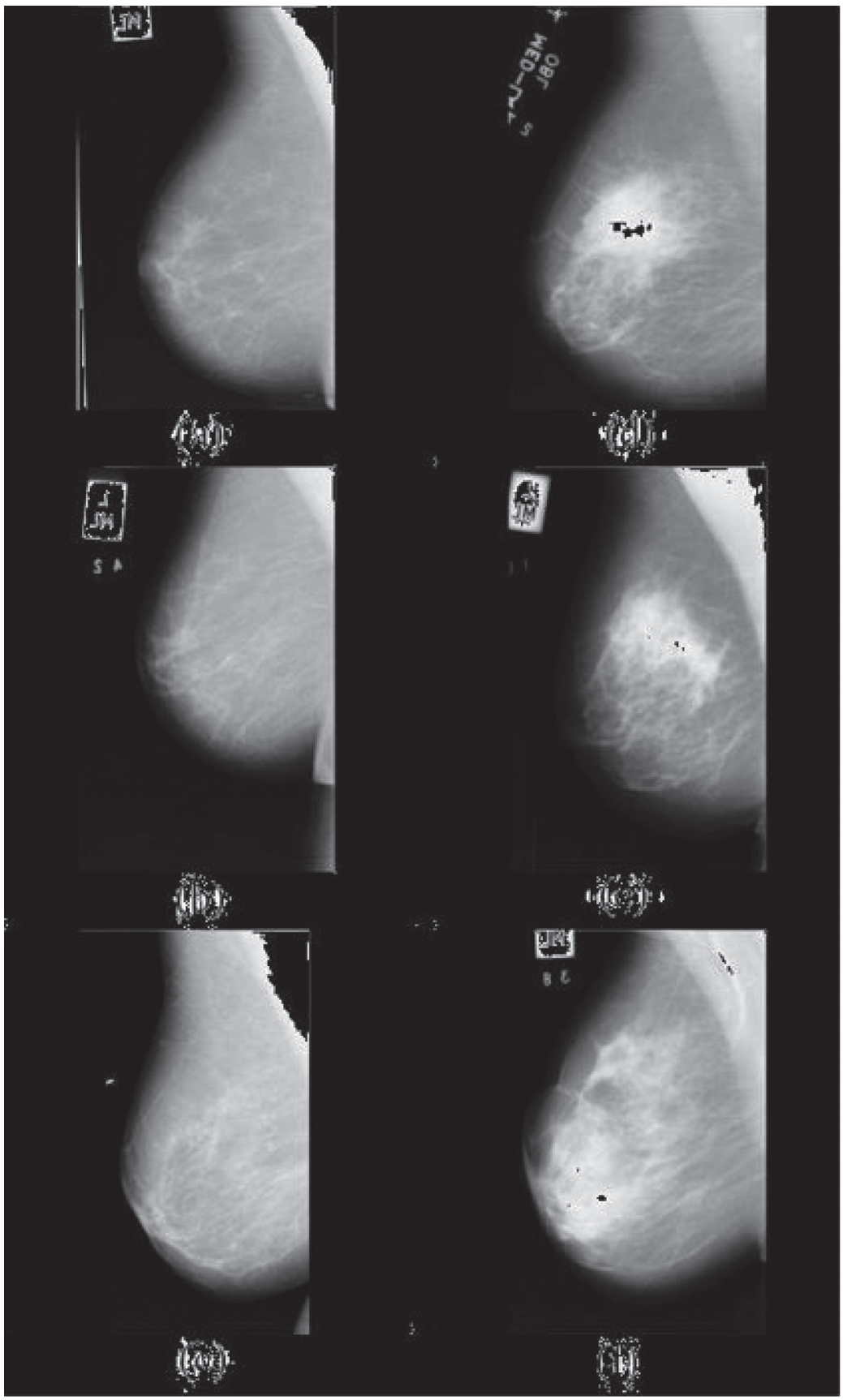


Figure 4: Sample image of normal and malignant tumor based on MIAS dataset [25].

Table 1: Comparison between existing methods for semantic segmentation.

Reference	Year	Method	Advantages	Disadvantages
Bourdev et al. [28]	2010	HOG	Ability to specify areas, especially at the edges, with high clarity and accuracy	High computational complexity, inability to be used in a wide range of applications
Xia et al. [27]	2005	HOG	Ability to specify areas, especially at the edges, with high clarity and accuracy	High computational complexity, inability to be implied in an extended variety of applications
Lowe [29]	2004	SIFT	Accurate identification of areas and edges	High computational complexity, inability to be implied in an extended variety of applications
He and Wang [30]	1990	LBP	High capability in cryptography and image data encryption and edge detection operations	The algorithm is slow
Bay et al. [31]	2008	SURF	Accurate identification of areas and edges	High computational complexity, inability to be used in a wide range of applications
Derpanis [32]	2004	Harris corner detection	Ability to find the corners of an image outside the edges, ability to be used in a wide range of high-sensitivity image processing systems	Low accuracy and slow method
Shi and Tomasi [33]	1994	Shi-Tomasi	Ability to specify areas, especially at the edges, with high clarity and accuracy	High computational complexity, inability to be used in a wide range of applications
Medioni and Yasumoto [34]	1987	Corner detection with subpixels	Ability to find the corners of an image outside the edges, ability to be used in a wide range of high-sensitivity image processing systems	Low accuracy and slow method
Smith and Brady [35]	1997	SUSAN corner detection	Accurate edge detection based on texture and brightness and better capabilities than classic operators such as Canny and Prewitt	Low accuracy in high-resolution images and slow method
Rosten and Drummond [36]	2005	FAST	Ability to specify areas, especially at the edges, with high clarity and accuracy	High computational complexity, inability to be used in a wide range of applications
Rosten et al. [37]	2010	FAST-ER	Ability to specify areas, especially at the edges, with high resolution and precision in multiscale modes	High computational complexity, inability to be used in a wide range of applications
Mair et al. [38]	2010	AGAST	Ability to specify areas, especially at the edges, with high clarity and accuracy	High computational complexity, inability to be used in a wide range of applications
Leutenegger et al. [39]	2011	Multiscale AGAST	Ability to specify areas, especially at the edges, with high resolution and precision in multiscale modes	High computational complexity, inability to be used in a wide range of applications
Venegas-Barrera and Manjarrez [40]	2004	BOW	Ability to specify areas, especially at the edges, with high clarity and accuracy	High computational complexity, inability to be used in a wide range of applications
Brox et al. [41]	2011	Poselets	Ability to specify areas, especially at the edges, with high clarity and accuracy	High computational complexity, inability to be used in a wide range of applications
Zhu et al. [42]	2005	Textons	Ability to specify areas, especially at the edges, with high clarity and accuracy	High computational complexity, inability to be used in a wide range of applications
Strauss and Hartigan [47]	1975	K-means	Ability to find clusters in images and cluster them	Slow method and need to combine with faster methods
Shen et al. [54]	2004	Edge detection and regional growth	Ability to specify areas, especially at the edges, with high clarity and accuracy	High computational complexity, inability to be used in a wide range of applications
Shan et al. [55]	2004	SVM	High capability in high-precision image classification operations in pairs and the ability to separate features with vectors	High computational complexity, slow method
Shotton et al. [56]	2006	MRF	High capability in high-precision image classification operations in pairs and the ability to separate features with vectors	High computational complexity, slow method

Table 1: Continued.

Reference	Year	Method	Advantages	Disadvantages
Hassantabar et al. [61]	2020	CNN	Ability to diagnose the COVID-19 infected lung tissue for segmentation and classification of patients	(i) Small numbers of images (ii) Unable to find illness severity
Dorosti et al. [62]	2020	Sensitivity analysis	This approach can help in the identification of beneficial parameters as well as the avoidance of patient mortality in all sorts of disease	(i) Data limit (ii) Ignoring other variables
Sharifi et al. [63]	2021	CNN	Diagnosis of fatigue foot using CNN	(i) High computational complexity (ii) Integrated only on CNN method
Laradji et al. [64]	2021	Supervised consistency learning	The best loss function for prediction	(i) Unable to detect patients uniquely (ii) Should be connected to other methods

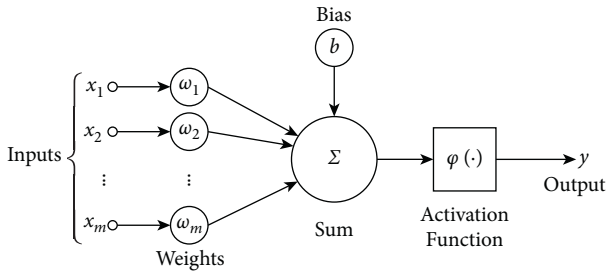


Figure 5: Model of artificial neurons [80].

4.2. Fully Convolutional Network (FCN). The article in [86] represents the first study in the image segmentation field to present ANNFCN. Replacing the utterly connected layer with the fully convolutional layer is the fundamental concept of this method. Using the interpolation layer, the network recognizes that the output size is just like the input required for segmentation. Most significantly, by successful inference and learning, the network is educated, takes on the required size, and produces the correct size output. The FCN was introduced in VGG-Net and has reached a substantial role in the segmentation of PASCAL VOC (20 percent relative increase to 62.2 percent of the average IU in 2012). However, the assumption takes less than one-fifth of a second for a regular image. The main FCN architecture is shown in Figure 7.

4.3. Interpolation against Parsing in Medical Image Semantic Segmentation. The parsing layer is also approved in the semantic segmentation of images and the FCN architecture. Degradation and sampling layers in the pooling layer that define pixel type labels and predict segmentation masks are the degradation network used in [88]. Unlike FCN in [88], this grid is used for proposed thing designs to get the synthetic parts as an example for the final semantic segmentation. The sampling method step adopts two-line interpolation, which can be found in [86]. The sampling stage of the samples has commonly approved two-line interpolation due to the computational efficiency and good retrieval of the original image. The decomposition operation is an inverse calculation of the convolution function that can also retrieve the input size. It can then be utilized to segment the function mapping size to the original input size to retrieve it. The architecture used in [88] is seen in Figure 8 . Some researchers still use the decomposition layer in multiple versions to introduce semantic segmentation, which can be found in [74, 89, 90].

CT images are also used to obtain information about COVID-19 patients. The CT image shows the condition of the patients' lungs and shows how much the disease has affected the lungs.

4.4. Connect FCN with CRF and Other Traditional Methods. Responses in the last layer of deep convolutional neural networks (DCNNs) are not sufficiently localized to effectively segment an entity, according to DeepLab research [92]. Mixing an ultimately linked random field or CRF in the DCNN end layer solves this poor localization property. Authors' method in the test determined in the semantic image segmentation work of PASCAL VOC-2012 reaches 71.6% IOU accuracy. After this, they implement another segmentation architecture by matching Domain Transform (DT) with DCNN [92]. Since the dense CRF inference is costly in terms of computation, Domain Transform (DT) relates to a modern method of maintaining edge filtering. A reference edge mapping governs the smoothing rate. The Domain Transform (DT) is several times more rapid than the dense CRF assumed. Finally, studies compare the effects of semantic segmentation and reliably document the boundaries of the object. Researchers also use superpixels to segment images in the domain of the semantic segmentation [93]. Reference [94] deals with semantic segmentation by merging rich details, including mixing label fields and high-order relationships, in the Markov random field (MRF).

4.5. Dilated Convolution. The majority of semantic segmentation methods focus on the compatibility of convolutional neural networks (CNNs), initially designed to classify images. Dense prediction, however, is structurally distinct from classification, as are semantic image segmentation tasks. An instance of an open structure of convolution can be seen in Figure 9. Reference [92] has previously used this technique in its work, which is named Atheros convolution or convolution hole [92] or open convolution [96]. Convolution was originally developed to efficiently calculate wavelet transform in an "algorithm à trous" scheme [97]. Reference [96] systematically presented

Table 2: Comparison between existing methods in the segmentation field in MRI images.

Reference	Year	Method	Advantages	Disadvantages
Darwiesh et al. [57]	2016	The method of Brownian motion of water molecules to produce contrast	Detecting edge areas to separate sections with tumors and nontumor sections	(i) Lack of detection of tumors in other tumors or other areas (ii) High computational complexity and slow method (iii) Lack of separation of areas with benign and malignant tumors
Aslam et al. [58]	2015	Edge detection	Detecting edge areas to separate sections with tumors and nontumor sections	(i) Lack of detection of tumors in other tumors or other areas (ii) High computational complexity and slow method (iii) Low accuracy (iv) Lack of separation of areas with benign and malignant tumors
Qiao et al. [59]	2021	Watershed and hierarchical clustering algorithm	Detecting edge areas to separate sections with tumors and nontumor sections	(i) Lack of diagnosis of tumors in other tumors or other areas (ii) High computational complexity and slow method (iii) Low accuracy (iv) Lack of separation of areas with benign and malignant tumors
Ain et al. [60]	2014	Concrete anisotropic emission based on group classification, support vector machine (SVM), and FCM	High accuracy in diagnosing and classifying areas with tumors	Lack of comparison with previous methods and lack of consideration for comparison with DL methods or other neural networks
Mobahi et al. [48]	2011	Genetic algorithm and discrete wavelet transform threshold method	Detecting edge areas to separate sections with tumors and nontumor sections	
Karnan, and Selvanayaki [65]	2010	The combined approach of ant colony optimization algorithms and genetic algorithm	Detecting edge areas to separate sections with tumors and nontumors sections	(i) Lack of diagnosis of tumors in other tumors or other areas (ii) Very high computational complexity and slowness of the method (iii) Low accuracy (iv) Lack of separation of areas with benign and malignant tumors
Ghosh et al. [66]	2018	FCM-based chaotic firefly algorithm	(i) Detecting edge areas for separating tumor and nontumor sections (ii) High execution speed with the complexity of the method (iii) Accurate detection of features	(i) High computational complexity (ii) Lack of separation of areas with benign and malignant tumors
Zhu et al. [67]	2018	Particle swarm optimization (PSO)	(i) Detecting edge areas for separating tumor and nontumor sections (ii) High execution speed with the complexity of the method (iii) Accurate detection of features	(i) Lack of diagnosis of tumors in other tumors or other areas (ii) High computational complexity and slow method (iii) Lack of separation of areas with benign and malignant tumors
Alagarsamy et al. [68]	2019	Bat algorithm	(i) Detecting edge areas for separating tumor and nontumor sections (ii) High execution speed with the complexity of the method (iii) Accurate detection of features	(i) Lack of detection of tumors in other tumors or other areas (ii) High computational complexity and slow method (iii) Lack of separation of areas with benign and malignant tumors
Memiş et al. [158]	2020	Deep CNN	Finding the head bone femoral and femur properties for low-quality MRI images	(i) Small volume of the dataset for validation and verification (ii) Unable to support any types of disease
Duran et al. [159]	2020	Self-attention model	End-to-end attention model with multiple classes	(i) Only unable to detect prostate cancer (ii) An additional mechanism for CAD models

Table 2: Continued.

Reference	Year	Method	Advantages	Disadvantages
Hu et al. [137]	2019	3D-DenseUNet-569	(i) Adaptable to depthwise separable convolution (ii) Drop the GPU processing time	(i) Low-level feature extraction (ii) Improper for big data (iii) Unable to adapt to 2D images
Karayegen & Feyzi [161]	2021	Deep learning models	(i) High prediction method (ii) Differing modality of MRI images (iii) 3D image analysis	(i) Limited dataset for verification (ii) Do not use all image area (iii) Needs ground truth
Ahmadi et al. [162]	2021	Deep spiking neural network	(i) Low computational complexity (ii) Used quantum filter (iii) High accuracy	(i) Multistep method (ii) Overfitting in some analysis
Ahmadi et al. [163]	2021	Robust PCA and CNN	(i) Clustering and segmentation method (ii) Automated clustering (iii) Used remove outliers (iv) High accuracy and sensitivity	(i) High complexity (ii) Do not support 3D images

Table 3: Comparison between existing methods in the field of segmentation in mammographic images.

Reference		Method	Advantages	Disadvantages
Rouhi et al. [69]	2015	Regional growth with a cellular neural network with a specific threshold	Ability to diagnose benign and malignant tumors, high accuracy in classification	High computational complexity (i) Lack of accurate detection of areas with tumor (ii) Lack of comparison with previous methods
Kaymak et al. [70]	2017	Back-Propagation (BP)	A convenient way to use in neural network training, highly fast execution speed in training	Uncertainty of the exact type of approach proposed and lack of comparison at the time of classification and uncertainty of benign and malignant tumors
Karabatak [71]	2015	Naïve Bayesian	Ability to diagnose benign and malignant tumors, high accuracy in classification	High computational complexity (i) Lack of accurate detection of areas with tumor (ii) Lack of comparison with previous methods
Wang et al. [72]	2018	Regression-based methods	Ability to estimate and predict remaining life based on tumor size, high accuracy in detection	
Pereira et al. [73]	2014	Wavelet analysis and genetic algorithm	Ability to diagnose benign and malignant tumors, high accuracy in classification	High computational complexity (i) Lack of accurate diagnosis of areas with tumor (ii) Lack of comparison with previous methods
Cordeiro et al. [75]	2016	Semisupervised adaptive algorithm GrowCut	Ability to diagnose benign and malignant tumors, high accuracy in classification	High computational complexity (i) Lack of accurate detection of areas with tumor (ii) Lack of comparison with previous methods
Ahmed et al. [76]	2020	Mask RCNN	(i) Increased AUC for transfer learning (ii) Use for X-ray mammographic image	(i) Low accuracy (ii) Used low volume dataset for verification (iii) High rate of oversampling
Lee et al. [77]	2020	Multiscale grid average pooling	(i) Utilizing global and local spatial feature (ii) Novel attention module (iii) Ultrasound image dataset	(i) Lower accuracy of segmentation (ii) High computational complexity (iii) A small volume of the dataset
Soulami et al. [78]	2021	UNet model	(i) High accuracy for breast cancer detection (ii) High f1-score and AUC	(i) High complexity model (ii) Overfitting in some models (iii) Lower volume of analysis
Huang et al. [79]	2021	Fuzzy fully CNN	(i) Fuzzy membership function (ii) Conditional random fields	(i) Low sensitivity (ii) Low intersection over union (iii) Low resolution and poor quality

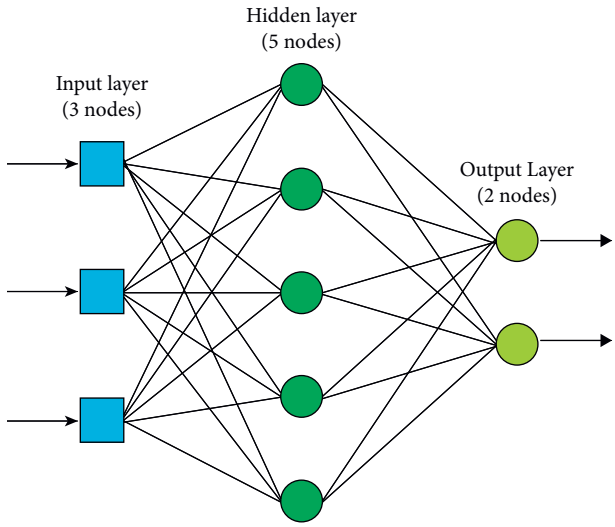


Figure 6: An example of a model of an artificial neural network [85].

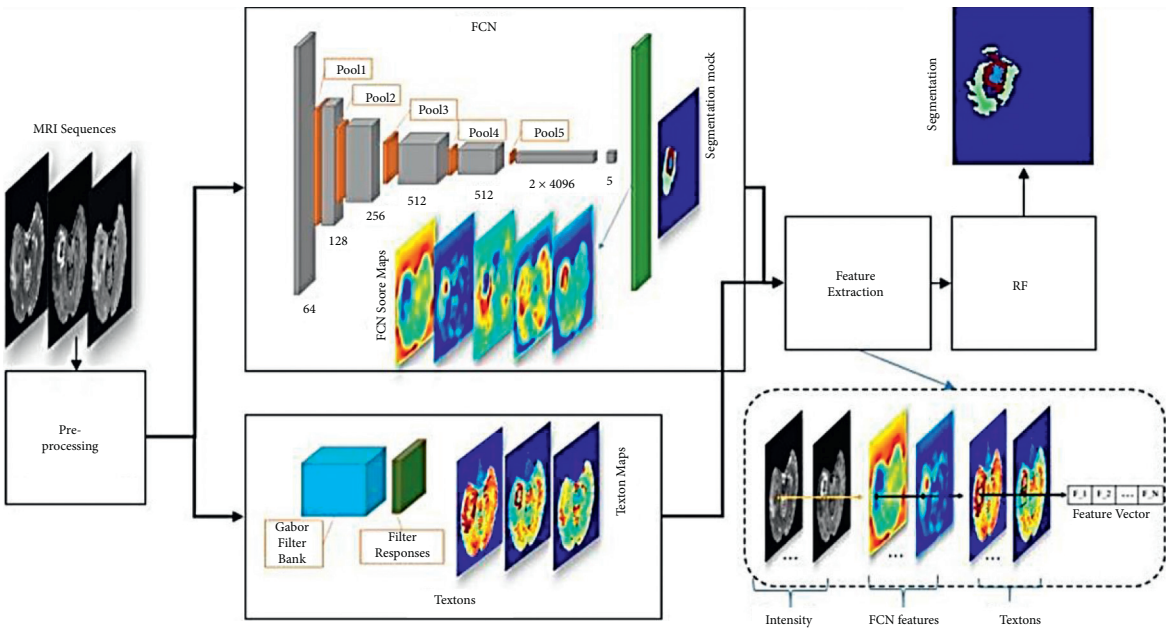


Figure 7: FCN architecture [87].

a module using discrete complexities for collecting multi-scale textual information. This architecture is based on open convolution, which supports the expansion of the exponential receiver field without losing sharpness or coverage, since the available convolution contains networked (or segmented) network artifacts in the input data.

Reference [98] created an approach called dilated residual networks (DRN) to eliminate these artifacts and further improve network efficiency.

4.6. Progress in the Main Pillar of the Network. The network's main column refers to the network's main structure. As it turns out from image classification tasks, the key source of image medical semantic segmentation is derived. The FCN [86] approved the VGG-16 net structure [99], which performed exceptionally well in ILSVRC14. Authors also considered the architecture of AlexNet [100], which won ILSVRC12, as well as GoogLeNet [101], and performed well in ILSVRC14. The VGG network has been validated in many previous studies [92, 94]. Following the release of ResNet or the deep residual network [102], DeepLab took its place in the ILSVRC 2015 classification work and implemented it, and semantic segmentation has made new progress. To reach a sufficient configuration, [103] evaluates the various changes of a fully complex residual network, including feature mapping resolution, number of layers, and field of view size. Also, [104] examines the remaining deep networks

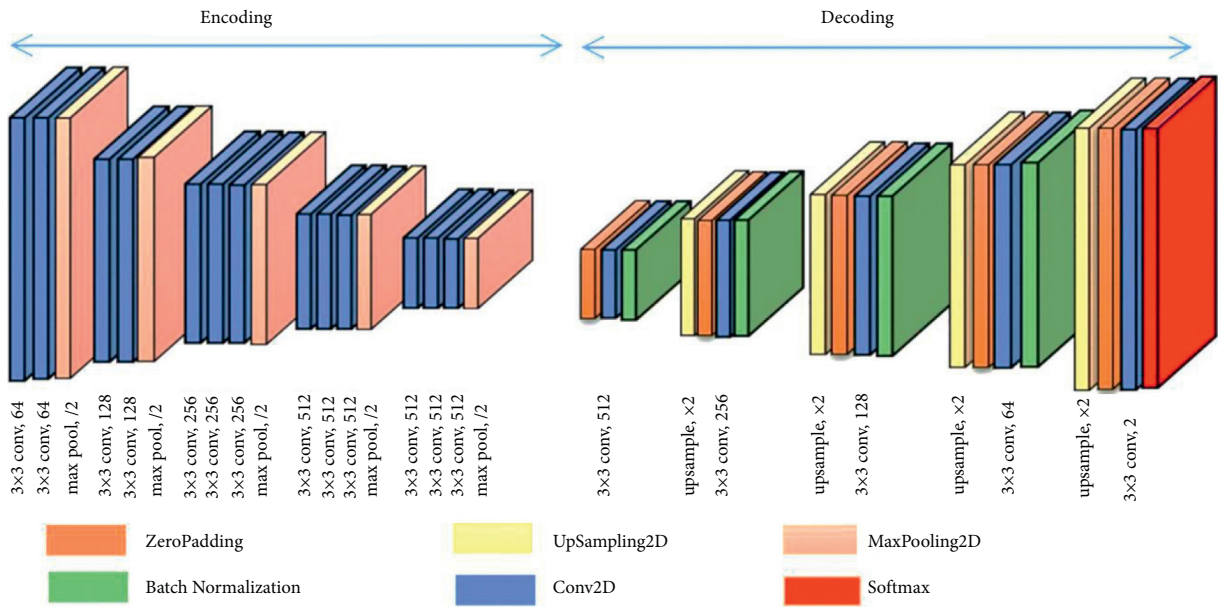


Figure 8: Convolution decomposition network architecture [91].

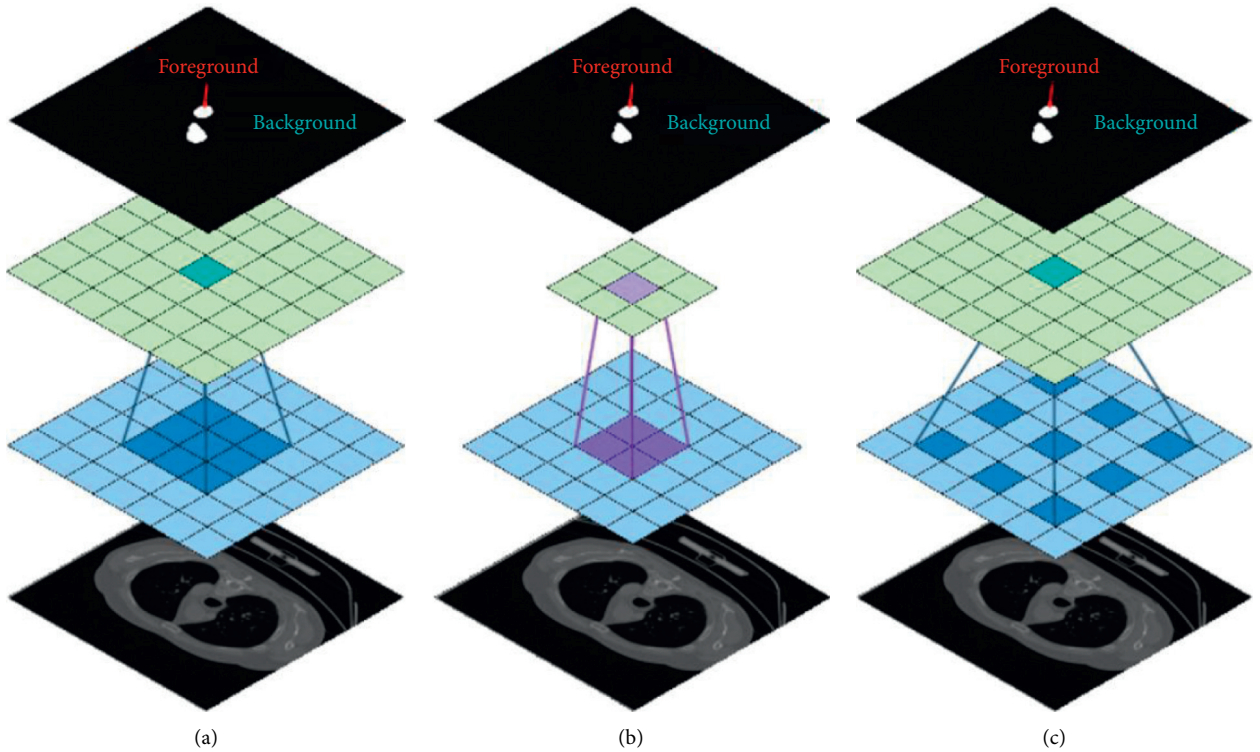


Figure 9: An example of an available convolution structure (Atrous convolution or hole convolution). (a) Convolution layer with 3 × 3 core size; a normal displacement operation with expansion parameter-1, (b) open convolution with expansion parameter-2, and (c) open convolution with expansion parameter-3 [95].

and explains some of the experimentally observed habits. As a result, authors get a shallow network architecture that in the ImageNet classification dataset is dramatically better than much deeper ones. Recently, ResNeXt [105] was introduced as the next generation of ResNet. This network is the basis for entering the ILSVRC 2016 classification work, in which it has won second place. GoogleNet also acquires extensions such as Inception-v2, Inception-v3 [106], Inception-v4, and Inception-ResNet [101], which has already been approved in the article [107].

5. Pyramid Methods in Semantic Segmentation

Aside from adopting networks with strong core columns, researchers are also trying to combine a pyramid strategy with CNN. A clear example of that is a pyramidal structure.

5.1. Image Pyramid. An image pyramid [108] is a series of sequentially segmented images before any of the desired stop criteria has been met. There are two different types of image pyramids: the Gaussian pyramid used for image sampling and the Laplacian pyramid used to recreate a scattered image from the lower image (with lower resolution). Three levels of the image pyramid can be seen in Figure 10.

In the field of image semantic segmentation, [110] establishes a network that can efficiently boost output with traditional multiscale image input and sliding pyramid mixing. This architecture captures the sense of the background patch. Similarly, by feeding input images of various sizes into a deep sharing network, DeepLab implements an image pyramid structure [111] that extracts multiscale functionality. The resulting features are combined for pixel classification at the end of each deep grid. You can see the picture pyramid used in the CNN system in Figure 11.

The Laplacian pyramid is also used to segment medical images semantically, and the reader can refer to the article in [113]. Authors have a multiresolution redevelopment architecture based on a Laplacian pyramid that utilizes higher-resolution map jump connections and a polygonal gate to change reconstructed boundaries with low-resolution function maps gradually. Reference [114] introduces a method of scene interpretation, and, through the Laplacian pyramid, the raw input image is transformed. In comparison, CNN creates a series of feature charts and generates each scale in two phases.

5.2. Atrous Spatial Pyramid Pooling (ASPP). Inspired by the image pyramid technique, [92] Atrous Spatial Pyramid Pooling (ASPP) is suggested to be done to provide robust object segmentation at different scales. ASPP explores the powerful fields-of-views (FOV) and the convolution feature layer with a multisampling rate filter and then captures the artifacts at various scales in the scene. ASPP architecture is seen in Figure 12.

5.3. Pyramid Pooling. According to the Pyramid Pooling shown in Figure 11, through gathering image data based on various regions, [115] exploits global knowledge capacity and calls its pyramid scene parsing network, known as PSPNet. The excellent results therein show that, with pyramidal pooling, a PSPNet brings a new mIoU score record of 85.4 percent in PASCAL VOC 2012 and brings 80.2 percent in Cityscapes dataset by experiments in [115]. Pyramid Pooling adopts multiple pooling size scales and applies the output to the original size for sampling processing. Finally, to shape a composite feature profile, it obtains the findings. Different scales of the size of the pool with different colors are marked in Figure 13. In general, pyramidal pooling can be used for any mapping of features. For example, the program in [115] applies pyramidal pooling in the pool5 layer.

5.4. Feature Pyramid. As mentioned in research backgrounds such as [117], the feature pyramid is an important component in image work for recognizing objects of various sizes. Object detectors have avoided displaying pyramids with recent DL methods because the computational volume and memory are compact. In [117], authors use the CNN multiscale pyramid hierarchy to build special pyramids at an additional cost. Also, a Feature Pyramid Network (FPN) has been created to construct high-level semantic maps at all scales. Machine learning also has many applications in the optimal selection of feature extraction [118–122].

6. Multilevel and Multistep Feature Methods

CNN may be known as a feature extractor [123], and, as a feature, CNN-based detection algorithms usually use the last-layer output. For dense forecasting, however, the data in this layer is too big. Instead, in localization, the primary layers can be correct but do not present the meaningful state. They describe hypercolumns as the activation vectors of all CNN units above that pixel to achieve both. You will see the form adopted as superstores in [123] in Figure 14. The FCN [86] has already approved jumps, as seen in Figure 5. The multilevel approach appears to have been used in the study [123], and multimodeling is a group approach to visualization ([125, 126]). In comparison to the multilevel approach, the multistage technique is utilized in semantic segmentation [107] to improve its accuracy and speed, which recommends the deep layer cascade (LC) method. The deep layer cascade (LC) method consists of multiple independent models, unlike the conventional model cascade (MC) ([125, 126]). As a multisubset model cascade, the LC system uses a single deep model, classifying several basic parts into the shallow stage and concentrating the deeper stage on several hard sections. This not only increases the productivity of segmentation but also accelerates both deep network training and research (Figure 12).

7. The Most Practical Deep Learning Methods in Medical Image Segmentation

The application of DL techniques to segmentation methods and MRI imaging aiming at brain tumors has been studied extensively. In general, different structures of a convolutional neural network can be acknowledged as the best DL technique in this research with studies on a convolutional neural network with a deeper layer [127], two-way convolutional neural network [128], cascaded CNN [129], multidimensional convolutional neural network [130], fully convolutional neural network (FCNN) for training with CRF [131], three-dimensional model of the convolutional neural network, two-dimensional model of the convolutional neural network [132], extreme learning machine (ELM) [133], Growing Deep Convolutional Network (GCNN) [134], complete convolutional neural network with

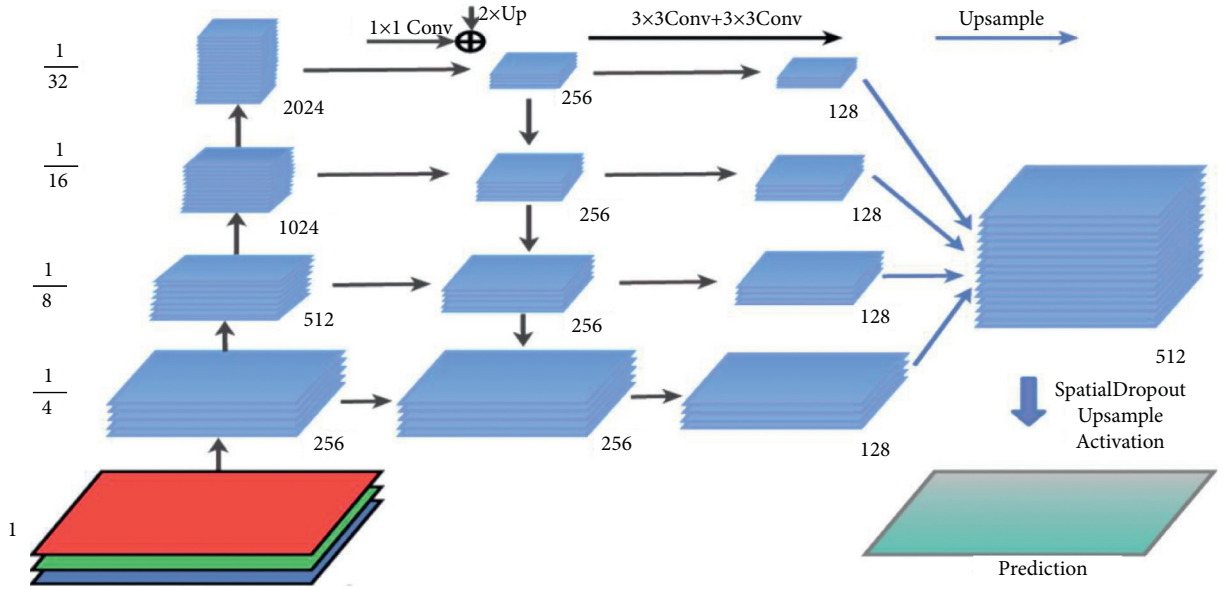


Figure 10: Three levels of the image pyramid [109].

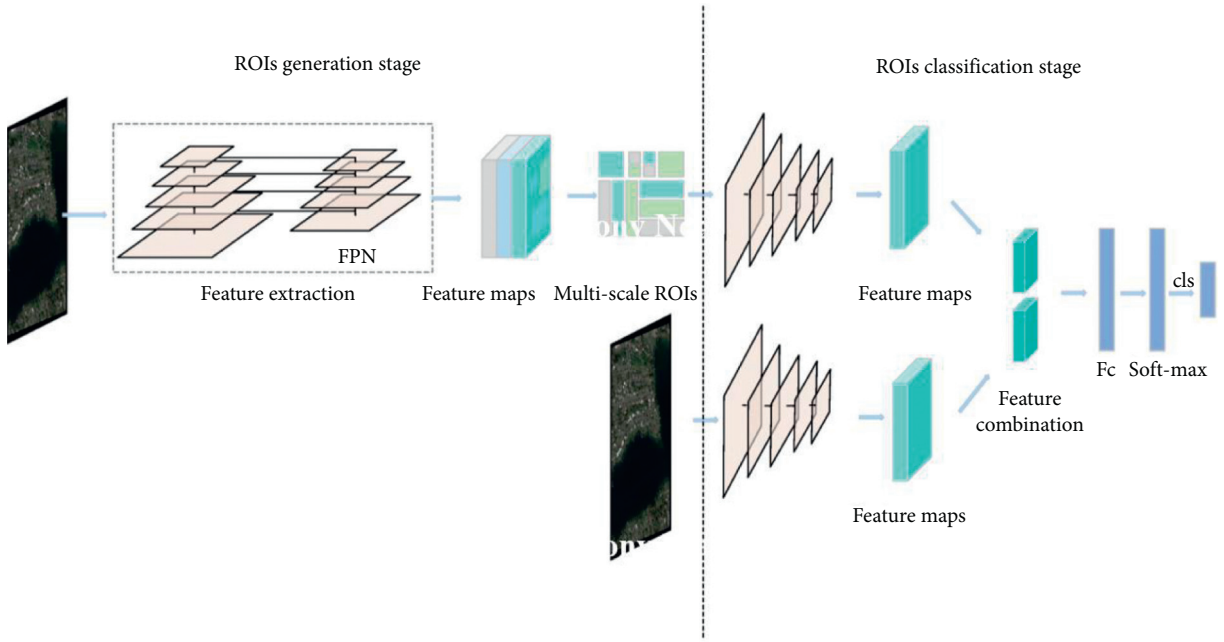


Figure 11: Image pyramid used in CNN structure [112].

Atheros convolution pyramid features [135], three-dimensional convolutional neural network test-time augmentation [136], and convolutional neural network referred to as CRF-based multicascade [137]. All the weaknesses along with the general application of all the advantages of the classical methods are presented in Table 2, and it is seen in the available and studied methods and also all of them have a wide range of applications. Yang et al. have used a portable evanescent wave sensor to detect SARS-CoV-2 using a CRISPR-based [138]. Reference [139] also uses DL methods to detect and classify breast tumors. Three different DL architectures, GoogLeNet, VGGNet, and ResNet, have been considered, and analysis has been performed between these methods. Visual detection and evaluation of breast tumors with DL principles are also presented in [140], which uses the combined methods of K-means and SURF algorithms in the structure of DL networks based on multiclass support vector machine. The detection of breast cancer using an extreme learning machine (ELM) based on feature fusion with deep convolutional neural network features is presented in [141]. Also, in [142], the extraction of a distinct pattern for the histopathological image classification of breast cancer has been done through an automated structure based on a convolutional neural network. All the weaknesses

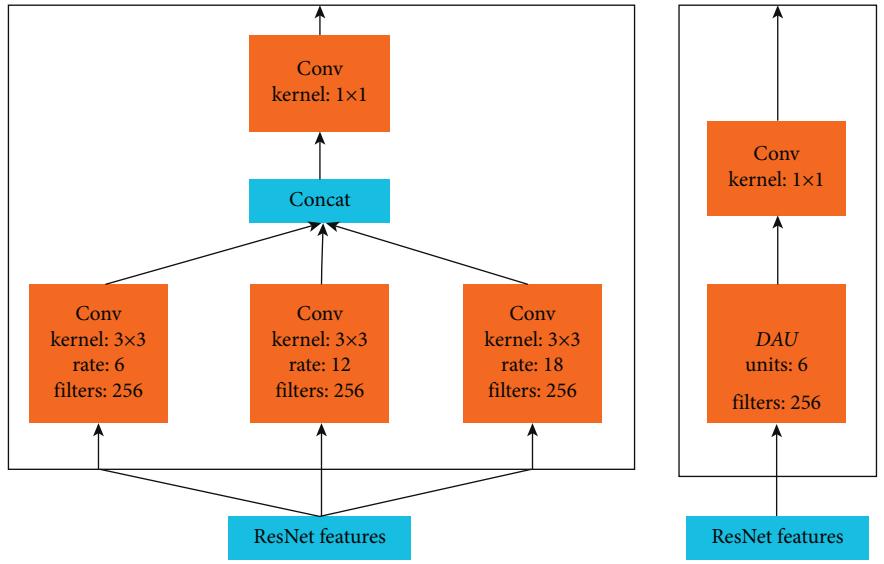


Figure 12: ASPP architecture (distance in convolution is not shown as the real rate in this image) [15].

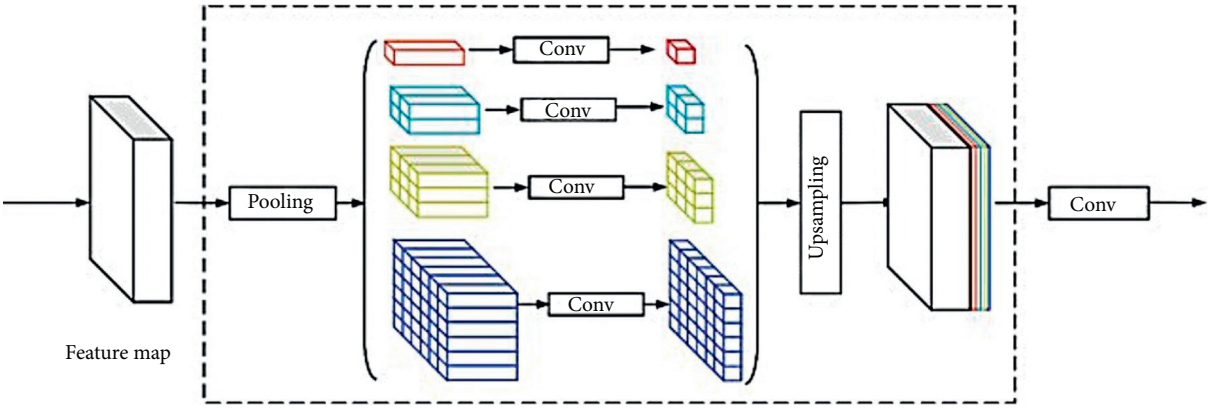


Figure 13: Demonstration of the structure of a pyramidal pooling [116].

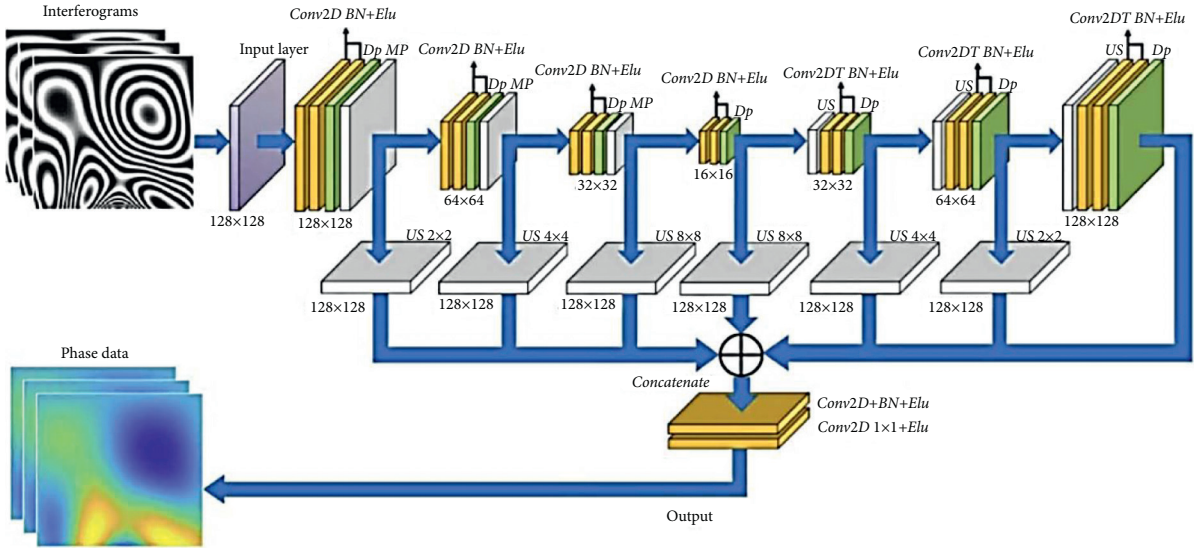


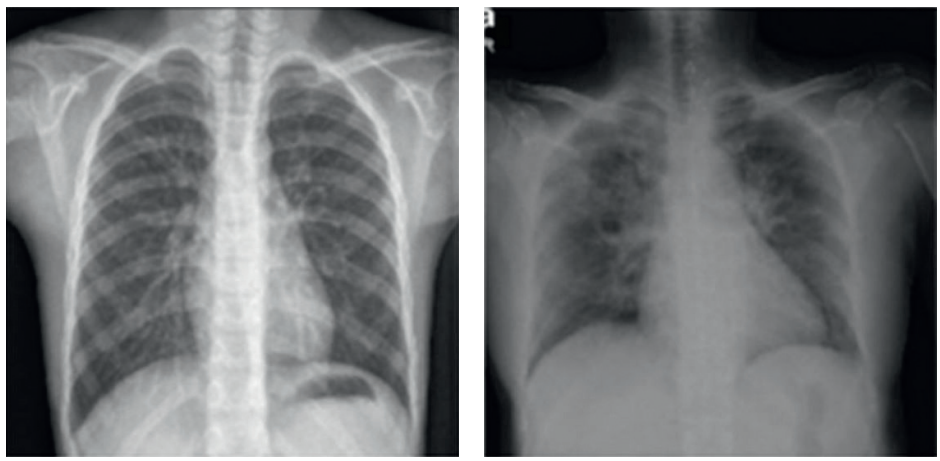
Figure 14: The structure adopted as hypercolumns [124].

along with the general application of all the advantages of the classical methods are presented in Table 3; and it is seen in the available and studied methods and all of them have a high range of applications. Figure 15 shows chest radiographs in healthy individuals and COVID-19 patients, respectively (see Table 4).

8. Discussion

A primary application in image processing and computer vision is image medical semantic segmentation. In addition to a brief overview of image semantic segmentation and traditional medicine, this article discusses recent advances in image semantic segmentation, particularly based on deep convolutional neural networks in the following aspects: (1) fully convolutional network, (2) method of sampling method, (3) combining FCN with CRF methods, (4) dilated convolution approaches, (5) progress in the main pillar of networks, (6) pyramid methods, and (7) multistage properties and multilevel methods. So far, there have been more and better ways to segment medical images semantically more accurately or faster or both with higher accuracy and speed, as well as better performance. Finally, the authors of this article hope that this review of recent advances in image medical semantic segmentation will help researchers in this field.

Maghdid et al. [21] reviewed a comprehensive, preprocessed dataset on X-rays and CT scan images from a variety of sources and provided an algorithm for accurate diagnosis of COVID-19 using DL and transmission learning tools. Also, a modified model was used by CNN and AlexNet as a pretested network on ready-made X-ray and CT scan datasets. After extensive experiments in both datasets, it has been shown that the proposed COVID-19 model predicts high accuracy and low response time. It is important to note that their proposed DL pattern has shown equivalent performance compared to that of a specialist radiologist. In addition, it can significantly improve the efficiency of radiologists while performing clinical practice [21]. Researchers are searching for new ways of screening, and the DL added to the chest X-rays of patients has shown positive outcomes. The computational cost of these approaches is still high considering their popularity, which causes difficulties with accessing them. Therefore the main purpose of this research is to accommodate the COVID-19 screening issue in chest X-ray with a reliable and successful approach in terms of memory and processing time. DL is a branch of artificial intelligence (AI) machine learning related to algorithms that are inspired as artificial neural networks by the structure and operation of the brain by using far higher-quality input images without any processing time. In addition, it is faster and cheaper to embed these versions in devices with more limited settings such as smartphones. To make use of embedded and large-scale devices, models can need little memory and carry out research quickly; and it encourages smartphones and emergency devices to work with them. DL models are complicated, so, to avoid inserting connections, a large number of things are necessary. In the training suite, for example, where the learning network performs well, to have less performance in the test suite, a large number of items are required. Unfortunately, there is not much data available for most real-world problems, even though the dataset is still small. Efficient training in deep neural networks has also been rendered possible through studying data transmission and amplification strategies in the small number of COVID-19-related images [144]. A popular approach for survival analysis and event prediction is the CPH model. This is therefore a semiparametric model, which suggests that the probability of misdiagnosis is a linear mixture of the clinical variables of the patient. In a fully data-driven way, the DL model can learn and infer high-order nonlinear interactions between clinical variables and disease effects. Data improvement techniques in DL will also make the model more robust to information noise and lost information, which usually happens in clinical datasets. It is also possible to expand the DL model to incorporate time-dependent factors such as vital signs and elevated visual attributes such as CT or X-ray images. It is inevitable to lose data on certain factors in reality and the real world. Data lost in less than three variables was then permitted in authors' online measurement tool, and risk evaluation based on DL methods can still be given by the field. In the clinical experience of Liang et al. [152], mild cases of COVID-19 are generally limited, and these are acute cases that need to be further investigated by physicians. Classified cases of their patients are clinically and economically expensive to manage COVID-19, especially due to the rapid outbreak of the disease which can happen and the high mortality rate related to acute disease, which has a high cost. By submitting clinical information online, medical personnel can use the predicted risk index to hospitalize patients and accordingly arrange patient treatment plans. In this way, medical resources can be appropriately allocated [152]. Arora et al. [26] suggested a DL model for estimating the number of patients who may have COVID-19 infection. They estimated the number of new cases of new coronaviruses in various states of the Indian Union for a span of one day to one week. For prediction, they utilized repetitive neural networks and Long Short-Term Memory (LSTM) and then tested several LSTM models in the Indian dataset and concluded that deeper LSTM models such as stacked LSTM, circular LSTM, and two-way LSTM were more accurate than simple LSTM models. To date, no research studies on COVID-19 cases have been reported from all Indian states, according to the authors. In one study, to predict the number of COVID-19-positive cases in the Indian states, Arora and colleagues suggested DL models. Because of the growing number of positive cases in India, exploratory data analysis has been undertaken. Depending on the number of cases and the daily growth rate, the government classifies states into mild, moderate, and severe areas to take strong action against the quarantine of the entire country, and this may cause economic and social problems. As predictive models, recurrent neural networks (RNN) are used based on long-term and short-



(a) (b)

Figure 15: (a) Radiographic images of a healthy person's chest. (b) Chest radiographs of patients with COVID-19 [143].

Table 4: DL methods used in COVID-19 detection and diagnosis.

Author	Purpose	Method	Advantages	Disadvantages	DL architecture	Results
Luz et al. [144]	To provide an accurate and efficient method for COVID-19 screening with chest X-rays for memory and processing time	Using EfficientNet artificial neural networks	High accuracy	Large and heterogeneous database	The use of algorithms, or artificial neural networks, is inspired by the brain's structure and function	From the hierarchical classification, experiments were performed to evaluate the performance of the neural network in the COVID-19 data. It became possible to use data transfer techniques and reinforce data. The accuracy was 93.9
Liu et al. [12]	Determining areas of infection and examining the lungs with the help of a chest CT scan	Use of CT scans to evaluate COVID-19, evaluation of system performance based on DL. This experiment was performed on 249 patients	Patient availability, high accuracy	Lack of sufficient information	Classification based on DL of VB-Net neural network	A DL system was developed for the segmentation and measurement of infection areas in CT scans of patients with COVID-19. The quantitative evaluation showed high accuracy for the infected area based on POI criteria
Li et al. [13]	Development of artificial intelligence CT imaging tools to diagnose coronavirus and isolate sick people away from healthy people	Uses powerful 2D and 3D DL models. Modifies and adapts existing AI models. This experiment was performed on 157 patients	High accuracy	Complexity	2D and 3D DL models were used	The AI-based analysis is rapidly evolving in the diagnosis of coronavirus, and the detection is being made with great accuracy

Table 4: Continued.

Author	Purpose	Method	Advantages	Disadvantages	DL architecture	Results
Wang and Wong [14]	Assist physicians in improving COVID-19 screening	Polymer reverse chain reaction screening from RT-PCR to diagnose COVID-19	Better understanding and character analysis by physicians, accelerating the development of high-precision DL solutions	Complexity, being time-consuming	Use of artificial intelligence systems based on DL, hardening and evaluation of COVID-Net prototype using Keras DL library with TensorFlow background	Assist physicians in improving screening, use of CXR images to diagnose COVID-19
Ghoshal and Tucker [145]	Evaluation of prop weights-based elliptic irritable neural networks (BCNN) to improve performance for COVID-19 diagnosis	Using the transmission learning method in COVID-19 X-ray images	Improvements in the diagnosis of COVID-19	Uncertainty in detection for radiologists	Use of DL for classified tasks, as well as chest radiographic diagnosis for COVID-19	Estimated uncertainty with DL can warn radiologists of incorrect predictions, which increases the use of DL in diagnosing the disease
Narin et al. [146]	Evaluation of the use of focal neural network-based methods to detect an infected patient using X-ray radiography of the chest	Use InceptionV3, ResNet50, and InceptionResNetV50 to diagnose infected patient	High-performance ResNet50 model (i) High accuracy (ii) Cost reduction	Ambiguity in matrices	Using COVID-19 X-ray images for DL models, using transition learning methods	Preliminary diagnosis of COVID-19 patients to prevent the spread of this disease in other people, using the ResNet50 model with 98% accuracy
Hemdon et al. [20]	The implementation of a DL framework for automated COVID-19 diagnosis of X-ray images	It was performed on 50 chest X-ray photographs of 25 positive COVID-19 cases. Seven distinct architectures from deep concealer neural network models are used in COVIDX-Net	Automatic diagnosis of COVID-19	Complexity	Use the COVID-19 classification to diagnose COVID-19 on X-ray images automatically using one of the DL frameworks	X-ray images based on the COVIDX-Net framework proposed
Zhang et al. [147]	Detection and differentiation of viral patient from the nonviral patient	Experiment on COVIDX data including 106 COVID-19 cases	Strengthen and improve the model, more efficiency in treatment	Complex calculations	Analysis of medical images including staging detection and drawing of pathological abnormalities, X-ray image change	Detection of the abnormality by viral pneumonia screening works well on chest X-ray images The learning model is useful for predicting job failure We have the CAAD model and we have never seen such cases in COVID-19, the data had 83.61% AUC, and the sensitivity was 70.71%

Table 4: Continued.

Author	Purpose	Method	Advantages	Disadvantages	DL architecture	Results
Maghdid et al. [21]	Provide artificial intelligence tools for fast and accurate detection of COVID-19, create a comprehensive set of X prototypes	Using X-rays and scanning CT images and using DL algorithms	Creating intelligent detection methods with higher efficiency (i) Increasing detection speed (ii) Increasing accuracy	Complexity	Build a DL-based detection system to detect COVID-19 pneumonia using DL algorithms	Accelerate the diagnosis of COVID-19 using the CNN model
Zeraati et al. [148]	Automatic classification of lung diseases including COVID-19 with X-ray images	Use of advanced convolutional neural network called MobileNet, use of 3905 X-ray images of more than 6 patients	Automatic diagnosis of COVID-19 from medical images (i) Low cost (ii) High speed	Limitations	Use DL to extract large-sized features from medical images	Low-cost, fast, and automatic diagnosis of COVID-19. Different infections may be differentiated by computer and detection using features extracted by DL
Li et al. [149]	Provide a fast and reliable way to diagnose COVID-19	Use 1020 CT images of 108 patients with COVID-19, use of ten confidential neural networks to diagnose COVID-19	Rapid diagnosis of COVID-19, being valid	High expenses	Use of CAD system based on DL to classify COVID-19 against other pneumonia	Using the textual CAD method on CT images to differentiate COVID-19 from other pneumococcal diseases. ResNet-101 can be used to diagnose COVID-19
Arora et al. [26]	Predict the number of new coronaviruses	Use LSTM-based RNN for forecasting	High accuracy of forecasting	High complexity and volume of data	Use the LSTM DL model	Two-way LSTM gives the best result and confidential LSTM gives the worst result. biLSTM gives very accurate results for short-term forecasts, such as 1 to 3 days, with less than 3% error
Huang et al. [150]	Quantitative evaluation of changes in lung tolerance in patients with COVID-19 using CT scan with an automated DL method	CT images show the entire lung and are measured and compared by commercial DL software	Classification of different groups and better detection	Loss of initial findings	Chest CT image evaluation using DL	The lung failure rate in COVID-19 was measured using a DL instrument based on a chest CT image and there was a significant difference between different groups

Table 4: Continued.

Author	Purpose	Method	Advantages	Disadvantages	DL architecture	Results
Oh et al. [151]	Use of the neural network to diagnose COVID-19	Inspired by CXR radiographic imaging	The usefulness of this method for the diagnosis of COVID-19 and patient triage	Difficulty in training deep neural network, difficulty in collecting big data	Use of X-ray chest images to classify COVID-19	Use artificial intelligence to improve CXR performance for detection
Liang et al. [152]	Use of a DL model to predict disease	Inclusion of 1590 patients from 575 medical centers, Use of DL models	Early detection	Complexity of calculations	Use of an integrated Cox model called Survival Cox DL	At least 60% of the data were used for prediction. A DL model was used to predict which was efficient

term memory (LSTM) cells. In 32 states/union states, LSTM types such as deep LSTM, circular LSTM, and bidirectional LSTM models have been tested, and the model is chosen with maximum accuracy based on absolute error. Based on estimation errors, the best outcome is the two-way LSTM, and the worst result is the hanging LSTM [26]. Ardakani et al. [149] suggested a DL-based CAD method to classify COVID-19 versus other pneumonic and abnormal pneumonia in research. They proposed that the DL method will assist radiologists to diagnose the disease associated with COVID-19, and they used ten convolutional neural networks (CNNs) to identify COVID-19-related diseases. Ten well-known CNNs were used in this analysis to provide a detailed view of the role of artificial intelligence in COVID-19 diagnosis. Data have shown that DL can distinguish COVID-19 with high accuracy from other pneumonia and viral diseases. For the ResNet-101 and Xeption networks, the best findings have been found. In the classification of COVID-19 and non-COVID-19 diseases, however, the Xeption network was most successful, but it did not have the highest sensitivity. In comparison, ResNet-101 was able to detect COVID-19 infection with the highest sensitivity and present fewer features compared to the Xeption network. In diagnosing patients with COVID-19, the trend is to incorporate a system with the greatest sensitivity. The advancement of DL programs helps researchers to do fast and deep X-ray scan analysis. DL is a mixture of methods of machine learning that focuses primarily on the automated extraction and classification of image characteristics, while its applications are commonly employed in medical work, medical detection, and classification. Machine learning and DL in the application of artificial intelligence for mining, pattern analysis from data, have been created as a discipline. To further evaluate the deep cognition method, Apostolopoulos et al. performed an experiment using six common lung diseases, including COVID-19. In this method, its capabilities in differentiating between different diseases are evaluated. Fine-tuning a deep network, in the context of DL, is a common approach for both learning the properties of depth and maintaining the method for extracting global properties, which exist in each image as different shapes. Specific research to detect potential trademarks focuses on X-ray images, and these biomarkers can be substantially correlated with COVID-19 disease. However, DL derives from images a large range of high-dimensional features, and some of these features may be known as real image markers. Li et al. studied the effect of self-assembly on fluorescent in magnetic fluid flow and its use for a new COVID-19 detection [149]. Recently, many studies have been done on various subjects about COVID-19, such as scheduling problems [153], climate change [154], sunspot assessment [155], disease severity and industry [156], energy after COVID-19 pandemic, travel-related risks among pandemics [157], and predictive modeling [148].

9. Limitations

The limitations of this research are mentioned in several aspects. First, the CT validation dataset is collected at one center, which may not represent all COVID-19 patients in other geographic areas. The generalization of the DL system must be approved in several centers. Second, the system is designed to determine the outbreak of the disease and may not be effective in measuring other pneumonia, such as bacterial pneumonia. Finally, in the next work, the authors will develop a system for quantifying the total intensity of pneumonia using transfer learning. Not ready for production, the researchers hope that the results obtained by COVID-Net in the COVIDX test dataset will be available as open source with descriptions of the open-source dataset. CXR images were used to accelerate the development of high-precision DL solutions for the diagnosis of COVID-19 patients. The future work will continue to include increasing accuracy and PPV for COVID-19 with the collection of new data, as well as the development of COVID-Net for risk classification for survival analysis, patient status prediction, and length of hospital stay.

10. Conclusion

In this study, the potential of deep learning methods in COVID-19 diagnosis is investigated. This study has reviewed the classification systems based on DL to assess the extent of the disease. This system not only automatically contours the

infected areas but also measures their shape, volume, and percentage of infection on a CT scan of patients with COVID-19. The methods involve radiologists to intervene effectively in the results of DL segmentation and repeatedly add more tutorials to update the model, thus accelerating the algorithm's development cycle. CT imaging has become an effective tool for screening patients with COVID-19 and for assessing COVID-19 levels. However, radiologists do not thave a computer tool to accurately determine the severity of COVID-19, for example, the percentage of infection in the lungs. DL has become a common method in medical image analysis and has been used in the analysis of lung diseases. Using this deep learning automated segmentation, many studies on imaging quantification and its association with syndromes, epidemiology, and therapeutic responses can provide further information on improving the diagnosis and treatment of COVID-19. An AI algorithm can be created quickly from one or more algorithms that do the same thing. This is in contrast to the standard method for generating a DL algorithm, which requires several steps. In order to review the data, expert annotations are needed at the data collection point at which a large number of samples need to be taken. The second is the process of training in which the data obtained is used to train network models. Every category should be well represented so that the training can be generalized during the test process to the new objects found by the network. A great number of network parameters (typically in the order of millions) are created automatically in this learning process. The third step is the experiment in which the network is presented with another collection of objects not included in the testing and the network performance is statistically evaluated to determine its classification. There is no solution that fits all; we hope that the positive results obtained by COVID-Net will be present in the COVIDX test dataset. Images are used to boost the advancement of highly accurate DL solutions for the diagnosis of COVID-19 patients and accelerate the treatment of patients. Future pathways including continuing to enhance sensitivity and PPV to COVID-19 disease by collecting new data as well as extending the suggested COVID-Net to risk classification for analysis, patient status prediction, and length of hospital stay will be useful. [158–164]

Data Availability

This is a review paper and data sharing is not applicable.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The funding sources had no involvement support in the study design, collection, analysis, or interpretation of data, in writing of the manuscript, or in the decision to submit the manuscript for publication.

References

[1] M. Thoma, "A survey of semantic segmentation," 2016, http://arxiv.org/abs/1602.06541.

[2] M. Rezaei, F. Farahanipad, A. Dillhoff, R. Elmasri, and V. Athitsos, "Weakly-supervised hand part segmentation from depth images," in *Proceedings of the 14th PErvasive Technologies Related to Assistive Environments Conference*, pp. 218–225, Corfu, Greece, June 2021.

[3] Y. Li, H. Qi, J. Dai, X. Ji, and Y. Wei, "Fully convolutional instance-aware semantic segmentation," in *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4438–4446, Honolulu, HI, USA, July 2017.

[4] S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gomez-Moreno, and F. Lopez-Ferreras, "Road-sign detection and recognition based on support vector machines," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 264–278, 2007.

[5] A. Cohen, E. Rivlin, I. Shimshoni, and E. Sabo, "Memory based active contour algorithm using pixel-level classified images for colon crypt segmentation," *Computerized Medical Imaging and Graphics*, vol. 43, pp. 150–164, 2015.

[6] C. Huang, L. S. Davis, and J. R. G. Townshend, "An assessment of support vector machines for land cover classification," *International Journal of Remote Sensing*, vol. 23, no. 4, pp. 725–749, 2002.

[7] N. Moon, E. Bullitt, K. van Leemput, and G. Gerig, "Automatic brain and tumor segmentation," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention-MICCAI 2002*, pp. 372–379, Springer Berlin Heidelberg, Tokyo, Japan, September 2002.

[8] G.-Q. Wei, K. Arbter, and G. Hirzinger, "Automatic tracking of laparoscopic instruments by color coding," in *Proceedings of the International Conference on Computer Vision, Virtual Reality, and Robotics in Medicine*, pp. 357–366, Springer Berlin Heidelberg, Grenoble, France, March 1997.

[9] H.-L. Chen, G. Wang, C. Ma, Z.-N. Cai, W.-B. Liu, and S.-J. Wang, "An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson's disease," *Neurocomputing*, vol. 184, pp. 131–144, 2016.

[10] W. Zhou, L. Yu, Y. Zhou, W. Qiu, M.-W. Wu, and T. Luo, "Local and global feature learning for blind quality evaluation of screen content and natural scene images," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2086–2095, 2018.

[11] Z. Lv, L. Qiao, A. K. Singh, and Q. Wang, "Fine-grained visual computing based on deep learning," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 17, no. 1s, pp. 1–19, 2021.

[12] R. Liu, X. Wang, H. Lu et al., "SCCGAN: style and characters inpainting based on CGAN," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 3–12, 2021.

[13] C. Li, L. Hou, B. Y. Sharma et al., "Developing a new intelligent system for the diagnosis of tuberculous pleural effusion," *Computer Methods and Programs in Biomedicine*, vol. 153, pp. 211–225, 2018.

[14] L. Wang and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," 2020, http://arxiv.org/abs/2003.09871.

[15] M. Zhang, Y. Chen, and W. Susilo, "PPO-CPQ: a privacy-preserving optimization of clinical pathway query for e-healthcare systems," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10660–10672, 2020.

[16] A. Ala, F. E. Alsaadi, M. Ahmadi, and S. Mirjalili, "Optimization of an appointment scheduling problem for healthcare systems based on the quality of fairness service using whale optimization algorithm and NSGA-II," *Scientific Reports*, vol. 11, no. 1, pp. 19816–19819, 2021.

[17] Q. Xu, Q. Guo, C.-X. Wang et al., "Network differentiation: a computational method of pathogenesis diagnosis in traditional Chinese medicine based on systems science," *Artificial Intelligence in Medicine*, vol. 118, Article ID 102134, 2021.

[18] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Kai Li, and L. Fei-Fei, "ImageNet: a large-scale hierarchical image database," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, Miami, FL, USA, June 2009.

[19] M. Ahmadi, T. Ali, D. Javaheri, A. Masoumian, S. Jafarzadeh Ghoushchi, and Y. Pourasad, "DQRE-SCnet: a novel hybrid approach for selecting users in federated learning with deep-Q-reinforcement learning based on spectral clustering," *Journal of King Saud University-Computer and Information Sciences*, 2021.

[20] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images," 2020, http://arxiv.org/abs/2003.11055.

[21] H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, and M. K. Khan, "Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms," 2020, http://arxiv.org/abs/2004.00038.

[22] D. Martin and C. Fowlkes, *The Berkeley Segmentation Database and Benchmark*, University of California, Berkeley, CA, USA, 2001, http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds.

[23] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, "LabelMe: a database and web-based tool for image annotation," *International Journal of Computer Vision*, vol. 77, no. 1-3, pp. 157–173, 2008.

[24] S. K. Hasan and C. A. Linte, "A modified U-Net convolutional network featuring a nearest-neighbor re-sampling-based elastic-transformation for brain tissue characterization and segmentation," in *Proceedings of the 2018 IEEE Western New York Image and Signal Processing Workshop (WNYISPW)*, pp. 1–5, IEEE, Rochester, NY, USA, October 2018.

[25] S. Charan, M. J. Khan, and K. Khurshid, "Breast cancer detection in mammograms using convolutional neural network," in *Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1–5, IEEE, Sukkur, Sindh, Pakistan, March 2018.

[26] P. Arora, H. Kumar, and B. K. Panigrahi, "Prediction and analysis of COVID-19 positive cases using deep learning models: a descriptive case study of India," *Chaos, Solitons & Fractals*, vol. 139, Article ID 110017, 2020.

[27] J. Xia, H. Chen, Q. Li et al., "Ultrasound-based differentiation of malignant and benign thyroid Nodules: an extreme learning machine approach," *Computer Methods and Programs in Biomedicine*, vol. 147, pp. 37–49, 2017.

[28] L. Bourdev, S. Maji, T. Brox, and J. Malik, "Detecting people using mutually consistent poselet activations," in *Proceedings of the Computer Vision-ECCV 2010*, pp. 168–181, Springer Berlin Heidelberg, Heraklion, Crete, Greece, September 2010.

[29] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[30] D. C. He and L. Wang, "Texture unit, texture spectrum, and texture analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 4, pp. 509–512, 1990.

[31] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[32] K. G. Derpanis, *The Harris Corner Detector*, pp. 1-2, York University, Toronto, Canada, 2004.

[33] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 593–600, Seattle, WA, USA, June 1994.

[34] G. Medioni and Y. Yasumoto, "Corner detection and curve representation using cubic B-splines," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 267–278, 1987.

[35] S. M. Smith and J. M. Brady, "Susan-a new approach to low level image processing," *International Journal of Computer Vision*, vol. 23, no. 1, pp. 45–78, 1997.

[36] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, vol. II, pp. 1508–1515, 2005.

[37] E. Rosten, R. Porter, and T. Drummond, "Faster and better: a machine learning approach to corner detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, 2010.

[38] E. Mair, G. D. Hager, D. Burschka, M. Suppa, and G. Hirzinger, "Adaptive and generic corner detection based on the accelerated segment test," in *Proceedings of the Computer Vision-ECCV 2010*, pp. 183–196, Springer Berlin Heidelberg, Heraklion, Crete, Greece, September 2010.

[39] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: binary Robust invariant scalable keypoints," in *Proceedings of the 2011 International Conference on Computer Vision*, pp. 2548–2555, Barcelona, Spain, November 2011.

[40] C. S. Venegas-Barrera and J. Manjarrez, "Visual categorization with bags of keypoints," *Revista Mexicana de Biodiversidad*, vol. 82, no. 1, pp. 179–191, 2011.

[41] T. Brox, L. Bourdev, S. Maji, and J. Malik, "Object segmentation by alignment of poselet activations to image contours," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2225–2232, Colorado Springs, CO, USA, June 2011.

[42] S.-C. Zhu, C.-E. Guo, Y. Wang, and Z. Xu, "What are Textons?" *International Journal of Computer Vision*, vol. 62, no. 1, pp. 121–143, 2005.

[43] F. Liu, G. Zhang, and J. Lu, "Multi-source heterogeneous unsupervised domain adaptation via fuzzy-relation neural networks," *IEEE Transactions on Fuzzy Systems*, vol. 29, 2020.

[44] Q. Xu, Y. Zeng, W. Tang et al., "Multi-task joint learning model for segmenting and classifying tongue images using a deep neural network," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, pp. 2481–2489, 2020.

[45] H. Liu, J. Liu, S. Hou, T. Tao, and J. Han, "Perception consistency ultrasound image super-resolution via self-supervised CycleGAN," *Neural Computing and Applications*, pp. 1–11, 2021.

[46] X. Zhao, X. Zhang, Z. Cai et al., "Chaos enhanced grey wolf optimization wrapped ELM for diagnosis of paraquat-

poisoned patients," *Computational Biology and Chemistry*, vol. 78, pp. 481–490, 2019.

[47] D. J. Strauss and J. A. Hartigan, "Clustering algorithms," *Biometrics*, vol. 31, no. 3, p. 793, 1975.

[48] H. Mobahi, S. R. Rao, A. Y. Yang, S. S. Sastry, and Y. Ma, "Segmentation of natural images by texture and boundary compression," *International Journal of Computer Vision*, vol. 95, no. 1, pp. 86–98, 2011.

[49] R. Kimmel and A. M. Bruckstein, "Regularized Laplacian zero crossings as optimal edge integrators," *International Journal of Computer Vision*, vol. 53, no. 3, pp. 225–243, 2003.

[50] H. Yu, W. Li, C. Chen et al., "Dynamic gaussian bare-bones fruit fly optimizers with abandonment mechanism: method and analysis," *Engineering with Computers*, 2020.

[51] X. Zhao, D. Li, B. Yang, C. Ma, Y. Zhu, and H. Chen, "Feature selection based on improved ant colony optimization for online detection of foreign fiber in cotton," *Applied Soft Computing*, vol. 24, pp. 585–596, 2014.

[52] C. Yu, M. Chen, K. Cheng et al., "SGOA: annealing-behaved grasshopper optimizer for global tasks," *Engineering with Computers*, pp. 1–28, 2021.

[53] T. Lindeberg and M.-X. Li, "Segmentation and classification of edges using minimum description length Approximation and complementary junction cues," *Computer Vision and Image Understanding*, vol. 67, no. 1, pp. 88–98, 1997.

[54] L. Shen, H. Chen, Z. Yu et al., "Evolving support vector machines using fruit fly optimization for medical data classification," *Knowledge-Based Systems*, vol. 96, pp. 61–75, 2016.

[55] W. Shan, Z. Qiao, A. Asghar Heidari, H. Chen, H. Turabieh, and Y. Teng, "Double adaptive weights for stabilization of moth flame optimizer: balance analysis, engineering cases, and medical diagnosis," *Knowledge-Based Systems*, vol. 214, Article ID 106728, 2020.

[56] J. Shotton, J. Winn, C. Rother, and A. Criminisi, "TextonBoost: joint appearance, shape and context modeling for multi-class object recognition and segmentation," in *Proceedings of the Computer Vision-ECCV 2006*, pp. 1–15, Springer Berlin Heidelberg, Graz, Austria, May 2006.

[57] A. M. N. Darwiesh, N. M. A.-E. Maboud, A. M. R. Khalil, and A. M. ElSharkawy, "Role of magnetic resonance spectroscopy & diffusion weighted imaging in differentiation of supratentorial brain tumors," *The Egyptian Journal of Radiology and Nuclear Medicine*, vol. 47, no. 3, pp. 1037–1042, 2016.

[58] A. Aslam, E. Khan, and M. M. S. Beg, "Improved edge detection algorithm for brain tumor segmentation," *Procedia Computer Science*, vol. 58, pp. 430–437, 2015.

[59] W. Qiao, M. Khishe, and S. Ravakhah, "Underwater targets classification using local wavelet acoustic pattern and Multilayer perceptron neural network optimized by modified whale optimization algorithm," *Ocean Engineering*, vol. 219, Article ID 108415, 2021.

[60] Q. Ain, M. A. Jaffar, and T.-S. Choi, "Fuzzy anisotropic diffusion based segmentation and texture based ensemble classification of brain tumor," *Applied Soft Computing*, vol. 21, pp. 330–340, 2014.

[61] S. Hassantabar, M. Ahmadi, and A. Sharifi, "Diagnosis and detection of infected tissue of COVID-19 patients based on lung X-ray image using convolutional neural network approaches," *Chaos, Solitons & Fractals*, vol. 140, Article ID 110170, 2020.

[62] S. Dorosti, S. Jafarzadeh Ghoushchi, E. Sobhrakhshankhah, M. Ahmadi, and A. Sharifi, "Application of gene expression programming and sensitivity analyses in analyzing effective parameters in gastric cancer tumor size and location," *Soft Computing*, vol. 24, no. 13, pp. 9943–9964, 2020.

[63] A. Sharifi, M. Ahmadi, M. A. Mehni, S. Jafarzadeh Ghoushchi, and Y. Pourasad, "Experimental and numerical diagnosis of fatigue foot using convolutional neural network," *Computer Methods in Biomechanics and Biomedical Engineering*, pp. 1–13, 2021.

[64] I. Laradji, P. Rodriguez, O. Manas et al., "A weakly supervised consistency-based learning method for covid-19 segmentation in ct images," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2453–2462, Waikoloa, HI, USA, 2021.

[65] M. Karnan and K. Selvanayaki, "Improved implementation of brain MR image segmentation using meta heuristic algorithms," in *Proceedings of the 2010 IEEE International Conference on Computational Intelligence and Computing Research*, pp. 702–705, Coimbatore, India, December 2010.

[66] P. Ghosh, K. Mali, and S. K. Das, "Chaotic firefly algorithm-based fuzzy C-means algorithm for segmentation of brain tissues in magnetic resonance images," *Journal of Visual Communication and Image Representation*, vol. 54, pp. 63–79, 2018.

[67] Q. Zhu, B. Du, P. Yan, H. Lu, and L. Zhang, "Shape prior constrained PSO model for bladder wall MRI segmentation," *Neurocomputing*, vol. 294, pp. 19–28, 2018.

[68] S. Alagarsamy, K. Kamatchi, V. Govindaraj, Y.-D. Zhang, and A. Thiyagarajan, "Multi-channeled MR brain image segmentation: a new automated approach combining BAT and clustering technique for better identification of heterogeneous tumors," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 4, pp. 1005–1035, 2019.

[69] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation," *Expert Systems with Applications*, vol. 42, no. 3, pp. 990–1002, 2015.

[70] S. Kaymak, A. Helwan, and D. Uzun, "Breast cancer image classification using artificial neural networks," *Procedia Computer Science*, vol. 120, pp. 126–131, 2017.

[71] M. Karabatak, "A new classifier for breast cancer detection based on Naïve Bayesian," *Measurement*, vol. 72, pp. 32–36, 2015.

[72] F. Wang, S. Zhang, and L. M. Henderson, "Adaptive decision-making of breast cancer mammography screening: a heuristic-based regression model," *Omega*, vol. 76, pp. 70–84, 2018.

[73] D. C. Pereira, R. P. Ramos, and M. Z. do Nascimento, "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm," *Computer Methods and Programs in Biomedicine*, vol. 114, no. 1, pp. 88–101, 2014.

[74] L. Xu, S. Jiang, J. Wu, and Q. Zou, "An in silico approach to identification, categorization and prediction of nucleic acid binding proteins," *Briefings in Bioinformatics*, vol. 22, no. 3, Article ID bbaa171, 2021.

[75] F. R. Cordeiro, W. P. Santos, and A. G. Silva-Filho, "An adaptive semi-supervised fuzzy growcut algorithm to segment masses of regions of interest of mammographic images," *Applied Soft Computing*, vol. 46, pp. 613–628, 2016.

[76] L. Ahmed, M. M. Iqbal, H. Aldabbas, S. Khalid, Y. Saleem, and S. Saeed, "Images data practices for semantic segmentation of breast cancer using deep neural network," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–7, 2020.

[77] H. Lee, J. Park, and J. Y. Hwang, "Channel attention module with multiscale grid average pooling for breast cancer segmentation in an ultrasound image," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 7, pp. 1344–1353, 2020.

[78] K. B. Soulami, N. Kaabouch, M. N. Saidi, and A. Tamtaoui, "Breast cancer: one-stage automated detection, segmentation, and classification of digital mammograms using UNet model based-semantic segmentation," *Biomedical Signal Processing and Control*, vol. 66, Article ID 102481, 2021.

[79] K. Huang, Y. Zhang, H. D. Cheng, P. Xing, and B. Zhang, "Semantic segmentation of breast ultrasound image with fuzzy deep learning network and breast anatomy constraints," *Neurocomputing*, vol. 450, pp. 319–335, 2021.

[80] R. M. S. d. Oliveira, R. C. F. Araújo, F. J. B. Barros et al., "A system based on artificial neural networks for automatic classification of hydro-generator stator windings partial discharges," *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, vol. 16, no. 3, pp. 628–645, 2017.

[81] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009.

[82] H. Larochelle and Y. Bengio, "Classification using discriminative restricted Boltzmann machines," in *Proceedings of the 25th International Conference on Machine Learning-ICML'08*, pp. 536–543, Helsinki, Finland, July 2008.

[83] M. Ahmadi, A. Sharifi, and S. Khalili, "Presentation of a developed sub-epidemic model for estimation of the COVID-19 pandemic and assessment of travel-related risks in Iran," *Environmental Science and Pollution Research*, vol. 28, no. 12, pp. 14521–14529, 2021.

[84] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[85] J. L. Olazagoitia, J. A. Perez, and F. Badea, "Identification of tire model parameters with artificial neural networks," *Applied Sciences*, vol. 10, no. 24, p. 9110, 2020.

[86] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2017.

[87] M. Soltaninejad, L. Zhang, T. Lambrou, N. Allinson, and X. Ye, "Multimodal MRI brain tumor segmentation using random forests with features learned from fully convolutional neural network," 2017, http://arxiv.org/abs/1704.08134.

[88] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1520–1528, Santiago, Chile, December 2015.

[89] R. Mohan, "Deep deconvolutional networks for scene parsing," 2014, http://arxiv.org/abs/1411.4101.

[90] J. B. De Monvel, E. Scarfone, S. Le Calvez, and M. Ulfendahl, "Image-adaptive deconvolution for three-dimensional deep biological imaging," *Biophysical Journal*, vol. 85, no. 6, pp. 3991–4001, 2003.

[91] N. Alipour and A. Behrad, "Semantic segmentation of JPEG blocks using a deep CNN for non-aligned JPEG forgery detection and localization," *Multimedia Tools and Applications*, vol. 79, no. 11-12, pp. 8249–8265, 2020.

[92] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[93] M. Mostajabi, P. Yadollahpour, and G. Shakhnarovich, "Feedforward semantic segmentation with zoom-out features," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3376–3385, Boston, MA, USA, June 2015.

[94] Z. Liu, X. Li, P. Luo, C.-C. Loy, and X. Tang, "Semantic image segmentation via deep parsing network," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1377–1385, Santiago, Chile, December 2015.

[95] X.-Y. Zhou, J.-Q. Zheng, and G.-Z. Yang, "Atrous convolutional neural network (ACNN) for biomedical semantic segmentation with dimensionally lossless feature maps," 2019, http://arxiv.org/abs/1901.09203.

[96] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2016, http://arxiv.org/abs/1511.07122.

[97] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Inverse Problems and Theoretical Imaging*, pp. 286–297, Springer Berlin Heidelberg, Berlin, Germnay, 1990.

[98] F. Yu, V. Koltun, and T. Funkhouser, "Dilated residual networks," in *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*, pp. 636–644, Honolulu, HI, USA, July 2017.

[99] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015, http://arxiv.org/abs/1409.1556.

[100] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[101] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proceedings of the 2017 Thirty-First AAAI Conference on Artificial Intelligence*, pp. 4278–4284, San Francisco, CA, USA, February 2017.

[102] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[103] Z. Wu, C. Shen, and A. van den Hengel, "High-performance semantic segmentation using very deep fully convolutional networks," 2016, http://arxiv.org/abs/1604.04339.

[104] Z. Wu, C. Shen, and A. van den Hengel, "Wider or deeper: revisiting the ResNet model for visual recognition," *Pattern Recognition*, vol. 90, pp. 119–133, 2019.

[105] S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5987–5995, Honolulu, HI, USA, July 2017.

[106] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, Las Vegas, NV, USA, June 2016.

[107] X. Li, Z. Liu, P. Luo, C. C. Loy, and X. Tang, "Not all pixels are equal: difficulty-aware semantic segmentation via deep layer cascade," in *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6459–6468, Honolulu, HI, USA, July 2017.

[108] E. H. Adelson and P. J. Burt, "Multi-resolution splining using A pyramid image representation," in *Proceedings of the 27th*

Annual Technical Symposium, pp. 204–210, San Diego, CA, USA, 1983.

[109] S. Seferbekov, V. Iglovikov, B. Alexander, and Alexey Shvets, "Feature pyramid network for multi-class land segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 272–275, Salt Lake City, UT, USA, June 2018.

[110] G. Lin, C. Shen, A. Van Den Hengel, and I. Reid, "Efficient piecewise training of deep structured models for semantic segmentation," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3194–3203, Las Vegas, NV, USA, June 2016.

[111] L.-C. Chen, Y. Yang, J. Wang, W. Xu, and A. L. Yuille, "Attention to scale: scale-aware semantic image segmentation," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3640–3649, Las Vegas, NV, USA, June 2016.

[112] C. Chen, W. Gong, Y. Chen, and W. Li, "Object detection in remote sensing images based on a scene-contextual feature pyramid network," *Remote Sensing*, vol. 11, no. 3, p. 339, 2019.

[113] G. Ghiasi and C. C. Fowlkes, "Laplacian pyramid reconstruction and refinement for semantic segmentation," in *Proceedings of the Computer Vision-ECCV 2016*, pp. 519–534, Amsterdam, The Netherlands, October 2016.

[114] C. Farabet, C. Couprie, L. Najman, and Y. Lecun, "Learning hierarchical features for scene labeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1915–1929, 2013.

[115] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6230–6239, Honolulu, HI, USA, July 2017.

[116] J. Li, C. Wang, L. Xu, F. Wu, H. Zhang, and B. Zhang, "Multitemporal water extraction of dongting lake and poyang lake based on an automatic water extraction and dynamic monitoring framework," *Remote Sensing*, vol. 13, no. 5, p. 865, 2021.

[117] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*, pp. 936–944, Honolulu, HI, USA, July 2017.

[118] Y. Xu, H. Chen, J. Luo, Q. Zhang, S. Jiao, and X. Zhang, "Enhanced Moth-flame optimizer with mutation strategy for global optimization," *Information Sciences*, vol. 492, pp. 181–203, 2019.

[119] X. Xu and H.-L. Chen, "Adaptive computational chemotaxis based on field in bacterial foraging optimization," *Soft Computing*, vol. 18, no. 4, pp. 797–807, 2014.

[120] Y. Zhang, R. Liu, X. Wang, H. Chen, and C. Li, "Boosted binary Harris hawks optimizer and feature selection," *Engineering with Computers*, 2020.

[121] Y. Zhang, Y. Zhang, R. Liu et al., "Towards augmented kernel extreme learning models for bankruptcy prediction: algorithmic behavior and comprehensive analysis," *Neurocomputing*, vol. 430, 2020.

[122] J. Hu, H. Chen, A. A. Heidari et al., "Orthogonal learning covariance matrix for defects of grey wolf optimizer: insights, balance, diversity, and feature selection," *Knowledge-Based Systems*, vol. 213, Article ID 106684, 2021.

[123] B. Hariharan, P. Arbelaez, R. Girshick, and J. Malik, "Hypercolumns for object segmentation and fine-grained localization," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 447–456, Boston, MA, USA, June 2015.

[124] Z. Zhao, B. Li, J. Lu, X. Kang, and T. Liu, "One-shot phase retrieval method for interferometry using a hypercolumns convolutional neural network," *Optics Express*, vol. 29, no. 11, pp. 16406–16421, 2021.

[125] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5325–5334, Boston, MA, USA, June 2015.

[126] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, Kauai, HI, USA, December 2001.

[127] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.

[128] G. Urban, M. Bendszus, F. Hamprecht, and J. Kleesiek, "Multi-modal brain tumor segmentation using deep convolutional neural networks," in *Proceedings of the Miccai-Brats*, pp. 31–35, Boston, MA, USA, 2014.

[129] M. Havaei, A. Davy, D. Warde-Farley et al., "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017.

[130] K. Hu, Z. Zhang, X. Niu et al., "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function," *Neurocomputing*, vol. 309, pp. 179–191, 2018.

[131] K. Kamnitsas, C. Ledig, V. F. J. Newcombe et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017.

[132] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks," in *Proceedings of the International MICCAI Brainlesion Workshop*, pp. 178–190, Quebec City, Canada, 2017.

[133] L. Hu, G. Hong, J. Ma, X. Wang, and H. Chen, "An efficient machine learning approach for diagnosis of paraquat-poisoned patients," *Computers in Biology and Medicine*, vol. 59, pp. 116–124, 2015.

[134] M. Mittal, L. M. Goyal, S. Kaur, I. Kaur, A. Verma, and D. Jude Hemanth, "Deep learning based enhanced tumor segmentation approach for MR brain images," *Applied Soft Computing*, vol. 78, pp. 346–354, 2019.

[135] Z. Zhou, Z. He, and Y. Jia, "AFPNet: a 3D fully convolutional neural network with atrous-convolution feature pyramid for brain tumor segmentation via MRI images," *Neurocomputing*, vol. 402, pp. 235–244, 2020.

[136] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks," in *Proceedings of the Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 178–190, Athens, Greece, October 2018.

[137] K. Hu, Q. Gan, Y. Zhang et al., "Brain tumor segmentation using multi-cascaded convolutional neural networks and conditional random field," *IEEE Access*, vol. 7, pp. 92615–92629, 2019.

[138] M. Kistler, S. Bonaretti, M. Pfahrer, R. Niklaus, and P. Büchler, "The virtual skeleton database: an open access repository for biomedical research and collaboration,"

Journal of Medical Internet Research, vol. 15, no. 11, p. e245, 2013.

[139] S. Khan, N. Islam, Z. Jan, I. Ud Din, and J. J. P. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognition Letters*, vol. 125, pp. 1–6, 2019.

[140] P. Kaur, G. Singh, and P. Kaur, "Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification," *Informatics in Medicine Unlocked*, vol. 16, p. 100151, 2019.

[141] Z. Wang, M. Li, H. Wang et al., "Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features," *IEEE Access*, vol. 7, pp. 105146–105158, 2019.

[142] Y. Yang, J. Liu, and X. Zhou, "A CRISPR-based and post-amplification coupled SARS-CoV-2 detection with a portable evanescent wave biosensor," *Biosensors and Bioelectronics*, vol. 190, Article ID 113418, 2021.

[143] J. Manokaran, F. Zabihollahy, A. Hamilton-Wright, and E. Ukwatta, "Deep learning-based detection of COVID-19 from chest X-ray images," in *Proceedings of the Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging*, Houston, TX, USA, February 2021.

[144] E. Luz, P. L. Silva, R. Silva et al., "Towards an effective and efficient deep learning model for covid-19 patterns detection in X-ray images," *Research on Biomedical Engineering*, vol. 7, pp. 1–14, 2021.

[145] B. Ghoshal and A. Tucker, "Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection," 2020, http://arxiv.org/abs/2003.10769.

[146] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," 2020, http://arxiv.org/abs/2003.10849.

[147] J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "COVID-19 screening on chest X-ray images using deep learning based anomaly detection," 2020, http://arxiv.org/abs/2003.12338.

[148] M. Zeraati, R. Pourmohamad, B. Baghchi, N. P. Singh Chauhan, and G. Sargazi, "Optimization and predictive modelling for the diameter of nylon-6,6 nanofibers via electrospinning for coronavirus face masks," *Journal of Saudi Chemical Society*, vol. 25, no. 11, Article ID 101348, 2021.

[149] X. Li, Z.-Q. Dong, P. Yu et al., "Effect of self-assembly on fluorescence in magnetic multiphase flows and its application on the novel detection for COVID-19," *Physics of Fluids*, vol. 33, no. 4, p. 042004, 2021.

[150] L. Huang, R. Han, T. Ai et al., "Serial quantitative chest CT assessment of COVID-19: a deep learning approach," *Radiology: Cardiothoracic Imaging*, vol. 2, no. 2, Article ID e200075, 2020.

[151] Y. Oh, S. Park, and J. C. Ye, "Deep learning COVID-19 features on CXR using limited training data sets," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2688–2700, 2020.

[152] W. Liang, J. Yao, A. Chen et al., "Early triage of critically ill COVID-19 patients using deep learning," *Nature Communications*, vol. 11, no. 1, 2020.

[153] M. Abadi, Q. Hasan, S. Rahmati, A. Sharifi, and M. Ahmadi, "HSSAGA: designation and scheduling of nurses for taking care of COVID-19 patients using novel method of hybrid salp swarm algorithm and genetic Algorithm," *Applied Soft Computing*, vol. 108, Article ID 107449, 2021.

[154] M. Ahmadi, A. Sharifi, S. Dorosti, S. Jafarzadeh Ghoushchi, and N. Ghanbari, "Investigation of effective climatology parameters on COVID-19 outbreak in Iran," *Science of the Total Environment*, vol. 729, Article ID 138705, 2020.

[155] Nasirpour, M. Hossein, A. Sharifi, M. Ahmadi, and S. Jafarzadeh Ghoushchi, "Revealing the relationship between solar activity and COVID-19 and forecasting of possible future viruses using multi-step autoregression (MSAR)," *Environmental Science and Pollution Research*, vol. 28, pp. 1–11, 2021.

[156] A. Davoudi, M. Ahmadi, A. Sharifi et al., "Studying the effect of taking statins before infection in the severity reduction of COVID-19 with machine learning," *BioMed Research International*, vol. 2021, Article ID 9995073, 12 pages, 2021.

[157] A. Sharifi, M. Ahmadi, and A. Ali, "The impact of artificial intelligence and digital style on industry and energy post-COVID-19 pandemic," *Environmental Science and Pollution Research*, vol. 28, pp. 1–21, 2021.

[158] A. Memiş, S. Varlı, and F. Bilgili, "Semantic segmentation of the multiform proximal femur and femoral head bones with the deep convolutional neural networks in low quality MRI sections acquired in different MRI protocols," *Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society*, vol. 81, Article ID 101715, 2020.

[159] A. Duran, P. M. Jodoin, and C. Lartizien, "Prostate cancer semantic segmentation by gleason score group in bi-parametric MRI with self attention model on the peripheral zone," in *Proceedings of the Medical Imaging with Deep Learning*, pp. 193–204, PMLR, Montreal, Canada, September 2020.

[160] N. Alalwan, A. Abozeid, A. A. ElHabshy, and A. Alzahrani, "Efficient 3d deep learning model for medical image semantic segmentation," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 1231–1239, 2021.

[161] G. Karayegen and M. F. Aksahin, "Brain tumor prediction on MR images with semantic segmentation by using deep learning network and 3D imaging of tumor region," *Biomedical Signal Processing and Control*, vol. 66, Article ID 102458, 2021.

[162] M. Ahmadi, A. Sharifi, S. Hassantabar, and S. Enayati, "QAIS-DSNN: tumor area segmentation of MRI image with optimized quantum matched-filter technique and deep spiking neural network," *BioMed Research International*, vol. 2021, Article ID 6653879, 16 pages, 2021.

[163] M. Ahmadi, A. Sharifi, M. Jafarian Fard, and N. Soleimani, "Detection of brain lesion location in MRI images using convolutional neural network and robust PCA," *International Journal of Neuroscience*, pp. 1–12, 2021.

[164] D. Tabernik, M. Kristan, and A. Leonardis, "Spatially-adaptive filter units for compact and efficient deep neural networks," *International Journal of Computer Vision*, vol. 128, no. 8, pp. 2049–2067, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 9528664, 13 pages
https://doi.org/10.1155/2021/9528664

Research Article

Skin Cancer Detection Based on Extreme Learning Machine and a Developed Version of Thermal Exchange Optimization

Shi Wang[1] **and Melika Hamian**[2]

[1]*Department of Computer Engineering, Dongguan Polytechnic, Dongguan 523808, Guangdong, China*
[2]*Department of Engineering, Payame Noor University (PNU), Tehran, Iran*

Correspondence should be addressed to Melika Hamian; mhamian92@gmail.com

Received 9 July 2021; Revised 17 August 2021; Accepted 31 August 2021; Published 3 November 2021

Academic Editor: Navid Razmjooy

Melanoma is defined as a disease that has been incurable in advanced stages, which shows the vital importance of timely diagnosis and treatment. To diagnose this type of cancer early, various methods and equipment have been used, almost all of which required a visit to the doctor and were not available to the public. In this study, an automated and accurate process to differentiate between benign skin pigmented lesions and malignant melanoma is presented, so that it can be used by the general public, and it does not require special equipment and special conditions in imaging. In this study, after preprocessing of the input images, the region of interest is segmented based on the Otsu method. Then, a new feature extraction is implemented on the segmented image to mine the beneficial characteristics. The process is then finalized by using an optimized Deep Believe Network (DBN) for categorization into 2 classes of normal and melanoma cases. The optimization process in DBN has been performed by a developed version of the newly introduced Thermal Exchange Optimization (dTEO) algorithm to obtain higher efficacy in different terms. To show the method's superiority, its performance is compared with 7 different techniques from the literature.

1. Introduction

The skin cancer has been known as one of the deadliest types of cancers which has recently witnessed an exponential growth worldwide. With all these interpretations, it can be cured if it is detected in the early stages, and, in most cases, a simple biopsy can prevent cancer from growing [1]. This cancer has been growing dramatically during the recent years and the importance of the initial therapy has been more to be seen. The melanoma skin cancer has been ranked as the 3rd most common malignant cancer among the skin cancer cases [2]. This cancer alters the skin color as a result of the irregular effects of the cells [3]. Notwithstanding this danger, it is known as a curable cancer if it is detected in the early stages. With all these interpretations, early detection of the melanoma from the other types of the benign skin moles is a thought-provoking mission. Any variations in the size, color, or the shape of a skin mole can be considered as initial signs for this cancer.

Based on the statistical information for 2019, skin cancer with 15000 cases is ranked as the fourth most common cancer, and with 1900 deaths cases are ranked as one of the 10th common cancers in the world [4]. The most common method for melanoma detection is by the physicians and their sampling and testing. However, in some cases, even experienced specialists make some mistakes in melanoma detection. Furthermore, laboratory sampling and testing is a time-consuming and expensive experiment that bothers the patients. Therefore, designing a (computer) system that can detect malignant lesions will be very useful. According to researches, the effectiveness of the computer-aided system is more accurate than that of a specialist doctor. Recently, many research works are accomplished for automatic early detection of skin cancers. Also, the high capability of the medical imaging assists the medical doctors to lessen the diagnosis difficulty of this work and to enhance the speed of the initial diagnosis of the disease.

The application of Artificial Neural Networks (ANNs) in different parts of image processing and machine vision has been increasing. The ANNs are imitated from the human brain actions and are used in a variety of applications, from medicine to economics and engineering [5].

The ANN is made up of lots of highly unified processing components called neurons which can solve problems with their interconnectedness. Artificial Neural Networks, like humans, learn by example to solve different kinds of problems from pattern recognition to information classification throughout a learning procedure.

Deep learning is a sort of ANNs which uses mathematical techniques to give a structure like the human brain. On the other hand, advances in technology have led to algorithms for optimizing normal neural networks so that we can count the number of neural layers. Neural networks are increased from multiple layers to thousands of layers and thousands of neurons in each layer, which until a few years ago could not have created such a structure [6]. This type of neural network is called a deep learning network. Several methodologies of using different kinds of ANNs and other machine learning techniques were introduced for skin cancer diagnosis [7, 8]. For instance, Li et al. [9] proposed a new data synthesis methodology to combine individual skin lesions images and heavily expand them to create noteworthy amounts of data. The study worked on a convolutional neural network (CNN) to provide superior efficiency to the old-style detection and tracking methods. Moreover, the system was trained by humans with simple criteria.

Esteva et al. [10] presented a supervised technique for skin lesions using deep learning. The method's efficiency was verified by twenty-one medical images with two critical binary classifications: malignant and benign groups. Two cases were analyzed here. The first one recognizes the most common cancers, and the second case signifies the deadliest skin cancer identification. The final results indicated the high performance of the proposed method.

Jafari et al. [11] proposed a deep learning-based procedure to provide a precise skin cancer area segmentation. After image denoising and preprocessing, a CNN was utilized for segmentation. They combined local and global structure information and outputted a label for the pixels to create a segmentation mask that indicates the lesion area. This mask was then improved with some postprocessing operators. The empirical results indicated the superiority of the presented method compared to the other techniques.

Mohamed et al. [12] proposed a learning technique to categorize skin lesions to determine melanoma. The method was based on applying a CNN with multiple outlines. The results of the suggested technique were performed by 14 layers into the International Skin Imaging Collaboration (ISIC) dataset. Final results indicated successful ratio for the method in the segmentation of the skin cancer area.

Xu et al. [2] proposed an automatic diagnosis method for melanoma. The presented technique was based on an optimized image segmentation methodology along with a CNN optimized by satin bowerbird optimization (SBO). After segmentation, feature extraction was implemented on the processed image. The main features were then selected using the SBO algorithm to trim additional information. At last, the results were presented by performing an SVM into two groups of normal and melanoma cases. The method's results were authenticated by comparison with other techniques from the literature to indicate its performance.

As the research background shows, the application of deep learning, including CNNs in medical images processing, especially in skin lesion diagnosis, has been exponentially increasing. The present study proposes an optimized technique of skin lesion segmentation using a new design of the Thermal Exchange Optimization algorithm which has been used to reach this purpose with higher performance. The overall technique for the proposed technique has been shown in Figure 1.

Therefore, the main contributions of the present study can be highlighted as follows:

(i) A new computer-aided method for the diagnosis of malignant melanoma

(ii) Segmenting the region of interest based on the Otsu method

(iii) Feature extraction to extract the beneficial features

(iv) A new optimized Deep Believe Network (DBN) to classify the melanoma cases

(v) Optimization based on a developed version of the newly introduced Thermal Exchange Optimization (dTEO) algorithm

2. Skin Cancer Preprocessing

Medical imaging is the method and process used to create images of the human body to examine or advance medical science and is widely used in medicine [13]. Diagnostic images are the subject of many scans, examinations, and images used in medicine. Researches have shown that medical images taken are highly susceptible to vulnerability to noise and other factors such as low contrast. Here are two simple ways to do this purpose.

2.1. Noise Reduction. Images noise removal is a significant part of all medical image processing. This process tries to reduce or remove the noise from the original image. Features that expose them to noise are available on all recording devices [14]. The noise may be white or random noise. Noise reduction recreates and restores the noisy image to its original model based on ideal models in which we analyze how this method is done. The noise has a major impact on the image in certain situations, particularly through the identification of the image edge that needs differentiation which increases the influence of high-frequency pixels that specifically contain noise [15, 16].

The noise in medical images is typically Gaussian noise. Gaussian noise is dispersed in the image and is located between the image main pixels, which makes some inconsistencies for the next operations, and because the processing must be done only on the image main pixels, it requires a proper noise elimination for this purpose. Noise pixels are usually recognized as a different pixel among a group of

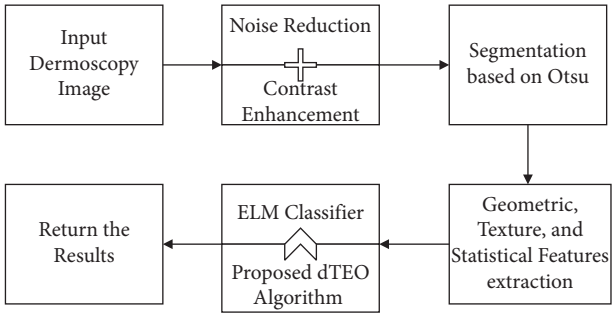


Figure 1: The overall procedure of the suggested method.

pixels in the original image that differs from adjacent pixels in different terms such as light intensity and transparency. Wang-Mendel algorithm is an efficient fuzzy-based noise removal tool [17]. Because of the simple conception of the fuzzy techniques, they are popular for different applications. Furthermore, because of their speediness as an early-stage process, they are very useful for producing the initial fuzzy model [18]. In this method, the input and the output datasets show the behavior of the solved state of the problem. The production of the rule database is performed as follows:

(1) Consider fuzzy isolation from the space that includes the input variables, which is achieved by the knowledge of an expert along with normalization. Then, it splits it into equivalent or unequal sections by doing a fuzzy-based method for parting the space of input variable. Then a form of the membership function is chosen and each component is given as a fuzzy package. Afterward, the membership and fuzzy set has been established to all the parts.

(2) A set of candidate language rules are produced, which are generated by choosing the most inclusive rules for the samples.

(3) Validity degree is assigned for the laws, which is achieved by multiplying the membership function values for all the components.

(4) The final rules database is obtained from the rules of the candidate language set, which is done by categorizing these rules into some groups, where the groups include the rules of candidates with the same assumption. The final rule is obtained by selecting the highest degree for all the groups.

2.2. Image Contrast Enhancement. Dermoscopy images typically lack high image contrast, which makes them impossible to process next stages. The concern arises from multiple circumstances, such as the poor quality of measuring instruments and cameras, the low degree of the user interface in photography, the environmental factors, and the prevalence of noise. Sometimes, any crucial details from the images disappear from the above examples, rendering the processing too complex [19]. Contrast enhancement is a method for addressing the problems of contrast consistency. In this research, image contrast enhancement is used to enhance and help demonstrate the specifics of the cancer regions. Global enhancement of contrast using the lookup table is used in the present research. For categorizing and storing the received images on the disk, an 8-bit lookup table is used. The strategy is usually developed as follows [20]:

$$\mathrm{PDF}_{\mathrm{out}} = \frac{\mathrm{PDF}_{\mathrm{In}} - \mathrm{PDF}_{\mathrm{min}}}{\mathrm{PDF}_{\mathrm{Max}} - \mathrm{PDF}_{\mathrm{min}}}, \tag{1}$$

where $\mathrm{PDF}_{\mathrm{In}}$ and $\mathrm{PDF}_{\mathrm{out}}$ represent the probability density function of the input and the output corrected images, respectively, and $\mathrm{PDF}_{\mathrm{min}}$ and $\mathrm{PDF}_{\mathrm{Max}}$ represent the minimum and the maximum probability density levels, respectively. Figure 2 shows an example of the preprocessed image for more clarification.

3. Skin Cancer Segmentation

3.1. Color Space. The RGB is the most common and basic color space in images processing, which usually is used for displaying images. Due to the dependence of these three colors, which are considered as base colors in the RGB color space, as well as their dependence on the intensity of ambient light, this color space is less used for main processing. To compensate for this shortcoming, the XYZ color model has been employed. This color space is a quantifiable link between the physiologically supposed colors in human color vision and wavelengths distributions in the electromagnetic visible spectrum . In this color space, the values of X and Z give color information, and Y defines the luminance. To convert color space from RGB to XYZ, the following equation has been used [21]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \times \begin{bmatrix} 0.49 & 0.31 & 0.2 \\ 0.17697 & 0.8124 & 0.01063 \\ 0 & 0.01 & 0.99 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{2}$$

The most significant benefit of XYZ color model is that it is completely independent of the device.

3.2. Method of Segmentation. In dermoscopy images, the red color space is the only dimension in RGB color space that gives nearly the strength of the images. As mentioned before, both X and Z values give related information from the color for the XYZ color space. Therefore, the red (R) dimension and the X dimension are just normalized for the segmentation; that is [21],

$$\begin{aligned} \widehat{R} &= \frac{R}{\sqrt{R^2 + G^2 + B^2}}, \\ \widehat{X} &= \frac{X}{\sqrt{X^2 + Y^2 + Z^2}}. \end{aligned} \tag{3}$$

On any pixel of the input images, these normalization values are performed. After this normalization, the Otsu thresholding is used for giving a low-cost segmentation in terms of time complexity.

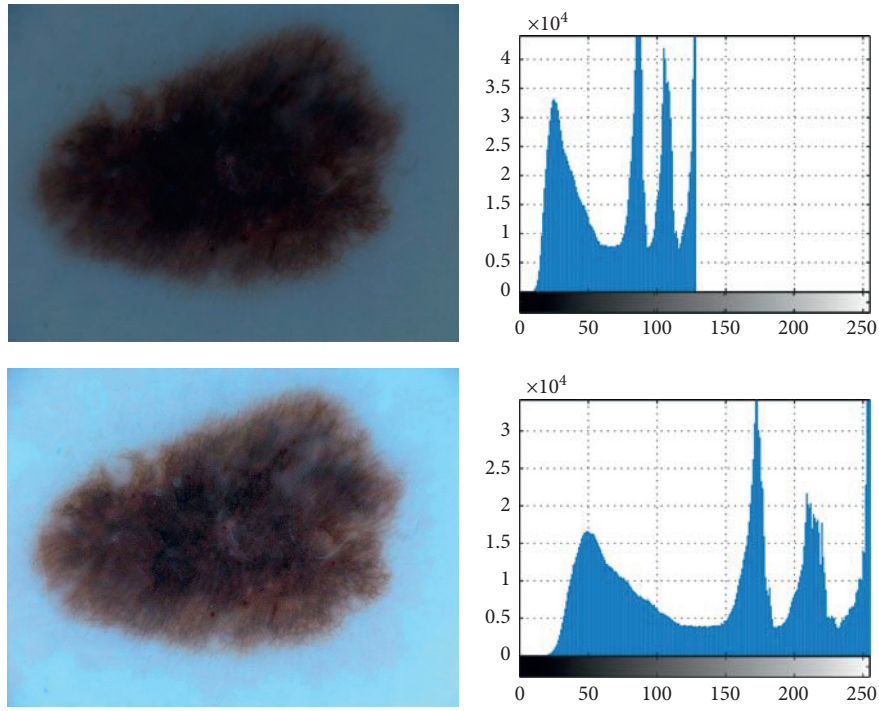


Figure 2: Two examples of the preprocessing of the image: (a) original image; (b) PDF of (a).

The Otsu technique is a popular method for proper thresholding of the images. The method is based on the conception of the variance of the intergroup maximization and the intragroup minimizing of the pixels. The global threshold has a problem when the image background resolution is inadequate. A local threshold can be used to eliminate the heterogeneity effect. This problem is solved by image preprocessing to eliminate inhomogeneities and apply a global threshold to the processed image. Based on Otsu technique, threshold level is comprehensively searched to provide minimum in-between variance class, which is expressed as follows:

$$v(t) = \omega_1(t)v_1^2(t) + \omega_2(t)v_2^2(t), \tag{4}$$

where ω_i denotes the probability of two separate classes with a threshold value of t. σ_i^2 in this case is the value of the variance of these classes. In fact, Otsu shows that minimizing the amount of variance in a class is like maximizing interclass variance:

$$v_b^2(t) = v^2 - v_\omega^2(t) = \omega_1(t)\omega_2(t)\left[m_1(t) - m_2(t)\right]^2, \tag{5}$$

where v_i^2 describes the variance and m_i signifies the mean value that can be updated intermittently. This technique is briefly given as follows:

(i) Evaluate image histogram and its intensity level probability

(ii) Give initial values of $\omega_i(0)$ and $v_i(0)$ for all levels of possible threshold ($t = 1, 2, \ldots$)

(iii) Renew ω_i and m_i

(iv) Calculate $v_b^2(t)$

The optimal threshold here is the maximum of $\sigma_b^2(t)$.

Then, a morphology operation including closing, opening, and filling is used for better results. In this stage, by applying the mathematical filling operator, the extra image holes are filled. This operator is as follows [22]:

$$X_k = \left(X_{k-1} \oplus D\right) \cap C^c, \quad k = 1, 2, 3 \ldots, \tag{6}$$

where A and B represent the area and the constructing component, respectively.

Then, the opening operation has been established on the filled image for eliminating the ignitor particulars with no changes on other gray surfaces. This operator is as follows:

$$C \circ D = (C \ominus D) \oplus D. \tag{7}$$

Finally, for connecting the narrow parts, the closing operation is performed based on the following equation:

$$C \cdot D = (C \oplus D) \ominus D. \tag{8}$$

Here, the structural element in this study is set as a 5×5 identity matrix. Algorithm 1 shows the pseudocode of the image segmentation.

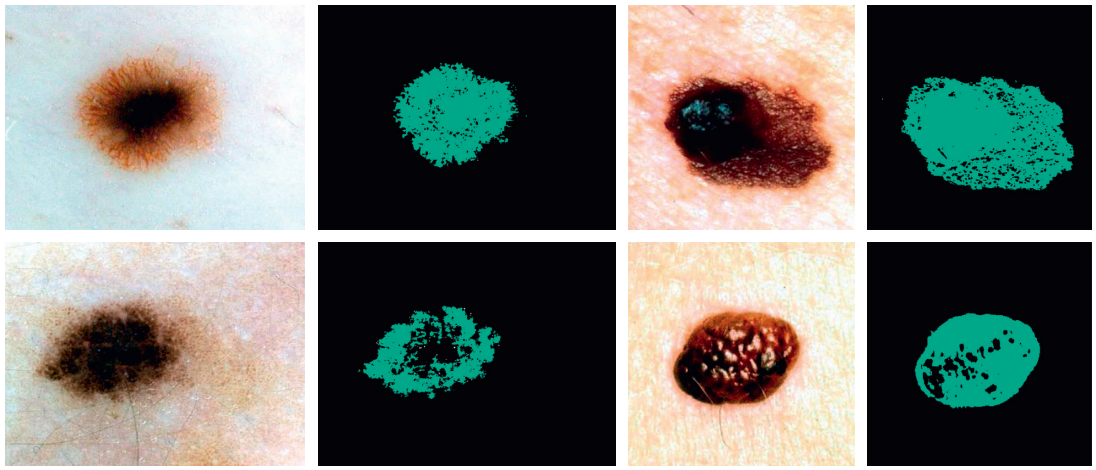
Figure 3 displays some samples of image segmentation using the suggested segmentation.

4. Features Extraction

Feature extraction is the process of collecting more detailed information from the image. The features extraction in this part of processing is utilized for extracting the main features

Input: the preprocessed skin dermoscopy image
Step 1. Normalize R dimension from RGB color space
Step 2. Convert RGB color space to XYZ
Step 3. Normalize X dimension from XYZ color space
Step 4. Apply Otsu threshold algorithm
Step 5. Apply postprocessing method based on mathematical morphology
Output: the segmented image

Algorithm 1: Pseudocode of the proposed image segmentation.



(a) (b) (a) (b)

Figure 3: Some samples of the skin cancer segmentation using the suggested technique: (a) the input image and (b) the segmented image.

of the segmented skin cancer area to simplify the diagnosis successful. There are numerous features for feature extraction. This research uses statistical features and texture features, and geometric features have been adopted for this purpose. Table 1 gives the features utilized here.

In the above table, MN describes the size of image, B_p describes the length of external side in the pixel boundary, $p(i, j)$ represents the intensity values of the pixels at position, μ defines the mean value, σ describes the standard deviation, and a and b define the major axis and the minor axis, respectively. However, some achieved features are not useful for the process and just increase the method's complexity. To prevent this issue, a feature selection procedure has been utilized. To do so, in this study, an optimized technique has been used. For giving optimal results of features selection, the following cost function has been considered:

$$\text{fitness} = \frac{(T_P \times T_N) - (F_P \times F_N)}{\sqrt{((T_N + F_P) \times (T_P + F_P) \times (T_P + F_N) \times (T_N + F_N))}}, \tag{9}$$

where F_P describes the false positive, F_N describes the false negative, T_P defines true positive, and T_N defines the true negative. The idea is to minimize the above equation based on a developed optimization method.

5. Extreme Learning Machine (ELM)

One of the neural network models which has recently received many considerations is the ELM model. Some of the advantages for ELM are their high speed in learning, their simple using, and their ability to be used in numerous activation functions and nonlinear kernel functions. The ELM model can provide an integrated template with a wide variety of feature transmissions that can be used in the hidden layer, which can be utilized directly for regression and categorization. This algorithm is a learning method for Single Hidden Neural Networks (SHNN) based on initializing the input biases and weights randomly and the output weights evaluation. Thus, the network can be trained in just a few steps. Figure 4 shows the arrangement of a simple ELM network.

For a classification problem with D dimension and N number of training samples [23],

$$\left(x^{(n)}, t^{(n)}\right), \quad n = 1: N, \tag{10}$$

where $x^{(n)} \in \mathbb{R}^D$ and $t^{(n)} \in \mathbb{R}^K$.

An ELM-based feed-forward neural network can be formulated as follows:

$$\sum_{m=1}^{M} \beta_m g\left(w_m^T x^{(n)} + b_m\right) = t^{(n)}, \tag{11}$$

where $g(\cdot)$ describes the activation function, b_m describes the bias of the m^{th} hidden neuron, M signifies the hidden neurons, $w_m = [w_{m_1}, w_{m_2}, \ldots, w_{\text{mD}}]$ determines the input weight vector which connects the input neurons to the m^{th} the neuron of the hidden layer, and $\beta_m = [\beta_{m_1}, \beta_{m_2}, \ldots, \beta_{\text{mK}}]$ determines the vector of weight that connects the m^{th} neuron.

This conception is given as follows:

$$H\beta = T, \tag{12}$$

where

$$H = \begin{bmatrix} g(w_m^T x^{(1)} + b_1) & \cdots & g(w_M^T x^{(1)} + b_1) \\ \vdots & \ddots & \vdots \\ g(w_1^T x^{(N)} + b_1) & \cdots & g(w_M^T x^{(N)} + b_M) \end{bmatrix}_{N\times M},$$
$$H = \left[\beta_1^T, \beta_2^T, \ldots, \beta_M^T\right]_{M\times N}^T, \tag{13}$$
$$T = \left[t_1^T, t_2^T, \ldots, t_M^T\right]_{N\times K}^T$$

Meanwhile, H defines a nonsquare matrix. The training samples quantity is bigger than the hidden neurons quantity. Therefore, to solve this problem, the following formulation has been used:

$$\widehat{\beta} = H^{\dagger}T, \tag{14}$$

where $H^{\dagger}$ describes the generalized Moore-Penrose matrix inverse.

Thus, the ELM networks can be shortened to three main steps:

(1) Initializing the input weights and bias with random values

(2) Evaluation of the hidden neurons' outputs matrix H

(3) Evaluation of the hidden neurons' outputs weights matrix by equation (16)

6. The Developed Thermal Exchange Optimization Algorithm

Optimization is the process of using specific techniques to solve the optimization problems. However, in recent years, the performance of the classic optimization methods due to increasing the complexity and nonlinearity of them has been decreased or even failed [24]. Recently, several methods have been introduced for resolving these problems [25]. Metaheuristics are one class of these methods that have been popular due to their simple and suitable results in various problems. The metaheuristic algorithms are derived from various phenomena from nature, human society, animal hunting behavior, and so forth. Several algorithms have been proposed in this field [26–29], for example, World Cup Optimizer [30], Ant Lion Optimizer (ALO) [31], Chimp Optimization Algorithm [32], Harris Hawks Optimization [33], and mayfly optimization algorithm [34]. In the present research, a novel modified model of the Thermal Exchange Optimization (TEO) algorithm has been presented to achieve optimal results for the considered methodology [5].

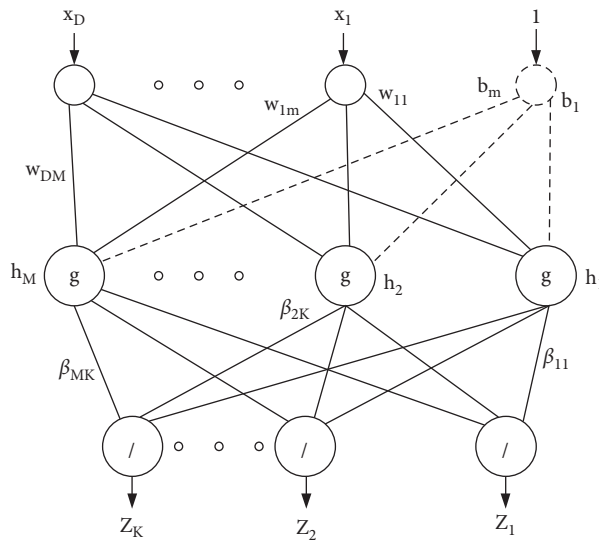


Figure 4: A sample configuration of the ELM network.

The TEO algorithm is a novel metaheuristic technique that is inspired by the temperature of the objects and their position that is switched between cold and warm places, indicating the updated positions. In the following, more explanation about this algorithm has been provided.

6.1. The Newton Cooling Law. The tests indicated that the rate of cooling was roughly related to the temperature difference between the environment and the heated object. This is formulated as in the following equation:

$$\frac{dQ}{dt} = \alpha \times A \times (T_s - T_b), \tag{15}$$

where A describes the area of the body surface that transmits heat, Q signifies the heat, α defines the coefficient of heat transfer that is reliant on different cases like object geometry, heat transfer mode, and surface state, and T_b and T_s represent the temperature of the body and the ambient temperature, respectively.

The heat loss in time dt is $\alpha \times A \times (T_s - T)\,dt$, which states the variation in kept heat as the temperature drops dT; in other words,

$$V \times d \times c \times dT = -\alpha \times A \times (T - T_b)\,dt, \tag{16}$$

where d signifies the density (kg/m^3), c defines the specific heat (J/kg/K), and V describes the volume (m^3).

So,

$$\frac{(T - T_b)}{(T_M - T_b)} = \exp\left(\frac{-\alpha \times t \times A}{V \times c \times d}\right), \tag{17}$$

where T_M describes the initial high temperature. Equation (19) is valid if when $\alpha \times A \times t/V \times d \times c$ is T-independant,

$$\zeta = \frac{(\alpha \times A)}{(V \times c \times d)}. \tag{18}$$

So, by considering a constant ζ,

$$\frac{T - T_b}{T_M - T_b} = \exp(-\zeta t). \tag{19}$$

So, consequently, by reformulating equation (19),

$$T = (T_M - T_b) \times \exp(-\gamma t) + T_b. \tag{20}$$

6.2. The Algorithm. In Thermal Exchange Optimization, some of the candidates have been assumed as the cooling objects and the other remaining candidates are assumed as the environment; afterward, the reverse operation has been done. In TEO algorithm, like any other metaheuristic algorithm, the candidates are first generated randomly. So, the initial temperature for each object is formulated by the following equation:

$$T_i^0 = T_{\min} + \text{rnd} \times (T_{\max} - T_{\min}), \tag{21}$$

where rnd signifies a random vector limited between 0 and 1, T_i^0 describes the initial solution vector of the i^{th}object, and$T_{\min}$ and $T_{\max}$ represent the minimum and maximum limitations of the candidates.

The randomly generated candidates are then performed to the cost fiction to evaluate their cost value. After that, the position of some best T candidate vectors is saved as *Thermal Memory* (TM) to be utilized then for improving the algorithm's efficacy by less computational cost. Afterward, some best TM candidates have been given and the equal numbers of population with the worst values have been eliminated.

The candidates are separated into two equal types that are shown in Figure 5.

For more clarification, T_1 determines the environment object for $T_{n/2+1}$ cooling object, and, inversely, if any object has lower value than ζ, the temperature exchanging is performed slowly. Therefore, ζ will be achieved by the following formula:

$$\gamma = \frac{\text{Cos (object)}}{\text{Cos (worst object)}}. \tag{22}$$

Another important term in the algorithm is time, which is associated with the iteration quantity. This parameter is formulated in the following:

$$t = \frac{\text{iteration}}{\text{max. iteration}}. \tag{23}$$

To give an exploration term to the algorithm, the environmental temperature changing has been modeled, which is given in the following equation:

$$T_i^e = (1 - (m_1 + m_2 \times (1 - t) \times \text{rand})) \times T_i^{\text{'e}}, \tag{24}$$

where m_1 and m_2 describe the control variables and $T_i^{\text{'e}}$ signifies the object earlier temperature which has been modified to T_i^e.

By considering the former models, the new temperature of the objects is updated as follows:

$$T_i^+ = T_i^e + \left(T_i^{\text{old}} - T_i^e\right)\exp(-\zeta t). \tag{25}$$

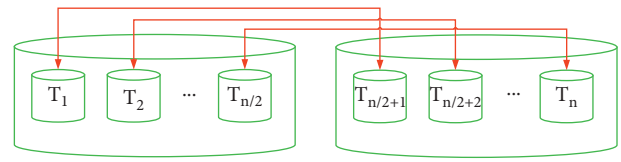


Figure 5: The cooling and heating transfer and environment objects pairs.

The next term is Pr, which is used to indicate whether the cooling objects' element should be altered or not.

The Pr components are compared with $R(i)(i = 1, 2, \ldots, n)$, which has a randomly distributed value between 0 and 1. If $R(i) < \text{Pr}$, the i^{th} candidate will be randomly chosen as 1-dimensional and the value is reformulated by the following:

$$T_{i,j} = T_j^{\min} + \text{rnd}\left(T_j^{\max} - T_j^{\min}\right)\exp(-\zeta t). \tag{26}$$

where $T_{i,j}$ determines the j^{th} variable of the i^{th} candidate and $T_j^{\min}$and $T_j^{\max}$ are the minor and the major limitations of the variable number j, respectively. At last, the algorithm is stopped when stopping criteria are checked to terminate the algorithm.

6.3. Developed TEO Algorithm. Although TEO algorithm has numerous benefits in providing the best global solution, it also has some drawbacks which could be resolved. The main drawback of the algorithm could be its local minimum results and low convergence in some cases. Numerous improvements were introduced for resolving the exploration of the metaheuristics. In this research, chaotic theory has been utilized as the first improvement. The reason for using chaos in this algorithm is to improve the algorithm's efficiency to resolve the local optimization followed by an advanced speed of convergence. In the considered dTEO algorithm, the modified randk is replaced with the "rnd," i.e.,

$$\begin{aligned} \text{rand}_k &= \alpha f_k^2 \sin(\pi f_k), \\ f_0 &\in [0, 1],\ \alpha \in (0, 4], \end{aligned} \tag{27}$$

where k signifies the number of iterations.

The other improvement for the algorithm to give a proper trade-off between exploration and exploitation in the algorithm using the Gaussian mutation mechanism is formulated below [35]:

$$g(x) = \frac{1}{\sqrt{2\pi\sigma}}\exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \tag{28}$$

where σ^2 signifies the variance of the Gaussian PDF and μ determines the Gaussian distribution expectation. This difference is applied to the considered location by renewing the TEO algorithm ($T_{i,j}^{\text{new}}$) as follows:

$$T_{i,j}^{\text{new}} = T_{i,j} \times (1 + \gamma \times g(0, 1)). \tag{29}$$

where γ describes a decreasing random value between 0 and 1 and $g(0, 1)$ describes the standard Gaussian distribution.

6.4. Algorithm Authentication. In this section, the proposed developed Thermal Exchange Optimization (dTEO) algorithm that is stated in the previous section has been programmed by the MATLAB 2017b environment and then validated based on a 16 GB RAM Core™ i7/4720HQ 1.60 GHz processor. The proposed dTEO algorithm validation is based on applying it to some standard test cases. The results are also compared with some other approaches, that is, Multiverse Optimizer (MVO) [36], Locust Swarm Optimization (LSO) [37], Spotted Hyena Optimizer (SHO) [38], and the original TEO [39], to show the proposed method's efficacy. Table 2 shows the details of the studied functions.

In Table 3, the parameters settings of all studied algorithms for use in this study are given below.

For determining the performance of the algorithm, they performed on the benchmark functions and two measurement indicators including the minimum value and the standard values are extracted from the results to show the algorithms' accuracy and reliability. It should be noted that, for all algorithms, the population size and the number of iterations are set as 120 and 100, respectively. The simulations have been repeated 20 times to give consistent results. Table 4 illustrates the simulation results.

As can be seen from Table 4, the suggested dTEO has the least value of the minimum value for the minimization problems. This means that the suggested method provides the best accuracy toward the other state-of-the-art methods for solving the studied functions. Plus, based on the results achieved for the standard deviation, it is definite that the suggested method has the minimum value of that. This shows that using the proposed method has the highest reliability against the comparative techniques in this study. So, we can conclude that using this algorithm in this study is a good idea in terms of accuracy and reliability.

7. Optimized ELM Network Based on Algorithm

The classification of the final extracted features is performed based on the proposed optimized ELM network. The main task of the ELM classifier here is to determine the skin cancer dermoscopy image into two groups of normal and malignant melanoma classes.

In this study, activation functions of the network are considered for enhancement. Generally, the significant effect of activation function on the ANNs is too clear, because a network with various activation functions gives good generalization capacities. This study considers a sigmoid activation function as one of the popular functions in this area. The formulation for this function is explained as follows:

$$g(X) = \frac{1}{1 + \exp(-(\text{as}.x + \text{bs}))}. \tag{30}$$

For improving the efficiency of the ELM-based classifier, different sigmoid functions have been utilized in the hidden neurons. In other words, different values are utilized foras and bs. Figure 6 displays the impact of as and bs on the function shape.

In other words, the problem of finding the suitable parameter of the sigmoid function can be considered as an optimization problem. Here, we used a developed Thermal Exchange Optimization algorithm to refine the issue. However, both "as" and "bs" have almost equal effect on the function shape; we used the "bs" for optimization of the network. Therefore, two optimization policies have been used.

The input weights have been first initialized. Then, the developed Thermal Exchange Optimization algorithm is utilized for optimal parameters selection by assuming the initial weights. The objective function of the developed TEO is based on the square error between the output of the desired value and the network by the following:

$$E = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k}\left(d_{ji} - y_{ji}\right)^2, \tag{31}$$

where n and k describe the numbers of training samples and the output layers, respectively, and y_{ji} and d_{ji} represent the output of the network and the desired value, respectively.

The primary target of the present study is to utilize the mentioned optimized ELM network for diagnosis of the skin cancer. The optimization is using a newly designed metaheuristic, namely, the developed Thermal Exchange Optimization algorithm.

8. Results and Discussion

In this paper, a new optimized methodology is used to detect the skin cancer from dermoscopy images.

8.1. The Database. For validating the proposed skin cancer identification system, there are many databases. This analysis uses a well-known dataset, called SIIM-ISIC Melanoma. The images are generated in DICOM format to be retrieved by popular libraries and include both image and metadata. Two formats of TFRecord and JPEG are considered for the dataset. The size of the images is uniform (1024×1024). Some examples of the utilized images in the study are displayed in Figure 7.

8.2. Results and Discussion. As mentioned before, the present study is authenticated by the ACS database and the results have been verified by some different methodologies. The simulations are performed on a laptop with MATLAB 2018b environment with the following information: 16 GB RAM and Core™ i7-4720HQ-1.60 GHz processor. The pipeline methodology of the method is explained in the following:

(1) Start
(2) Apply preprocessing
(2.1) Apply noise reduction
(2.2) Apply image contrast enhancement
(3) Perform skin cancer segmentation

Table 1: The utilized features of this study.

Feature name	Formula	Feature name	Formula
Mean	$1/MN\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)$	Contrast	$\sum_{i=1}^{M}\sum_{j=1}^{N}p^2(i,j)$
Variance	$1/MN\sum_{i=1}^{M}\sum_{j=1}^{N}(p(i,j)-\mu)$	Perimeter	$\sum_{i=1}^{M}\sum_{j=1}^{N}B_p(i,j)$
Standard deviation	$\sqrt{\text{variance}}$	Entropy	$-\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)\log\ p(i,j)$
Area	$\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)$	Solidity	Area/Convex Area
Rectangularity	Area/$b\times a$	Correlation	$\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)-\mu_r\mu_c/\sigma_r\sigma_c$
Irregularity index	$4\pi\times$ Area/Perimeter^2	Invariant moments	$\varphi_1=\eta_{20}+\eta_{02}$
Form factor	Area/a^2		$\varphi_2=(\eta_{20}-\eta_{02})^2+4\eta_{11}^2$
			$\varphi_3=(\eta_{30}-3\eta_{12})^2+(3\eta_{21}-\mu_{03})^2$
Eccentricity	$2/a(a^2-b^2)^{0.5}$	Energy	$\sum_{i=1}^{M}\sum_{j=1}^{N}p^2(i,j)$
Elongation	$2\text{Area}^{0.5}/a\sqrt{\pi}$	Homogeneity	$\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)/1+\lvert i-j\rvert$

Table 2: The details of the studied functions.

#	Test function	Minimum value	Boundary
1	$F_1=\sum_{n=1}^{N}x_n^2$	0	$-\infty\le x\le\infty$
2	$F_2=\sum_{n=1}^{N-1}(100\times[x_{n+1}-x_n^2]^2+[1-x_n]^2)$	0	$-\infty\le x_n\le\infty$
3	$F_3=\sum_{n=1}^{N}\lvert x_n\rvert-10\cos(\sqrt{\lvert 10x_n\rvert})$	0	$-\infty\le x_n\le\infty$
4	$F_4=x\sin(4x)+1.1y\sin(2y)$	−18.5547	$0\le x,\ y\le 10$
5	$F_5=[\sum_{n=1}^{N}nx_n^4]+N_n(0,1)$	Varies	$-\infty\le x\le\infty$
6	$F_6=10N+\sum_{n=1}^{N}[x_n^2-10\cos(2\pi x_n)]$	0	$-\infty\le x_n\le\infty$
7	$F_7=1+\sum_{n=1}^{N}\frac{x_n^2}{4000}-\prod_{n=1}^{N}\cos(x_n)$	0	$-\infty\le x_n\le\infty$
8	$F_8=1/2+\sin^2\sqrt{x^2+y^2}-0.5/1+0.1(x^2+y^2)$	−0.5231	$-\infty\le x,\ y\le\infty$

Table 3: The utilized parameters settings of all studied algorithms.

Algorithm	Parameter	Value	Algorithm	Parameter	Value
LSO [37]	F	0.6	ALO [31]	w	[2, 6]
	L	1		No. of search agents	50
	g	20		Traveling distance rate	[0.6, 1]
SHO [38]	$\overrightarrow{M}$	[0.5, 1]	MVO [36]	Wormhole existence probability	[0.2, 1]
	$\overrightarrow{h}$	[5, 0]			

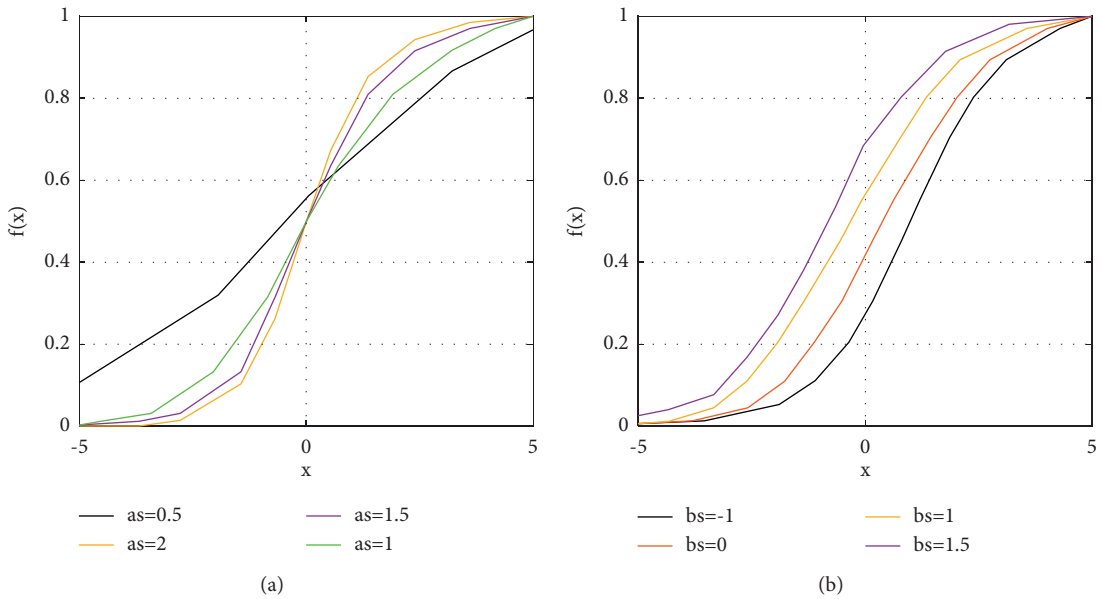


Figure 6: The impact of (a) as and (b) bs on the function shape.

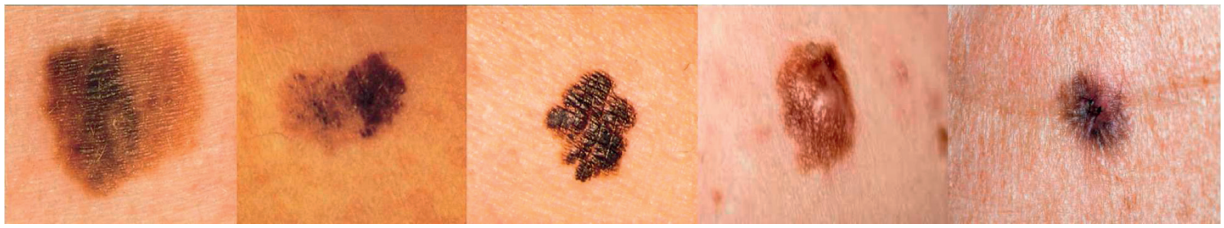


Figure 7: Some examples of the utilized images in the study.

Table 4: The comparison results of the studied algorithms implemented on the test functions.

Function		Algorithm					
		LSO [37]	SHO [38]	MVO [36]	ALO [31]	TEO	dTEO
f_1	Min	$3.837e-24$	$2.1201e-28$	$-2.2800e-27$	$3.3197e-28$	$3.4511e-28$	$8.2193e-31$
	Std	$2.559e-19$	$2.4937e-27$	$4.0865e-28$	$2.9938e-29$	$2.0073e-30$	$2.37928e-32$
f_2	Min	$7.2874e-5$	$7.564e-4$	$4.824e-4$	$2.56384e-5$	$4.5463e-6$	$6.7811e-5$
	Std	$3.2084e-6$	$4.18297e-5$	$2.1167e-5$	$3.5918e-6$	$5.0648e-7$	$2.1253e-5$
f_3	Min	−7.2553	−8.15754	−10.00	−7.19376	−10.00	−10.00
	Std	0.35	0.46	0.27	0.15	0.15	0.10
f_4	Min	−5.2846	−18.053	−17.1146	−16.3927	−18.1106	−18.1683
	Std	3.126	1.294	2.391	4.190	1.630	1.01
f_5	Min	$11.46e-10$	$2.597e-16$	$2.18769e-7$	$5.0913e-9$	$2.8096e-23$	$3.7938e-22$
	Std	$8.942e-11$	$4.0973e-17$	$2.89430e-8$	$6.5514e-10$	$2.8897e-25$	$7.1937e-24$
f_6	Min	$4.276e-11$	$2.1953e-12$	$2.1967e-21$	$3.1957e-10$	$2.0054e-18$	$5.6124e-23$
	Std	$7.297e-12$	$1.5374e-14$	$4.2849e-23$	$6.8091e-11$	$2.3169e-19$	$4.6498e-24$
f_7	Min	$2.623e-15$	$3.3733e-10$	$5.0406e-8$	$3.7628e-13$	$3.6781e-9$	$6.39482e-17$
	Std	$1.167e-16$	$4.1967e-12$	$4.9364e-9$	$3.1936e-15$	$4.1945e-10$	$2.22864e-19$
f_8	Min	0.0067	−0.1472	−0.3492	−0.5846	−0.5790	−0.5273
	Std	0.653	0.467	0.385	0.805	0.152	0.100

Table 5: The performance validation of the suggested technique compared with other methods.

Method	Performance metric				
	Accuracy	Sensitivity	Specificity	PPV	NPV
Fractal analysis [40]	88.23	91.18	88.23	80.88	92.65
CNN [41]	82.35	83.82	61.76	66.18	83.82
Delaunay Triangulation [42]	86.76	89.70	86.76	77.94	88.23
Side-by-Side method [43]	67.64	83.82	58.82	58.82	83.82
Genetic Algorithm [44]	80.88	79.412	79.41	75.00	85.29
Fusion method [45]	83.82	82.35	76.47	70.58	80.88
SVM [46]	83.82	83.82	76.47	75.00	89.70
Proposed method	92.65	91.18	89.70	86.76	94.12

(4) Apply features extraction based on texture features, statistical features, and the geometric features

(5) Perform dTEO-based ELM classifier for the final diagnosis

(6) Return the results

(7) End

In the presented dTEO-based ELM classifier, 80% of data are set to train data, and the remaining 20% are set to test the data. The training step of the network is assumed by 700 iterations. The training process has been iterated 15 times to achieve a confidential result. The validation has been performed using 5 measurement indexes: specificity, sensitivity, accuracy, PPV, and NPV, which are given as follows:

$$\begin{aligned}
\text{accuracy} &= \frac{\text{correctly detected cases}}{\text{total cases}},\\
\text{specificity} &= \frac{\text{correctly detected healthy skin cases}}{\text{total healthy skin cases}},\\
\text{sensitivity} &= \frac{\text{correctly detected skin cancer cases}}{\text{total skin cancer cases}},\\
\text{PPV} &= \frac{\text{correctly detected skin cancer cases}}{\text{detected skin cancer cases}},\\
\text{NPV} &= \frac{\text{correctly detected healthy skin cases}}{\text{detected healthy skin cases}}.
\end{aligned} \tag{32}$$

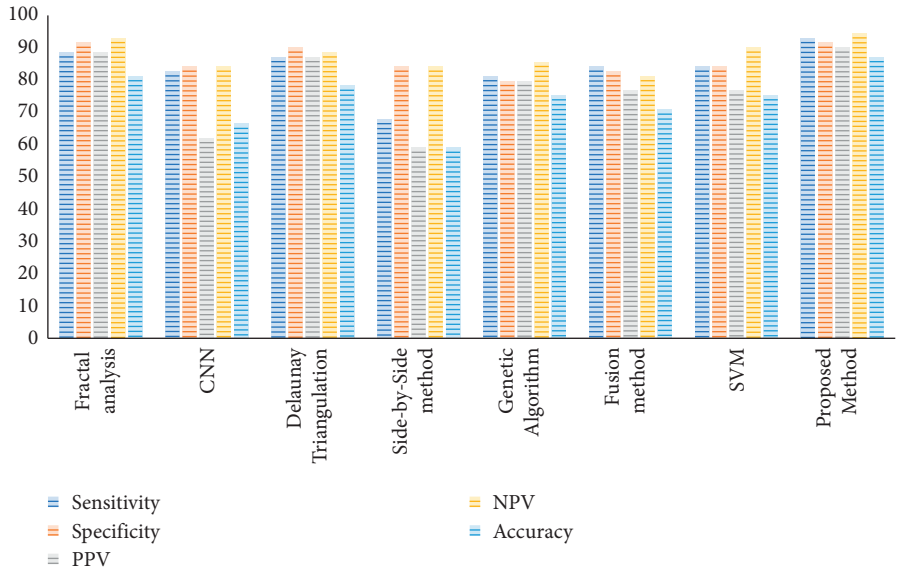


Figure 8: The bar plot of the performance validation for the proposed method compared with other state-of-the-art methods.

The results are compared with 7 other methods, fractal analysis [40], CNN [41], Delaunay Triangulation [42], Side-by-Side method [43], Genetic Algorithm [44], fusion method [45], and SVM [46], to verify the method's performance in different terms.

Table 5 tabulates the performance validation for the suggested technique compared with other methods. The results show the mean value of 30 runs for the methods.

As is concluded from Table 5, the suggested method with 95% accuracy provides the uppermost precision among the other compared techniques. Also, with 95% of sensitivity, which is the highest against other methods, the reliability of the proposed method has been proved. This can be proved also for NPV, PPV, and specificity toward the others. The better results of NPV and PPV indicate higher commonness of the condition to determine the probability of a test diagnosing cancer. Also, better results of the specificity and sensitivity for the suggested technique show higher prevalence-independent results for the algorithm. For more clarification, the results of Table 5 are shown in Figure 8.

9. Conclusions

In recent years, skin cancer has been recognized as the most dangerous and common type of cancer in humans. There are different types of skin cancer. Melanoma is a common type of skin cancer, where early detection can be helpful in its treatment and can suggestively stop death from this deadly skin cancer. Designing an approach that facilitates the skin cancer detection in the early stages is very useful and valuable. In this study, an optimized pipeline procedure was utilized for the optimal detection of melanoma from dermoscopy images. In the proposed method, after preprocessing of the input dermoscopy images based on noise reduction and contrast enhancement, the region of interest was segmented. Afterward, feature extraction was implemented on the segmented images to extract useful features from them. Finally, an optimized Deep Believe Network (DBN) was utilized to separate images into two classes of healthy and cancerous. The optimization of the DBN was using a new metaheuristic method called the developed Thermal Exchange Optimization algorithm to improve the network efficiency in terms of reliability and accuracy. So, the main contribution of the proposed method is to use a newly developed version of the newly introduced Thermal Exchange Optimization for the diagnosis of malignant melanoma. The main advantage of using this technique is that the results showed that it improved the system's efficiency in both terms of accuracy and precision or, in other words, its reliability during different runs. The performance of the proposed technique was authenticated by comparing it with 7 other methods: fractal analysis, CNN, Delaunay Triangulation, Side-by-Side method, Genetic Algorithm, fusion method, and SVM. Simulation results specified that, based on some various performance indexes, the suggested procedure gave the best results against the compared techniques. The main limitation of the proposed method is that, due to different soft techniques, it needs a lot of time. In the future study, we will work on designing a method to simplify the method from a theoretical methodology to a real-time system for practical applications.

Data Availability

The data are available at https://www.kaggle.com/c/siim-isic-melanoma-classification.

Conflicts of Interest

The authors declare that there are no conflicts of interest in this research.

Acknowledgments

This study was supported by specialty construction project of high level higher vocational colleges and Technical Skills Expert Project of Dongguan Polytechnic (Y17040323).

References

[1] S. Bao, A. Ebadi, M. Toughani et al., "A new method for optimal parameters identification of a PEMFC using an improved version of monarch butterfly optimization algorithm," *International Journal of Hydrogen Energy*, vol. 45, no. 35, pp. 17882–17892, 2020.

[2] Z. Xu, F. R. Sheykhahmad, N. Ghadimi, and N. Razmjooy, "Computer-aided diagnosis of skin cancer based on soft computing techniques," *Open Medicine*, vol. 15, no. 1, pp. 860–871, 2020.

[3] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.

[4] A. Costa, Y. Kieffer, A. Scholer-Dahirel et al., "Fibroblast heterogeneity and immunosuppressive environment in human breast cancer," *Cancer Cell*, vol. 33, no. 3, pp. 463.e10–479.e10, 2018.

[5] D. Yu, Y. Wang, H. Liu, K. Jermsittiparsert, and N. Razmjooy, "System identification of PEM fuel cells using an improved elman neural network and a new hybrid optimization algorithm," *Energy Report*, vol. 5, pp. 1365–1374, 2019.

[6] V. Rajinikanth and M. S. Couceiro, "RGB histogram based color image segmentation using firefly algorithm," *Procedia Computer Science*, vol. 46, pp. 1449–1457, 2015.

[7] U. R. Acharya, S. L. Fernandes, J. E. WeiKoh et al., "Automated detection of alzheimer's disease using brain MRI images-a study with various feature extraction techniques," *Journal of Medical Systems*, vol. 43, no. 9, p. 302, 2019.

[8] N. S. M. Raja, S. L. Fernandes, N. Dey, S. C. Satapathy, and V. Rajinikanth, "Contrast enhanced medical MRI evaluation using tsallis entropy and region growing segmentation," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–12, 2018.

[9] Y. Li, A. Esteva, B. Kuprel, R. Novoa, J. Ko, and S. Thrun, "Skin cancer detection and tracking using data synthesis and deep learning," 2016, https://arxiv.org/abs/1612.01074.

[10] A. Esteva, B. Kuprel, R. A. Novoa et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[11] M. H. Jafari, N. Karimi, E. Nasr-Esfahani et al., "Skin lesion segmentation in clinical images using deep learning," in *Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR)*, December 2016.

[12] A. Mohamed, W. A. Mohamed, and A. H. Zekry, "Deep learning can improve early skin cancer detection," *International Journal of Electronics and Telecommunications*, vol. 65, 2019.

[13] R. Pugalenthi, M. P. Rajakumar, J. Ramya, and V. Rajinikanth, "Evaluation and classification of the brain tumor MRI using machine learning technique," *Journal of Control Engineering and Applied Informatics*, vol. 21, no. 4, pp. 12–21, 2019.

[14] S. C. Satapathy, N. Sri Madhava Raja, V. Rajinikanth, A. S. Ashour, and N. Dey, "Multi-level image thresholding using otsu and chaotic bat algorithm," *Neural Computing and Applications*, vol. 29, no. 12, pp. 1285–1307, 2018.

[15] D. Anoraganingrum, "Cell segmentation with median filter and mathematical morphology operation," in *Proceedings of the 10th International Conference on Image Analysis and Processing*, September 1999.

[16] T. Loupas, W. N. McDicken, and P. L. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images," *IEEE Transactions on Circuits and Systems*, vol. 36, no. 1, pp. 129–135, 1989.

[17] V. Rajinikanth and S. C. Satapathy, "Segmentation of ischemic stroke lesion in brain MRI based on social group optimization and fuzzy-tsallis entropy," *Arabian Journal for Science and Engineering*, vol. 43, no. 8, pp. 4365–4378, 2018.

[18] N. Dey, V. Rajinikanth, A. Ashour, and J. M. Tavares, "Social group optimization supported segmentation and evaluation of skin melanoma images," *Symmetry*, vol. 10, no. 2, p. 51, 2018.

[19] S. R. Fernandes, J. T. de Assis, V. V. Estrela et al., "Nondestructive diagnosis and analysis of computed microtomography images via texture descriptors," in *Advances in Multidisciplinary Medical Technologies—Engineering, Modeling and Findings*, pp. 249–261, Springer, Berlin, Germany, 2020.

[20] K. B. Singh, T. V. Mahendra, R. S. Kurmvanshi, and C. V. Rama Rao, "Image enhancement with the application of local and global enhancement methods for dark images," in *Proceedings of the 2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC)*, April 2017.

[21] M. N. Carney and W. M. Johnston, "A novel regression model from RGB image data to spectroradiometric correlates optimized for tooth colored shades," *Journal of Dentistry*, vol. 51, pp. 45–48, 2016.

[22] P. Maragos, "Differential morphology and image processing," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp. 922–937, 1996.

[23] J. Cao, K. Zhang, M. Luo, C. Yin, and X. Lai, "Extreme learning machine and adaptive sparse representation for image classification," *Neural Networks*, vol. 81, pp. 91–102, 2016.

[24] M. Ramezani, D. Bahmanyar, and N. Razmjooy, "A new optimal energy management strategy based on improved multi-objective antlion optimization algorithm: applications in smart home," *SN Applied Sciences*, vol. 2, no. 12, pp. 1–17, 2020.

[25] M.-W. Tian, S.-R. Yan, S.-Z. Han, S. Nojavan, K. Jermsittiparsert, and N. Razmjooy, "New optimal design for a hybrid solar chimney, solid oxide electrolysis and fuel cell based on improved deer hunting optimization algorithm," *Journal of Cleaner Production*, vol. 249, Article ID 119414, 2020.

[26] L. Yanda, Z. Yuwei, and N. Razmjooy, "Optimal arrangement of a micro-CHP system in the presence of fuel cell-heat pump based on metaheuristics," *International Journal of Ambient Energy*, pp. 1–12, 2020.

[27] Z. Yuan, W. Wang, H. Wang, and N. Razmjooy, "A new technique for optimal estimation of the circuit-based PEMFCs using developed sunflower optimization algorithm," *Energy Report*, vol. 6, pp. 662–671, 2020.

[28] Y. Guo, X. Dai, K. Jermsittiparsert, and N. Razmjooy, "An optimal configuration for a battery and PEM fuel cell-based hybrid energy system using developed krill herd optimization algorithm for locomotive application," *Energy Report*, vol. 6, pp. 885–894, 2020.

[29] Z. Yang, Q. Liu, L. Zhang, J. Dai, and N. Razmjooy, "Model parameter estimation of the PEMFCs using improved barnacles mating optimization algorithm," *Energy*, vol. 212, Article ID 118738, 2020.

[30] R. Navid, M. Khalilpour, and M. Ramezani, "A new metaheuristic optimization algorithm inspired by FIFA world cup competitions: theory and its application in PID designing for AVR system," *Journal of Control, Automation and Electrical Systems*, vol. 27, no. 4, pp. 419–440, 2016.

[31] M. Mani, O. Bozorg-Haddad, and X. Chu, "Ant lion optimizer (ALO) algorithm," in *Advanced Optimization by Nature-Inspired Algorithms*, pp. 105–116, Springer, Berlin, Germany, 2018.

[32] M. Khishe and M. R. Mosavi, "Chimp optimization algorithm," *Expert Systems with Applications*, vol. 149, Article ID 113338, 2020.

[33] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: algorithm and applications," *Future Generation Computer Systems*, vol. 97, pp. 849–872, 2019.

[34] K. Zervoudakis and S. Tsafarakis, "A mayfly optimization algorithm," *Computers & Industrial Engineering*, vol. 145, Article ID 106559, 2020.

[35] N. Higashi and H. Iba, "Particle swarm optimization with gaussian mutation," in *Proceedings of the 2003 IEEE Swarm Intelligence Symposium. SIS'03 (Cat. No. 03EX706)*, April 2003.

[36] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-verse optimizer: a nature-inspired algorithm for global optimization," *Neural Computing and Applications*, vol. 27, no. 2, pp. 495–513, 2016.

[37] E. Cuevas, F. Fausto, and A. González, "The locust swarm optimization algorithm," in *New Advancements in Swarm Algorithms: Operators and Applications*, pp. 139–159, Springer, Berlin, Germany, 2020.

[38] G. Dhiman and V. Kumar, "Spotted hyena optimizer: a novel bio-inspired based metaheuristic technique for engineering applications," *Advances in Engineering Software*, vol. 114, pp. 48–70, 2017.

[39] A. Kaveh and A. Dadras, "A novel meta-heuristic optimization algorithm: thermal exchange optimization," *Advances in Engineering Software*, vol. 110, pp. 69–84, 2017.

[40] R. Dobrescu, M. Dobrescu, S. Mocanu, and D. Popescu, "Medical images classification for skin cancer diagnosis based on combined texture and fractal analysis," *WISEAS Transactions on Biology and Biomedicine*, vol. 7, no. 3, pp. 223–232, 2010.

[41] N. Zhang, Y.-X. Cai, Y.-Y. Wang, Y.-T. Tian, X.-L. Wang, and B. Badami, "Skin cancer diagnosis based on optimized convolutional neural network," *Artificial Intelligence in Medicine*, vol. 102, Article ID 101756, 2020.

[42] A. Pennisi, D. D. Bloisi, D. Nardi, A. R. Giampetruzzi, C. Mondino, and A. Facchiano, "Skin lesion image segmentation using delaunay triangulation for melanoma detection," *Computerized Medical Imaging and Graphics*, vol. 52, pp. 89–103, 2016.

[43] G. Babino, A. Lallas, M. Agozzino et al., "Melanoma diagnosed on digital dermoscopy monitoring: a side-by-side image comparison is needed to improve early detection," *Journal of the American Academy of Dermatology*, vol. 85, no. 3, pp. 619–625, 2020.

[44] C. Salem, D. Azar, and S. Tokajian, "An image processing and genetic algorithm-based approach for the detection of melanoma in patients," *Methods of Information in Medicine*, vol. 57, no. 1, pp. 74–80, 2018.

[45] R. Das, S. De, S. Bhattacharyya, J. Platos, V. Snasel, and A. E. Hassanien, "Data augmentation and feature fusion for melanoma detection with content based image classification," in *Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications*, Cairo, Egypt, January 2020.

[46] H. Alquran, I. A. Qasmieh, A. M. Alqudah et al., "The melanoma skin cancer detection and classification using support vector machine," in *Proceedings of the 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, October 2017.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 6509982, 10 pages
https://doi.org/10.1155/2021/6509982

Research Article

Improving the Security and Confidentiality in the Internet of Medical Things Based on Edge Computing Using Clustering

Anita Hatamian,[1] Mohammad Bagher Tavakoli,[1] and Masoud Moradkhani[2]

[1]*Department of Electrical Engineering, Arak Branch, Islamic Azad University, Arak, Iran*
[2]*Department of Electrical Engineering, Ilam Branch, Islamic Azad University, Ilam, Iran*

Correspondence should be addressed to Mohammad Bagher Tavakoli; m-tavakoli@iau-arak.ac.ir

Received 13 July 2021; Revised 4 September 2021; Accepted 14 September 2021; Published 26 October 2021

Academic Editor: Navid Razmjooy

Families, physicians, and hospital environments use remote patient monitoring (RPM) technologies to remotely monitor a patient's vital signs, reduce visit time, reduce hospital costs, and improve the quality of care. The Internet of Medical Things (IoMT) is provided by applications that provide remote access to patient's physiological data. The Internet of Medical Things (IoMT) tools basically have a user interface, biosensor, and Internet connectivity. Accordingly, it is possible to record, transfer, store, and process medical data in a short time by integrating IoMT with the data communication infrastructure in edge computing. (Edge computing is a distributed computing paradigm that brings computation and data storage closer to the sources of data. This is expected to improve response times and save bandwidth. A common misconception is that edge and IoT are synonymous.) But, this approach faces problems with security and intrusion into users' medical data that are confidential. Accordingly, this study presents a secure solution in order to be used in the IoT infrastructure in edge computing. In the proposed method, first the clustering process is performed effectively using information about the characteristics and interests of users. Then, the people in each cluster evaluated by using edge computing and people with higher scores are considered as influential people in their cluster, and since users with high user interaction can publish information on a large scale, it can be concluded that, by increasing user interaction, information can be disseminated on a larger scale without any intrusion and thus in a safe way in the network. In the proposed method, the average of user interactions and user scores are used as a criterion for identifying influential people in each cluster. If there is a desired number of people who are considered to start disseminating information, it is possible to select people in each cluster with a higher degree of influence to start disseminating information. According to the research results, the accuracy has increased by 0.2 and more information is published in the proposed method than the previous methods.

1. Introduction

The Internet of Things is a new concept in the world of information and communication technology; in fact, it is a modern technology that provides the ability to send and receive data through communication networks, whether the Internet or intranet for any entity [1]. Smart devices are all part of an emerging category called the Internet of Things, or IoT for short. At a basic level, the Internet of Things deals with connecting different objects through the Internet and communicating with each other to achieve its goal of providing a more efficient and intelligent experience.

In the meantime, improving health is considered to be the ultimate goal of economic, social, and technical progress, and rapid population growth and aging will change the world significantly. These factors have put a lot of pressure on food suppliers and health systems around the world, and the advancement of new IoT technology is expected to provide appropriate and potential solutions to these challenges. Therefore, the use of IoT for health can certainly be effective in strategic research roadmap in medicine and the use of IoT. Nowadays, the advent of various wearable devices has enabled people to monitor their health anywhere and anytime through the Internet of Medical Things (IoMT) [2].

Accordingly, IoT-based wearable devices will become a key part of human health in the not-too-distant future. IoMT not only is able to monitor physiological information about the human body through wearable devices but also can send the results of this information to a remote monitoring center or family doctor, or even generate emergency alerts. It should be noted that IoMT can help physicians to suggest treatment programs through recommender systems.

This method can significantly reduce the time of medical examination, improve diagnostic efficiency, and save human resources [3]. But now, it is impractical to transfer information from wearable devices (where they were created) to the cloud infrastructure (where they are processed) because the cloud-centric processing method results in high communication delays and reduces data transfer rates between IoT devices and IoT devices with potential users.

In order to accomplish the previously stated goal, the concept of edge computing has recently been proposed, which makes it possible to process IoT services to process data near data sources and data sinks instead of engaging in the cloud [4]. This solution reduces communication delays and makes better use of the computing, storage, and network resources that are currently available. It can also reduce runtime and power consumption, which can be very useful for using the Internet of Things in health and medicine.

Given that edge computing is a new research topic, robust integrated solutions to its various challenges, including security, and secure routing, have not yet been proposed. However, one of the problems with using such systems is that unauthorized people infiltrate the system and use its information or perform unauthorized manipulations. This feature allows unauthorized access to existing data and performing manipulations in these networks [5]. Therefore, this study examines the process of detecting intrusion into the Internet of Medical Things using clustering methods in the edge computing infrastructure and the results will be compared with some other methods in this field.

The most important contributions of this study include the following:

Providing a clustering-based routing method for use in the Internet of Medical Things (IoMT).

Using edge computing to reduce delays in sending important medical information to emergency centers safely

An interactive and structural approach is used to increase the accuracy within the clusters in order to increase efficiency and increase security. In this way, the influential nodes can be identified with higher accuracy.

2. Research Background

The study in [6] reviewed security and confidentiality in the IoMT. Accordingly, firstly security and privacy problems in various parts of the IoMT have been investigated. The existing solutions in the field of security and privacy problems, which provide the suitable context for solving these challenges for the IoT infrastructure, were reviewed. According to the results of this study, there is a need for new methods in the process of providing security and intrusion detection in IoMT.

The challenges and methods of detecting malware intrusion in the Internet of Things were reviewed in [7]. This study aimed to conduct a comprehensive review of existing problems and solutions. The presence of various botnets in IoMT infrastructures has led to attacks on the security of an infrastructure, including comprehensiveness, confidentiality, and availability. These can reduce the spread of new infrastructure such as IoMT. Accordingly, firstly, the types of malware attacks and how they work have been examined.

Then, an IoT-based architecture with various security protocols is examined for use in deploying and securing the IoMT infrastructure. The study in [8] proposed a secure framework using the Internet of Medical Things that also improves energy consumption. This method is designed to reduce communication overhead and energy consumption between medical sensors that process and transmit health information.

On the other hand, the proposed method is able to secure the medical data of patients against unwanted access. According to the evaluation results, energy consumption is reduced by 29%, and productivity is increased by 18%. A fog computing-based (Fog computing or fog networking, also known as fogging, is an architecture that uses edge devices to carry out a substantial amount of computation, storage, and communication locally and routed over the Internet backbone.) attack detection system using MOA-WMA algorithm is presented in [9].

The proposed model consists of several parts. In preprocessing, the input data is first processed and converted into a format that can be used in machine learning, and the required features are extracted from them and then sent to the feature selection section, where the required features are selected for accurate detection of intrusion and enter the next section after passing a filter to remove additional information, for modeling in this phase using a multilayer neural network. Finally, the next steps are done based on this modeling and learning that is done with the help of training data. The evaluation results indicate an increase in accuracy and efficiency with the help of this method. An in-depth learning-based solution is presented in [10] to extract IoT data features and intrusion detection in IoMT. The deep learning model is combined with intrusion detection technology to help with this approach. This method solves the problem of distributive incompatibility of source and target data. The proposed solution can be used in various neural networks. Accordingly, the problem of data learning can be solved with its help, and abnormal data can be detected more accurately. Thus, evaluations have shown that intrusion into the IoMT can be detected with high accuracy. A Markov model-based approach is presented in [11] to detect anomalies and secure health monitoring systems.

According to the results of this study, due to the continuity of the applied traffic in the network and their interdependence, as a result, abnormal movements in the network can be detected by using the Markov model. This

solution is based on data dependence on the network, so it can operate without supervision based on the learning model and be used on wearable sensors with low energy.

Also, unlike other methods that only operate linearly, this method can be performed in nonlinear dimensions as well, which in turn will be able to detect intrusion and anomalies in the network with more power. Finally, the solution is evaluated and it is concluded that, on average, it significantly increases the accuracy of network anomaly detection. A combination of different machine learning methods to detect intrusion into the IoT infrastructure is presented in [12]. Accordingly, six types of machine learning methods (Random Forest, K-Nearest Neighbor Support Vector Machine, Neural Networks, J48, and Decision Table) were used.

This method uses unlabeled data along with monitored methods to increase the accuracy of the IoMT intrusion detection system. This solution teaches a single-layer neural network to categorize samples on unlabeled data. Next, each of the created categories is combined with the initial data set, and retraining is applied to the classification method. According to the results of the research, the accuracy of the solution is high, and also the execution time of the intrusion detection process is reduced. Table 1 provides a summary of previous methods and the advantages and disadvantages of each method.

3. Clustering of the Proposed Method

An interactive approach is used to identify influential nodes in IoT networks, that is, paying attention to interactions between users. User interaction is reflecting a node that receives more tweets but responds less (has fewer retweets) and is a better option to be happy, and in fact, such a node is selected as the influential node. That is, the influential nodes in each association are identified using an analytical relationship based on a combination of structural and interactive approaches. It should be noted that the number of tweets received and responding to each user is used within the database.

3.1. Utilizing an Interactive and Structural Approach to Increase Accuracy within Clusters. After identifying the effective users in the previous step, the impact point of each of them is calculated by the following formula:

$$V(u_i, u_j) = \frac{nr_{ij+nr_{ji}}}{NR_i + NR_j} + \frac{T_i}{T_{ij}}, \quad (1)$$

where nr_{ij} is the number of times u_i retweets u_j tweets, nr_{ji} indicates the number of times u_j retweets u_i tweets, NR_i is the number of times u_i retweets other users' tweets, NR_j is the number of times u_j retweets the tweets of other users, T_i is the total number of messages received by i, and T_{ij} is the total number of messages sent from i to j, each of which is normalized individually.

3.2. Research Plan. Given that some of the nodes of a large database are studied (case study) in this study, the research method in this dissertation is a case study.

First, information and theoretical foundations of the research are extracted from valid internal and external sources in the present study. Then the IoT general database Mirai and Bashlite which is available at http://www.archive.ics.uci.edu/mI/datasets/detection_of_iot_botnet_attacks_n_balot is selected for implementation and execution. Data are first categorized by the k-means clustering algorithm; then, intrusion detection and routing are optimized by edge calculations in MATLAB and the results are presented in the form of tables and graphs. Finally, the obtained results are compared with the results of other previous methods.

3.3. Problem Solving Method. In the present study, K-means clustering is used to classify the database after selecting the appropriate database. Different forums are identified in the IoT network. In the next step, special attention is paid to the demographic characteristics of users, including their age, gender, and occupation, to identify the forums in the IoT networks. A code between 0 and 1 is assigned to each of these attributes. A kind of clustering algorithm should be used to continue the process of identifying associations based on the mentioned demographic characteristics; the k-means algorithm is selected in the present study.

Then, influential nodes are identified in the IoT networks. Edge computing is used to identify the influential nodes in the IoT networks. The interactive approach in edge computing, that is, interactions between users, has been considered. Interaction between users means a node that receives more information but responds less (has less response) is a better option to be head-cluster, and in fact, such a node is selected as the influential node.

That is, the influential nodes in each association are identified using an analytical relationship, based on a combination of structural and interactive approaches. It should be noted that the amount of information received and answered for each user is used in the data. By identifying the influential users in the previous step, the impact score of each of them is now calculated and at the end, the obtained values are placed in the objective function which is presented in the next section. Figure 1 shows the alpha, beta, and gamma wolves movements.

Figure 2 provides the optimal solutions by alpha, beta, and gamma wolves in the proposed method.

The pseudocode of the gray wolf algorithm is given in Figure 3.

Table 2 tabulates a part of the study data set.

3.4. Objective Function in Edge Computation Algorithm. This study was carried out to increase the accuracy and secure routing of the IoT network using edge computing. Therefore, the objective function is as follows:

$$F(u_i, u_j) = \left(\frac{nr_{ij+nr_{ji}}}{NR_i + NR_j}\right) * V + 1, \quad (2)$$

where F is the accuracy in the IoT network and is calculated by edge computing.

Table 1: An overview of previous models and the advantages and disadvantages of each.

Disadvantages	Advantages	Model (method)
Lack of privacy in some cases.	More accuracy and expansion of information in less time.	A model based on fuzzy logic, artificial intelligence, and gray wolf [9].
The process of searching and exchanging data is associated with problems in relatively large volumes of data.	More precision in the cluster social network.	Clustering model of social networks using multilayer perceptron network and gray wolf [10].
Dissemination of confidential information that owners are reluctant to disclose.	Wider use of social media, viral marketing, in which various forms of marketing message related to a company, brand, or product are transmitted with exponential growth, often through the use of social media applications.	A hybrid model of standard algorithm (naive k-means) and bee and gray wolf algorithm [11].
Time-consuming and repetitive results in the third stage and later, which is associated with high error.	Drawing a decision tree from data, expand the dissemination of information in social networking associations, break down a resource set into subcategories based on testing the value of an attribute.	Decision tree and gray wolf model [13].
There is a mapping of all related edges among the studied heads that the genetic algorithm is unable to examine.	Utilizing the capacity of social networks at many individual and social levels to identify problems and determine their solutions, establishing social relations, managing organizational affairs, policy-making, and guiding individuals on the path to achieving goals.	Model based on genetic algorithm and gray wolf [14].
Increased computational complexity compared to previous methods.	High accuracy and efficiency of the algorithm.	A model based on new genetic algorithm and gray wolf [15].
Occurrence of shilling attacks and loss of users' trust in recommender systems.	More speed in social networking forums.	Model based on participatory refinement algorithm and gray wolf [12].
Precise settings of input parameters that may be associated with errors.	Improving the quality of community discovery.	Model based on F algorithm and gray wolf [16].
Identifying edges statically, disregarding user interactions.	Maintaining data privacy.	Model based on a cuckoo optimization algorithm and gray wolf algorithm [17].
When the number of users is large, the number of errors in the algorithm exceeds the allowable limit defined in the default algorithm, and the results become invalid.	Improving the structure of users' social network, providing different services to users spatially.	Model based on a new algorithm based on spatial data and gray wolf [18].
Loss of time and place and more role of space in the social network and shaking the foundations of users' identities.	The social networks studied in the communication process can have positive effects.	Model based on genetic algorithm with a generational division and gray wolf [19].
The complexity of the implementation process and its time-consumption, the occurrence of human error in computing.	High accuracy and speed in social network processing.	A hybrid model of clustering and artificial neural network and gray wolf [20].
Lack of coverage of all nodes in the forum on large social networks.	Identifying the intermediate centrality of communication channels with anonymity and the characteristics of each anonymous node and effective communication channels, maintaining data privacy.	Combined model of anonymity clustering and genetic algorithm and gray wolf [21].
Inability to check all nodes and community members on social networks.	Identifying how information is disseminated on social networks, identifying key nodes in this type of network, providing quality control of information dissemination on social networks.	Model based on the evolutionary algorithm of bees and gray wolf [22].
Central network nodes cannot be checked by this method.	Identification and mapping of information exchange networks in social networking associations, revealing the main nodes.	Model based on modified genetic algorithm and gray wolf [23].

In this study, the community includes all users of the database. The sample group includes 1000 users of the database. By definition, the random sampling method is simple [25]. MATLAB (version 2017) was used for data analysis, and results are presented. In this section, simulation results on IoT network data were presented. The simulation consists of two stages.

In the first stage, network forums or clustering of 1000 users is found based on the three characteristics of gender, age, and occupation, and in the next step, the

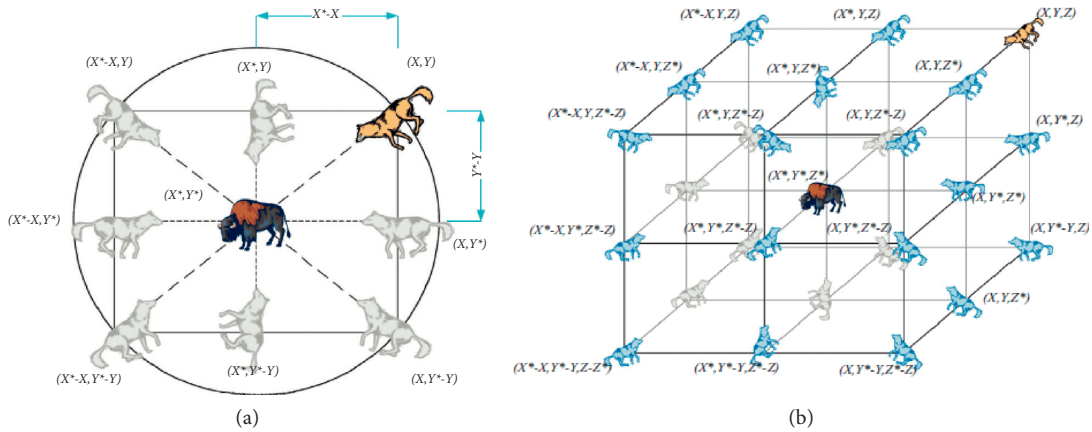


Figure 1: The alpha, beta, and gamma wolf movements.

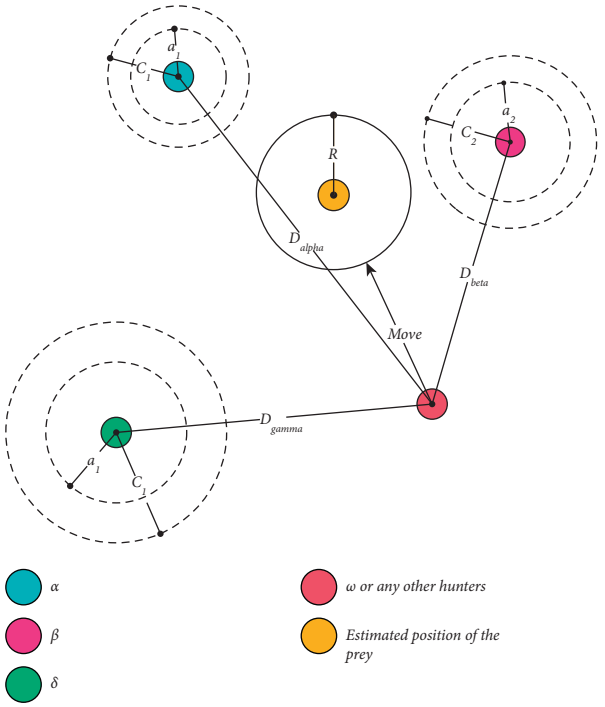


Figure 2: Optimal solutions by alpha, beta, and gamma wolves in the proposed method.

Initiailze the grey wolf population X_i *(i = 1, 2, ..., n)*
Initiailze a, A, and C
Calculate the fitness of each search agent
X_α*=the best search agent*
X_β*=the second best search agent*
X_δ*=the third best search agent*
while *(t < Max number of iterations)*
  for *each search agent*
    Update the position of the current search agent by equation (3.7)
  end for
  Update a, A, and C
  Calculate the fitness of all search agents
  Update X_α, X_β, *and* X_δ
  t=t+1
end while
return X_α

Figure 3: Pseudocode of the gray wolf algorithm.

Table 2: A part of the study data set.

ID	Gender
1	1
2	0
3	0
4	0
5	1
6	1
7	1
8	1
9	0
10	1

degree of interaction and ranking of clustered users are obtained. The data in this section includes information about these 1000 users and their neighbors, in addition to information about user interactions, shared posts, and user comments about posts. In this method, users are selected with a higher degree of interaction and rank to start the process of disseminating information in each cluster, and information is disseminated through these users to other users.

MATLAB software was used for simulations related to user clustering, and the degree of interaction and user rank were used to perform the relevant simulations.

3.5. Platform Used. The platform used Windows 7, 64-bit CPU, 4 GHZ, Core 7 Duo.

3.6. Clustering, Number of Clusters, Clustering Effect, Clustering Criteria, and Reason for Using Cummins Algorithm for Clustering. Data clustering is done in this section. Clustering leads to the categorization and ordering of large volumes of information. As a result, it is much easier to study and analyze data in a particular order than to examine large volumes of information without categorization. There are several methods for clustering. The K-means method is one of the data clustering methods in data mining.

This method is a basic method for many other clustering methods (such as fuzzy clustering) although it is a simple method. This method is considered an exclusive method. As a result, this algorithm is used for clustering in this study due to its ease of use. The parameters to be set in the K-means clustering algorithm are given in Table 3.

Figure 1 shows the set parameters, user features, and ignored properties as follows:

Number of repetitions: 250

Number of samples: 1000

Number of clusters: 25

Distance function: Euclidean distance

Number of features: 4

Information about the number of users in each forum, the percentage of users in each forum, and a comparison between accuracy and number of clusters as well as a comparison between the number of clusters and the time of their next analysis to find effective users are provided in this section. In Figure 4, the columns represent the number of each cluster and the rows the percentage of network users in each cluster. Then, the accuracy of the K-means algorithm was calculated for different numbers. The accuracy of this algorithm increases with the number of different clusters, as the number of clusters increases. The number of 25 clusters is relatively accurate among the number of clusters tested.

Despite the help of user profile information-based clustering in IoT networks, it has the disadvantage of identifying groups with common characteristics because individuals are not interested in disclosing personal information publicly for security reasons.

In this section, the cosine similarity measure was used to perform clustering. (A cosine similarity measure for fuzzy sets [26] is defined as the inner product of two vectors divided by the product of their lengths. This is nothing but the cosine of the angle between the vector representations of the two fuzzy sets.) The cosine similarity measure used by Collins et al. is the similarity measure between two vectors of an interior multiplication space based on the cosine of the angle between them. The cosine of angle of zero degrees is equal to one and is less than one for the other angles. Therefore, it is thus a judgment of orientation and not magnitude. Two vectors with the same direction have a cosine similarity criterion equal to one, two vectors with a 90° angle have a similarity of zero, and two vectors with the opposite direction have a similarity of −1.

The cosine similarity measure is used especially in a positive space with a range of [0, 1]. It should be noted that this range is used for any number of dimensions, typically used in positive dimensions with high dimensions.

First, the cosine distance is defined as

$$D_C(A, B) = 1 - S_C(A, B), \tag{3}$$

where A and B are two arbitrary vectors and S is the square of the distance between them.

Table 3: Factors related to set the parameters in MATLAB.

English name	Variable
Max Iteration	Number of iterations for clustering
Num Clusters	Number of clusters
Distance Function	Distance function

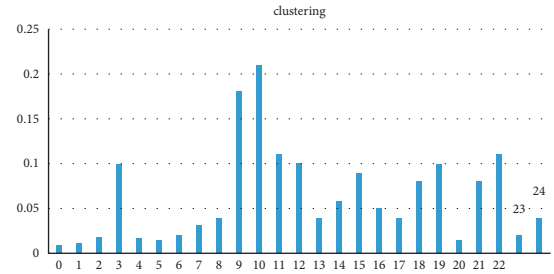


Figure 4: Percentage of users in each cluster.

However, this criterion is not an appropriate distance because it does not guarantee the principle of triangular inequality and goes beyond the principle of superposition. To compensate for the principle of triangular inequality, while maintaining the same order as before, it is necessary to convert it to an angular distance.

The cosine distance of two vectors can be converted into equation (4) with the Euclidean inner product formula:

$$A.B = \|A\|\|B\|\cos\,\theta, \tag{4}$$

where A and B are two arbitrary vectors and θ is the angle between them.

And finally, by applying a series of changes and inner product, the criterion of cosine similarity used in this study can be expressed as

$$\begin{aligned} \text{similarity} = \cos(\theta) &= \frac{A.\,B}{\|A\|\|B\|}, \\ &= \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}, \end{aligned} \tag{5}$$

where A and B are two arbitrary vectors, θ are the angles between them, and Ai and Bi are components of the two vectors. The results of the above formula are between −1 and 1, in which −1 represents completely different, 1 represents completely similar, 0 indicates noncorrelation or orthogonality, and the values between them indicate the degree of similarity or dissimilarity.

After clustering, there is now a regular data set that is divided into 25 categories (clusters). Using the edge calculation algorithm, three larger values are calculated in each cluster, that is, 25 values of A, 25 values of B, and 25 values of C. The largest value of A in every 25 clusters is called the total A. Also, the largest value of B in a total of 25 clusters is called the value of total B, and the largest value of C in a total of 25 clusters is called the value of total C (Table 4).

Table 4: Calculating the values of A, *B*, and *C* in each cluster.

C in each cluster	B in each cluster	A in each cluster	Cluster number
9.73	10.26	12.11	1
8.21	10.53	11.65	2
9.66	10.11	10.94	3
8.29	9.41	10.23	4
8.43	8.59	9.66	5
7.96	8.71	9.58	6
9.98	10.26	10.50	7
9.44	10.02	10.31	8
10.65	12.69	13.62	9
10.89	11.00	11.26	10
9.06	9.41	13.87	11
7.08	7.25	8.31	12
8.21	9.43	9.62	13
7.29	8.71	9.00	14
9.59	10.00	10.13	15
9.36	10.41	11.30	16
5.41	5.41	7.21	17
7.96	8.23	8.97	18
6.50	7.01	7.26	19
9.65	11.43	12.63	20
8.45	10.23	11.85	21
7.69	9.32	10.26	22
9.05	10.02	10.33	23
10.63	11.52	12.54	24
9.032	10.66	12.26	25
		= 13.87	Calculating total *A*, *B*, and C

As can be seen, total *A* (13.87) is created in cluster number 11, total *B* in cluster number 9, and total *C* in cluster number 10.

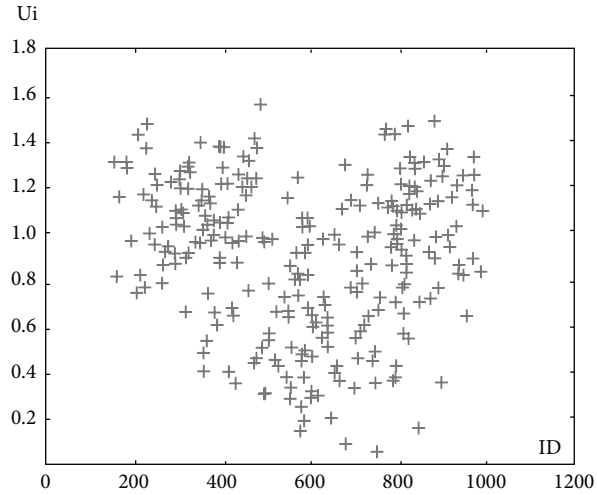


Figure 5: Degree of user interaction in the first cluster.

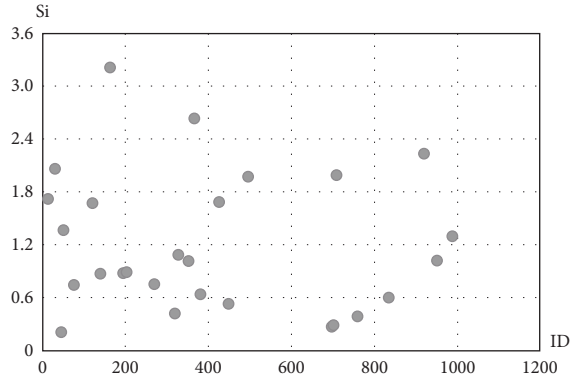


Figure 6: Score of users in the first cluster.

3.7. Degree of User Interaction in Each Cluster. Here, the values of the UIi formula for nodes with different ID are shown. It can be said that whichever has the highest UIi value is selected as the most important node. The amount of user interactions in the first cluster is shown in Figure 5.

3.8. Score of Users in Each Cluster. The score of users was obtained according to the community algorithm using the corresponding formula, for nodes with different IDs. In each cluster, each user with the highest score (Si) is considered the most important and effective nodes. Figure 6 shows the score of users in the first cluster.

3.9. Impact of Users on the Network. The final results of the proposed method for identifying users who are effective in each community are presented in this section, which is a hybrid measure of two criteria of user interaction and the score of the user. Figure 7 shows the users in the first cluster based on the amount of impact (Mi). Given how many users are considered to start diffusion, we can select the users who have the most impact (and are at the top of the chart) from each community and start diffusion through them.

3.10. Calculating the Accuracy of the Proposed Method and Comparing It with Existing Methods. The proposed method has an accuracy of 86.081, which is about 0.2% better than the basic research method with an accuracy of 66.081%. It should be noted that, in the proposed method, the community structure of the IoT is considered, the community is first selected, and then the proposed method was applied; this issue was not considered in the basic research.

The accuracy of the proposed method and the basic research for different numbers of users were compared, and as shown in Figure 8 and Table 5, the accuracy of the proposed method for different numbers has improved compared to the basic research.

It is observed that the accuracy of existing methods to find effective publishers decreases with the increasing number of users in IoT networks.

In order to start diffusion and diffusion information in basic methods, for example, 20 users are selected as initial acceptors to start diffusion information. All selected users may be in specific forums, and other forums may not share from publication; eventually the published messages may never reach them. However, in the proposed method, all forums are considered to start and at least one user is selected in

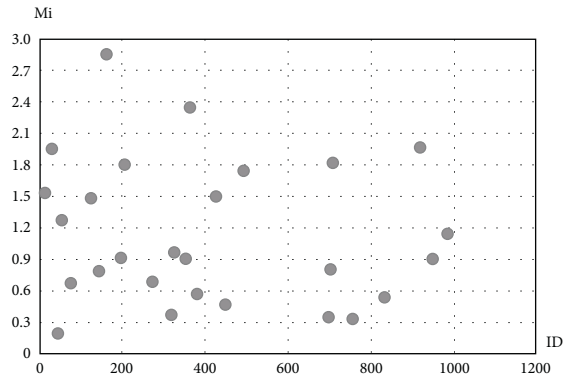


Figure 7: Degree of effectiveness of the first cluster's users.

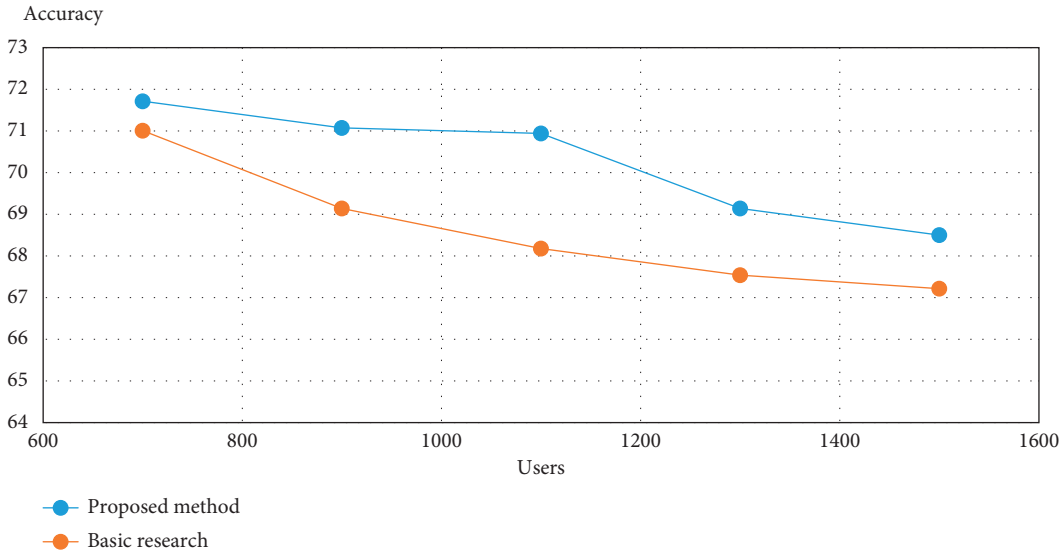


Figure 8: Comparing the accuracy of the proposed method and method [26].

Table 5: Details of comparison accuracy of the proposed method and method [26].

User number	Proposed method (%)	Basic research (%)
800	72	70.9
1000	71	68
1200	70.8	66.5
1400	68	65.5
1500	67	65

each forum, and since this user is more influential and due to the similarity of his characteristics and desires with the users of the forum, this user can publish information and prevent the infiltration of information.

4. Conclusion

The IoMT is a connected infrastructure of medical devices, applications, and health systems and services. The rise of IoMT is driven by an increase in the number of connected medical devices that are able to generate, collect, analyze, or transmit health data or images and connect to healthcare provider networks, transmitting data to either a cloud repository or internal servers. This connectivity between medical devices means that security can no longer be examined within the neat, physical walls as it was considered before. Attacks on IoMT devices can potentially cause significant physical harm and life-threatening damage to the patients.

For addressing this problem, many researches have been done to provide a way for preventing IoMT intrusion and secure routing, in many of which IC and LT diffusion models start the diffusion process from nodes and have no assumptions about starting them. The publication process is not considered, and there is also a need to consider the possibility of accepting and receiving the publication which was done randomly. In other methods, it was based on publication or on popular topics, as they are limited in publication time. In later periods, they may lose their effectiveness and may not be popular in the future. So the methods used by popular users have become more popular for using in IoMT.

In degree-based methods, K-core and other ranking algorithms, including association algorithms, have no presuppositions about networks with association structure and cannot be deduced in networks with association structure. In the proposed method, users were placed in their respective forums, taking into account network forums that are based on individual characteristics in the user profile. Forums based on the characteristics of users' profiles, despite their advantages, are weak because users are reluctant to share their information with the public for security reasons.

Data Availability

Datasets of Mirai and Bashlite can be obtained from the following site: https://archive.ics.uci.edu/ml/datasets/detection_of_IoT_botnet_attacks_N_BaIoT.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. Forghani and S. Hakimpour, "Inform ation discovery in location-based social networks (LBSN) with the aim of providing a spatial recommender system," *Journal of Surveying Science and Technology*, vol. 3, no. 3, pp. 87–100, 2014.

[2] R. Hosseini, "Investigation of user interface elements and components in clustering information retrieval systems," *Information Science and Technology*, vol. 26, no. 1, pp. 40–45, 2011.

[3] M. Mirzaei, J. Hamidzadeh, and S. Savadi, "A new fuzzy method for clustering heterogeneous nodes in wireless sensor networks," in *Proceedings of the International Conference on Nonlinear Systems and Electrical and Computer Engineering Optimization*, Shillong, India, November 2015.

[4] R. Eslami, Samaneh, H. Golpayegani, and S. Alireza, "Identifying associations in social networks based on user behavior and using game theory approach," in *Proceedings of the 8th International Conference on Information and Knowledge Technology, Hamedan, Iran Information and Communication Technology Association*, p. 12, Bu Ali Sina University, Hamedan, Iran, September 1395.

[5] N. Hooshmandnejad and M. Noushfar, "Recognizing the association in the social media graph using structural similarities between individuals," in *Proceedings of the First International Conference on New Perspectives in Electrical and Computer Engineering, Tehran, International Federation of Inventors' Associations (IFIA), Comprehensive University of Applied Sciences*, p. 9, London, UK, July 2016.

[6] F. Sharifzadeh and M. Barari, "Presenting a new index to identify active and influential nodes in social networks using imperialist competitive algorithms," in *Proceedings of the International Conference on Information Engineering and Technology*, UAE-Dubai, Culture and Technology Research Institute, Research Center For Culture, Art, Dalian, China, August 2017.

[7] J. Vernamkhasti, Mohammad, and J. Neda, "Identification and ranking of social networks in Iran and their role in business management," in *Proceedings of the First International Conference on Systems Optimization and Business Management*, Babol, Noshirvani University of Technology - Iranian Association for operations research, Jaipur, India, August 2017.

[8] M. Dastani and M. Mokhtarzadeh, "Motivation to use virtual social networks in Iranian users: a review of texts," in *Proceedings of the 12th Annual Research Congress of Medical Students in the East, Gonabad*, Student Research Committee of Gonabad University of Medical Sciences and Health Services, Gonabad, Iran, September 2017.

[9] A. Saberi, "Graph in social networks," *Journal of Science and Engineering Elites*, vol. 2, no. 2, p. 47, 2017.

[10] H. Mirzarza and M. Javanmard, "Clustering social networks using multilayer perceptron neural network," in *Proceedings of the Third National Conference on Distribution Computing and Big Data Processing*, vol. 10, Azerbaijan Shahid Madani University, Tabriz, Iran, May 2017.

[11] A. B. Lavasani, S. Shamila, and H. Hamidi, "The impact of social networks on information dissemination through viral marketing," in *Proceedings of the Second International Conference on Management and Accounting, Tehran*, vol. 8, Salehan Institute of Higher Education, Bandung, Indonesia, August 2017.

[12] U. J. Lee, "A new mechanism to reduce the error of detecting shill attacks in recommender systems," *Journal of Management the Information Technology, Shuanghai, China*, vol. 4, no. 3, p. 12, 2016.

[13] M. Dosti, "Application of decision tree algorithm in data mining," in *Proceedings of the Fifth International Conference on Computer, Electrical and Electronics Engineering*, vol. 6, Pendar Hamayesh Pars Institute, IUMW University of Kuala Lumpur, Kuala Lumpur - Malaysia, January 2016.

[14] A. Yar and Alireza, "Presenting the problem of optimizing a three-method system and a path generation algorithm based on graph theory to find optimal paths in a social network," in *Proceedings of the Second National Conference on Computer and Electrical Sciences, Hamedan, Permanent Conference Secretariat*, vol. 11, Shahid Mofteh University, Prayagraj, India, October 2014.

[15] M. Donaldson, *Provide a Genetic Algorithm to Identify Hierarchical Clustering in Social Networks, Master's Degree in Computer Engineering*, University of Berlin, Berlin, Germany, 2016.

[16] R. Anderson, "The discovery of information in the location-based social networks with the aim of providing a spatial advisory system," M. S Thesis of Computer- Software, p. 42, Texas University, USA, Austin, TX, USA, 2015.

[17] F. Manson, "Identify effective communication channels in social networks based on message diffusion," *Computer Science Quarterly*, vol. 5, no. 3, pp. 8-9, 2015.

[18] S. Fitz, "Discover communities on social networks by exploring repetitive patterns," MS Thesis of Information Technology, p. 73, Berlin University, Berlin, Germany, 2012.

[19] R. Benson, "Virtual social networks and identity crisis (with an emphasis on identity crisis)," MS Thesis of Information Technology, p. 41, Oxford University, Oxford, UK, 2016.

[20] T. Ashley, "Investigating factors affecting information dissemination in social networks with clustering and artificial neural networks methods," *Computer Engineering Magazine*, vol. 7, no. 2, pp. 12-13, 2016.

[21] S. Alex, "Violation of security in the dissemination of information in social networking associations," *Computer Magazine*, vol. 8, no. 1, pp. 15-16, 2015.

[22] F. Broni, "Investigate the dissemination of information on social networks," MS Thesis of Information Technology, p. 53, Paris University, Paris, France, 2015.

[23] T. Kloni, "Study and review the main data exchange in the intranet networks," *Computer Quarterly*, vol. 4, no. 1, pp. 3-4, 2015.

[24] W. De Nooy, A. Mrvar, and V. Batagelj, Exploratory SocialNetwork Analysis with Pajek, vol 27, Cambridge UniversityPress, Cambridge, UK, 2011.
[25] A. Bayani, *Research Methods in Various Technical and Engineering Sciences*, Amirkabir Publications, Tehran, Iran, 2011.
[26] M. Abdur Rahman and M. Shamim Hossain, "An internet of medical things-enabled edge computing framework for tackling COVID-19," *IEEE Internet of Things Journal*, p. 1, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 1360414, 13 pages
https://doi.org/10.1155/2021/1360414

Research Article

Improved ECG-Derived Respiration Using Empirical Wavelet Transform and Kernel Principal Component Analysis

Shuxin Zhuang,[1,2] **Fenlan Li**,[2] **Zhemin Zhuang**,[1] **Wenbin Rao**,[2] **Alex Noel Joseph Raj**,[1,2] **and Vijayarajan Rajangam**[3]

[1]*Key Laboratory of Digital Signal and Image Processing of Guangdong Province, Shantou, Guangdong, China*
[2]*Department of Electronic Engineering, Shantou University, Shantou, Guangdong, China*
[3]*Centre for Healthcare Advancement, Innovation and Research, VIT, Chennai, Tamil Nadu, India*

Correspondence should be addressed to Alex Noel Joseph Raj; jalexnoel@stu.edu.cn

Received 14 June 2021; Revised 11 September 2021; Accepted 29 September 2021; Published 15 October 2021

Academic Editor: José Alfredo Hernández-Pérez

Many methods have been developed to derive respiration signals from electrocardiograms (ECGs). However, traditional methods have two main issues: (1) focusing on certain specific morphological characteristics and (2) not considering the nonlinear relationship between ECGs and respiration. In this paper, an improved ECG-derived respiration (EDR) based on empirical wavelet transform (EWT) and kernel principal component analysis (KPCA) is proposed. To tackle the first problem, EWT is introduced to decompose the ECG signal to extract the low-frequency part. To tackle the second issue, KPCA and preimaging are introduced to capture the nonlinear relationship between ECGs and respiration. The parameter selection of the radial basis function kernel in KPCA is also improved, ensuring accuracy and a reduction in computational cost. The correlation coefficient and amplitude square coherence coefficient are used as metrics to carry out quantitative and qualitative comparisons with three traditional EDR algorithms. The results show that the proposed method performs better than the traditional EDR algorithms in obtaining single-lead-EDR signals.

1. Introduction

Respiratory signals are important physiological signals commonly used in clinical monitoring. They are used in the detection of sleep apnoea and in stress tests; moreover, they play an important role in the clinical diagnosis of diseases [1]. Respiratory signal detection methods can be divided into two main categories. The first is to detect the air flow from the human nose, and the second is to detect thoracic deformation or the change in thoracic impedance caused by respiration [2]. Both methods require additional sensors and may interfere with natural breathing.

The idea of soft sensors is a one of the solutions to overcome the issues of detecting respiratory signals. Soft sensor is an inferential model that uses easily accessible variables to estimate the variables, which are difficult to be obtained. At present, soft sensors have been widely adopted in industrial processes [3]. The Luenberger observer [4] used state differential equations, with which the dynamic behaviour of the bioprocess is described with a mechanistic model. Yan et al. [5] proposed a framework of data driven soft sensor modeling based on semisupervised regression to estimate the total Kjeldahl nitrogen in a wastewater treatment process.

Obtaining respiratory signals from the ECG is a typical application of soft sensor in the medical field. The ECG signal is obtained noninvasively using a few electrodes and recorded conveniently without interfering natural breath. Respiration affects ECG signals mainly through mechanical interactions and respiratory sinus arrhythmia (RSA) [6]. Mechanical interaction is caused by the displacement of the electrodes relative to the heart and the change in thoracic impedance caused by variations in lung volume [7]. RSA is caused by breath-induced changes in the autonomic nervous system, which in turn causes changes in the heart rate. Heart rate increases during inspiration and decreases during

expiration [8]. Respiration affects the heart rate and ECG in the aforementioned two ways, and such a signal modulation phenomenon forms the theoretical basis for obtaining respiratory signals from ECGs, called ECG-derived respiratory signals.

Owing to the advantages of the EDR algorithm, scientists have conducted multiple studies in this field. Most EDR methods are divided into two categories [9]. One is the EDR method based on the morphological characteristics of the ECG signal. The other is by directly processing the ECG signal. Vargas-Luna et al. [10] obtained the EDR signals through the *R* peak amplitude of ECG signals. Bailón et al. [11] proposed an EDR method based on singular value decomposition of the intervals between the *R* peaks of ECG signals. Chazal et al. [12] obtained EDR signals by calculating the area under the QRS complexes. The EDR methods are based on a single morphological characteristic that provides a rather unsatisfactory accuracy and robustness. Nemati et al. [13] proposed a data fusion method for estimating respiratory frequency based on Kalman filtering, which involves many other physiological signals, and only the respiratory rate can be obtained. Widjaja et al. [14] used kernel principal component analysis to calculate the QRS complexes in the ECG signal and considered the eigenvector as the EDR signal. This method performs well but requires manual deletion of ectopic QRS complexes that involve considerable calculations.

To resolve the limitations of the existing methods and realize an accurate and fully automatic EDR signal obtaining method, an improved EDR algorithm based on EWT and KPCA is proposed. The ECG signal is decomposed to obtain the low-frequency component. Multiple signal decomposition methods, such as wavelet approaches or empirical mode decomposition (EMD) [15], are available at present. However, the disadvantages of this method cannot be ignored. Traditional adaptive wavelet approaches often use prescribe scale subdivision scheme, which is hard to achieve an ideal adaptability. For example, the wavelet packets used a constant prescribe ratio, leading to a limited adaptability. The Brushlet method [16] decomposed the signal on Fourier spectrum, and it is also based on a prescribe subdivisions. EMD shows an ideal adaptability, but its main issue is that it lacks mathematical theory. EWT incorporates the advantages of the above two methods. It not only has rigid mathematical basis but also can decompose signal adaptively.

After using EWT to decompose the ECG signal into five modes, three modes with the lowest frequency are selected to form a new signal. Meanwhile, the *R* peak positions are determined using the Pan–Tompkins algorithm to help locating the QRS complexes. Then, the new signal is sampled based on the position of the QRS complexes. However, a few ectopic samples are captured during sampling. To address this challenge, a method based on variance is developed to delete ectopic samples automatically. Finally, to capture the nonlinear relationship between respiratory and ECG, the processed samples are processed using KPCA and preimaging to obtain the EDR signal. The radial basis function (RBF) is adopted as the kernel function of KPCA; hence, considerable calculations are required when selecting the parameters of the RBF kernel function [17]. Therefore, the parameter selection algorithm is improved in this study to reduce the calculation load.

Our contributions in this paper are as follows. (1) The EDR algorithm framework of EWT + KPCA is proposed to overcome the disadvantages of traditional EDR algorithm based on morphological characteristics of ECG signals, but also captured the nonlinear relationship between respiratory and ECG. (2) A new method based on variance to automatically delete the abnormal samples is introduced during sampling procedure. (3) The selection of RBF kernel parameters in KPCA algorithm is improved to reduce the computational requirement.

The remaining sections of this paper are organized as follows. In Section 2, the EDR algorithm based on EWT and KPCA is described in detail. In Section 3, the proposed method is compared with three traditional EDR algorithms. The qualitative and quantitative experimental results are presented. The results are discussed in Section 4, and conclusions are presented in Section 5.

2. Methodology

The EDR method proposed in this study is divided into two parts and shown in Figure 1. Part 1 involves the decomposition of the ECG signal based on EWT. The ECG signal is decomposed into five modes with different spectral sizes based on EWT; three modes with the lowest frequencies are selected to form a new signal. Part 2 describes the steps for obtaining the EDR signal based on the KPCA. First, the Pan–Tompkins algorithm is used to find the *R* peaks of the ECG signal, and then the QRS complexes are located. Second, the new signal formed by the three modes is sampled based on the locations of the QRS complexes, while some ectopic samples are deleted automatically. These samples serve as the input matrix for KPCA. Third, the input matrix is mapped to a higher-dimensional space through KPCA. The eigenvalues and eigenvectors of the kernel matrix are calculated. Finally, the eigenvector corresponding to the maximum eigenvalue is selected for preimaging to obtain the EDR signal.

2.1. Decomposition of ECG Signal Based on EWT. In general, the human respiratory rate is approximately 0.1–0.5 Hz. To extract the modes with a low-frequency ECG signal completely and adaptively, EWT is used to decompose the ECG signal. The low-frequency modes are reconstructed to form a new signal.

EWT is a mode decomposition algorithm proposed by Gilles [18]. The main concept is to extract the different modes of a signal by designing an appropriate wavelet filter bank, including a low-pass filter and several band-pass filters. The low-pass filter is used to extract the approximate component, and the band-pass filter is used to extract the component details. The number of decomposed modes is selected adaptively in the traditional EWT algorithm. Different ECG signals may be decomposed into different

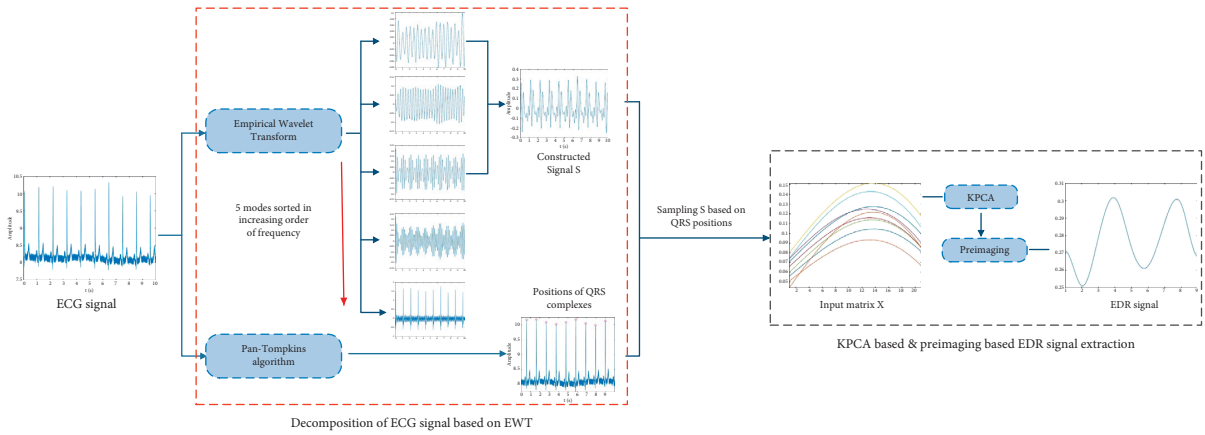


Figure 1: The proposed EDR method based on EWT and KPCA.

numbers of modes, which affects the following calculation. To unify the number of decomposed modes while ensuring the performance of EWT, the number of decomposed modes is set to five based on the experiments. The specific steps to implement the EWT algorithm are as follows:

(1) The local maxima in the spectrum of the ECG signal $f(t)$ are obtained and sorted out in decreasing order after normalization. Next, the first six local maxima are selected, and the boundaries of each mode $\omega_n (n = 1, 2, \ldots, 5)$ are defined as the center of two consecutive maxima.

(2) After determining the boundaries, the empirical scaling function $\varphi_n(\omega)$ and empirical wavelet $\phi_n(\omega)$ are constructed using the Littlewood–Paley–Meyer wavelet [19]. Here, $\varphi_n(\omega)$ and $\phi_n(\omega)$ are expressed as

$$\varphi_n(\omega) = \begin{cases} 1, & |\omega| \le \omega_n(1-\gamma), \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}\left(|\omega| - \omega_n + \tau_n\right)\right)\right], & (1-\gamma) \le |\omega| \le (1+\gamma), \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

$$\phi_n(\omega) = \begin{cases} 1, & (1+\gamma)\omega_n \le |\omega| \le (1-\gamma)\omega_{n+1}, \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}\left(|\omega| - \omega_{n+1} + \tau_{n+1}\right)\right)\right], & (1-\gamma)\omega_{n+1} \le |\omega| \le (1+\gamma)\omega_{n+1}, \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}\left(|\omega| - \omega_n + \tau_n\right)\right)\right], & (1-\gamma)\omega_n \le |\omega| \le (1+\gamma)\omega_n, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $\tau_n = \gamma\omega_n n (0 < \gamma < 1)$. Here, $\beta(x)$ should satisfy the following condition:

$$\beta(x) = \begin{cases} 1 & \text{if } x \le 0 \text{ and } \beta(x) + \beta(1-x) = 1 \\ 0 & \text{if } x \ge 0 \end{cases}, \quad \forall x \in [0, 1]. \tag{3}$$

Since numerical functions satisfy the above condition, we choose $\beta(x) = x^4(35 - 84x + 70x^2 - 20x^3)$ according to [18].

(3) The different modes of $f(t)$ are obtained by $\varphi_n(\omega)$ and $\phi_n(\omega)$. The detail coefficient $W_f^{\varepsilon}(n, t)$ and approximate coefficient $W_f^{\varepsilon}(0, t)$ are defined as

$$W_f^{\varepsilon}(n,t) = f, \psi_1 = \int f\overline{\psi_1(\tau - t)}\mathrm{d}\tau = \left(\hat{f}(\omega)\overline{\psi_n(\omega)}\right)^{\vee}, \tag{4}$$

$$W_f^{\varepsilon}(0,t) = f, \varphi_1 = \int f\overline{\varphi_1(\tau - t)}\mathrm{d}\tau = \left(\hat{f}(\omega)\overline{\hat{\varphi}_1(\omega)}\right)^{\vee}, \tag{5}$$

where $\wedge$ and $\vee$ refer to the Fourier transform and its inverse transform. From equations (1)–(5), the empirical mode $f_k(t)$ can be obtained as

$$\begin{aligned} f_0(t) &= W_f^{\varepsilon}(0,t) * \varphi_1(t), \\ f_k(t) &= W_f^{\varepsilon}(n,t) * \psi_n(t). \end{aligned} \tag{6}$$

After the three steps of processing, the ECG signal is decomposed into five modes. Figure 2 shows the results of the time domain and frequency domain of a 10 s ECG signal after EWT.

As shown in Figure 2(b), the spectra of the five modes are sorted out in increasing order. To extract the low-frequency part of $f(t)$ completely and adaptively, the first three modes are selected to form a new signal. The new signal $f_s(t)$ is shown in Figure 3.

Figure 3 shows that $f_s(t)$ only preserves the low-frequency part and abandons the high-frequency part of $f(t)$, which prevents the influence of high-frequency noise on subsequent calculations. Here, $f_s(t)$ serves as the input for the following KPCA algorithm.

2.2. EDR Signal Acquisition Based on KPCA. KPCA is a generalization, proposed by Scholkopf et al. [20], of principal component analysis in high-dimensional feature space. In KPCA, the data are mapped to a high-dimensional feature space that is nonlinear to the input space. Using KPCA, the EDR acquisition algorithm can describe the nonlinear interaction between the ECG signals and respiratory signals. The steps of KPCA in the proposed method are described in detail in this section.

Before performing the KPCA algorithm, the input matrix X should be determined. The evaluation of X consists of the following steps:

(1) The first step is detection of R peaks: the positions of all the R peaks in $f(t)$ are obtained using the Pan–Tompkins algorithm [21], denoted as $X_R = \{x_i\}_{k=1}^{n}$. The parameter n is the number of R peaks in $f(t)$. The Pan–Tompkins results are shown in Figure 4.

(2) The second step is sampling of the signal $f_s(t)$: after the detection of the R peaks, a fixed window is selected to sample the signal $f_s(t)$. In this study, x_i is regarded as the window center, and $f_s(t)$ is sampled in the range of 40 ms before and after x_i. Then, the samples are used to construct the matrix X' with dimensions $m \times n$, where m is the length of the fixed window and n is the number of R peaks. Because the sampling interval of the ECG signal in this study is 4 ms, the value of m is fixed at 21.

(3) The third step is deletion of ectopic samples: as shown in Figure 5(a), there might be some ectopic samples in X' that affect the accuracy of subsequent calculations. Therefore, an adaptive method based on variance is proposed to delete ectopic samples automatically. The specific steps are as follows:

(1) First, $\{\alpha_i\}_{i=1}^{n}$ is denoted as the result of sampling, and X' can be written as $X' = [\alpha_1, \alpha_2, \ldots, \alpha_k, \ldots, \alpha_n]$. The average sample is defined as

$$\alpha_{\text{mean}} = \frac{1}{n}\sum_{i=1}^{n}\alpha_i. \tag{7}$$

(2) The matrix Y is defined as $Y = [\alpha_1 - \alpha_{\text{mean}}, \ldots, \alpha_k - \alpha_{\text{mean}}, \ldots, a_n - \alpha_{\text{mean}}]$, and the variance of each column vector in Y is calculated. The results are expressed by the vector $V = (v_1, v_2, \ldots, v_k, \ldots, v_n)$.

(3) It is assumed that the ectopic samples are α_p and the normal samples are α_q, and in the equation, the condition, $v_p \gg v_q$, is satisfied. The ectopic samples are removed according to this property, and an input matrix X without ectopic samples is obtained.

The outline of the matrix X is shown in Figure 5(b).

After the input matrix X is determined, KPCA is introduced. The essence of KPCA is to solve the following equation:

$$\lambda v = Cv, \tag{8}$$

where λ and v are the eigenvalues and eigenvectors of matrix C, respectively. Here, $X = [x_1, x_2, \ldots, x_k, \ldots x_r]$, where r is the number of samples in X, and an implicit nonlinear mapping is defined as φ. Then, the mapped data of x_k in the high-dimensional feature space F can be defined as $\varphi(x_k)$. In equation (8), C is the covariance matrix of $\varphi(x_k)$, which is defined as

$$C = \frac{1}{n}\sum_{j=1}^{r}\varphi\left(x_j\right)\varphi\left(x_j\right)^T. \tag{9}$$

Equation (8) is equivalent to the following equation:

$$\lambda\left(\varphi\left(x_k\right) * v\right) = \left(\varphi\left(x_k\right) * Cv\right), \quad k = 1, 2, \ldots, n. \tag{10}$$

When $v = \sum_{i=1}^{r}\alpha_i\varphi(x_i)$ and the RBF is introduced as the kernel function $k(x, y)$,

$$k(x, y) = e^{-\left(x - y^2/2\sigma^2\right)}. \tag{11}$$

After the kernel function is determined, equation (10) can be written as

$$r\lambda\alpha = K\alpha. \tag{12}$$

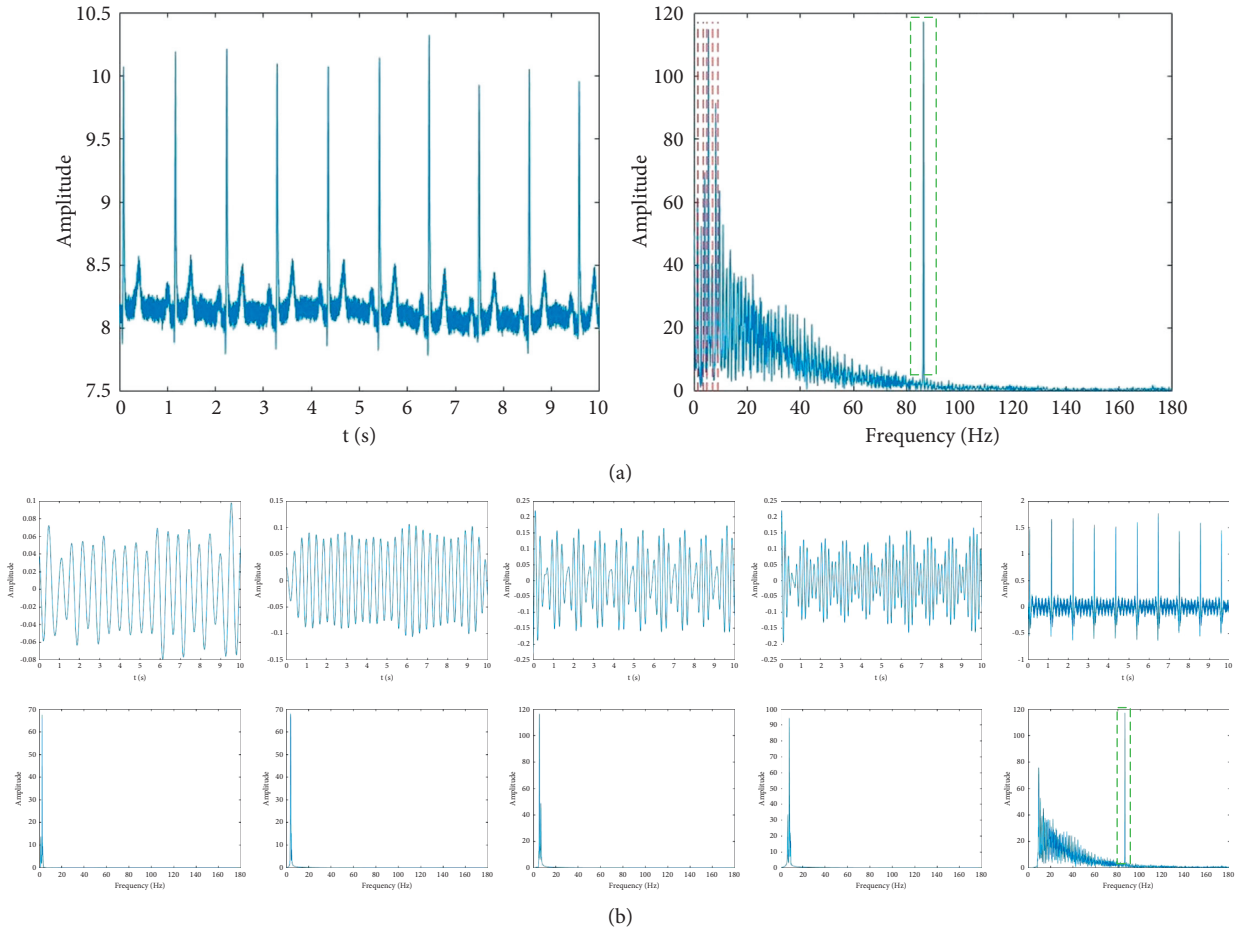


Figure 2: ECG signal of 10 s in time and frequency domains: (a) ECG signal and the corresponding Fourier spectrum, in which the red dotted line in the spectrum graph is the boundary of different modes, and the green boxes represent the high-frequency noise and (b) the five modes of EWT in time and frequency domains.

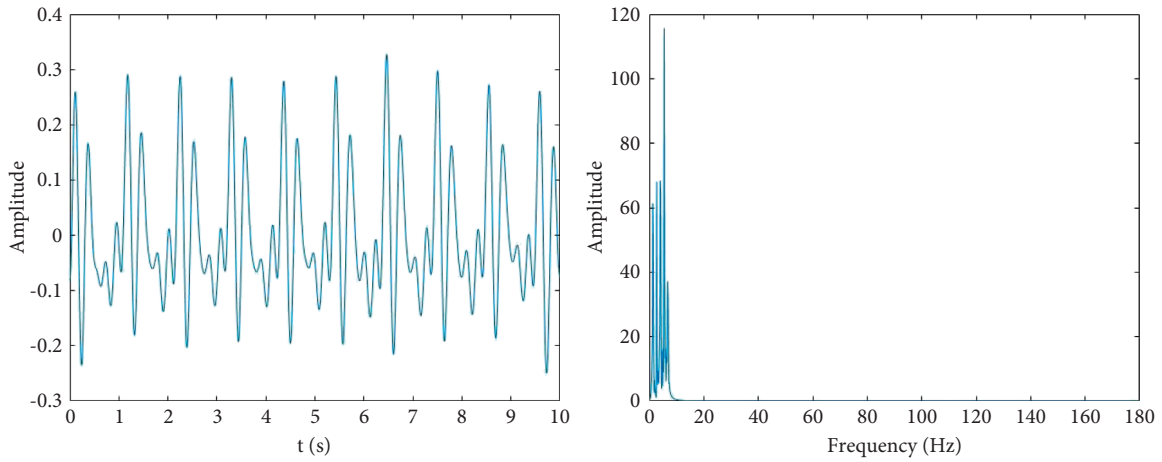


Figure 3: Distribution of $f_s(t)$ in time and frequency domains.

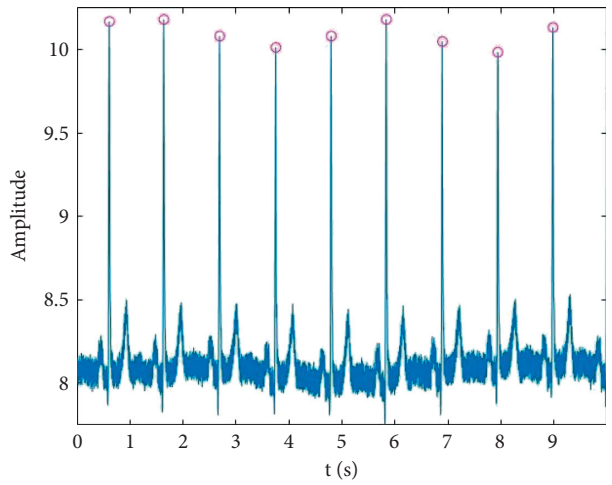


Figure 4: Results of Pan–Tompkins algorithm (the red circles are the locations of the R peaks).

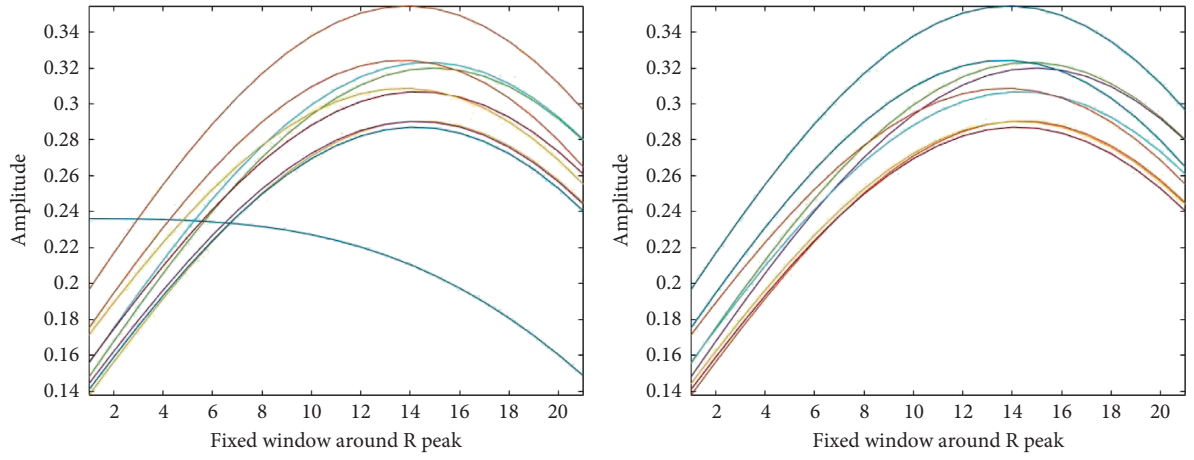


Figure 5: Outline of (a)X' and (b) X: the abscissa is the length of the fixed window.

where α is the vector constituted by parameter α_i and K is the kernel matrix corresponding to $k(x, y)$. To extract the principal component, the projection of a test point $\varphi(x)$ on the eigenvector Vk is calculated using

$$\left(v^k * \varphi(x)\right) = \sum_{i=1}^{r} \alpha_i^k \left(\varphi(x_i)\varphi(x)\right). \tag{13}$$

The aforementioned computation is carried out in the high-dimensional feature space F, whereas the construction of the EDR signal is based on the first eigenvector of the input space. The eigenvalues and eigenvectors obtained in F cannot be directly used for constructing the EDR signal. To solve this problem, a limited number of eigenvectors can be used to find approximations of the data in the input space. This process is called 'preimaging' [22]. Therefore, the first eigenvector of the input space is reconstructed by preimaging the first eigenvector of F. The EDR signal can be obtained by performing cubic spline interpolation on the reconstructed first eigenvector of the input space.

During the process of KPCA, the parameter σ^2 must be carefully selected so that KPCA can deliver a better performance. First, σ^2 is roughly determined as $\mathrm{var}(z)$, which is denoted as $\widehat{\sigma}^2 = \mathrm{var}(z)$. Parameter z represents the vector transformed by X. Then, σ^2 is further tuned in the range, $(0, \widehat{\sigma}^2 \times 100)$, with a step of $\sigma^2/10$. The steps are as follows [14]:

(1) KPCA is applied to the range $(0, \widehat{\sigma}^2 \times 100)$ for σ^2 to obtain the eigenvalues denoted as $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_i)$

(2) Here, $d = \gamma_1 - (\gamma_2 + \cdots + \gamma_i)$ and is calculated for each σ^2. Then, σ^2 is selected, which achieves the maximum d

Although the aforementioned method can determine the appropriate σ^2, it requires high computational effort because the eigenvalues of the kernel matrix are calculated for each σ^2. However, d reaches its maximum early and decreases monotonically thereafter, as shown in Figure 6(a). Thus, the calculation of the monotonically decreasing part is redundant. Therefore, in this study, d is determined when the

aforementioned two steps are done. When d reaches its maximum, the selection process of σ^2 is terminated, as shown in Figure 6(b), and the σ^2 corresponding to the maximum of d is selected for subsequent calculations. Figure 6 shows an example of a curve graph of d. As shown in the figure, if not terminated at the maximum d, the algorithm calculates the eigenvalues of the kernel matrix 1000 times. If terminated at the maximum d, only 180 calculations are required. In this way, not only the accuracy of σ^2 is ensured, but also the computational effort is reduced.

3. Results

In this section, the proposed EDR methods are compared with three traditional EDR methods, including KPCA-based [14], R-peaks-interval-based [23], and R-peaks-amplitude-based [24] EDR methods. The experimental results and the metrics of morphological similarity are presented to evaluate the performance of the aforementioned EDR methods.

3.1. Material. The ECG signals and reference RESP signals were provided by the Fantasia database and Shantou Institute of Ultrasonic Instruments Co., Ltd. (SIUI). The Fantasia database [25] was collected from healthy subjects in a supine posture at a sampling rate of 250 Hz.

3.2. Morphological Similarity Metrics. To measure quantitatively, the morphological similarity between the EDR signal and the reference respiratory signal, the correlation coefficient (C) and magnitude squared coherence coefficient (MSC) were introduced [24]. C is expressed as

$$C = \frac{1/(m-1)\sum_{k=1}^{m-n}\left(R_{\text{ref}}[k] - \overline{R_{\text{ref}}[k]}\right)\left(R_{\text{EDR}}[k+n] - \overline{R_{\text{EDR}}(k+n)}\right)}{\sqrt{1/(m-1)^2\sum_{k=1}^{m}\left(R_{\text{ref}}[k] - \overline{R_{\text{ref}}[k]}\right)^2\sum_{k=1}^{m}\left(R_{\text{EDR}}[k] - \overline{R_{\text{EDR}[k]}}\right)^2}}, \tag{14}$$

where m is the length of the EDR signal and R_{ref} and R_{EDR} represent the reference RESP and EDR signals, respectively. The MSC is defined as

$$\text{MSC}_{xy}(f) = \frac{\left|P_{xy}(f)\right|^2}{P_{xx}(f)P_{yy}(f)}, \tag{15}$$

where $P_{xx}(f)$ and $P_{yy}(f)$ represent the power spectral densities of x and y, respectively, and P_{xy} is the cross-power spectral densities of x and y. The spectra were calculated using Welch's method, a periodic Hamming window, and an overlap of 50%.

3.3. Experimental Results. To compare the proposed method to the three traditional EDR methods in an intuitive manner, some of the experimental results are presented in this section.

As shown in Figure 7, the EDR signal obtained by the algorithm in this study has a competitive similarity to the reference respiratory signal. As Figures 7(f) and 7(g) show, the proposed method performs well in extracting some of the RESP signals with poor quality.

Figure 8 shows that the proposed method performs better than the three traditional EDR methods. In addition, the proposed method achieves a better performance than the three traditional EDR methods in extracting poor-quality RESP signals. Figure 8(d) shows that the proposed method maintains a relatively high morphological similarity with the poor-quality RESP signals. However, the EDR signals obtained by the three traditional EDR methods differ significantly from the reference RESP signals in terms of morphology.

In addition to the aforementioned qualitative comparison results, the performances of the proposed method and three traditional EDR methods were evaluated using C and MSC.

Figure 9 shows the experimental results of the four EDR methods based on the Fantasia database and SIUI. The box plots in Figure 9 specify the median values and interquartile ranges (IQRs). As shown in Figure 8 and Table 1, the EDR methods based on a single morphology characteristic of the ECG signal (RR-interval-based and RPA-based EDR algorithm) show poor results for C and MSC; they also have the disadvantage of poor robustness. Although the KPCA-based EDR method is good, there is still a gap between it and the proposed method in terms of accuracy and robustness. In general, the proposed method is superior to the three traditional EDR methods.

To measure the performance of the proposed method for different age groups, young (21–34 years old) and elderly (68–85 years old) subjects are chosen for the experiment. The experimental results are shown in Figure 10.

As shown in different age groups, methods based on single ECG signal morphological characteristics have the disadvantages of poor robustness and low accuracy. The proposed method exhibits better performance in samples of different ages.

4. Discussion

4.1. Preprocessing. In this study, the ECG signal was decomposed into five different modes, and the three modes with low frequencies were selected to construct a new signal for EDR signal acquisition. No denoising algorithm is introduced in the proposed method. This is because the RESP signal with a relatively low frequency causes a baseline drift to the ECG signal. The correction of baseline drift by the denoising algorithm affects the extraction of the RESP signal to a certain extent. The influence of high-frequency noises is

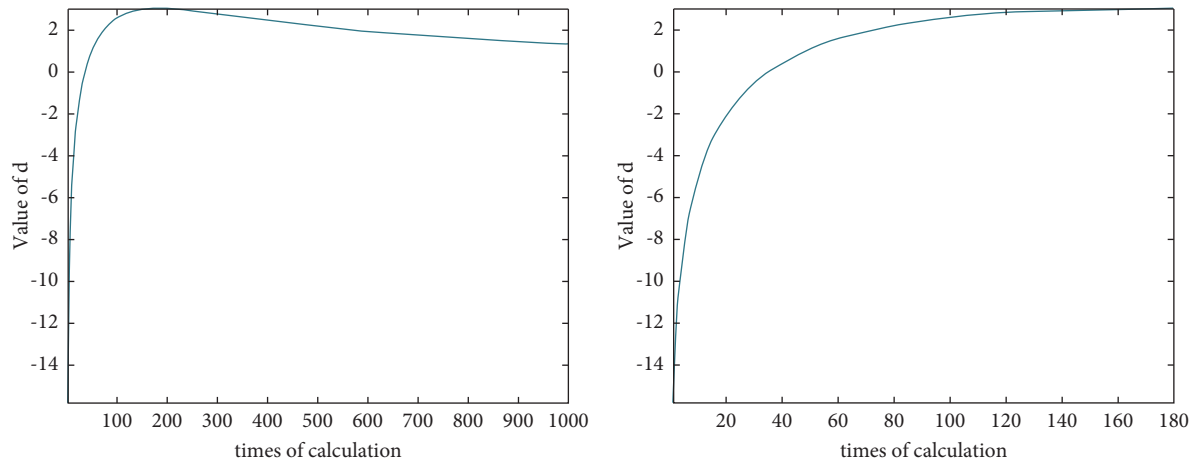


Figure 6: Curve graphs of d: (a) applying KPCA to the range $(0, \widehat{\sigma}^2 \times 100)$ without terminating at the maximum of d and (b) terminating at the maximum of d.

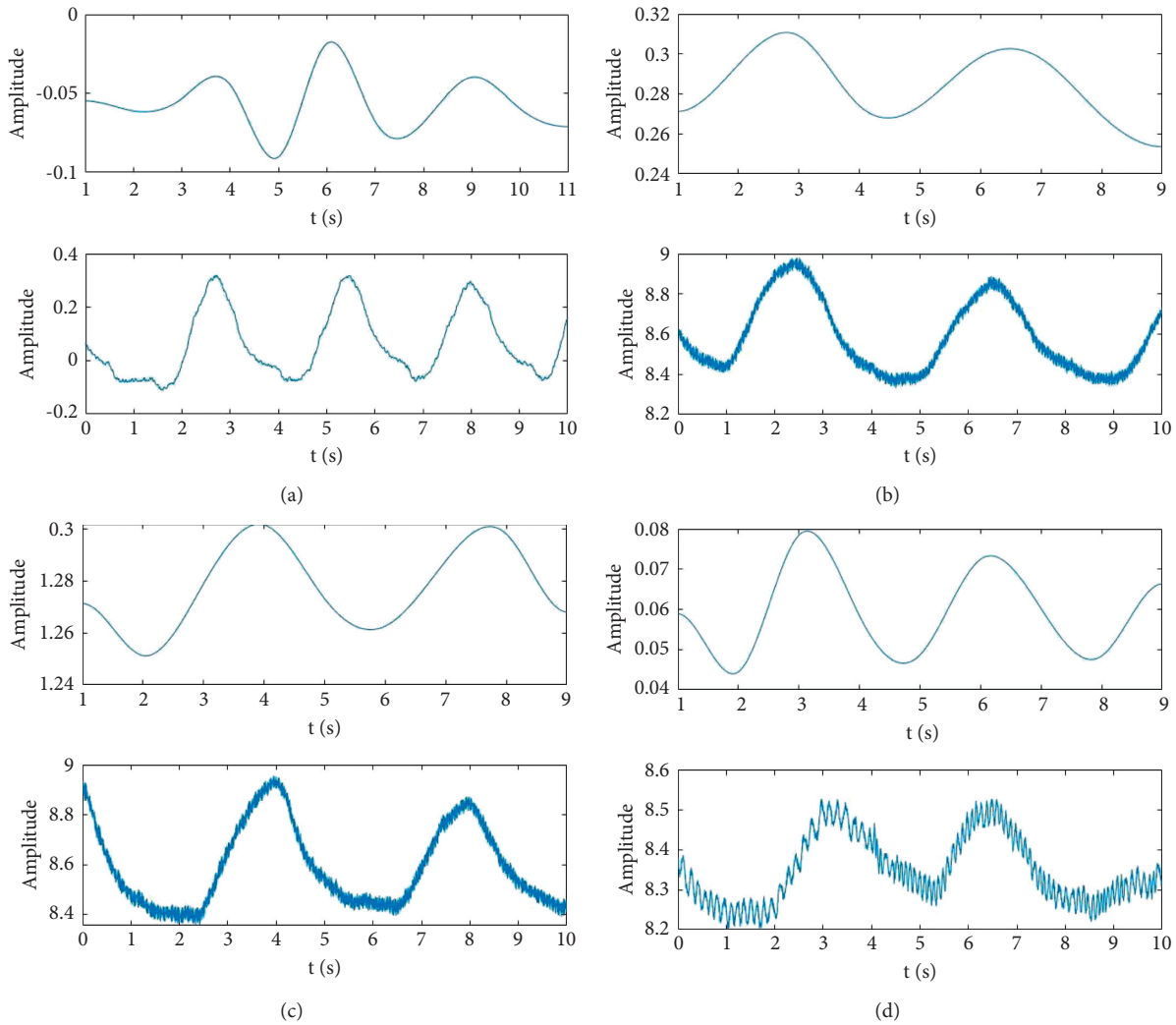


Figure 7: Continued.

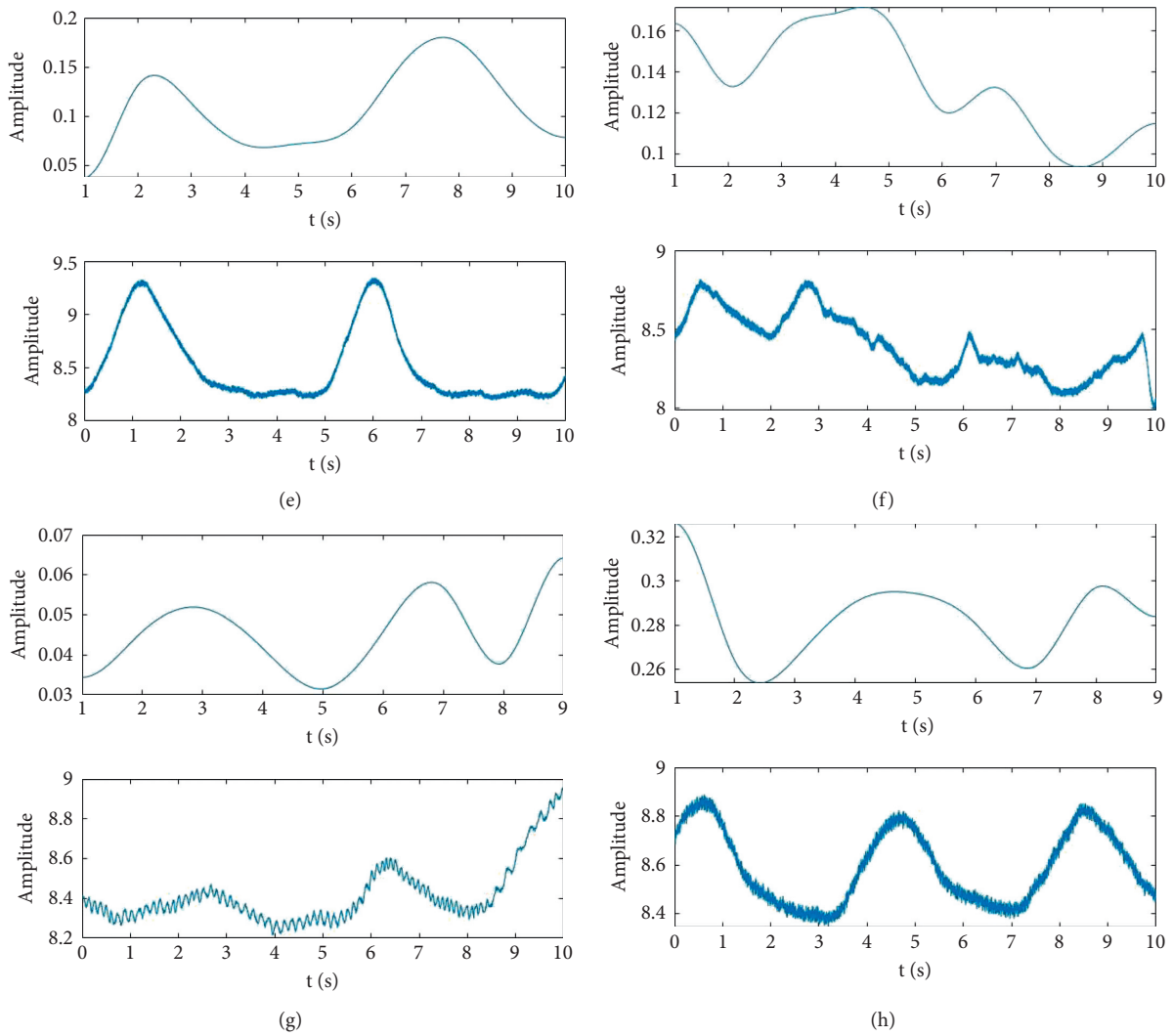


Figure 7: Comparison of the EDR signal obtained by the proposed method and the reference RESP signal. In each picture, the subplot above is the EDR signal, and the subplot below is the reference RESP signal.

also mitigated because the three modes with lower frequency are chosen for subsequent calculation, as mentioned in Section 2.1.

4.2. The Effect of the Number of the Extracted Modes. In this paper, EWT is introduced to adaptively decompose the ECG signal into 5 modes, and low-frequency modes are reconstructed to form a new signal. In this section, we carried out experiments on the effect of the number of modes extracted from EWT. Here, 30 ECG signals with a length of 10 s were randomly selected and are decomposed into 4, 5, and 6 modes, respectively. C and MSC are introduced to evaluate the performance. The results are presented in Figure 11, revealing the better performance when the number of modes extracted from EWT is 5.

4.3. Computational Effort. In general, EDR methods involving KPCA are computationally intensive. To evaluate the extra computational effort of the proposed method with respect to the three traditional EDR methods, the computation time is recorded. Because all the mentioned EDR methods must locate the R peak through the Pan–Tompkins algorithm, the time required by the PTK algorithm is deducted. This experiment is based on Intel Core i7-9750H 2.60 GHz and runs on MATLAB 2018A. Five ECG signals with a length of 10 s are randomly selected from the Fantasia database for five tests. The results show that the average computation time of the proposed method and the EDR algorithm based on KPCA, R peaks-interval, and RPA is 0.161 s, 0.312 s, 0.006 s, and 0.013 s, respectively. The data reveal that the EDR method based on the morphological characteristics of ECG signals has a higher computational speed than the EDR method using KPCA, but at the expense of accuracy and robustness. There are two main reasons for this difference in computational complexity:

(1) The EDR algorithms based on the morphological characteristics of ECG signals only process the specific characteristics (such as R wave amplitude

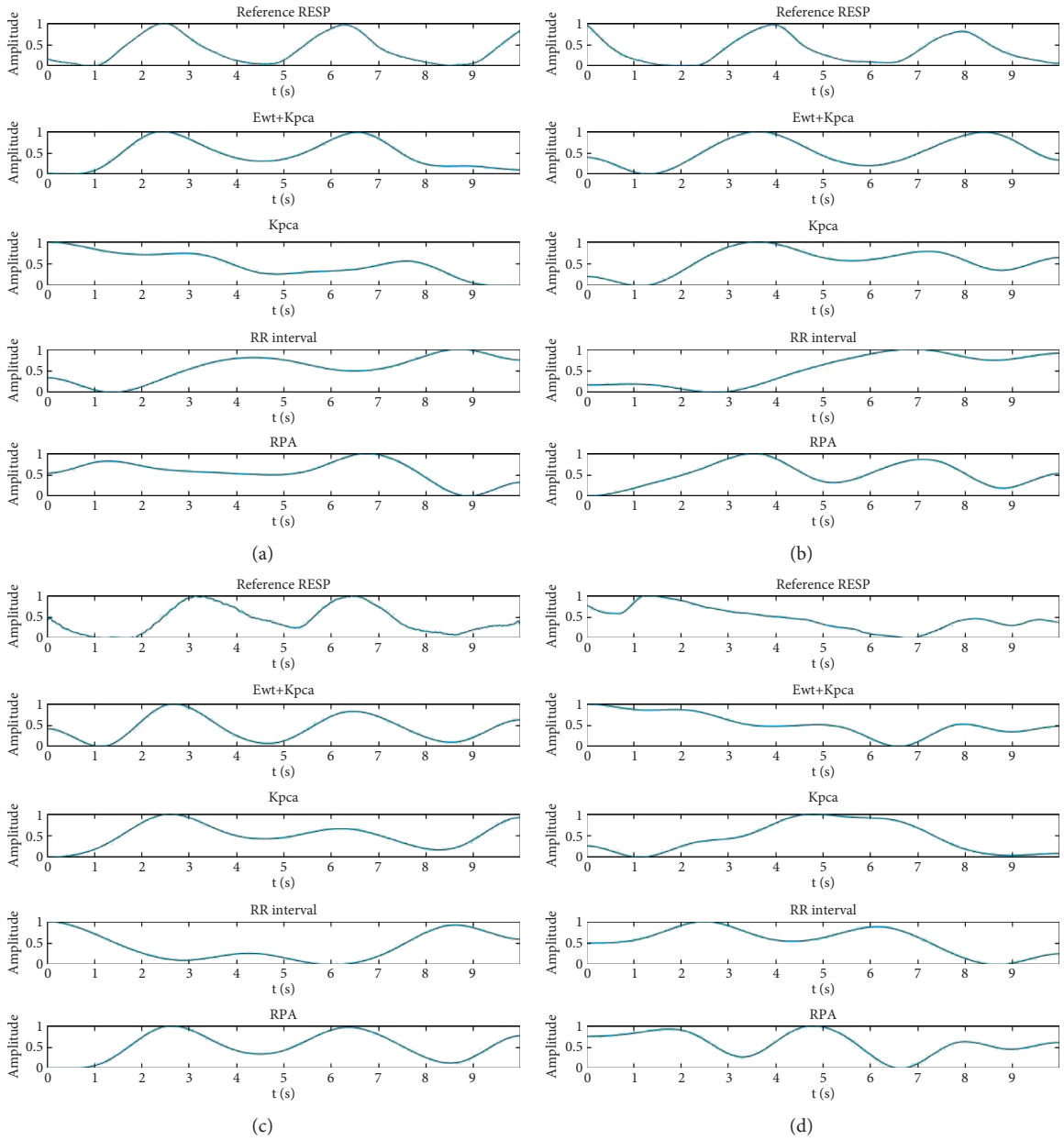


Figure 8: Performance of the proposed method compared with those of the KPCA-based, R-peaks-interval-based (RR interval), and R-peaks-amplitude-based (RPA) EDR methods.

and R wave interval) of each cardiac cycle, while the EDR algorithm using KPCA should process the entire QRS complex of each cardiac cycle. There is a huge gap in the amount of data to be processed by the two methods.

(2) KPCA algorithm is complex. In addition, repeated calculation is required when choosing the parameter δ of RBF kernel function. It is found through experiments that in traditional KPCA-based EDR algorithm, the computational time of searching for appropriate parameter δ accounts for more than 90% of the total computational time of KPCA.

In the proposed method, the average computation time of the KPCA part is only 0.074 s owing to the mechanism of determining parameter d, whereas the average time required to compute the KPCA part in the traditional KPCA-based EDR algorithm is 0.267 s. The time saving rate reached 72.3%, which significantly reduces the computational effort.

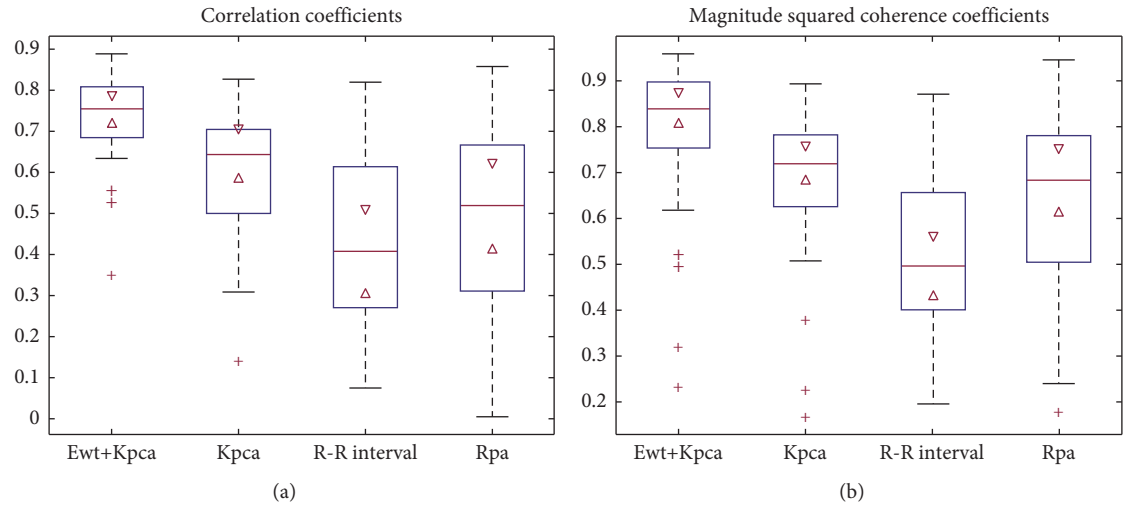


Figure 9: Comparison between the proposed method and three traditional EDR methods: (a) correlation coefficients and (b) magnitude squared coherence coefficients.

Table 1: C, MSC, and IQR of the four EDR methods.

Method	EWT + KPCA	KPCA	RR interval	RPA
C	0.730	0.600	0.439	0.485
IQR (C)	0.685–0.808	0.706–0.498	0.6130–0.267	0.311–0.665
MSC	0.784	0.680	0.516	0.655
IQR (MSC)	0.755–0.897	0.626–0.782	0.657–0.400	0.504–0.780

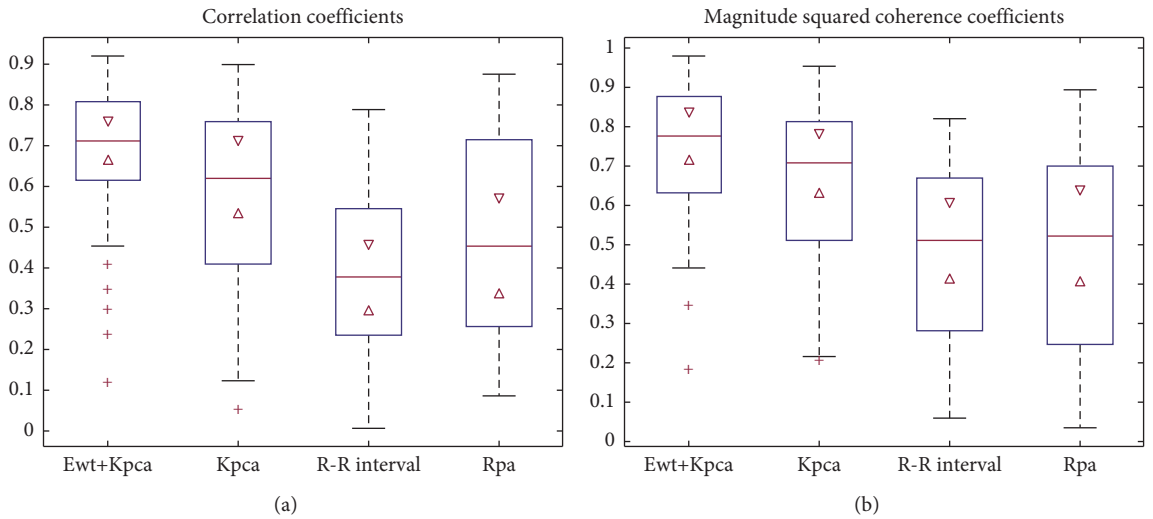


Figure 10: Continued.

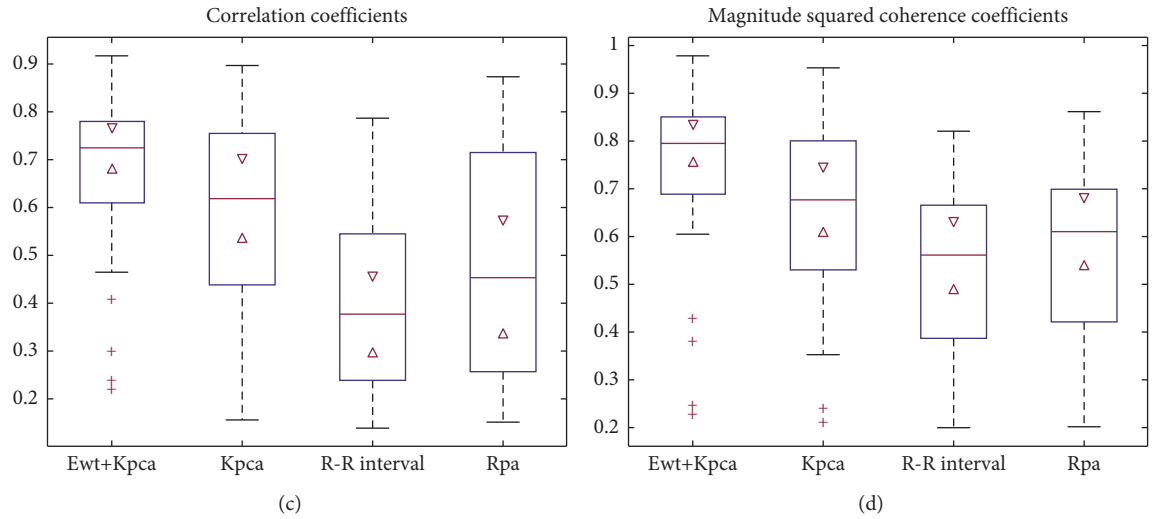


Figure 10: Comparison of four EDR algorithms based on ECG signals of different ages: (a, b) the results of elderly samples and (c, d) the results of young samples.

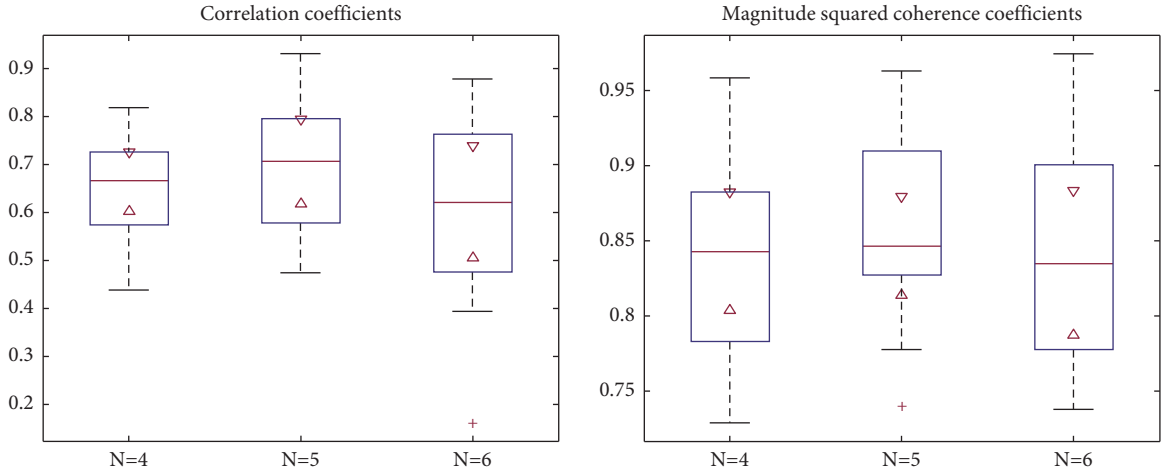


Figure 11: Comparison of different numbers (N) of the modes extracted from EWT.

5. Conclusions

The proposed EDR method based on EWT and KPCA shows good performance than the traditional EDR methods in the extraction of EDR signals from single-lead-ECG signals. The ECG signal is decomposed into five different modes through the EWT, and a new signal is formed by constructing the three components with a low frequency. Then, the new signal is sampled to form the input matrix based on the location of the QRS complex, and an ectopic sample removal method is used to delete the ectopic samples. Subsequently, KPCA is used to obtain the eigenvectors and eigenvalues. Finally, the EDR signal can be obtained by processing the results using preimaging and cubic spline interpolation. After the selection method of the parameters of the RBF kernel in KPCA is improved, the computation time is significantly reduced, alleviating the problem of high computational effort in the EDR method with KPCA to a certain extent. Experimental results show that the proposed method performs better than the three traditional EDR methods and is suitable for monitoring respiration through single-lead-ECG signals without additional sensors.

Data Availability

Part of the data used in this paper can be found in the website https://physionet.org/about/database/, and the other part is the data provided by the SIUI, which is not open to the public because it involves privacy.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

The authors would like to thank Mr. Bin Li of the Shantou Institute of Ultrasonic Instruments Co., Ltd. (SIUI) for his contributions to the data and professional suggestions. This work was supported by the Basic and Applied Basic Research Foundation of Guangdong Province (Grant no. 2020B1515120061), National Natural Science Foundation of China (Grant no. 82071992), the Guangdong Province University Priority Field (Artificial Intelligence) Project (Grant no. 2019KZDZX1013), and National Key R&D Program of China (Grant no. 2020YFC0122103).

References

[1] H. Singh, R. K. Tripathy, and R. B. Pachori, "Detection of sleep apnea from heart beat interval and ECG derived respiration signals using sliding mode singular spectrum analysis," *Digital Signal Processing*, vol. 104, 2020.

[2] H. Sharma and K. K. Sharma, "ECG-derived respiration based on iterated Hilbert transform and Hilbert vibration decomposition," *Australasian Physical and Engineering Sciences in Medicine*, vol. 41, no. 2, pp. 429–443, 2018.

[3] S. Guo, X. Zhang, Y. Du, Y. Zheng, and Z. Cao, "Path planning of coastal ships based on optimized DQN reward function," *Journal of Marine Science and Engineering*, vol. 9, no. 2, p. 210, 2021.

[4] D. Krämer and R. King, "A hybrid approach for bioprocess state estimation using NIR spectroscopy and a sigma-point Kalman filter," *Journal of Process Control*, vol. 82, pp. 91–104, 2019.

[5] W. Yan, P. Guo, Y. Tian, and J. Gao, "A framework and modeling method of data-driven soft sensors based on semisupervised Gaussian regression," *Industrial and Engineering Chemistry Research*, vol. 55, no. 27, pp. 7394–7401, 2016.

[6] T. Jin, H. Ding, H. Xia, and J. Bao, "Reliability index and Asian barrier option pricing formulas of the uncertain fractional first-hitting time model with Caputo type," *Chaos, Solitons and Fractals*, vol. 142, Article ID 110409, 2021.

[7] M. Deviaene, I. D. Castro, P. Borzee et al., "Capacitively-coupled ECG and respiration for the unobtrusive detection of sleep apnea," *Physiological Measurement*, vol. 42, 2021.

[8] A. Zarei and B. M. Asl, "Automatic classification of apnea and normal subjects using new features extracted from HRV and ECG-derived respiration signals," *Biomedical Signal Processing and Control*, vol. 59, 2020.

[9] W. Deng, S. Shang, X. Cai et al., "Quantum differential evolution with cooperative coevolution framework and hybrid mutation strategy for large scale optimization," *Knowledge-Based Systems*, vol. 224, Article ID 107080, 2021.

[10] J. L. Vargas-Luna, W. Mayr, and J. A. Cortés-Ramírez, "Amplitude modulation approach for real-time algorithms of ECG-derived respiration," *Revista Mexicana de Ingeniería Biomédica*, vol. 35, pp. 53–69, 2014.

[11] R. Bailón, L. Sornmo, and P. Laguna, "A robust method for ECG-based estimation of the respiratory frequency during stress testing," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 7, pp. 1273–1285, 2006.

[12] P. De Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 6, pp. 686–696, 2003.

[13] S. Nemati, A. Malhotra, and G. D. Clifford, "Data fusion for improved respiration rate estimation," *EURASIP Journal on Applied Signal Processing*, vol. 2010, Article ID 926305, 2010.

[14] D. Widjaja, C. Varon, A. Dorado, J. A. K. Suykens, and S. Van Huffel, "Application of kernel principal component analysis for single-lead-ECG-derived respiration," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 1169–1176, 2012.

[15] T. Jin, X. Yang, H. Xia, and H. Ding, "Reliability index and option pricing formulas of the first-hitting time model based on the uncertain fractional-order differential equation with caputo type," *Fractals*, vol. 29, no. 1, Article ID 2150012, 2021.

[16] F. G. Meyer and R. R. Coifman, "Brushlets: a tool for directional image analysis and image compression," *Applied and Computational Harmonic Analysis*, vol. 4, no. 2, pp. 147–187, 1997.

[17] W. Deng, J. Xu, X.-Z. Gao, and H. J. I. T. O.S. Zhao, "An enhanced MSIQDE algorithm with novel multiple strategies for global optimization problems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020.

[18] J. Gilles, "Empirical wavelet transform," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 3999–4010, 2013.

[19] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, USA, 1992.

[20] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[21] K. K. Kim and A. rana, "Cardiac disease detection using modified PanTompkins algorithm," *Journal of Sensor Science and Technology*, vol. 28, pp. 13–16, 2019.

[22] A. García-González, A. Huerta, S. Zlotnik, and P. J. A. P. A. Díez, "A kernel principal component analysis (kPCA) digest with a new backward mapping (pre-image reconstruction) strategy," *Research Square*, 2020, https://arxiv.org/abs/2001.01958.

[23] N. Sadr and P. de Chazal, "A comparison of three ECG-derived respiration methods for sleep apnoea detection," *Biomedical Physics and Engineering Express*, vol. 5, no. 2, Article ID 025027, 2019.

[24] C. Varon, J. Morales, J. Lázaro et al., "A comparative study of ECG-derived respiration in ambulatory monitoring using the single-lead ECG," *Scientific Reports*, vol. 10, pp. 1–14, 2020.

[25] N. Iyengar, C. K. Peng, R. Morin, A. L. Goldberger, and L. A. Lipsitz, "Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 271, no. 4, pp. R1078–R1084, 1996.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 4196241, 21 pages
https://doi.org/10.1155/2021/4196241

Review Article

Osteolysis: A Literature Review of Basic Science and Potential Computer-Based Image Processing Detection Methods

Soroush Baseri Saadi,[1] Ramin Ranjbarzadeh,[2] Ozeir kazemi,[3] Amir Amirabadi,[1] Saeid Jafarzadeh Ghoushchi,[4] Oveis Kazemi,[5] Sonya Azadikhah,[6] and Malika Bendechache[7]

[1]*Department of Electrical Engineering, Islamic Azad University, South Tehran Branch, Tehran, Iran*
[2]*Department of Telecommunications Engineering, Faculty of Engineering, University of Guilan, Rasht, Iran*
[3]*PPD - Global Pharmaceutical Contract Research Organization, Central Lab, Zaventem, Belgium*
[4]*Faculty of Industrial Engineering, Urmia University of Technology, Urmia, Iran*
[5]*Mehr Parand Clinic, Parand, Tehran, Iran*
[6]*R.E.D. Laboratories N.V./S.A., Z.1 Researchpark, Zellik, Belgium*
[7]*School of Computing, Faculty of Engineering and Computing, Dublin City University, Dublin, Ireland*

Correspondence should be addressed to Ramin Ranjbarzadeh; ranjbar.ramin24@gmail.com

Received 10 June 2021; Revised 30 July 2021; Accepted 14 September 2021; Published 4 October 2021

Academic Editor: Navid Razmjooy

Osteolysis is one of the most prominent reasons of revision surgeries in total joint arthroplasty. This biological phenomenon is induced by wear particles and corrosion products that stimulate inflammatory biological response of surrounding tissues. The eventual responses of osteolysis are the activation of macrophages leading to bone resorption and prosthesis failure. Various factors are involved in the initiation of osteolysis from biological issues, design, material specifications, and model of the prosthesis to the health condition of the patient. Nevertheless, the factors leading to osteolysis are sometimes preventable. Changes in implant design and polyethylene manufacturing are striving to improve overall wear. Osteolysis is clinically asymptomatic and can be diagnosed and analyzed during follow-up sessions through various imaging modalities and methods, such as serial radiographic, CT scan, MRI, and image processing-based methods, especially with the use of artificial neural network algorithms. Deep learning algorithms with a variety of neural network structures such as CNN, U-Net, and Seg-UNet have proved to be efficient algorithms for medical image processing specifically in the field of orthopedics for the detection and segmentation of tumors. These deep learning algorithms can effectively detect and analyze osteolytic lesions well in advance during follow-up sessions in order to administer proper treatments before reaching a critical point. Osteolysis can be treated surgically or nonsurgically with medications. However, revision surgeries are the only solution for the progressive osteolysis. In this literature review, the underlying causes, mechanisms, and treatments of osteolysis are discussed with the main focus on the possible computer-based methods and algorithms that can be effectively employed for the detection of osteolysis.

1. Background

Basically, osteolysis is a biological process initiated by induced particles at the interface of bone metal or bone cement of prosthetic implants which is radiographically manifested as linear endosteal radiolucencies or scalloped focal. This phenomenon in long-term results in bone loss, periprostatic fractures, and finally loosening of implants. From early observations, it was generally believed that, in cemented implants, osteolysis is due to acrylic cement and the term "cement disease" was introduced because of this belief. Nevertheless, osteolysis is now defined as "particle disease" with the demonstration of lytic lesions in implants without cement in the interface [1–3].

Osteolysis is one of the main causes for late reoperation in patients with total joint arthroplasty (TJA). This operation is the most effective therapeutic solution for patients enduring end-stage degenerative arthritis. It is believed that the demands for TJAs will gradually increase worldwide in subsequent years where total hip arthroplasty (THA) and total knee arthroplasty (TKA) are the most frequents types. Although the lifetime of THA has improved, such that approximately 90% of the implants function appropriately at 15 years [4], osteolysis is the main cause of at least 50% of all THA revision surgeries based on the majority of national registries. Despite the fact that the effectiveness of surgery may depend on various variables including the level of physical activity of patients, type of implant material, time of THA in service, model of the implant and placement [5], more than 100,000 patients for each million THA procedures might experience a revision surgery within a 15-year period of service. Osteolysis has been noted in both cemented and cementless implants, with 0% to 16% for cemented and 6% to 30% for cementless TJA.

It is indicated by clinical experience that there is a strong correlation between the likelihood of osteolysis and the magnitude of polyethylene [6]. The majority of the wear particles are ultrahigh molecular weight polyethylene (UHMWPE) which is mainly implicated in the development of lesions [7–9]. The appearance of osteolysis at shorter times of follow-up will become more apparent as the higher wear rates cause higher rates of osteolysis [10] (Figure 1). Considering a uniform distribution of wear particles through the periprosthetic tissue, the wear volume can be directly related to the number of particles per unit volume of tissue [11]. Nevertheless, it is improbable that the distribution of wear particles is uniform as the permeability of tissues varies, and there are also a small number of pathways for particle access. Local osteolysis can be produced as a result of local accumulation of wear debris even if the general density of particles is low. This makes it difficult to determine the relationship of appearance of osteolysis to wear.

Osteolysis is a progressive medical condition. Thus, regular follow-up sessions are advised. During these sessions, orthopedists check the progress of the condition with the use of different imaging modalities depending on the severity of the condition. Nevertheless, regardless of imaging techniques, detection and analysis of osteolysis are difficult tasks to perform as osteolytic lesions are not easy to be distinguished especially at the first years after implantation. This is the main reason that researchers and biomedical engineers in the field of medical image processing are making a great effort to devise new methods and algorithms to overcome these barriers.

In recent years, artificial intelligence (AI) has become highly popular and is developing into every aspect of the modern life by its advances in large-data retrieval and explicit evaluation of features that are ideal for medical image processing [12–15]. In comparison to traditional image processing methods, deep learning is more efficient and reliable, as it can automatically extract features of the images instead of hand-crafted features [16, 17]. Thanks to the deep neural networks (DNNs), the methods of computation can allow an algorithm to self-programme through learning from a large set of examples that show the desired behavior, eliminating the need for specifying the rules explicitly [18,19]. Deep learning (DL) techniques have developed fast and have been shown to exceed the performance of human beings. Deep convolutional neural networks (DCNNs) demonstrate a great benefit in image processing. CNN is a subclass of deep, feedforward neural networks, in which image data moves in a forward direction from the input to the output nodes. The advantageous feature of CNNs is that they can learn directly useful image features and other structured data, whereas the task of feature extraction has been carried out by machine learning models or by hand before CNNs. These types of neural network architecture proved to be powerful deep learning models in the field of image analysis, judging from the existence of specific features in their structure [20–22]. A typical CNN is composed of one or several filters that are called filters, kernels, or convolutional layers, accompanied by several layers (an aggregation and pooling layer) that are employed for classification purposes. As the characteristics of CNN are similar to those of the standard artificial neural network (ANN), backpropagation and gradient descent are used for training tasks, while it is comprised of additional pooling layers along with kernels. The final results are derived from the vector that is sited at the end of the network architecture. CNNs can have various structures and methods. In medical image analysis, the most popular structures are U-Net [23], SegNet [24], and V-Net [25] as well as the conventional CNN structures. Additionally, the most common CNNs methods are the following: ZFNet (2013) [26], VGGNet-16 (2014) [27], AlexNet (2012) [28], GoogLeNet [29], DenseNet (2017) [30], and ResNet (2015) [31]. There are several studies that prove the effectiveness of CNN algorithms in medical image processing, such as breast cancer detection [30], diagnosis of breast ultrasonography images [32], liver cancer detection and segmentation [33], differentiation of liver masses [34], lung infection segmentation [35,36], and classification of interstitial lung disease [37]. However, one key drawback of most deep learning approaches is that they need a lot of training data to learn features properly.

The intent of this study was not to investigate comprehensively the biological and medical aspect of osteolysis. An attempt was made, however, to give a guideline for the main factors and issues involved in this medical phenomenon and to introduce computer-based methods and algorithms especially in the field of artificial neural networks used successfully for similar purposes.

2. Discussion

2.1. The Biological Mechanisms of Osteolysis. Osteolysis is an active biological response to particulate wear debris [38], primary bone tumors [39], and metastatic diseases [40]. In fact, this phenomenon is a cell-mediated biological process leading to bone loss as a direct reaction of stimulation of macrophages by biologically active particles. It is noted that the mechanical wear of the articulating surface that releases wear debris begins this process. This key debris is

Figure 1: Periprosthetic femoral osteolysis caused by wear of an extended-chain crystallite polyethylene [10].

phagocytosed, which activates osteoclasts and macrophages and results in resorption of the bones [3] (Figure 2).

The cells which are mainly involved in creating a response to the particles are fibroblasts, lymphocytes, macrophages, osteoclasts, osteoblasts, and foreign body giant cells. The main cellular mediator of osteolysis is macrophages that take part individually in the resorption of bones by transforming into osteoclasts leading to much faster resorption. These cells secrete a variety of cytokines and are present in the pseudomembrane. Among them, the most important cells are interleukin 6, interleukin 1α, interleukin 1β, TNF-α, prostaglandin E2, RANKL, gelatinase, and collagenases [1, 41–43].

Generally, we can distinguish between nonwear-related and wear-related byproducts. At each stage, wear particles are released from the surface of the softer material by abrasion and adhesion of bearing surface [44]. During the time of service, it is recognized that all lower-limb arthroplasties generate wear particles [45]. Based on the biotribological experiments, materials which are hard similar to ceramics that create nanosized wear particles can cause inflammation [46]. Bone cement debris and metallic debris were also noted to have relation to hypersensitivity and inflammation [47–49]. Nevertheless, there is limited evidence on the influence of the prosthetic particle dimension, surface charge, shape, and osteoclast capacity. The experimental evidence also demonstrates that erosive bone resorption can cause prosthetic joint infection if not diagnosed and treated early [50–52].

It has also been recognized that local hydrodynamics factors can contribute to osteolysis. Local fluid pressure gradients around implants are considered to create a mechanism for the particle to transport and force fluid and to surround the bone. Very high intra-articular pressures, caused by fluid pressure, changes during motion can mount osteocyte death and subsequently induce osteoclast bone resorption [53].

Finally, it is also noted that the type of metal debris has a profound consequence for the extent of osteolysis. Titanium has a greater destructive stimulatory impact than cobalt-chromium (CoCr). Early death of the macrophages caused by CoCr particles reduces the inflammatory mediators, leading to osteolysis. On the other hand, titanium debris is less toxic to the macrophage, thus triggering upper levels of the inflammatory mediation [54]. In addition, experiments have demonstrated that the mean particle dimension of less than 1.7 ± 0.7 um of wear particulate debris will increase the rate of osteolysis after THA [55].

2.2. Types of Osteolysis. Generally, there is no agreement on the classification of osteolysis as different methods of measurements are used in experiments and analysis. Although many studies have been conducted on this phenomenon, normally the incidence is reported. In fact, for the classification of osteolysis, lesion volume is measured. This is carried out by measuring the area of the osteolytic regions on lateral and anteroposterior radiographs. In some cases, only one view is used. On plain radiographs, normally femoral lesions can be easily viewed, whereas this is not the case for acetabulum lesions. Revision surgeries have shown that the lesions behind the acetabulum are larger than the dimension expected from preoperative radiographs [56, 57].

Not all osteolytic lesions (softened section of a patient's bone) lead to failure although they are considered as an unfavourable medical issue and are a predictive factor for later adverse consequences. There are some osteolytic lesions that are stable and do not expand, whereas there are others which have a balloon-like features and can expand and propagate [58, 59]. Osteolytic lesions can diffuse or be localized. The diffuse lesion can progress along the interface or be stable. Depending on the location, the prognosis for a lesion can vary. Although the lesions behind the acetabular component may not progress, they are naturally expansive at this location [60].

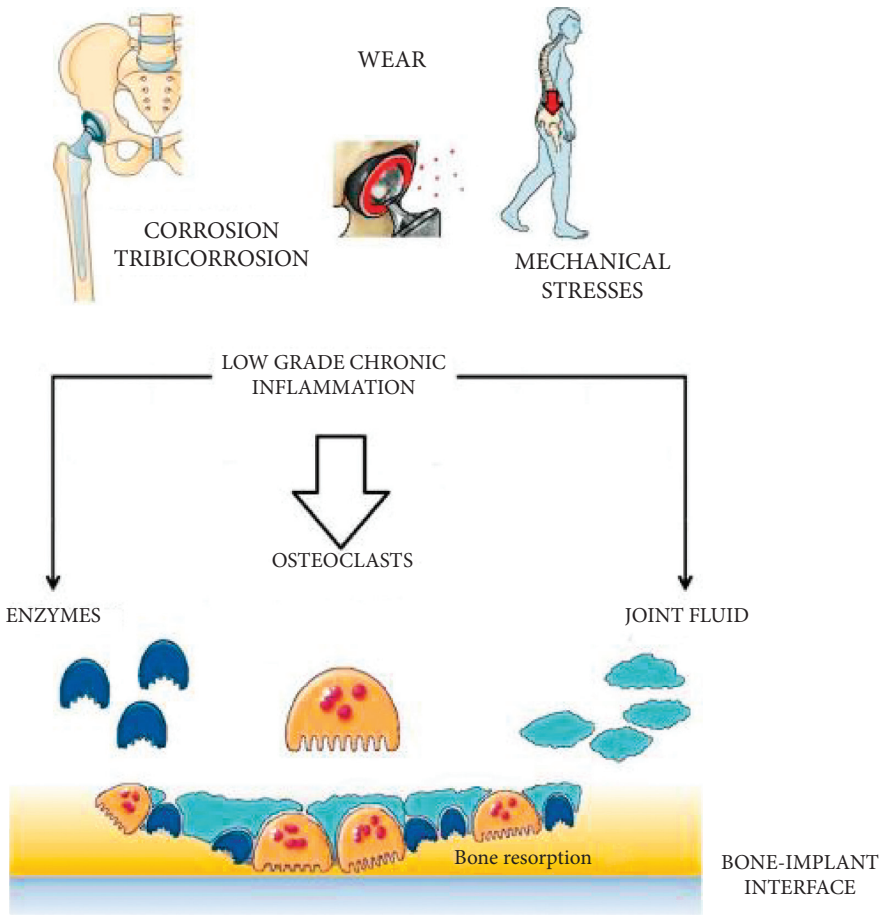


Figure 2: Amplification and translation of biological signals start with the interaction of prosthetic bearing wears with the human innate immune system cells resulting in bone resorption in bone multicellular units at the interface of the bone implant [1].

Furthermore, it seems also that cemented and cementless implants give rise to different biological processes [61]. In cemented acetabular components, there is a tendency to show a pattern of osteolysis leading to losing the implant, whereas cementless components tend more to exhibit expansile, localized lesions with the cup of the implant remaining stable (Figure 3) [62]. Additionally, although loosening of the component of the implant may precede the progress of an expansile lesion, it is less often (Figure 4).

2.3. Diagnosis. The most popular method of detection of the extent of osteolytic lesions is computed tomography whereas the most sensitive method is magnetic resonance imaging (MRI). The studies have demonstrated that the sensitivity of CT and plain radiography in the detection of osteolysis is 74.7% and 51.7%, respectively, while the corresponding figure for MRI is 95.4% [63]. Furthermore, thanks to MRI images, it is not only possible to detect intracapsular synovial disease well before discernible loss of bone, but also to detect the granulomatous tissues caused by the wear debris, which are actually the merits of this tomographic modality [64].

Conversely, it is a difficult task to identify osteolytic lesions on plain radiographic images because of two reasons. First, the visualization of the cancellous bone adjacent to the prosthesis can be blocked by the tibial and femoral components. Second, before detecting any changes on the quantity of the skeletal calcium on the radiographic images, 50% of them are normally lost [65].

2.4. Methods of Diagnosis. Currently, various methods are employed for the detection and analysis of the osteolytic lesions. However, the application of these methods is highly dependent on the knowledge of the medical profession, available facilities, and severity of the condition. Diagnosis of the osteolysis can be implemented manually using radiographic, CT, and MRI images or can be carried out with semiautomatic or automatic computer-based techniques, algorithms, toolkits, and software. Recently, thanks to the advances in the artificial neural networks, deep learning algorithms are widely used by orthopedists and medical profession with the assistance of biomedical research for the detection of a variety of bone tumors and fractures, analysis of prosthetic movements, etc. Although fewer computer-based studies have been carried out specifically on the topic of osteolysis as there are very limited public datasets for this topic, because of strict resemblance between anatomical regions of interest, tissue properties, biological behaviors, detection, and analytic modalities, the majority of these

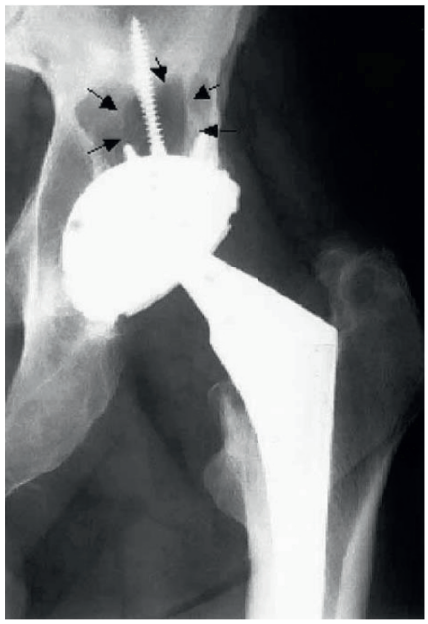


Figure 3: Osteolytic lesion superior to the well-fixed modular acetabular component [62].

methods are applicable for the analysis and detection of osteolysis.

Following the most recent and state-of-the-art potential computer-based methods from semiautomatic to automatic methods, software and algorithms that can be employed to assist medical profession in the analysis of osteolytic lesions are briefly discussed.

2.4.1. Deep Learning Methods.

(1) Method 1: *AnoGAN*. It is important to detect osteolytic lesions in advance during the follow-up sessions before they reach to a level leading to surgical treatments. Since osteolytic lesions are very small and have abnormal shape from unpredictable adjacent areas in the bone, the image collection with osteolytic lesions is challenging compared to major cancers. In this case, it can be determined that abnormality recognition is a proper idea. In the anomaly detection, the network is trained with a large number of images without lesions and then judges whether images show lesions or not based on a determined abnormality score [66]. As the number of images with lesions is limited in comparison to the number of images without lesions, the anomaly technique could be an efficient solution. Although labeled images are used for training the network, it is a heavy burden for clinicians to annotate the images. In order to reduce the labor for medical professionals, unsupervised learning methods are normally considered. This is due to the fact that all CAD systems can be more helpful in detecting a target more accurately in the image. Watanabe et al. [67] carried out a study using a similar modality and called it AnoGAN, which can be useful in detecting osteolytic lesions. They designed a classification method that functions based on adversarial learning but used bone tumor datasets to detect the metastatic bone tumors. Tumors lesions are distinguished based on computed abnormality scores. These scores are defined by comparing a generated image with a test image in the image level and the future level. The suggested strategy takes only nonmetastatic bone tumor images and learns the normal image distributions based on a generative adversarial network (GAN). Although using unsupervised learning method is quite challenging in medical image processing, it is a significant task for clinical applications. Figure 5 demonstrates the anomaly detection framework proposed by Watanabe et al. [67].

As can be seen from Figure 5, the proposed method is composed of two steps: training step and test step. During the training phase, the generative G and discriminator D networks compete to gain ability. After adversarial training, the generator network is able to produce nonmetastatic bone image from a latent noise vector z. This trained generator and the noise vector z are then used for the evaluation of the test samples in the next phase. The test phase employs the latent space to evaluate test samples. When a test image is imported, the method finds the finest z corresponding to the image $G(z)$ generated by the generator network. The z value should be visually similar to X in iterative backpropagation stages. In the latent space, the best noise vector z can be found. The anomaly score is obtained with the following formula:

$$A(X) = (1 - \lambda).R(X) + \lambda.D(X), \quad (1)$$

where $D(X)$ represents the discrimination loss that calculates the distinction of extracted features using the trained discriminator D in the features level. $R(X)$ denotes the residual loss that computes the visual dissimilarity between $G(z)$ and X in the image level. λ represents a weighted coefficient. The larger the value for the anomaly score, the greater the probability of anomaly detection. Figure 6 shows the experimental results using this method.

(2) Method 2: *SG-CNN*. CNNs are very popular for classification tasks. There are various classic models of CNN such as VGGNet [27] and AlexNet [68] that can be selected for categorization tasks. Although these networks demonstrate good performance for classification of natural images, they are not very efficient for classification of medical images especially when the goal is to detect tumor lesions like osteolysis. The reason is that, in natural images, most objects are normally at the center of image and the variation between objects is apparent, whereas medical image categorization demands fine-grained visual classification that cannot be done with classic CNNs with high classification accuracy. Although several methods are proposed to overcome the fine-grained labeling problem such as a mask-CNN structure based on annotations of part of images by Wei et al. [69] and part-based R-CNN structure for fine-grained labeling by Zhang et al. [70], these methods are time consuming on making datasets. To solve this problem, a novel CNN is proposed by Li et al. [71]; it can produce ROI areas automatically by network independently without employing annotated images (Figure 7).

This model is a superlabel-guided convolutional neural network (SG-CNN) that classifies CT images of bone tumor. SG-CNN is composed of two subnetworks that are

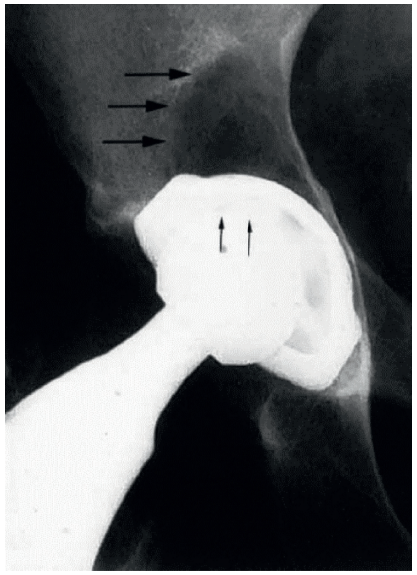


Figure 4: Osteolytic lesion superior to the loose press-fit modular acetabular component [62].

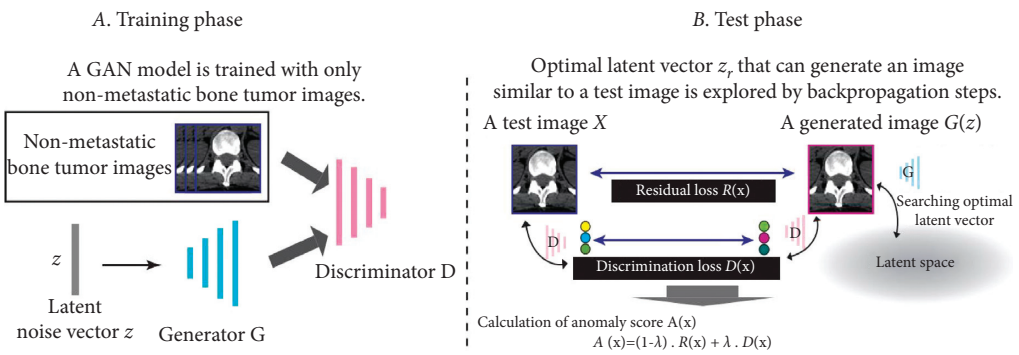
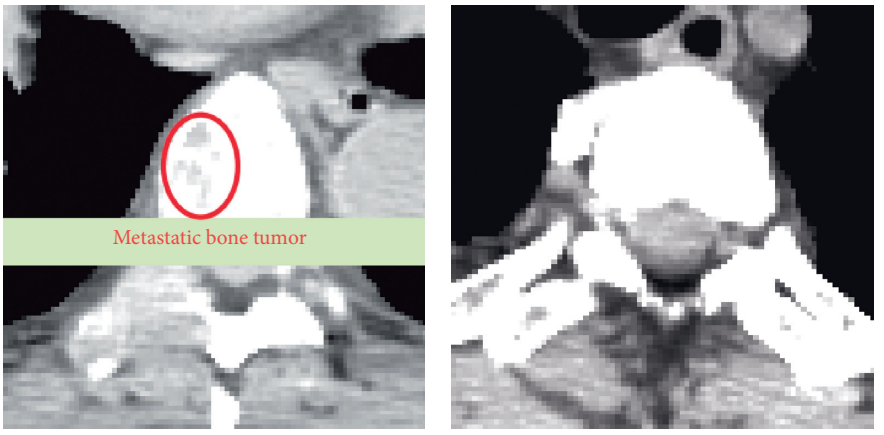


Figure 5: An overview of the generative adversarial network with anomaly detection of the bone tumor [67]. (a) Training phase. (b) Test phase.



(a) (b)

Figure 6: Image samples of true positive (a) with metastatic bone tumor and true negative without tumor (b) detected by the AnoGAN method [67].

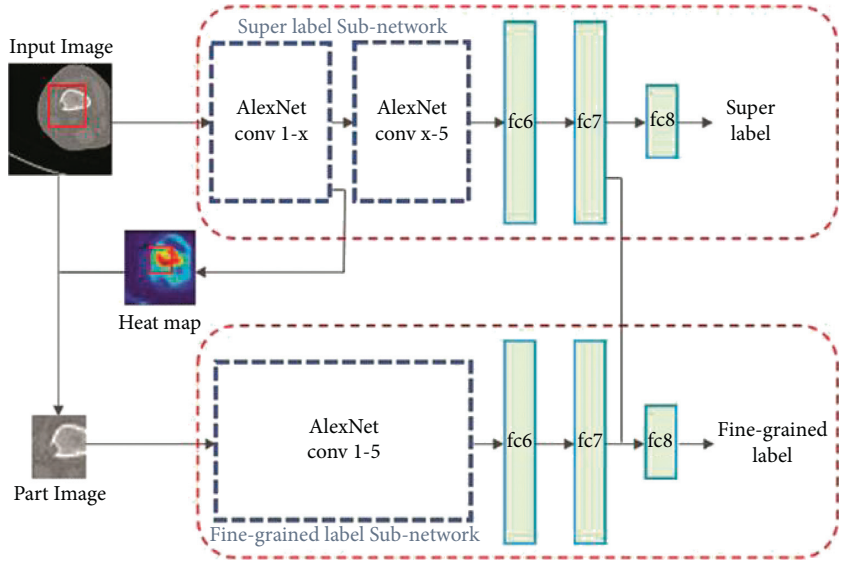


Figure 7: The superlabel-guided convolutional neural network (SG-CNN) structure. A raw CT image with two labels without any annotations is fed into the model. The input image is cropped under the guide of the heatmap generated by the first convolutional layer of the superlabel subnetwork and then inserted into the other network branch. The network output provides two labels [71].

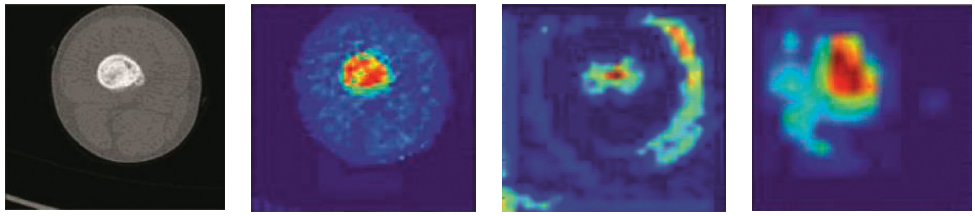
responsible for learning the whole image and focus on bone tumor lesions to learn more detailed data. The inputs of these networks are CT images with two labels in hierarchical relationship, and the outputs are fine-grained labels and superlabels that both are used for training the SG-CNN to gain classification accuracy. AlexNet CNN model was used as the network architecture. During training phase, images are fed into the superlabel subnetwork. When all the feature maps of the guide convolution layer of the subnetwork are added together, a heatmap is generated (Figure 8). The heatmap generated during this phase, superimposed on the points corresponding to the image parts, is used as the input of the fine-grained label subnetwork. As the image is cropped, most of the background parts of the image are removed. Thus, the model emphasizes the potential tumor areas. Finally, the output of the network predicts fine-grained label whose classification accuracy is considered by two structure branches simultaneously.

Employing this artificial neural network model could effectively be used for the classification of osteolytic lesions, as the same obstacles are involved in the detection and classification of bone tumors, and the appearance of osteolytic lesions in CT images is very similar to that of bone tumors. Considering the successful results obtained using a bone tumor dataset on the SG-CNN by Li et al. [71], this method could be highly reliable.

(3) Method 3: *U-Net*. Convolutional CNNs used for the classification are composed of convolutional layers followed by several fully connected layers that map the feature image produced by kernels into a fixed-size feature vector [72]. Nevertheless, one of the defects of CNN is that each time a convolutional operation is accomplished for the classification, feature map resolution is diminished by half. This can reduce the accuracy of the classification if the feature map achieved by the final classification is utilized. Therefore, if convolutional neural networks are used for the classification of osteolysis, CNN may cause transition during the convolution operation. Although the performance of the model is rarely affected, the position of osteolytic lesions may simply be affected in the recognition step. Considering the fact that osteolytic lesions have abnormal shape and low contrast, this defect can visually affect the images by metallic artifacts and image noise [12, 25, 73, 74].

In comparison to the CNN, the fully convolutional layers in the fully convolutional networks (FCN), which accept input data (images) at any dimension, replace all the convolutional layers. In this network, a deconvolution layer is applied in order to perform upsampling for the feature map of the last convolutional layer to be at the same dimension of the input 2D data (image). Thus, while preserving the spatial information in the original input image, a prediction can be made for each pixel. Finally, the pixel-by-pixel classification can be performed on the upsampled feature map (Figure 9). Based on the residual network strategy, the FCN can solve the contradiction between the translation invariance in the target detection step and the translation invariance in the classification network. Thus, a semantic segmentation structure U-Net based on fully convolutional neural network can be an efficient method for the detection and classification of osteolytic legions.

In a study carried out by Jian et al. [75], a U-Net neural network consisting of a contraction path and an expanding path (Figure 10) was used for the diagnosis of osteoporosis, which is the most common bone disease [76, 77]. For the detection of osteoporosis of the femoral neck, if a conventional CNN is employed, it could cause translation during the convolution operation. This could easily affect the position of the osteoporosis boundary box in the detection step that directly leads to a reduction in the segmentation accuracy of the femoral neck.



(a) (b) (c) (d)

Figure 8: The presentation of raw data and heatmap. (a) A typical input image. (b) The heatmap generated by conv1. (c) The heatmap produced by conv2. (d) The heatmap generated as the output of conv3. It is determined by the images that the more we go deeper into the network, the more the abstract and semantic meanings contained in the heatmap [71].

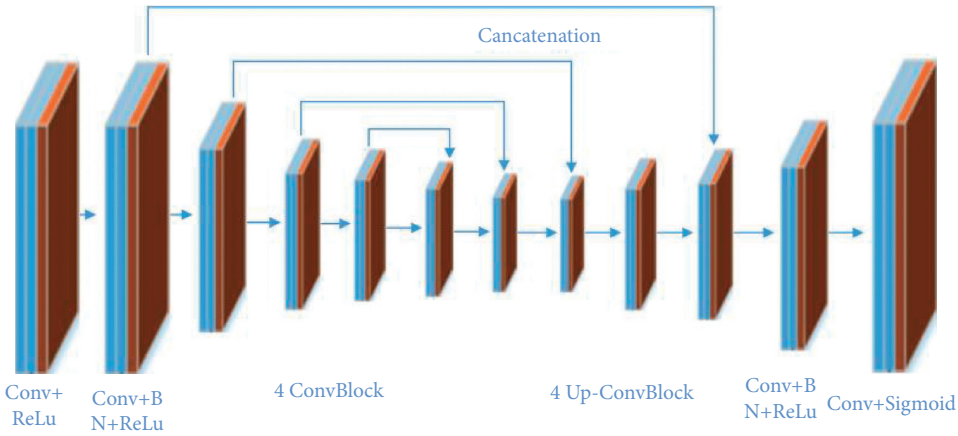


Figure 9: End-to-end network structure [75].

Osteoporosis is normally specified by a reduction in the bone density, thinning of cortical bone, and thinning of trabecular bone. The U-Net network proposed by Jian et al. [75] was implemented on an image dataset of patients having undergone pelvic X-ray imaging. Among the patients, 30 had normal bone mass, 28 had lower bone mass, and 31 had osteoporosis. Normal bone mass X-ray images demonstrate thick cortical bone and high bone density (Figure 11).

The recognition results and classification of the U-Net network proposed by Liu et al. are as follows: the total recognition rate of lower bone mass images from normal bone mass is 83.88%, the corresponding figure for the osteoporosis from normal bone mass is 86.74%, and the one for osteoporosis from lower bone mass is 79.55%. The experimental results of the network demonstrate that it can successfully solve the influence of image interference for the bone density analysis. The proposed U-Net network with the recognition rate of above 81% for the detection of osteoporosis could be a highly functional solution to solve a similar problem for the detection of osteolysis.

(4) Method 4: *Seg-UNet*. Do et al. [78] proposed a novel method for the detection of knee bone tumors from X-ray images using a multilevel Seg-UNet model with global- and patch-based techniques. This network is used as a computer-based assistive tool for the segmentation and classification of tumor regions into three labels: normal, benign, and malignant. Although this network is designed only for bone tumor detection around knee regions, as the anatomical region is at the same place where osteolysis will occur after TKA and there are similar visual characteristics in X-ray images, developing this network can effectively solve the problem of segmentation and classification of osteolytic lesion.

The proposed Seg-UNet architecture is illustrated in Figure 12. This multilevel network uses a combined global- and patch-based approach in order to not only detect small tumor regions but also achieve a high improvement in malignant tumor detection.

This model has an encoder-decoder architecture to exploit the mutual advantage of segmentation and classification branches to learn the local texture features and global geometric context at every pixel. The encoding block $\mathbf{E}(\mathbf{X}_{\text{enc}})$ with the global encoding features X_{enc} from the input image X is located at the left side of the model. The input of the model can be either a down scale image X_G or an image patch X_p from the original image X_o with high-resolution. The classification branch in the middle of the model employs the global average pooling for the extraction of the encoding features, followed by dense and soft max layers for the classification of the input image. The network is composed of three outputs: $\hat{Y}_{\text{clas}}$ denotes the classification result, $\hat{Y}s_{\text{eg}}$ presents the tumor segmentation result, and finally $\hat{Y}_{\text{dist}}$ is the multilevel high-risk tumor result. The $\hat{Y}_{\text{clas}}$ determines

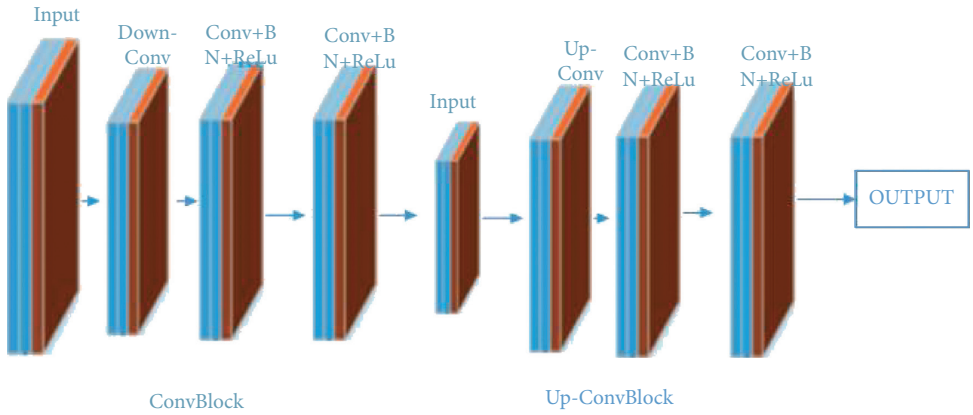
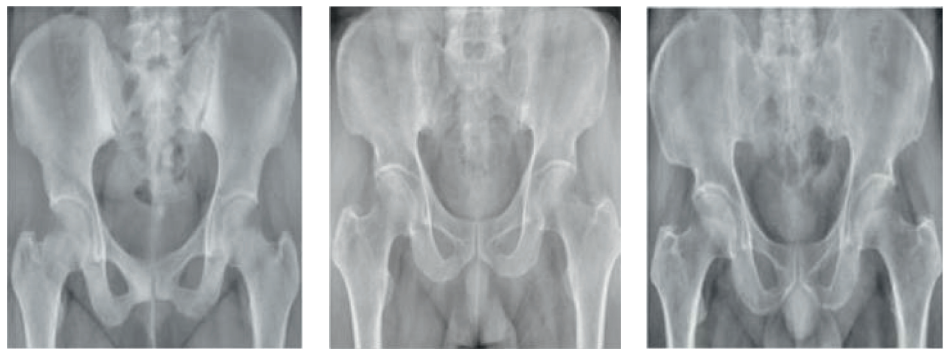


Figure 10: The U-Net network architecture. The contracting path is composed of two 3×3 convolution layers for repetitive processing, each accompanied with a modified linear unit (ReLU). The downsampling is composed of a single 2×2 maximum pooling operation [75].



(a) (b) (c)

Figure 11: Some examples of pelvis acquired by X-Ray imaging: (a) normal bone mass, (b) lower bone mass, and (c) osteoporosis [75].

whether the input belongs to the normal, benign, or malignant labels. Because of the complexity and challenging conditions of the X-ray images of the knee bone, this separate classification branch at the global-context level was designed. The decoding block $\mathbf{D}$ ($\mathbf{X}_{enc}$) on the right side of the model is the expanding path that maps the encoding feature into decoding feature X_{map} at the pixel level. For outputting tumor segmentation mask $\hat{Y}s_{eg}$ and multilevel distance features $\hat{Y}_{dist}$, the 2D extracted features are improved by multitask learning at the pixel level between the high-risk tumor segmentation H_{dis} and the pixel-tumor segmentation $H_{seg.}$ $\hat{Y}_{seg}$ with the vector size of W×H×2 is used for the classification of pixels of the input image into normal or tumor groups. $\hat{Y}_{dist}$ with the vector size of W×H×5 decides the level of attention, i.e., normal, tumor, or high-risk, based on the distance to tumor in three levels of 1 to 3.

As, in many cases, tumors are very small in comparison to the background regions, this model made an attempt to detect small tumors by learning mutual information from adjacent feature maps around tumors. The size of the high-resolution image compared to the very small size of the tumors is one of the challenges in knee bone tumor detection. Due to the memory limits, the input image is normally resized to be suitable for the global-based patch, leading to a loss of some image texture important for tumor recognition, especially for small tumors. The patch-based model that learns image texture detailed from image patches can solve this problem. The global- and patch-based models' data used to train the network are demonstrated in Figures 13 and 14. The effectiveness and the ability of each model (global-based model and patch-based model) in the detection of bone tumors are illustrated in Figures 15 and 16. Noise in small tumors, nontumor detection in variant pose, and noise in larger tumors are the failures of the patch-based model (Figure 15), whereas it assists global-based model for the detection of small, long, and large tumors (Figure 16). The Seg-UNet network proposed by Do et al. [78] demonstrated that the fusion model of the patch- and global-based models could provide mean classification accuracy of 99.05% and segmentation mean IoU of 84.84% for bone tumor segmentation and classification when both global- and patch-based models are used.

2.4.2. Temporal Radiographic Texture Analysis (tRTA). tRTA was developed as an alternative to the conventional method of diagnosis of osteolysis, in which the condition of patients is postsurgically monitored by several X-ray images

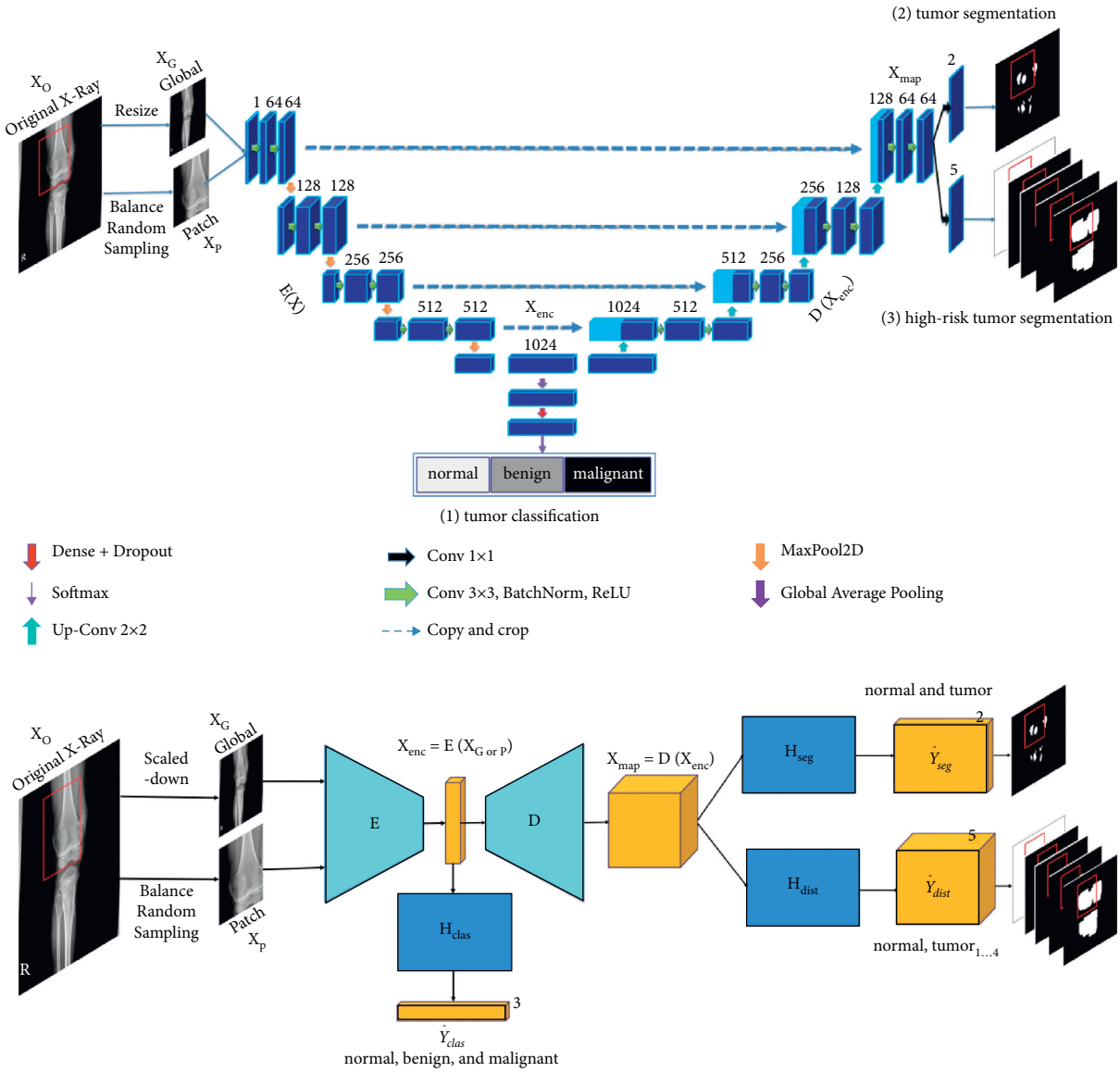


Figure 12: The multilevel Seg-UNet model with patch- and global-based approaches (top) and the block demonstration of the model (bottom). In the block diagram, E represents the encoding block; D depicts the decoding block; the three branch blocks H_{seg}, H_{dist}, and H_{clas} denote multitask learning [78].

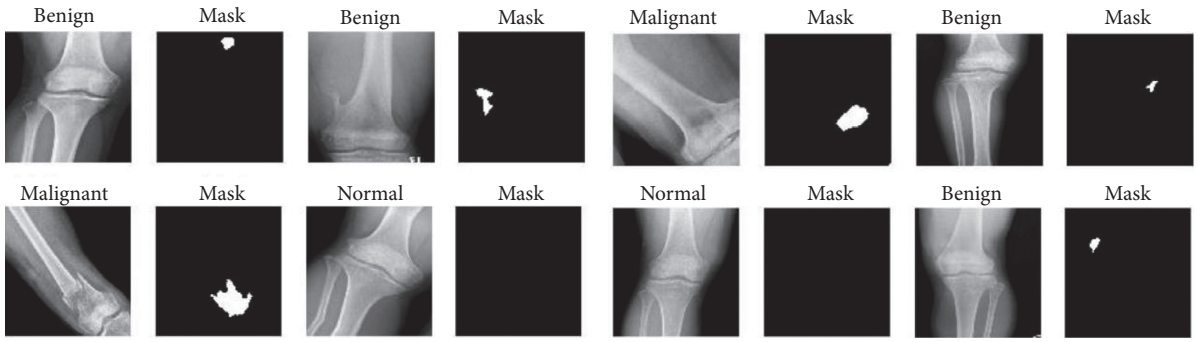


Figure 13: The augmented data in the global-based model underwent transform operations such as rotating, center cropping, resizing, and random flipping [78].

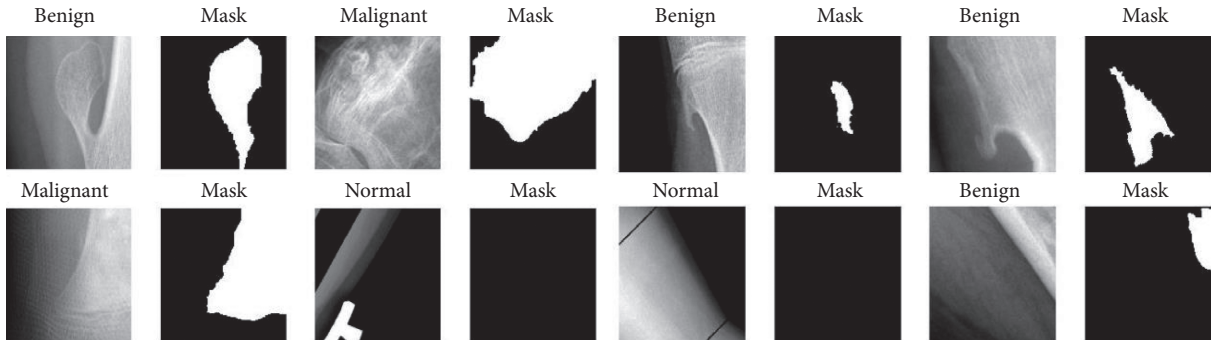


Figure 14: The augmented data in the patch-based model underwent transform operations such as rotating, center cropping, resizing, and random flipping [78].

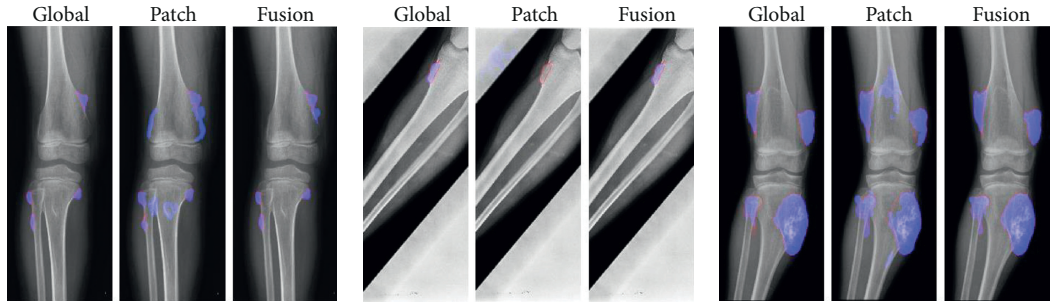


Figure 15: The successful ability of the global-based model to enhance the fusion results for the detection of (a) small tumors, (b) variant poses, and (c) large tumors. The red lines and the blue regions demonstrate the ground truth and tumor detection, respectively [78].

[79]. As osteolysis is a biological process that evolves and appears slowly, early follow-up and diagnosis of osteolysis in the radiographic images are cumbersome and in some cases impossible tasks. Furthermore, the symptoms of trabecular texture changes caused by osteolysis are difficult to observe normally until the lesions progress. Although CT images could ameliorate the detection process of osteolysis, the cost and exposure issues make them impractical for regular follow-up. tRTA is a computerized radiographic texture analysis method used as an alternative to RTA previously used for the measurement of the patterns of trabecular bone to assist the detection of osteolytic lesions [79–81]. In the tRTA method, ROIs including the potential osteolytic lesions are selected from the image database, taken during follow-up, and visually compared with the images of the previous sessions. Then, texture features are computed from the selected ROIs to perform trend analysis with simple linear regression technique, BANN temporal analysis technique, and a LDA merging features technique. This method has the advantage of incorporating the absolute texture measures, as well as how these measures alter over time.

2.4.3. Morphometry. In this method, the area of osteolytic lesions is measured based on the idea of cross intersect counting [82]. A morphometric grid is superimposed over the region of interest on the radiographic images, and then the number of test points overlapping the area of interest is counted (Figure 17).

Based on the study carried out by Smith et al. [82], the application results of morphometric method are compared with the estimations of professional orthopedists and proved to be reliable for the measurement of the area of osteolytic lesions if applied by trained orthopedic observers (non-medical or medical health professionals).

2.4.4. Software and Toolkits. Various software packages and toolkits have been developed for the purpose of medical image processing. Although they are not specifically developed for a single medical purpose, a variety of them are used by the medical profession and researchers for tumor detection and analysis using a variety of imaging modalities. The most popular and commonly used software and toolkits are mentioned as follows.

(1) ImageJ. This is an open-source image processing program developed in Java language inspired by NIH Image [83]. This program is currently available on all the current computer platforms and can be used with variety of plugins and macros. Although ImageJ is not specialized software, it is sometimes used by professionals and researchers for medical purposes. It performs its tasks in eight steps: dataset resizing, component hollowing, volume rendering, model slicing, image slicing, image cropping, standardization of the image size, and analysis. Nevertheless, ImageJ suffers from a variety of limitations such as a lake of 3D data analysis and requiring heavy interuser variability.

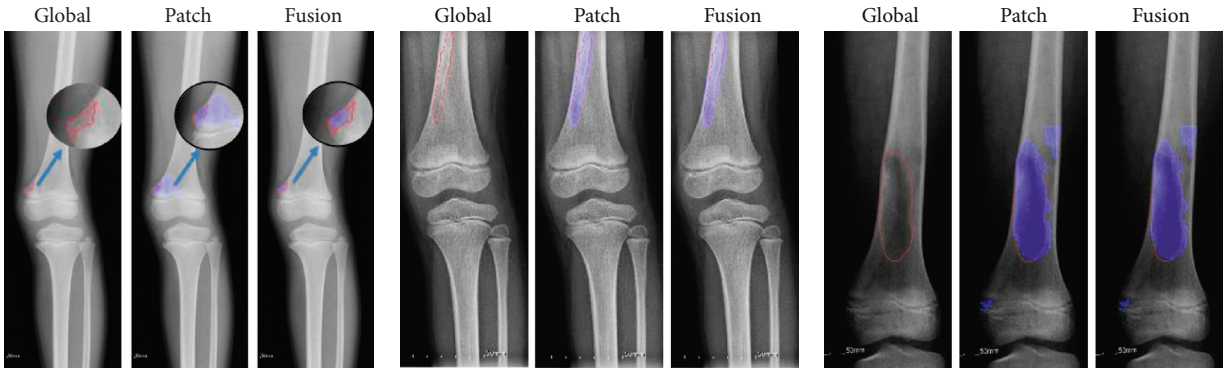


Figure 16: The successful ability of the patch-based model to enhance the fusion results for the detection of (a) small tumors, (b) long tumors, and (c) large tumors. The red lines and the blue regions demonstrate the ground truth and tumor detection, respectively [78].

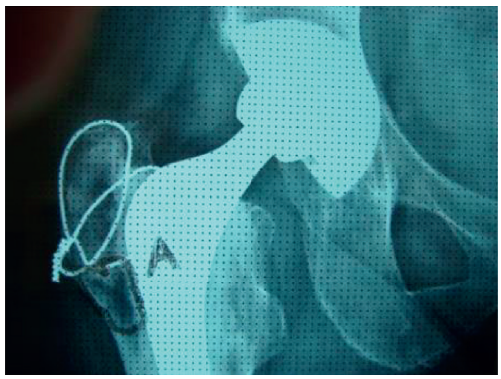


Figure 17: The superimposition of the morphometric grid onto the morphometry grid on a radiographic image for the measurement of osteolytic lesion areas [82].

(2) Osteolytica. It is image processing software specifically designed for the measurement of lytic bone lesions [84]. It employs novel graphic card acceleration, and it is capable of 3D rendering to make a rapid analysis and reconstruction of osteolysis. It is designed to be faster, more user-friendly, and less biased compared to manual osteolytic lesion measuring methods or ImageJ 2D analysis. The principal goal of Osteolytica is to measure the dimension and areas affected by osteolysis in a 3D bone analysis. It is composed of four operating processes: dataset resizing, dataset loading, selecting the maximum lesion size, and finally lesion analysis. Osteolytica employs the process of reconstructing the surface of a volume sample and then subtracting the reconstructed volume from the original surface through volumetric diffusion method (Figure 18) [85].

(3) 3D Slicer. 3D Slicer is an open-source, free, and multiplatform software package that can be widely used for various medical applications such as virtual reality, real-time 3D ultrasound reconstruction, adaptive radiation therapy, tracked ultrasound for needle guidance, robot assistance intervention, surgical navigation, and image segmentation (Figure 19) [85, 87]. Furthermore, this software provides multiorgan analysis from head to toe and supports various imaging modalities including ultrasound, computed tomography, magnetic resonance imaging, microscopy, and nuclear medicine imaging. It also provides real-time analysis, which is highly useful during surgical navigation. Nevertheless, 3D Slicer is not approved for clinical use, and the distribution is intended for research use although there is no restriction on its employment.

(4) ITK. The Insight Toolkit (ITK) is an open-source, cross-platform library developed by the he US National Library of Medicine of the National Institutes of Health to provide developers with an extensive suite of software tools for the segmentation and registration of medical images (Figure 20) [89, 90]. ITK with its extreme programming methodologies and spatially oriented architecture allows processing medical images in two, three, or more dimensions.

(5) MITK. The Medical Imaging Interaction Toolkit (MITK) is free, open-source software for the development of interactive medical image processing software [91, 92]. MITK is a class library based on ITK that provides leading edge segmentation and registration techniques and forms the basis of algorithms. It executes the visualization commands with the Visualization Toolkit (VTK) [93]. MITK workbench has a highly customizable and extensible end-user application providing all steps of a clinical workflow such as

(a) (b) (c)

Figure 18: The method of calculation of osteolytic lesions in Osteolytica. (a) Original volume of the sample. (b) Reconstruction of the surface of the sample and filling lesions, diffusing the volumetric surface outwards. (c) The volume of the osteolytic lesions measured through the subtraction of the original surface from the reconstructed one. The expanded volume will diffuse inwards through a variable over the surface [84].

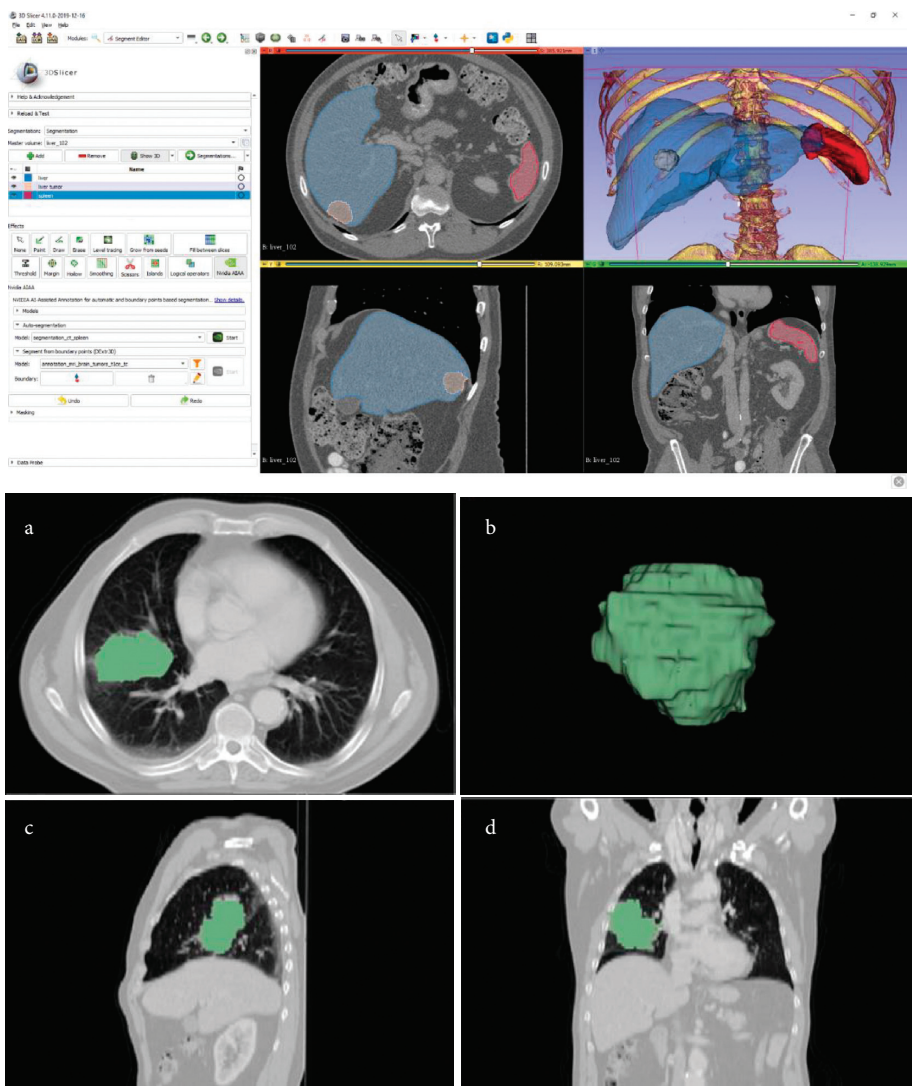


Figure 19: The top figure illustrates the workbench of the 3D Slicer with a variety of toolkits and tools that can be used for different purposes. The bottom figure demonstrates a typical tumor segmentation capability of the 3D Slicer [86].

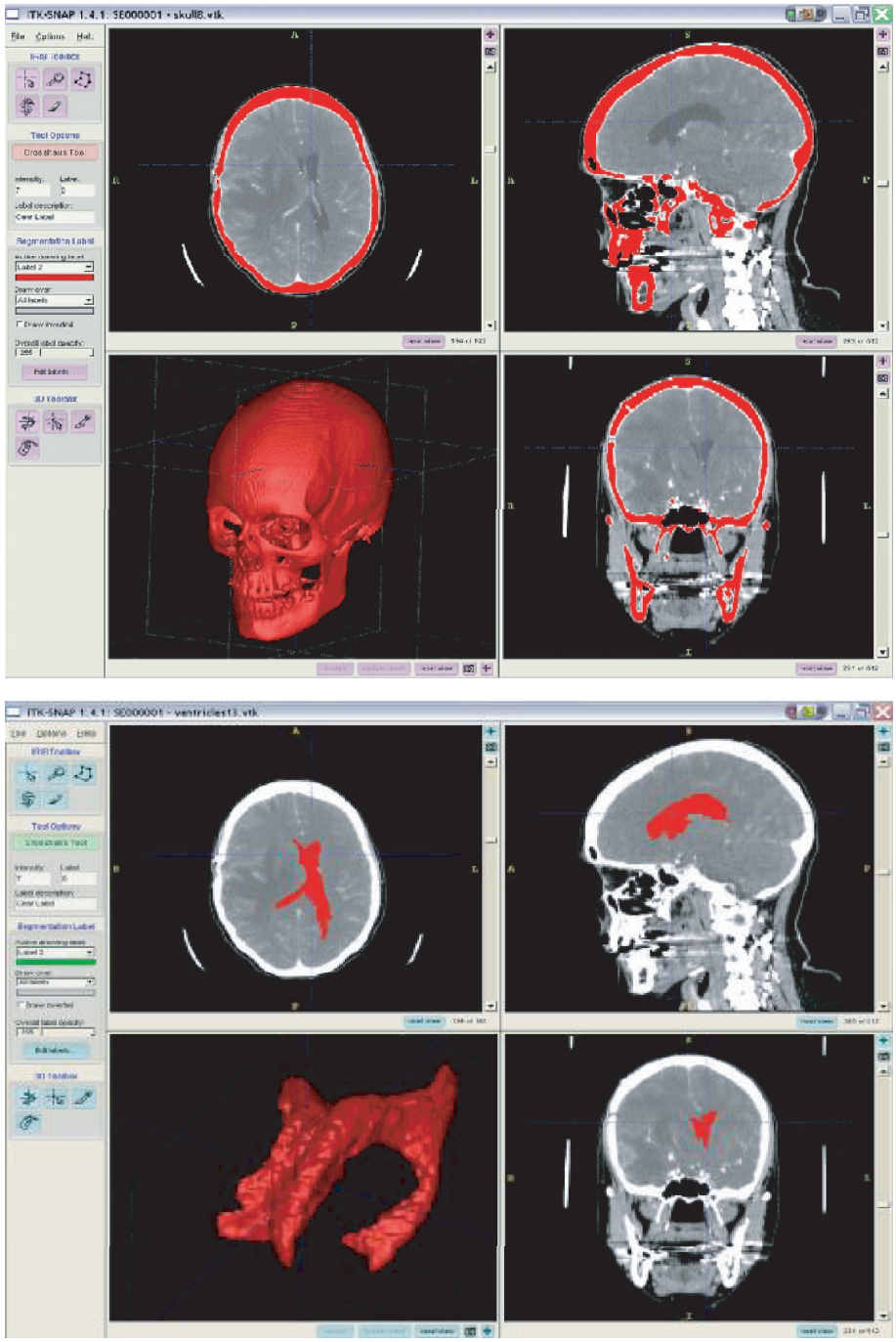


Figure 20: The workbench of ITK software and its segmentation ability (in this typical case, brain bone and tumor segmentation) [88].

data retrieval, image analysis, image-guided therapy, diffusion imaging treatment planning, tool tracking, diagnosis, intervention support, and treatment control (Figure 21).

2.5. Overview of the Image Processing Diagnosis Methods. Various techniques and methods that can be automatic, semiautomatic, or manual can be employed for the detection and segmentation of bone tumors, to be more specific osteolytic lesions. Depending on the objective of the research, they can provide different accuracies. In the previous section, the state-of-the-art automatic and semiautomatic computer-based methods, as well as manual methods with the use of software and toolkits, were mentioned for this purpose. However, the methods and techniques are not limited to this. An overview of the mentioned methods with their key points, in order to provide a well comparison for selection and application, is presented in Table 1.

2.6. Bone Tumor Datasets. Datasets play a prominent role in the performance of neural networks. Comprehensive

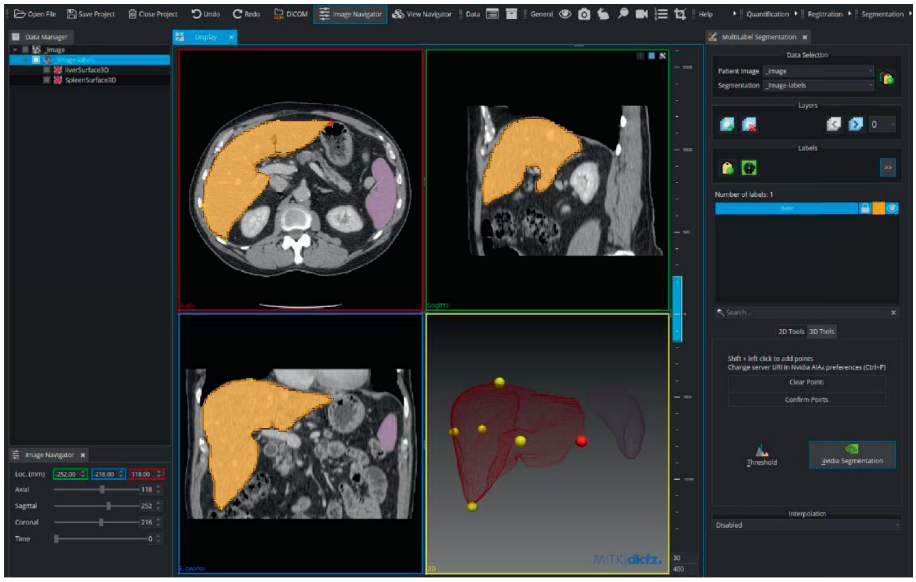


Figure 21: Workbench of MITK for medical image processing purposes [91].

Table 1: A brief overview of the computer-based methods mentioned in this study for the purpose of analysis of medical images with osteolytic lesions.

Method name	Technique	Purpose of the application	Mechanism of the application
AnoGAN	Adversarial learning (generative adversarial neural network (GAN))	This unsupervised learning method is suitable when the dataset is limited	This method performs anomaly detection by generating a large number of nonlesion images by GAN to detect images with lesions.
SG-CNN	Convolutional neural network (CNN)	This method can automatically produce ROI areas independently through a superlabel-guided CNN	This method can improve classification accuracy by generating fine-grained labels and superlabels of the region of interest in medical images whose lesions of interest are not well apparent.
U-Net	U-Net structure deep neural network	This method is suitable for the segmentation of lesions when they have abnormal shape and low contrast and are susceptible to transition during classification	The U-Net structure performs semantic segmentation of the osteolytic lesions on the input image by concatenating the convolutional layers in the encoder path with the deconvolutional layers in the decoder path.
Seg-UNet	Multilevel Seg-UNet	This method is suitable for segmentation of lesions of interest on the input image when the lesion has abnormal shape and low contrast and the size of the lesion is very small compared to the input image size	The Seg-UNet exploits U-Net structure as well as the global- and patch-based approach in order to improve the classification accuracy.
tRTA	Mathematical computation	This method is a manual image processing method for the segmentation of the lesions of interest from dataset that requires trained medical practitioners	The tRTA is a computerized radiographic texture analysis method for the evaluation of ROI through linear regression, BANN temporal analysis technique, and a LDA merging features technique.
Morphometry	Manual computation	The method employs the cross-intersect counting approach	In this method, with the use of a morphometric grid that is superimposed onto the region of interest on the radiographic images, computation is performed.
ImageJ	Manual image processing	General-purpose image processing software	ImageJ can take the advantage of different plugins and macros for various image processing goals.
Osteolytica	Manual image processing	Specifically designed for the measurement of lytic bone lesions	This image processing software is designed for 3D analysis of lesions and requires trained staff.

Table 1: Continued.

Method name	Technique	Purpose of the application	Mechanism of the application
3D Slicer	Medical image processing software	Manual image processing	3D Slicer is medical software designed only for research purposes that can perform various image analyses using variety of packages on different anatomical positions.
ITK	Open-source medical library	Manual image processing	This medical library is suitable for developers for medical image processing purposes.
MITK	Open-source medical library	Manual image processing	This class medical library is based on the ITK library and provides segmentation and registration techniques. It also has a highly customizable workbench.

Table 2: The names and properties of the potential datasets that can be used for training deep learning neural networks for bone tumor detection and segmentation. Each dataset is also labeled as either public or private.

Name of the dataset	Description	Type of the data	Number of images
CNUH [78] (private)	Provided by the Chonnam National University Hospital (CNUH)	CT datasets focusing on benign and malignant tumors in two regions of knee bone of distal femur and proximal tibia	Benign tumor: 1061, malignant tumor: 134, normal: 381
Shenzhen No. 2 People's Hospital [71] (private)	Derived from patients diagnosed with bone tumors in the years 2014–2017	CT images stored in DICOM	6422 images
Hokkaido University [67] (private)	Provided by the Hokkaido University in Japan	CT images of metastatic and nonmetastatic tumor images	8790 images
Sichuan University [94] (private)	Provided by the Institutional Ethics Committee of West China Hospital in Sichuan University	Bone scintigraphy (BS) images using SPECT/CT taken from patients diagnosed with bone metastasis and having undergone whole-body BS	13477 images
Public datasets [95] (public)	CT images corresponding to different body parts	CT images manually segmented	270 images
In-house dataset [96] (private)	Containing bone images of different body parts: head, chest, abdomen, neck	CT images	16218 images
Bone tumor [97] (private)	Containing benign and malignant bone tumors	Plain radiography	2899 images
CNUH [98] (private)	Provided by the Chonnam National University Hospital initially for bone segmentation on a deep learning approach containing both malignant and benign tumors	Plain radiography	963 total images, benign tumor: 329, malignant tumor: 134
DIAGNOSTIKO IATRIKI A.E. [63] (private)	Collected from prostate cancer patients with suspected bone metastatic disease, who underwent whole-body scintigraphy	Scintigraphy images (SPECT)	908 total images, 778 bone scan, 328 bone metastasis, 271 benign scan, 179 normal
Chittagong University [99] (private)	Consisting of 60 MRI images of patients diagnosed with bone cancer and their grand truth images	MRI images	60 total images, benign tumor: 30, malignant tumor: 30

datasets can well train the networks during the training phase. Subsequently, they can have a great impact on the effectiveness and accuracy of the proposed computer-based methods for the detection and segmentation of bone tumors. The larger and the more comprehensive the datasets are, the more valuable the provided data will be. It is important to note that not all the datasets are publicly available for use. Some datasets are private and require formal requests and legal permission for use. Furthermore, the images could be acquired through different imaging modalities such as MRI, CT, PET, SPECT, or conventional X-rays and with different anatomical positions that should be implemented for the evaluation of the networks based on the objective's similarities. Table 2 presents a brief overview of the private and public datasets that can provide highly valuable data regarding bone tumors that can be used for training and evaluation of the proposed deep learning networks for the detection of osteolytic lesions.

2.7. Treatments for Osteolysis. Osteolytic lesions can be treated though two ways: surgical and nonsurgical treatment.

2.7.1. Surgical Treatments. Operative treatments are recommended to address prior or ongoing osteolytic lesions and to correct failing articulation. Surgical treatments of osteolytic regions of THA, TJA, and shoulder arthroplasties

follow a similar philosophy [100]. Nevertheless, the methods are varying depending on the structure, anatomical location, function, design, and materials employed.

If the prosthetic components are aligned, well-fixed, functional, and proper modular replacement parts can be obtained, the operative surgical treatment is then focused on the revision of the bearing surfaces and potential grafting of bone.

The osteolytic lesions caused by metal corrosion and byproducts from metal-on-metal implants are debrided, and modular metal-on-metal bearings are replaced with other ones, normally with ceramic on polyethylene [101].

2.7.2. Nonsurgical Treatments. Nonoperative treatment of osteolysis is a possible option for patients who are either not able to tolerate reoperation immediately or under low loads of prosthetic byproducts and susceptible to osteolysis. The main goal of nonoperative treatment is merely to postpone the need for reoperation and also to keep the size of bone defects limited. The common nonsurgical treatment methods are mentioned as follows.

(1) Bisphosphonates. Bisphosphonates are the drugs that should be consumed orally or parenterally to treat metabolic disease, osteolysis associated with metastatic and osteoporosis. This medication is a synthetic analogue of pyrophosphate and is advised for osteolysis treatment in a physician-directed or off-label manner [102]. The action mechanism of bisphosphonates is mainly on the osteoclast that undergoes apoptosis leading to the inhibition of bone resorption [103, 104].

(2) Cell Therapy. Delivery of cells as a local therapeutic platform can indirectly or directly affect osteolysis. This can heal the bone and provide paracrine and autocrine factors [105]. Autologous bone grafting is one type of local cell therapy. Crosstalk between MSCs and macrophages is also an ongoing process in all the inflammatory bone disorders and bone healing process [106, 107].

3. Conclusion

Osteolysis is a progressive and biological reaction to particulate wear debris. It is the most common indication for revision surgeries after total joint arthroplasty in long-term reviews. The biological mechanisms leading to osteolytic lesions are only now beginning to be understood. Although many research studies have been carried out to characterize the complex cellular interactions that result in the bone loss, it is still not apparent why some patients undergo early osteolysis and in others this phenomenon is postponed for many years or even never occurs. It is more probable that the observed differences between the conditions of patients indicate the differential sensitivity or differences in ability to mount a wear response and generate wear particles. To expand our knowledge about the process of osteolysis, the basic science studies must be translated into clinical studies and eventually clinical practice. Although in general there is no specific classification for osteolysis, it is categorized based on the measurement of volume of the affected lesions. Studies on the genetic profile may help to explain the variability in the rate of development and the extent of osteolysis in different patients.

To detect this asymptomatic process, regular radiographic follow-up is necessary. Imaging modalities like MRI and CT scans are utilized where the extent of bone loss is uncertain. Nevertheless, it is important to note that each imaging modality has its pros and cons. In addition, various computer-based image processing methods have been developed as clinically assisted tools to facilitate this process. Among them, artificial neural networks have proved to be highly efficient for tumor detection and segmentation. The research results of the popular deep learning networks used for tumor detection such as CNN, U-Net, and Seg-UNet proved that these deep learning algorithms could assist early detection and analysis of osteolytic lesions. Further developments and studies in this field are highly crucial for the development of a solid and specialized network for the detection and segmentation of osteolysis.

There are numerous nonsurgical therapeutic interventions for the treatment of the osteolytic lesion, but they still require clinical approval to be verified. Some of them have already passed limited clinical evaluations. Nevertheless, it must be determined that progressive osteolysis resulting from wear debris is not only a biological process, but also related to modular interfaces, material issues, or failure of the bearing surface. In this respect, currently, there is no evidence to prove that nonsurgical treatment methods can clinically treat osteolysis except for delaying the process. Although it is hoped that changes in the design of implants can positively affect the reduction of wear particles and osteolysis development, only time will prove this. Fortunately, there are some strong research results supporting this issue. Nevertheless, long-term clinical follow-up is desperately required to make any new benchmarks.

Data Availability

The authors did not use any specific dataset.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. T. Manley, J. A. M. D'Antonio, W. N. M. Capello, and A. A. P. Edidin, "A disease of access to fixation interfaces: clinical orthopaedics and related research®," 2021, https://journals.lww.com/clinorthop/Fulltext/2002/12000/Osteolysis__A_Disease_of_Access_to_Fixation.15.aspx.

[2] P. A. Banaszkiewicz, "The synovial-like membrane at the bone-cement interface in loose total hip replacements and its proposed role in bone lysis," *Classic Papers in Orthopaedics*, vol. 65, no. 5, pp. 117–119, 2014.

[3] P. A. Banaszkiewicz, "Reactions of the articular capsule to wear products of artificial joint prostheses," in *Classic Papers*

*in Orthopaedics*Springer-Verlag London Ltd, Berlin, Germany, 2014.

[4] A. Hart, V. Janz, R. T. Trousdale, R. J. Sierra, D. J. Berry, and M. P. Abdel, "Long-Term Survivorship of Total Hip Arthroplasty with Highly Cross-Linked Polyethylene for Osteonecrosis," *Journal of Bone and Joint Surgery*, vol. 101, no. 17, pp. 1563–1568, 2019.

[5] J. T. Evans, J. P. Evans, R. W. Walker, A. W. Blom, M. R. Whitehouse, and A. Sayers, "How long does a hip replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up," *Lancet (London, England)*, vol. 393, pp. 647–654, Article ID 10172, 2019.

[6] J. E. M. D. Dowd, C. J. M. S. Sychterz, A. M. B. S. Young, and C. A. M. D. Engh, "Characterization of long-term femoral-head-penetration rates," 2021, https://journals.lww.com/jbjsjournal/Abstract/2000/08000/Characterization_of_Long_Term.6.aspx.

[7] C. Klapperich, J. Graham, L. Pruitt, and M. D. Ries, "Failure of a metal-on-metal total hip arthroplasty from progressive osteolysis," *The Journal of Arthroplasty*, vol. 14, no. 7, pp. 877–881, 1999.

[8] T. R. M. D. YOON, "Osteolysis in association with a total hip arthroplasty with," 2021, https://journals.lww.com/jbjsjournal/Abstract/1998/10000/Osteolysis_in_Association_with_a_Total_Hip.7.aspx.

[9] H.-G. Willert, H. Bertram, and G. H. Buchhorn, "Osteolysis in alloarthroplasty of the hip," *Clinical Orthopaedics and Related Research*, vol. 258, pp. 108–121, 1990.

[10] B. J. M. D. Livingston, M. J. M. D. Chmell, M. M. D. Spector, R. M. D. Poss, and M. Boston, "Complications of total hip arthroplasty associated with the," 2021, https://journals.lww.com/jbjsjournal/Abstract/1997/10000/Complications_of_Total_Hip_Arthroplasty_Associated.10.aspx.

[11] A. Kobayashi, M. A. R. Freeman, W. Bonfield et al., "Number OF polyethylene particles and osteolysis IN total joint replacements," *The Journal of Bone and Joint Surgery. British volume*, vol. 79, no. 5, pp. 844–848, 1997.

[12] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Scientific Reports*, vol. 11, no. 1, Article ID 20210, 2021.

[13] W. A. Ali, M. K. N, M. Aljunid, M. Bendechache, and P. Sandhya, "Review of current machine learning approaches for anomaly detection in network traffic," *Journal of Telecommunications and the Digital Economy*, vol. 8, no. 4, pp. 64–95, 2020.

[14] A. Aghamohammadi, R. Ranjbarzadeh, F. Naiemi, M. Mogharrebi, S. Dorosti, and M. Bendechache, "TPCNN: two-path convolutional neural network for tumor and liver segmentation in CT images using a novel encoding approach," *Expert Systems with Applications*, vol. 183, Article ID 115406, 2021.

[15] S. J. Ghoushchi, R. Ranjbarzadeh, A. H. Dadkhah, Y. Pourasad, and M. Bendechache, "An extended approach to predict retinopathy in diabetic patients using the genetic algorithm and fuzzy C-means," *BioMed Research International*, vol. 2021, Article ID 5597222, 13 pages, 2021.

[16] Z. Liu, Y.-Q. Song, V. S. Sheng et al., "Liver CT sequence segmentation based with improved U-Net and graph cut," *Expert Systems with Applications*, vol. 126, pp. 54–63, 2019.

[17] G. Chen, C. Li, W. Wei et al., "Fully convolutional neural network with augmented atrous spatial pyramid pool and fully connected fusion path for high resolution remote sensing image segmentation," *Applied Sciences*, vol. 9, no. 9, 2019.

[18] F. Nensa, A. Demircioglu, and C. Rischpler, "Artificial intelligence in nuclear medicine," *Journal of Nuclear Medicine*, vol. 60, no. 2, pp. 29S–37S, 2019.

[19] D. Visvikis, C. Cheze Le Rest, V. Jaouen, and M. Hatt, "Artificial intelligence, machine (deep) learning and radio(geno)mics: definitions and nuclear medicine imaging applications," *European Journal of Nuclear Medicine and Molecular Imaging*, Springer, vol. 46, no. 13, pp. 2630–2637, 2019.

[20] A. Mahmood, M. Bennamoun, S. An et al., "Deep learning for coral classification," in *Handbook of Neural Computation*, Elsevier, Amsterdam, Netherland, 2017.

[21] S. Hussein, P. Kandel, C. W. Bolan, M. B. Wallace, and U. Bagci, "Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches," *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1777–1787, 2019.

[22] Y. Xu, Y. Wang, J. Yuan, Q. Cheng, X. Wang, and P. L. Carson, "Medical breast ultrasound image segmentation by machine learning," *Ultrasonics*, vol. 91, pp. 1–9, 2019.

[23] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," *Lecture Notes in Computer Science*, vol. 9351, pp. 234–241, 2015.

[24] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: a deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[25] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings - 2016 4th International Conference on 3D Vision*, Stanford,CA, USA, December 2016.

[26] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," *Computer Vision - ECCV 2014*, vol. 8689, pp. 818–833, 2014.

[27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015, http://www.robots.ox.ac.uk/.

[28] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[29] C. Szegedy, "Going deeper with convolutions," 2015.

[30] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," 2017, https://github.com/liuzhuang13/DenseNet.

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2016, http://image-net.org/challenges/LSVRC/2015/.

[32] X. Qi, L. Zhang, Y. Chen et al., "Automated diagnosis of breast ultrasonography images using deep neural networks," *Medical Image Analysis*, vol. 52, pp. 185–198, 2019.

[33] R. Ranjbarzadeh and S. B. Saadi, "Automated liver and tumor segmentation based on concave and convex points using fuzzy c-means and mean shift clustering," *Measurement*, vol. 150, Article ID 107086, 2020.

[34] K. Yasaka, H. Akai, O. Abe, and S. Kiryu, "Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study," *Radiology*, vol. 286, no. 3, pp. 887–896, 2018.

[35] R. Ranjbarzadeh, S. Jafarzadeh Ghoushchi, M. Bendechache et al., "Lung infection segmentation for COVID-19

pneumonia based on a cascade convolutional network from CT images," *BioMed Research International*, vol. 2021, Article ID 5544742, 16 pages, 2021.

[36] M. Ahmadi, A. Sharifi, S. Dorosti, S. Jafarzadeh Ghoushchi, and N. Ghanbari, "Investigation of effective climatology parameters on COVID-19 outbreak in Iran," *The Science of the Total Environment*, vol. 729, Article ID 138705, 2020.

[37] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1207–1216, 2016.

[38] A. Lübbeke, G. Garavaglia, C. Barea, R. Stern, R. Peter, and P. Hoffmeyer, "Influence of patient activity on femoral osteolysis at five and ten years following hybrid total hip replacement," *The Journal of Bone and Joint Surgery. British volume*, vol. 93, no. 4, pp. 456–463, 2011.

[39] L. Ye, M. D. Mason, and W. G. Jiang, "Bone morphogenetic protein and bone metastasis, implication and therapeutic potential," *Frontiers in Bioscience*, vol. 16, no. 1, pp. 865–897, 2011.

[40] S. B. Goodman, E. Gibon, and Z. Yao, "The basic science of periprosthetic osteolysis," *Instructional Course Lectures*, vol. 62, pp. 201–206, 2013, http://pmc/articles/PMC3766766/.

[41] S. M. Horowitz, S. B. Doty, J. M. Lane, and A. H. Burstein, "Studies of the mechanism by which the mechanical failure of polymethylmethacrylate leads to bone resorption," *The Journal of Bone & Joint Surgery*, vol. 75, no. 6, pp. 802–813, 1993.

[42] J. H. Dumbleton, M. T. Manley, and A. A. Edidin, "A literature review of the association between wear rate and osteolysis in total hip arthroplasty," *The Journal of Arthroplasty*, vol. 17, no. 5, pp. 649–661, 2002.

[43] W. H. Harris, "The lysis threshold: an erroneous and perhaps misleading concept?" *The Journal of Arthroplasty*, vol. 18, no. 4, pp. 506–510, 2003.

[44] H. A. McKellop, "The lexicon of polyethylene wear in artificial joints," *Biomaterials*, vol. 28, no. 34, pp. 5049–5057, 2007.

[45] G. Singh, T. Reichard, R. Hameister et al., "Ballooning osteolysis in 71 failed total ankle arthroplasties," *Acta Orthopaedica*, vol. 87, no. 4, pp. 401–405, 2016.

[46] J. Bertrand, D. Delfosse, V. Mai, F. Awiszus, K. Harnisch, and C. H. Lohmann, "Ceramic prosthesis surfaces induce an inflammatory cell response and fibrotic tissue changes," *The Bone & Joint Journal*, vol. 100, no. 7, pp. 882–890, 2018.

[47] W. Z. W. Teo and P. C. Schalock, "Metal hypersensitivity reactions to orthopedic implants," *Dermatology and Therapy*, vol. 7, no. 1, pp. 53–64, 2017.

[48] J. H. Ingram, R. Kowalski, J. Fisher, and E. Ingham, "The osteolytic response of macrophages to challenge with particles of Simplex P, Endurance, Palacos R, and Vertebroplastic bone cement particlesin vitro," *Journal of Biomedical Materials Research Part B: Applied Biomaterials*, vol. 75B, no. 1, pp. 210–220, 2005.

[49] F. Vega, R. Bazire, M. T. Belver, M. V. Mugica, A. Urquia, and C. Blanco, "Aseptic loosening of a total knee prosthesis caused by delayed hypersensitivity to bone cement," *Annals of Allergy, Asthma, & Immunology*, vol. 117, no. 1, pp. 89–91, 2016.

[50] S. Grosse, H. K. Haugland, P. Lilleng, P. Ellison, G. Hallan, and P. J. Høl, "Wear particles and ions from cemented and uncemented titanium-based hip prostheses-A histological and chemical analysis of retrieval material," *Journal of Biomedical Materials Research Part B: Applied Biomaterials*, vol. 103, no. 3, pp. 709–717, 2015.

[51] A. Sabokbar, R. Pandey, and N. A. Athanasou, "The effect of particle size and electrical charge on macrophage-osteoclast differentiation and bone resorption," *Journal of Materials Science: Materials in Medicine*, vol. 14, no. 9, pp. 731–738, 2003.

[52] T. R. Green, J. Fisher, J. B. Matthews, M. H. Stone, and E. Ingham, *Effect of Size and Dose on Bone Resorption Activity of Macrophages by in Vitro Clinically Relevant Ultra High Molecular Weight Polyethylene Particles*, Wiley Online Library, Hoboken, NJ, USA, 2000.

[53] H. M. Van Der Vis, P. Aspenberg, R. K. Marti, W. Tigchelaar, and C. J. F. Van Noorden, "Fluid pressure causes bone resorption in a rabbit model of prosthetic loosening," *Clinical Orthopaedics and Related Research*, vol. 350, pp. 201–208, 998.

[54] D. R. Haynes, S. D. Rogers, S. Hay, M. J. Pearcy, and D. W. Howie, "The differences in toxicity and release of bone-resorbing mediators induced by titanium and cobalt-chromium-alloy wear particles," *The Journal of Bone & Joint Surgery*, vol. 75, no. 6, pp. 825–834, 1993.

[55] T. P. Schmalzried, M. Jasty, A. Rosenberg, and W. H. Harris, "Polyethylene wear debris and tissue reactions in knee as compared to hip replacement prostheses," *Journal of Applied Biomaterials*, vol. 5, no. 3, pp. 185–190, 1994, Sep. 1994.

[56] M. O. Soto, J. A. Rodriguez, and C. S. Ranawat, "Clinical and radiographic evaluation of the harris-galante cup," *The Journal of Arthroplasty*, vol. 15, no. 2, pp. 139–145, 2000.

[57] T. P. Schmalzried, I. C. Brown, H. C. Amstutz, C. A. Engh, and W. H. Harris, "The role of acetabular component screw holes and/or screws in the development of pelvic osteolysis," *Proceedings of the Institution of Mechanical Engineers - Part H: Journal of Engineering in Medicine*, vol. 213, no. 2, pp. 147–153, 1999.

[58] H. E. Rubash, R. K. Sinha, W. Paprosky, C. A. Engh, and W. J. Maloney, "A new classification system for the management of acetabular osteolysis after total hip arthroplasty," *Instructional Course Lectures*, vol. 48, pp. 37–42, 1999, https://europepmc.org/article/med/10098026.

[59] F. S. Haddad, B. A. Masri, D. S. Garbuz, and C. P. Duncan, "Femoral bone loss in total hip arthroplasty: classification and preoperative planning," *Instructional Course Lectures*, vol. 49, pp. 83–96, 2000, https://europepmc.org/article/med/10829164.

[60] D. G. Campbell, B. Masri, D. Garbuz, and C. P. Duncan, "Acetabular bone loss during revision total hip replacement: preoperative investigation and planning - pubmed," 2011, https://pubmed.ncbi.nlm.nih.gov/10098027/.

[61] S. B. Goodman, P. Huie, Y. Song et al., "Cellular profile and cytokine production at prosthetic interfaces," *The Journal of Bone and Joint Surgery. British volume*, vol. 80, no. 3, pp. 531–539, 1998.

[62] B. Zicat, C. A. Engh, and E. Gokcen, "Patterns of osteolysis around total hip components inserted with and without cement," *The Journal of Bone & Joint Surgery*, vol. 77, no. 3, pp. 432–439, 1995.

[63] N. Papandrianos, E. Papageorgiou, A. Anagnostis, and K. Papageorgiou, "Efficient bone metastasis diagnosis in bone scintigraphy using a fast convolutional neural network architecture," *Diagnostics*, vol. 10, no. 8, p. 532, 2020.

[64] H. G. Potter, B. J. Nestor, C. M. Sofka, S. T. Ho, L. E. Peters, and E. A. Salvati, "Magnetic resonance imaging after total hip arthroplasty: eva," 2013, https://journals.lww.com/jbjsjournal/Abstract/2004/09000/Magnetic_Resonance_Imaging_After_Total_Hip.13.aspx.

[65] M. M. Madkour and H. Sharif, "Bone and joint imaging," *Madkour's Brucellosis*, vol. 32, pp. 90–132, 2021.

[66] D. Sato, S. Hanaoka, Y. Nomura et al., "A primitive study on unsupervised anomaly detection with an autoencoder in emergency head CT volumes," *Medical Imaging 2018: Computer-Aided Diagnosis*, vol. 10575, p. 60, 2018.

[67] H. Watanabe, R. Togo, T. Ogawa, and M. Haseyama, "Bone metastatic tumor detection based on AnoGAN using CT images," *2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech)*, vol. 65, pp. 235-236, 2019.

[68] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural Networks*, vol. 12, no. 1, pp. 145–151, 1999.

[69] X.-S. Wei, C.-W. Xie, and J. Wu, "Mask-CNN: localizing parts and selecting descriptors for fine-grained image recognition," 2016, http://arxiv.org/abs/1605.06878 Accessed: Jun. 11, 2021. [Online]. Available:.

[70] N. Zhang, J. Donahue, R. Girshick, and T. Darrell, "Part-based R-CNNs for fine-grained category detection," *Computer Vision - ECCV 2014*, vol. 8689, pp. 834–849, 2014.

[71] Y. Li, W. Zhou, G. Lv, G. Luo, Y. Zhu, and J. Liu, "Classification of bone tumor on CT images using deep convolutional neural network," *Artificial Neural Networks and Machine Learning - ICANN 2018*, vol. 11140, pp. 127–136, 2018.

[72] S. R. Fernandes, J. T. de Assis, V. V. Estrela et al., "Non-destructive diagnosis and analysis of computed micro-tomography images via texture descriptors," *Advances in Multidisciplinary Medical Technologies — Engineering, Modeling and Findings*, vol. 32, pp. 249–261, 2021.

[73] M. Ahmadi, A. Sharifi, M. Jafarian Fard, and N. Soleimani, "Detection of brain lesion location in MRI images using convolutional neural network and robust PCA," *International Journal of Neuroscience*, vol. 46, pp. 1–12, 2021.

[74] R. Ranjbarzadeh, S. B. Saadi, and A. Amirabadi, "LNPSS: SAR image despeckling based on local and non-local features using patch shape selection and edges linking," *Measurement*, vol. 164, Article ID 107989, 2020.

[75] J. Liu, J. Wang, W. Ruan, C. Lin, and D. Chen, "Diagnostic and gradation model of osteoporosis based on improved deep U-net network," *Journal of Medical Systems*, vol. 44, no. 1, pp. 1–7, 2020.

[76] P. Qian, Y. Jiang, Z. Deng et al., "Cluster prototypes and fuzzy memberships jointly leveraged cross-domain maximum entropy clustering," *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 181–193, 2016.

[77] P. Qian, Y. Jiang, S. Wang et al., "Affinity and penalty jointly constrained spectral clustering with all-compatibility, flexibility, and robustness," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 5, pp. 1123–1138, 2017.

[78] N.-T. Do, S.-T. Jung, H.-J. Yang, and S.-H. Kim, "Multi-level seg-unet model with global and patch-based X-ray images for knee bone tumor detection," *Diagnostics*, vol. 11, no. 4, p. 691, 2021 Apr. 2021.

[79] J. R. Wilkie, M. L. Giger, C. A. Engh, Sr., Martell ., and J. M. Martell, "Investigation of temporal radiographic texture analysis for the detection of periprosthetic osteolysis," *Medical Imaging 2006: Image Processing*, vol. 6144, Article ID 61446Z, 2006.

[80] P. Caligiuri, M. L. Giger, M. J. Favus, H. Jia, K. Doi, and L. B. Dixon, "Computerized radiographic analysis of osteoporosis: Preliminary evaluation," *Radiolog*, vol. 186, no. 2, pp. 471–474, 1993, Feb. 01.

[81] M. R. Chinander, M. L. Giger, J. M. Martell, C. Jiang, and M. J. Favus, "Computerized radiographic texture measures for characterizing bone strength: a simulated clinical setup using femoral neck specimens," *Medical Physics*, vol. 26, no. 11, pp. 2295–2300, 1999, Nov. 1999.

[82] L. K. Smith, F. Cramp, S. Palmer, N. Coghill, and R. F. Spencer, "Use of morphometry to quantify osteolysis after total hip arthroplasty," *Clinical Orthopaedics and Related Research*, vol. 468, no. 11, pp. 3077–3083, 2010.

[83] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nature Methods*, vol. 9, no. 7, pp. 671–675, 2012, Jul. 28.

[84] H. R. Evans, T. Karmakharm, M. A. Lawson et al., "Osteolytica: an automated image analysis software package that rapidly measures cancer-induced osteolytic lesions in in vivo models with greater reproducibility compared to other commonly used methods," *Bone*, vol. 83, pp. 9–16, 2016, Feb. 2016.

[85] J. Davis, S. R. Marschner, M. Garr, and M. Levoy, "Filling holes in complex surfaces using volumetric diffusion," *Proceedings. First International Symposium on 3D Data Processing Visualization and Transmission*, vol. 18, pp. 428–441, 2002.

[86] E. R. Parmar, M. Jermoumi, R. H. Mak et al., "Volumetric CT-based segmentation of NSCLC using 3D-Slicer," *Scientific Reports*, vol. 3, no. 1, 2013.

[87] A. Fedorov, R. Beichel, J. Kalpathy-Cramer et al., "3D slicer as an image computing platform for the quantitative imaging network," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012.

[88] S. H. Rizzi, P. P. Banerjee, and C. J. Luciano, "Automating the extraction of 3D models from medical images for virtual reality and haptic simulations," *2007 IEEE International Conference on Automation Science and Engineering*, vol. 23, pp. 152–157, 2007.

[89] M. Mccormick, X. Liu, J. Jomier, C. Marion, and L. Ibanez, "Itk: enabling reproducible research and open science," *Frontiers in Neuroinformatics*, vol. 8, p. 13, 2014.

[90] T. S. Yoo, "International news - diversity of technical papers offered at SAMPE 2002," *Materials & Design*, vol. 23, no. 6, pp. 586–592, 2002.

[91] I. Wolf, M. Vetter, I. Wegner et al., "The medical imaging interaction toolkit," *Medical Image Analysis*, vol. 9, no. 6, pp. 594–604, 2005.

[92] M. Nolden, S. Zelzer, A. Seitel et al., "The medical imaging interaction toolkit: challenges and advances," *International Journal of Computer Assisted Radiology and Surgery*, vol. 8, no. 4, pp. 607–620, 2013.

[93] W. J. Schroeder, L. S. Avila, and W. Hoffman, "Visualizing with VTK: a tutorial," *IEEE Computer Graphics and Applications*, vol. 20, no. 5, pp. 20–27, 2000.

[94] Z. Zhao, Y. Pi, L. Jiang et al., "Deep neural network based artificial intelligence assisted diagnosis of bone scintigraphy for cancer bone metastasis," *Scientific Reports*, vol. 10, no. 1, pp. 1–9, 2020.

[95] J.-A. Pérez-Carrasco, B. Acha, C. Suárez-Mejías, J.-L. López-Guerra, and C. Serrano, "Joint segmentation of bones and muscles using an intensity and histogram-based energy

minimization approach," *Computer Methods and Programs in Biomedicine*, vol. 156, pp. 85–95, 2018, Mar. 2018.

[96] S. Noguchi, M. Nishio, M. Yakami, K. Nakagomi, and K. Togashi, "Bone segmentation on whole-body CT using convolutional neural network with novel data augmentation techniques," *Computers in Biology and Medicine*, vol. 121, 2020 Jun. 2020, Article ID 103767.

[97] Y. He, I. Pan, B. Bao et al., "Deep learning-based classification of primary bone tumors on radiographs: a preliminary study," *EBioMedicine*, vol. 62, Article ID 103121, 2020.

[98] N.-H. Ho, H.-J. Yang, S.-H. Kim, S. T. Jung, and S.-D. Joo, "Regenerative semi-supervised bidirectional w-network-based knee bone tumor classification on radiographs guided by three-region bone segmentation," *IEEE Access*, vol. 7, pp. 154277–154289, 2019.

[99] E. Hossain, M. F. Hossain, and M. A. Rahaman, "An approach for the detection and classification of tumor cells from bone MRI using wavelet transform and KNN classifier," " *2018 International Conference on Innovation in Engineering and Technology (ICIET)*, vol. 14, 2018.

[100] J. G. Dalling, K. Math, and G. R. Scuderi, "Evaluating the progression of osteolysis after total knee arthroplasty," *Journal of the American Academy of Orthopaedic Surgeons*, vol. 23, no. 3, pp. 173–180, 2015.

[101] M. H. L. Liow and Y.-M. Kwon, "Metal-on-metal total hip arthroplasty: risk factors for pseudotumours and clinical systematic evaluation," *International Orthopaedics*, vol. 41, no. 5, pp. 885–892, 2017.

[102] R. L. Smith and E. M. Schwarz, "Are biologic treatments a potential approach to wear- and corrosion-related problems?" *Clinical Orthopaedics and Related Research*, vol. 472, no. 12, pp. 3740–3746, 2014.

[103] G. Holt, J. Reilly, and R. M. D. Meek, "Effect of alendronate on pseudomembrane cytokine expression in patients with aseptic osteolysis," *The Journal of Arthroplasty*, vol. 25, no. 6, pp. 958–963, 2010.

[104] J. M. Wilkinson and D. G. Little, "Bisphosphonates in orthopedic applications," *Bone*, vol. 49, no. 1, pp. 95–102, 2011, Jul. 01.

[105] J. Pajarinen, T. Lin, E. Gibon et al., "Mesenchymal stem cell-macrophage crosstalk and bone healing," *Biomaterials*, vol. 196, pp. 80–89, 2019, Mar. 2019.

[106] T. Lin, J. Pajarinen, A. Nabeshima et al., "Establishment of NF-κB sensing and interleukin-4 secreting mesenchymal stromal cells as an on-demand drug delivery system to modulate inflammation," *Cytotherapy*, vol. 19, no. 9, pp. 1025–1034, 2017.

[107] T. Lin, Y. Kohno, J.-F. Huang et al., "Preconditioned or IL4-secreting mesenchymal stem cells enhanced osteogenesis at different stages," *Tissue Engineering Part A*, vol. 25, no. 15-16, pp. 1096–1103, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 2214971, 11 pages
https://doi.org/10.1155/2021/2214971

Research Article

Evaluation of Ergonomics-Related Disorders in Online Education Using Fuzzy AHP

Hemant Upadhyay,[1] **Sapna Juneja**,[2] **Abhinav Juneja**,[3] **Gaurav Dhiman**,[4] **and Sandeep Kautish**[5]

[1]*BMIET, Sonepat, India*
[2]*IMS Engineering College, Ghaziabad, India*
[3]*KIET Group of Institutions, Delhi NCR, Ghaziabad, India*
[4]*Government Bikram College for Commerce, Patiala, India*
[5]*LBEF Campus, Kathmandu, Nepal*

Correspondence should be addressed to Sapna Juneja; sapnajuneja1983@gmail.com and Sandeep Kautish; dr.skautish@gmail.com

Received 7 July 2021; Accepted 18 September 2021; Published 27 September 2021

Academic Editor: Navid Razmjooy

The aim of the presented work is to analyze the ergonomics-related disorders in online education using the fuzzy AHP approach. A group dialogue with online education academicians, online education students, biotechnologists, and sedentary computer users has been performed to spot ergonomics-related disorders in online education. Totally eight ergonomics-related disorders in online education have been identified, and the weight of each disorder has been computed with triangle-shaped fuzzy numbers in pairwise comparison. Furthermore, the ergonomics-related disorders in online education are kept in four major categories such as afflictive disorders, specific disorders, psychosocial disorders, and chronic disorders. These four categories of ergonomics-related disorders in online education are evaluated and compared using fuzzy analytical hierarchical process methodology to get ranked in terms of priorities. The results may be instrumental for taking appropriate corrective actions to prevent ergonomics-related disorders.

1. Introduction

The international associations have termed ergonomics to be "the design of work, in such a manner that human competencies can be utilized in the best possible manner without overcoming human constraints" [1]. Ergonomics is the scientific know-how of the man at work, with the numerous psychosocial and medical characteristics of human work. The practical objective of ergonomics is the conditioning and justification of the adaptation of work to man [2]. The ergonomics literature such as that of Hünting et al. [3], Sauter et al. [4], and Berqvist et al. [5] have been mostly on the basis of desktop and laptop computers.

Several documents of reputed research and academic institutions [6] indicate common ergonomics-based documentation. These days, a large number of studies on computer-human interfaces have introduced newer ergonomics criteria and recommendations [2]. It has been observed that if essential precautions were not considered for inappropriate and very frequent computer utilization in our day-to-day lives, a considerable enhancement is noted in the number of disorder practicing persons as a result of the screen time with laptop or desktop for a longer duration [7]. Heiden et al. [8] suggest that by accommodating subjective rankings by the employees as well as objective computation of the work surroundings, an analysis can be made to develop an exhortation regarding light availability and further emphasize the relevance that visual surrounding factors like glare are the reason for headache and eyestrain. The presented paper is an investigation work in ergonomics of computer-human interface that recommends the evaluation of ergonomics-related disorders in online education

according to four main categories such as afflictive disorders, special disorders, psychosocial disorders, and chronic disorders and presents the ability of fuzzy AHP methodology in selecting and prioritizing of above criteria.

The presented paper has been set as follows. Section 2 is about online education and types of ergonomics-related disorders in online education. Fuzzy AHP methodology is described in Section 3. A proposed hierarchical model for categorization of ergonomics-related disorders has also been discussed in Section 3. Section 4 brings the results. Finally, in Section 5, conclusions are drawn, and future scope is mentioned.

2. Literature Review

People who have been permanent stakeholders of the information world can explore a virtual-based life available in the World Wide Web at the computer system for data searching. Various in-service learning events are conducted for making the teachers be well-versed with computer engineering and information and communication technologies for playing a significant role in the execution of these activities [9]. A variety of commercial companies are linked with the growth of the progression of e-learning [10]. A teacher facilitates a knowledge analyst by demonstrating the subject's concept [11].

The epidemic caused by lockdown has enhanced the use of the online mode of learning and forced academicians to adjust education mode [12]. It is notable that online education surroundings are developed and analyzed in academic patterns by considering both utility and erudition aspects [13]. The unparallel magnification of web-based engineering is leading towards the emergence of several methods for the area of academic visibility in the online education scenario [14]. The objective for the utilization of online education technology is to modify the way of teaching in academic institutions from the conventional approach towards the more participated and interactive [15]. The recent changes in online learning have witnessed unprecedented growth in the last couple of years; furthermore, the current pandemic condition has accelerated the process of online learning management [16]. Recently, the pandemic has changed all aspects of our lives. Social isolation has deranged conventional educational practices and has influenced traditional schooling and training. There is an intense requirement to innovate and implement alternative educational and evaluation procedures. Next, ergonomics-based adverse effects related to sedentary sitting have been shortlisted in different categories. Fuzzy analytical hierarchical process-based hierarchical model of eight chosen ergonomics-related disorders in four categories has been shown in Figure 1.

2.1. Afflictive Disorders. Afflictive disorders are a set of circumstances affecting the joints, tendons, and muscles [17]. These can impact the lower back, shoulders, neck, and wrists. These are usually because of a poor working posture [18]. The afflictive disorders can be visualized in Figure 2 as how poor working posture with sedentary sitting can impact the lower back and neck.

2.1.1. Lower Back Pain. The application of computers and laptops and sedentary sitting may develop ascendancy on the lower part of the back of desktop users, as a result of inadequate posture [19]. Low back pain is practiced by computer users who have to work with a keyboard while in a sedentary position [20]. Lower back pain simultaneously creates effects on other muscles resulting in pain [21].

2.1.2. Neck Pain. There is a strong relation between sedentary sitting and pain in the neck, with an affirmative association of pain in neck and neck flexion [22]. There is a larger decline to be linked with less menace of musculoskeletal strain in the neck [23].

2.2. Specific Disorders

2.2.1. Eyestrain. Eyestrain is a family of vision or eyesight-related issues that result from a longer use of computer system, tablet, or mobile phone [24]. Eye discomfort and visual problems have been linked to computerized tasks [25]. The abundance of CVS in computer system regular users is indeed linked to the screen time duration of the computer system [26].

2.2.2. Hearing Loss. Enhancing measures of unprecedented occupational and surrounding commotion have become the prime reasons for commotion-generated hearing loss and form a vital public health issue [27]. The application of individual instruments for listening has been recognized as a prime cause in the expansion of noise-induced ear issues [28]. The stereotypically used devices are the earbud fashioned headphones that have been related to a highly demanded and absolute measure of listening [29].

2.3. Psychosocial Disorders. The management of psychosocial factors is often not psychosocial in nature. At times, it may be very tedious to correct psychosocial conditions directly [30].

2.3.1. Mental Stress. The definition of "stress" may be a general reaction to a stressor, consisting of several physiological reactions. Many modeling frameworks demonstrate that detrimental psychosocial issues become the reason behind mental stress [23]. Several studies emphasize the importance of stress [31]. The longer the sedentary period, the higher the perceived stress score; it is significant up to a great extent [32].

2.3.2. Fatigue. To spot the manifestation of ergonomic disorders, a questionnaire [33] was used enabling recognition of parts of the body where individuals practiced symptoms of fatigue. It is unfounded that interventions that focus on sitting less and moving more often increase fatigue [34].

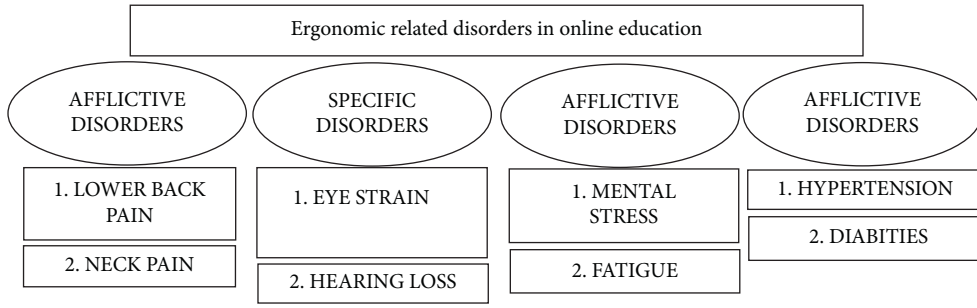


Figure 1: The proposed fuzzy AHP-based hierarchical model.

Figure 2: Poor working posture impacting lower back pain and neck pain.

2.4. Chronic Disorders

2.4.1. Hypertension. Numerous studies have demonstrated correlations between significant Internet use and health issues such as hypertension. Screen time has been related to enhanced blood pressure independent of body composition (American Heart Association). Interrupting sedentary sitting on the computer may considerably decrease hypertension [35].

2.4.2. Diabetes. Statements of internationally recognized agencies show that diabetes has enclosed particular recommendations to decrease sedentary sitting of computer users [36]. Continuous and everyday stable sitting hours on computers may be related to poorer health outcomes in those with diabetes. Dempsey et al. [37] demonstrated a relation between diabetes and a day of sedentary sitting on a computer.

3. Methodology

3.1. Fuzzy AHP. The current work has proposed the use of fuzzy analytical hierarchical process methodology in the analysis of ergonomics-related disorders in online education. The AHP, firstly discussed by Saaty [38], proposed a method for the calculation of the relative importance in an MCDM problem [39]. The traditional analytical hierarchical process is not sufficient to manage the ambiguous characteristic of lingual analysis [40]. In analytical hierarchical process, all comparisons are not included [41]. Analytical hierarchical process is unable to deal with its constraints such as (a) analytical hierarchical process generates the decisions in crisp nature and has been considered as an approach with less accuracy and (b) assessment of alternatives, done by the group of experts, is based on approximation and is the intuitive analysis of approximate computation [42]. The utilization of fuzzy set theory [43] may permit the decision-makers to accommodate qualitative data, improper data, and ignorant facts in the decision model [44]. The evolution of the fuzzy analytical hierarchical process has shown its capabilities for the resolution of complex problems with considerable accuracy [45]. The fuzzy analytical hierarchical process model is very instrumental in the ranking of nonobjective factors [46]. Even after complex calculations, fuzzy AHP is able to deal with the general appraisal of ambiguity because of human tendency [47]. The advantage of fuzzy logic is much accurate computation of the factors prioritizing and hierarchy creation [48]. Popular technique, fuzzified analytical hierarchical process is capable of solving the complicated problems

related to selection of vendors [49]. The utilization of the fuzzy analytical hierarchical process has been becoming much popular in several specializations due to its ability to work as disintegration technique [50].

The aim of F-AHP is to cope with complicated decision science issues using a hierarchy with main criteria and the computation of eigenvectors [51]. Fuzzy AHP is instrumental in assessing pairwise comparative study, alternatives, and factors [52]. Triangle fuzzy numbering structures are utilized for pairwise comparative tables [53]. The geometric mean technique assists in computing the fuzzy weightages and hierarchic rank system of the factors [54]. Numerous mathematical methods are available for the deriving process from fuzzy pairwise comparative matrices into crisp weights [55]. Fuzzy AHP is a simple and presentiment technique with appropriate validation of consistent index [56]. Past relevant work exploration shows that fuzzy analytical hierarchical process has the potential to get several complex multicriteria decision-making problems solved as some research works are mentioned in Table 1.

3.2. Proposed Model. Fuzzy AHP is a modified version of an analytic hierarchy process with the theory based on fuzzified logic. Fuzzy analytic hierarchy process frames the fuzzy triangular scale consisting of lower, middle, and upper values for computation of priorities. We have three frequently employed fuzzy analytical hierarchical process methods but probably the first fuzzy AHP technique presented by van Laarhoven and Pedrycz [60]. Mikhailov [61] suggested a fuzzy programming process to determine optimized crisp priorities that are achieved with fuzzy PWC judgments. Chang [62] suggested a unique method of integers and fractions in triangular structure in pairwise CMs. Afterwards, Buckley [63] extended the study by computing the fuzzy priority of each comparative ratio. Pairwise comparative matrices are allocated to elements of the analysis hierarchy of the Saaty scale [64, 65]. Buckley's technique [63] is instrumental in computing the relative weightages for solutions as well as factors in fuzzy AHP. This methodology suggests the utilization of geometric means achieving fuzzy weightages that enhance the ease for computation of the local weightages [66]. The research framework of the present study has been shown in Figure 3.

We have different types of fuzzy numbers, but the triangle shape fuzzy numbers have been mostly applied, and the number structure is a set of three numbers (lower, medium, and upper). Real number values (lower, medium, and upper) have the triangle structure as "l," "m," and "n" as the least probability number, the most probability value, and the largest probability number, respectively [21]. Linguistic-based scaling for triangle fuzzy numbers has been indicated in Table 2.

3.3. Proposed Hierarchy. Abbreviations of the chosen ergonomics-related disorders have been category-wise proposed in Table 3.

4. Results

For the validation of our proposed model, a group dialogue with online education academicians, online education students, biotechnologists, and sedentary computer users has been performed in detail. We demonstrate the formation of fuzzy pairwise comparison of ergonomics-related disorders in online education (Table 4).

Furthermore, the GM of fuzzy compared values of each ergonomics-related disorder in online education is calculated as shown in Table 5.

Furthermore, the fuzzy weight of every ergonomics-related disorder in online education is computed in Table 6. Lower, medium, and upper fuzzy weights have been calculated for each ergonomics-related disorder by multiplying respective geometric means with the specified factor (as calculated in the last row of Table 5).

Furthermore, the crisp values of the weight of every ergonomics-related disorder (Mi) have been computed as the arithmetic mean of fuzzified values for every disorder. Finally, the crisp weights of every ergonomics-related disorder are put after normalization as Ni as tabulated in Table 7.

Table 7 shows that neck pain (D2) has the least-normalized relative weight among the chosen ergonomics-related disorders in online education.

It has been shown in Figure 4 that mental stress has the greatest normalized relative weight among the discussed ones. Hypertension has the second-highest normal weightages, and eyestrain has the third-highest normal weightages. On the other hand, neck pain has the least normal weight out of all eight considered ergonomics-related disorders in online education.

Table 8 depicts the priority computation by fuzzy analytical hierarchical process method and the rank calculation of the different subcriteria. It has been found that mental stress is the most important, followed by hypertension and eyestrain as an ergonomics-related disorder in online education.

It has been shown in Figure 5 that priority of "neck pain" is the least in selected ergonomics-related disorders. Figure 5 is the demonstration of priorities of all eight considered ergonomics-related disorders with the help of a pie chart. Different colors are presenting priorities of different ergonomics-related disorders.

Table 9 expresses the priority computed using the fuzzy analytical hierarchical process method.

Based on the values of Table 9, psychosocial disorders have the first place, followed by chronic disorders, specific disorders, and then afflictive disorders.

Figure 6 signifies that the weightage of "psychosocial disorders" is the most (45.7%), and this is the most important ergonomics-related disorder category among the discussed ones.

It has been shown in Figure 7 that the priority of "afflictive disorders" category is the least (5.8%) ranked ergonomics-related disorder category in the chosen ones.

The outcome of the work by some of the similar efforts has been summarized in Table 10.

The prime objective of the new specialization of "human factors and ergonomics" is its exploration for applying of

Table 1: Recent works on applications of fuzzy AHP.

S. No.	Author	Objective and outcome
1	Akbar et al.	The authors [57] prioritized the factors about the scaling procedure of agile methodology for the GSD industry of a particular nation and developed their taxonomy using fuzzy analytical hierarchical process approach.
2	Sunday Oyinlola Ogundoy and Ismaila Adeniyi Kamil	The authors [58] used the fuzzy AHP to rank trust factors in fog computing. They observed the most- and least-preferred factors. Moreover, they considered the subcategory of every criterion for ranking them using the fuzzy analytical hierarchical process approach.
3	Kumar et al.	The authors [59] developed the preference-based categorization of 21 software processing development success attributes using the fuzzified AHP.

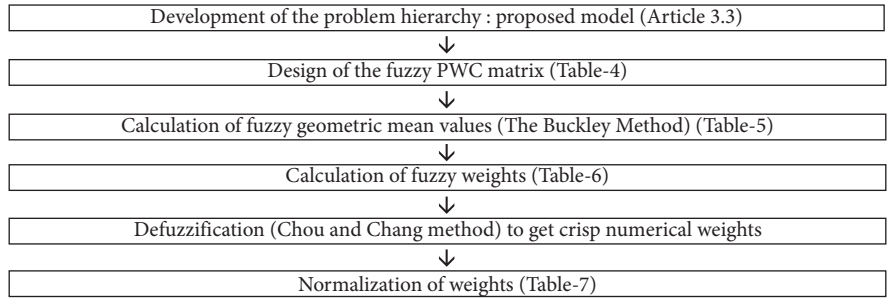


Figure 3: The research framework deployed in this study.

Table 2: Satty-scale-based linguistic-based scaling for triangle fuzzy numerals.

Scale	Lingual scale for relative weightage	Fuzzified values	Respective values
1	Identical superiority	(1, 1, 1)	(1, 1, 1)
2	Identical to weak superiority	(1, 2, 3)	(1/3, 1/2, 1)
3	Weak superiority	(2, 3, 4)	(1/4, 1/3, 1/2)
4	Weak to moderate superiority	(3, 4, 5)	(1/5, 1/4, 1/3)
5	Moderate superiority	(4, 5, 6)	(1/6, 1/5, 1/4)
6	Moderate to strong superiority	(5, 6, 7)	(1/7, 1/6, 1/5)
7	Strong superiority	(6, 7, 8)	(1/8, 1/7, 1/6)
8	Very strong superiority	(7, 8, 9)	(1/9, 1/8, 1/7)
9	Extreme superiority	(9, 9, 9)	(1/9, 1/9, 1/9)

Table 3: The abbreviations of ergonomics-related disorders.

Criteria	Disorder	Abbreviation
Afflictive disorders	Lower back pain	Di1
	Neck pain	Di2
Specific disorders	Eyestrain	Di3
	Hearing loss	Di4
Psychosocial disorders	Mental strain	Di5
	Fatigue	Di6
Chronic disorders	Hypertension	Di7
	Diabetes	Di8

Table 4: Fuzzy pairwise comparison of ergonomics-related disorders in online education.

Disorder	Di1	Di2	Di3	Di4	Di5	Di6	Di7	Di8
Di1	(1, 1, 1)	(3, 4, 5)	(1/6, 1/5, 1/4)	(2, 3, 4)	(1/8, 1/7, 1/6)	(1/5, 1/4, 1/3)	(1/7, 1/6, 1/5)	(1/4, 1/3, 1/2)
Di2	(1/5, 1/4, 1/3)	(1, 1, 1)	(1/8, 1/7, 1/6)	(1/4, 1/3, 1/2)	(1/9, 1/9, 1/9)	(1/7, 1/6, 1/5)	(1/9, 1/8, 1/7)	(1/6, 1/5, 1/4)
Di3	(4, 5, 6)	(6, 7, 8)	(1, 1, 1)	(5, 6, 7)	(1/5, 1/4, 1/3)	(2, 3, 4)	(1/4, 1/3, 1/2)	(3, 4, 5)
Di4	(1/4, 1/3, 1/2)	(2, 3, 4)	(1/7, 1/6, 1/5)	(1, 1, 1)	(1/9, 1/8, 1/7)	(1/6, 1/5, 1/4)	(1/8, 1/7, 1/6)	(1/5, 1/4, 1/3)

Table 4: Continued.

Disorder	Di1	Di2	Di3	Di4	Di5	Di6	Di7	Di8
Di5	(6, 7, 8)	(9, 9, 9)	(3, 4, 5)	(7, 8, 9)	(1, 1, 1)	(4, 5, 6)	(2, 3, 4)	(5, 6, 7)
Di6	(3, 4, 5)	(5, 6, 7)	(1/4, 1/3, 1/2)	(4, 5, 6)	(1/6, 1/5, 1/4)	(1, 1, 1)	(1/5, 1/4, 1/3)	(2, 3, 4)
Di7	(5, 6, 7)	(7, 8, 9)	(2, 3, 4)	(6, 7, 8)	(1/4, 1/3, 1/2)	(3, 4, 5)	(1, 1, 1)	(4, 5, 6)
Di8	(2, 3, 4)	(4, 5, 6)	(1/5, 1/4, 1/3)	(3, 4, 5)	(1/7, 1/6, 1/5)	(1/4, 1/3, 1/2)	(1/6, 1/5, 1/4)	(1, 1, 1)

Table 5: Computation of geometric means of different disorders.

Disorders	(Lower, medium, and upper) geometric means		
Di1	$(1/1120)^{1/8}$	$(1/210)^{1/8}$	$(1/36)^{1/8}$
Di2	$(1/544320)^{1/8}$	$(1/181440)^{1/8}$	$(1/45360)^{1/8}$
Di3	$(36)^{1/8}$	$(210)^{1/8}$	$(1120)^{1/8}$
Di4	$(1/30240)^{1/8}$	$(1/6720)^{1/8}$	$(1/1260)^{1/8}$
Di5	$(45360)^{1/8}$	$(181440)^{1/8}$	$(544320)^{1/8}$
Di6	$(1)^{1/8}$	$(6)^{1/8}$	$(35)^{1/8}$
Di7	$(1260)^{1/8}$	$(6720)^{1/8}$	$(30240)^{1/8}$
Di8	$(35)^{1/8}$	$(6)^{1/8}$	$(1)^{1/8}$
Sum	10.293	12.616	15.116
Reciprocal	1/(10.293)	1/(12.616)	1/(15.116)
Increasing order	**1/(15.116)**	**1/(12.616)**	**1/(10.293)**

Table 6: Computation of fuzzified weights.

Disorders	Relative fuzzy weightage		
Di1	$(1/15.116)\ (1/1120)^{1/8}$	$(1/12.616)\ (1/210)^{1/8}$	$(1/10.293)\ (1/36)^{1/8}$
Di2	$(1/15.116)\ (1/544320)^{1/8}$	$(1/12.616)\ (1/181440)^{1/8}$	$(1/10.293)\ (1/45360)^{1/8}$
Di3	$(1/15.116)\ (36)^{1/8}$	$(1/12.616)\ (210)^{1/8}$	$(1/10.293)\ (1120)^{1/8}$
Di4	$(1/15.116)\ (1/30240)^{1/8}$	$(1/12.616)\ (1/6720)^{1/8}$	$(1/10.293)\ (1/1260)^{1/8}$
Di5	$(1/15.116)\ (45360)^{1/8}$	$(1/12.616)\ (181440)^{1/8}$	$(1/10.293)\ (544320)^{1/8}$
Di6	$(1/15.116)\ (1)^{1/8}$	$(1/12.616)\ (6)^{1/8}$	$(1/10.293)\ (35)^{1/8}$
Di7	$(1/15.116)\ (1260)^{1/8}$	$(1/12.616)\ (6720)^{1/8}$	$(1/10.293)\ (30240)^{1/8}$
Di8	$(1/15.116)\ (35)^{1/8}$	$(1/12.616)\ (6)^{1/8}$	$(1/10.293)\ (1)^{1/8}$

Table 7: Average and normal weightage.

Disorders	Mi	Ni
Di1	0.043	0.04115
Di2	0.018	0.01722
Di3	0.163	0.15598
Di4	0.028	0.02679
Di5	0.372	0.35598
Di6	0.105	0.10048
Di7	0.249	0.23828
Di8	0.067	0.06412

much ultramodern design structures for current era workplaces. Engineering will witness the evolution of the novel "human factors and ergonomics" domain in upcoming decades [70].

Upcoming developments will emphasize on the aligning of human-machine interface and ergonomics areas for sustainable progress, user safety, and psychophysical issues [71].

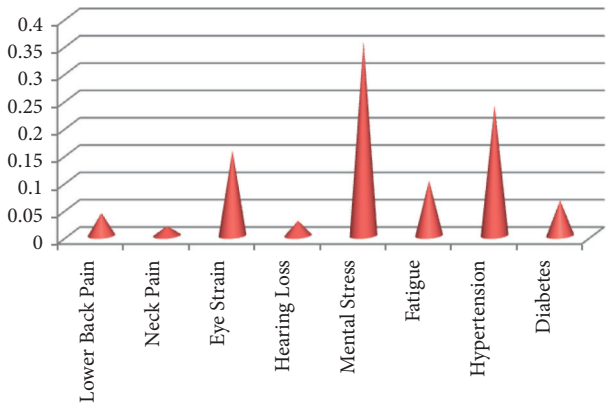


Figure 4: Normalized relative weights of ergonomics-related disorders.

Table 8: Subcriteria priority computation.

Criteria	Subcriteria	Priority	Rank
Afflictive disorders	Lower back pain	0.04115	6
	Neck pain	0.01722	8
Specific disorders	Eyestrain	0.15598	3
	Hearing loss	0.02679	7
Psychosocial disorders	Mental stress	0.35598	1
	Fatigue	0.10048	4
Chronic disorders	Hypertension	0.23828	2
	Diabetes	0.06412	5

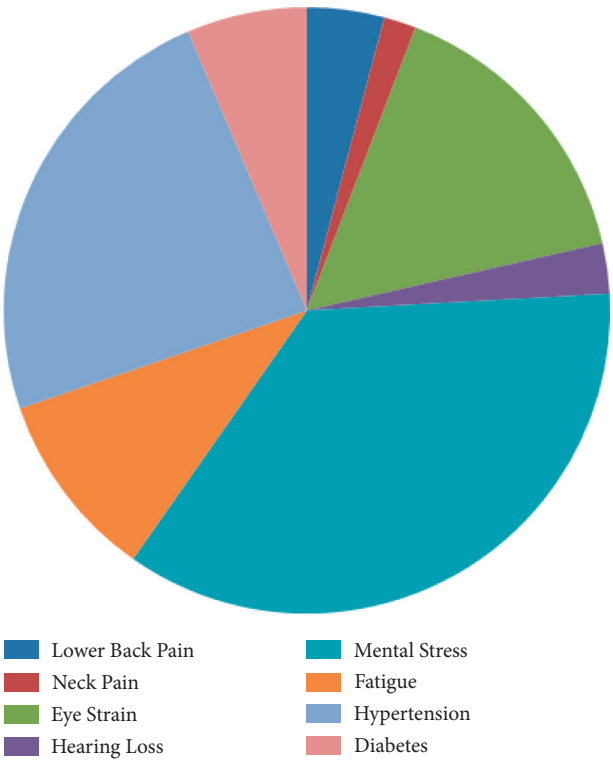


Figure 5: Pie chart of priorities of ergonomics-related disorders.

Table 9: Criteria priority computation.

Criteria	Priority	Rank
Afflictive disorders	0.05837	IV
Specific disorders	0.18277	III
Psychosocial disorders	0.45646	I
Chronic disorders	0.30240	II

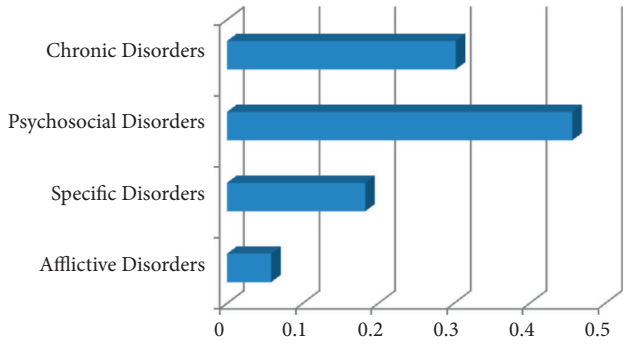


Figure 6: Priority of categories of ergonomics-related disorders.

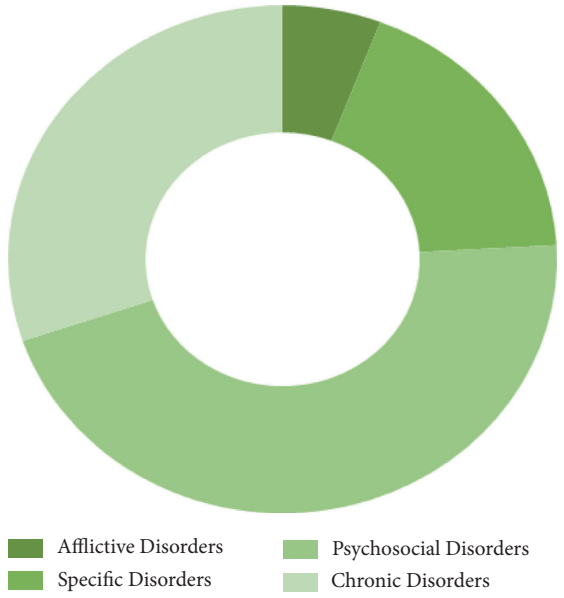


Figure 7: Weightage of categories of ergonomics-related disorders.

Table 10: Comparison of our work with some recent contributions.

S. No.	Paper	Outcome	Present work
1	Daneshmandi et al. [67]	The authors observed that sedentary sitting is linked to MSD. They concluded that prolonged sitting behavior had adverse effects and suggested active workstations for improving the working environment.	The ergonomics-related disorders in online education such as lower back pain, neck pain, eyestrain, hearing loss, mental strain, fatigue, hypertension, and diabetes have been evaluated and compared.

Table 10: Continued.

S. No.	Paper	Outcome	Present work
2	Golabchi et al. [68]	The authors presented a fuzzy-based methodology for ergonomics evaluation by incorporating perceived differences in distinguishing human poses in the analysis system of ergonomics tactics.	Total eight ergonomics-related disorders (in categories such as afflictive disorders, specific disorders, psychosocial disorders, and chronic disorders) are compared using fuzzy analytical hierarchical process methodology to get ranked in terms of priority.
3	Metin and Yuksel [69]	The authors developed a model on the basis of calculating the best influencing parameters causing malfunctioning and applying precautionary measures in the correction of the parameters, using the fuzzy AHP method.	The present research work has demonstrated an MCDM modeling and assessment of total of eight ergonomic disorders (in four categories) in online education applying the fuzzy AHP technique.

5. Conclusion

This research work has demonstrated an MCDM modeling for the evaluation of ergonomics disorders analysis in online education applying the fuzzy AHP technique. Eight ergonomics-related disorders in online education were recognized as lower back pain, neck pain, eyestrain, hearing loss, mental stress, fatigue, hypertension, and diabetes. Mental stress has emerged as the most influential ergonomics-related disorder in online education, followed by hypertension and eyestrain. These ergonomics-related disorders in four categories (afflictive disorders, specific disorders, psychosocial disorders, and chronic disorders) have been compared and ranked. Based on calculated mathematical values, "psychosocial disorders" have been ranked the first position followed by "chronic disorders" and "specific disorders."

The results may be instrumental for taking appropriate corrective actions to prevent ergonomics-related disorders. For numerous industries and corporate organizations, the discussed analysis can play a vital role to enhance the productivity of manpower especially related to sedentary computer screen work for long hours.

This study could consider ergonomics-related disorders in online education with only one methodology. The prospect of further work can be the development of a similar model by applying other multiple criteria decision-making methods available such as best-worst method and so on. Modern MCDM techniques such as case-based reasoning and data envelopment analysis may be used for much-detailed problems in this domain.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Australian Information Industry Association, *Computers in the Workplace, a Guide to Screens and Keyboards for Managers and Users*, Australian Information Industry Association, Melbourne, Australia, 1988.

[2] K. Benmoussa, M. Laaziri, S. Khoulji, M. L. Kerkeb, and A. E. Yamami, "AHP-based approach for evaluating ergonomic criteria," *Procedia Manufacturing*, vol. 32, pp. 856–863, 2019.

[3] W. Hünting, T. Läubli, and E. Grandjean, "Postural and visual loads at VDT workplaces II. Lighting conditions and visual impairments," *Ergonomics*, vol. 24, no. 12, pp. 933–944, 1981.

[4] S. L. Sauter, L. M. Schleifer, and S. J. Knutson, "Work posture, workstation design, and musculoskeletal discomfort in a VDT data entry task," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 33, no. 2, pp. 151–167, 1991.

[5] U. Bergqvist, E. Wolfgast, B. Wolgast, and M. Voss, "Musculoskeletal disorders among visual display terminal workers: individual, ergonomic, and work organizational factors," *Ergonomics*, vol. 38, no. 4, pp. 763–776, 1995.

[6] C. Alppay and A. Hedge, "Development of an ergonomics checklist for the evaluation of medical tablet personal computers," *Procedia Manufacturing*, vol. 3, pp. 21–28, 2015.

[7] S. Bayir and H. Keser, "Information and communication technologies coordinator teachers' evaluations of computer working environments in terms of ergonomics," *Procedia Social and Behavioral Sciences*, vol. 1, no. 1, pp. 335–341, 2009.

[8] M. Heiden, C. Zetterberg, P. Lindberg, P. Nylén, and H. Hemphälä, "Validity of a computer-based risk assessment method for visual ergonomics," *International Journal of Industrial Ergonomics*, vol. 72, pp. 180–187, 2019.

[9] H. Keser and S. Bayir, "Information and communication technologies coordinator teacher trainers' evaluations of computer working environments in terms of ergonomics," *Procedia Social and Behavioral Sciences*, vol. 1, no. 1, pp. 2371–2379, 2009.

[10] E. Lavrov, O. Kupenko, T. Lavryk, and N. Barchenko, "Organizational approach to the ergonomic examination of E-learning modules," *Informatics in Education*, vol. 12, no. 1, pp. 107–124, 2013.

[11] T. Gavrilova, R. Farzan, and P. Brusilovsky, "One practical algorithm of creating teaching ontologies," in *Proceedings of the Network-Based Education 2005*, Rovaniemi, Finland, September 2005.

[12] C. Dams, M. Langford, D. Uehara, and R. Scherer, "Teachers' agency and online education in times of crisis," *Computers in Human Behavior*, vol. 121, Article ID 106793, 2021.

[13] C. Ardito, M. F. Costabile, M. D. Marsico et al., "An approach to usability evaluation of E-learning applications," *Universal Access in the Information Society*, vol. 4, no. 3, pp. 270–283, 2006.

[14] D. R. Garrison, *E-learning in the 21st Century: A Framework for Research and Practice*, Routledge, London, UK, 2nd edition, 2011.

[15] D. Persico, S. Manca, and F. Pozzi, "Adapting the Technology Acceptance Model to evaluate the innovative potential of e-learning systems," *Computers in Human Behavior*, vol. 30, 2013.

[16] M. Pikhart, "Human-computer interaction in foreign language learning applications: applied linguistics viewpoint of mobile learning," *Procedia Computer Science*, vol. 184, pp. 92–98, 2021.

[17] C. Jensen, C. U. Ryholt, H. Burr, E. Villadsen, and H. Christensen, "Work-related psychosocial, physical and individual factors associated with musculoskeletal symptoms in computer users," *Work & Stress*, vol. 16, no. 2, pp. 107–120, 2002.

[18] M. Rajinder Kumar, P. S. Lakhwinder, and M. Neelam, "Prevalence of Musculoskeletal Disorder among Computer Bank Office Employees in Punjab (India): A Case Study," *Procedia Manufacturing*, vol. 3, pp. 6624–6631, 2015.

[19] L. Korpinen, R. Pääkkönen, and F. Gobba, "Self-reported ache, pain, or numbness in hip and lower back and use of computers and cell phones amongst Finns aged 18-65," *International Journal of Industrial Ergonomics*, vol. 48, pp. 70–76, 2015.

[20] A. Mozafari, M. Vahedian, S. Mohebi, and M. Najafi, "Prevalence and risk factors of Musculoskeletal Disorders among official workers in Qom," *Afinidad*, vol. 80, no. 567, 2014.

[21] D. Sharan, P. Parijat, A. P. Sasidharan, R. Ranganathan, M. Mohandoss, and J. Jose, "Workstyle risk factors for work related musculoskeletal symptoms among computer professionals in India," *Journal of Occupational Rehabilitation*, vol. 21, no. 4, pp. 520–525, 2011.

[22] G. A. M. Ariens, P. M. Bongers, M. Douwes et al., "Are neck flexion, neck rotation, and sitting at work risk factors for neck pain? Results of a prospective cohort study," *Occupational and Environmental Medicine*, vol. 58, no. 3, pp. 200–207, 2001.

[23] M. Marcus, F. Gerr, C. Monteilh et al., "A prospective study of computer users: II. Postural risk factors for musculoskeletal symptoms and disorders," *American Journal of Industrial Medicine*, vol. 41, no. 4, pp. 236–249, 2002.

[24] A. Chawla, T. C. Lim, S. N. Shikhare, P. L. Munk, and W. C. G. Peh, "Computer vision syndrome: darkness under the shadow of light," *Canadian Association of Radiologists Journal*, vol. 70, no. 1, pp. 5–9, 2019.

[25] C. Blehm, S. Vishnu, A. Khattak et al., "Computer vision syndrome: a review," *Survey of Ophthalmology*, vol. 50, no. 3, pp. 253–262, 2005.

[26] J. E. Sheedy, "Vision problems at video display terminals: a survey of optometrists," *Journal of the American Optometric Association*, vol. 63, no. 10, pp. 687–692, 1992.

[27] J. Kil, E. Lobarinas, C. Spankovich et al., "Safety and efficacy of ebselen for the prevention of noise-induced hearing loss: a randomised, double-blind, placebo-controlled, phase 2 trial," *The Lancet*, vol. 390, 2017.

[28] D. Twardella, U. Raab, C. Perez-Alvarez, T. Steffens, G. Bolte, and H. Fromme, "Usage of personal music players in adolescents and its association with noise-induced hearing loss: a cross-sectional analysis of Ohrkan cohort study data," *International Journal of Audiology*, vol. 56, no. 1, pp. 38–45, 2016.

[29] W. E. Hodgetts, J. M. Rieger, and R. A. Szarko, "The effects of the listening environment and earphone style on preferred listening levels of normal hearing adults using an MP3 player," *Ear and Hearing*, vol. 28, no. 3, pp. 290–297, 2007.

[30] H. Laitinen, J. Saari, M. Kivistö, and P.-L. Rasa, "Improving physical and psychosocial working conditions through a participatory ergonomic process A before-after study at an engineering workshop," *International Journal of Industrial Ergonomics*, vol. 21, no. 1, pp. 35–45, 1998.

[31] J. Frank, I. Pulcins, M. S. Kerr, H. Shannon, and S. Stansfeld, "Occupational back pain—an unhelpful polemic," *Scandinavian Journal of Work, Environment & Health*, vol. 21, no. 1, pp. 3–14, 1995.

[32] G. Ashdown-Franks, A. Koyanagi, D. Vancampfort et al., "Sedentary behavior and perceived stress among adults aged ≥50 years in six low- and middle-income countries," *Maturitas*, vol. 116, 2018.

[33] F. Serranheira, M. Pereira, C. S. Santos, and M. Cabrita, "Auto referencia de sintomas de lesoes musculo esqueleticas ligadas ao trabalho numa grande empresa em Portugal," *Saude Ocupacional*, vol. 21, no. 2, 2003.

[34] P. Wongngam, B. S. Rajaratnam, and P. Janwantanakul, "The SitLess with MS program: intervention feasibility and change in SedentaryBehavior," *Arch Rehabil Res Clin Transl*, vol. 2, Article ID 100083, 2020.

[35] P. Dempsey, J. Sacre, N. Owen et al., "Interrupting prolonged sitting reduces resting blood pressure in adults with type 2 diabetes," *Heart Lung & Circulation*, vol. 24, pp. S127–S128, 2015.

[36] S. R. Colberg, R. J. Sigal, J. E. Yardley et al., "Physical activity/exercise and diabetes: a position statement of the American diabetes association," *Diabetes Care*, vol. 39, no. 11, pp. 2065–2079, 2016.

[37] P. C. Dempsey, D. W. Dunstan, R. N. Larsen, G. W. Lambert, B. A. Kingwell, and N. Owen, "Prolonged uninterrupted sitting increases fatigue in type 2 diabetes," *Diabetes Research and Clinical Practice*, vol. 135, 2017.

[38] T. L. Saaty, *The Analytic Hierarchy Process*, McGraw-Hill, New York, NY, USA, 1980.

[39] M. A. Badri, "A combined AHP-GP model for quality control systems," *International Journal of Production Economics*, vol. 72, no. 1, pp. 27–40, 2001.

[40] P. Srichetta and W. Thurachon, "Applying fuzzy analytic hierarchy process to evaluate and select product of notebook computers," *International Journal of Modeling and Optimization*, vol. 2, no. 2, 2012.

[41] T. K. Biswas, S. M. Akash, and S. Saha, "A fuzzy-AHP method for selection best apparel item to start-up with new garment factory: a case study in Bangladesh," *International Journal of Research in Industrial Engineering*, vol. 7, no. 1, pp. 32–50, 2018.

[42] A. Gnanavel Babu and P. Arunagiri, "Ranking of MUDA using AHP and Fuzzy AHP algorithm," *Materials Today: Proceedings*, vol. 5, no. 5, pp. 13406–13412, 2017.

[43] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[44] O. Kulak, B. Durmusoglu, and C. Kahraman, "Fuzzy multi-attribute equipment selection based on information axiom," *Journal of Materials Processing Technology*, vol. 169, no. 3, pp. 337–345, 2005.

[45] A. Özdagoglu and G. Özdagoglu, "Comparison of AHP and fuzzy AHP for the multi criteria decision making processes with linguistic evaluations," *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, vol. 6, no. 11, 2007.

[46] S.-W. Chou and Y.-C. Chang, "The implementation factors that influence the ERP (enterprise resource planning)

benefits," *Decision Support Systems*, vol. 46, no. 1, pp. 149–157, 2008.

[47] Y. C. Erensal, T. Öncan, and M. L. Demircan, "Determining key capabilities in technology management using fuzzy analytic hierarchy process: a case study of Turkey," *Information Sciences*, vol. 176, no. 18, pp. 2755–2770, 2006.

[48] D. Dsilva Winfred Rufuss, V. Raj Kumar, L. Suganthi, S. Iniyan, and P. A. Davies, "Techno-economic analysis of solar stills using integrated fuzzy analytical hierarchy process and data envelopment analysis," *Solar Energy*, vol. 159, pp. 820–833, 2018.

[49] M. Batuhan, "A fuzzy AHP approach for supplier selection problem: a case study in a gear motor company," *International Journal of Managing Value and Supply Chains*, vol. 4, no. 3, 2013.

[50] O. M. Olabanji and K. Mpofu, "Hybridized fuzzy analytic hierarchy process and fuzzy weighted average for identifying optimal design concept," *Heliyon*, vol. 6, Article ID e03182, 2020.

[51] S. M. Saad, N. Kunhu, and A. M. Mohamed, "A fuzzy-AHP multi-criteria decision-making model for procurement process," *International Journal of Logistics Systems and Management*, vol. 23, no. 1, pp. 1–24, 2016.

[52] A. Çebi and H. Karal, "An application of fuzzy analytic hierarchy process for evaluating students' project," *Educational Research and Reviews*, vol. 12, no. 3, pp. 120–132, 2017.

[53] Y. Deng, "Fuzzy analytical hierarchy process based on canonical representation on fuzzy numbers," *Journal of Computational Analysis and Applications*, vol. 22, no. 2, 2017.

[54] B. Das and S. C. Pal, "Combination of GIS and fuzzy-AHP for delineating groundwater recharge potential zones in the critical Goghat-II block of West Bengal, India," *HydroResearch*, vol. 2, pp. 21–30, 2019.

[55] D. P. Mochammad Sobandi, F. SeptiAndryana, and A. Gunaryati, "Fuzzy analytical hierarchy process method to determine the quality of gemstones," *Advances in Fuzzy Systems*, vol. 2018, Article ID 9094380, 6 pages, 2018.

[56] M. Güler and G. Büyüközkan, "Analysis of digital transformation strategies with an integrated fuzzy AHP-axiomatic design methodology," *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 1186–1191, 2019.

[57] A. A. Khan, M. Shameem, M. Nadeem, and M. A. Akbar, "Agile trends in Chinese global software development industry: fuzzy AHP based conceptual mapping," *Applied Soft Computing*, vol. 102, Article ID 107090, 2021.

[58] O. O. Sunday and A. K. Ismaila, "A Fuzzy-AHP based prioritization of trust criteria in fog computing services," *Applied Soft Computing*, vol. 97, Article ID 106789, 2020.

[59] A. A. Khan, M. Shameem, R. R. Kumar, S. Hussain, and X. Yan, "Fuzzy AHP based prioritization and taxonomy of software process improvement success factors in global software development," *Applied Soft Computing*, vol. 83, p. 105648, 2019.

[60] P. J. M. Van Laarhoven and W. Pedrycz, "A fuzzy extension of Saaty's priority Theory," *Fuzzy Sets and Systems*, vol. 11, no. 1-3, pp. 199–227, 1983.

[61] L. Mikhailov, "A fuzzy approach to deriving priorities from interval pairwise comparison judgements," *European Journal of Operational Research*, vol. 159, no. 3, pp. 687–704, 2004.

[62] D.-Y. Chang, "Applications of the extent analysis method on fuzzy AHP," *European Journal of Operational Research*, vol. 95, no. 3, pp. 649–655, 1996.

[63] J. J. Buckley, "Fuzzy hierarchical analysis," *Fuzzy Sets and Systems*, vol. 17, no. 1, pp. 233–247, 1985.

[64] T. L. Saaty, *Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World*, RWS Publications, Pittsburgh, PA, USA, 3rd edition, 1999.

[65] T. L. Saaty, *Decision Making with Dependence and Feedback: The Analytic Network Process*, RWS Publications, Pittsburgh, PA, USA, 1996.

[66] A. Radionovs and O. Uzhga-Rebrov, "Comparison of different fuzzy ahp methodologies in risk assessment," *Environment. Technology. Resources. Proceedings of the International Scientific and Practical Conference*, vol. 2, pp. 137–142, 2017.

[67] H. Daneshmandi, A. Choobineh, H. Ghaem, and M. Karimi, "Adverse effects of prolonged sitting behavior on the general health of office workers," *Journal of Lifestyle Medicine*, vol. 7, no. 2, pp. 69–75, 2017.

[68] A. Golabchi, S. U. Han, and A. Robinson Fayek, "A fuzzy logic approach to posture-based ergonomic analysis for field observation and assessment of construction manual operations," *Canadian Journal of Civil Engineering*, vol. 43, 2015.

[69] D. Metin and I. Yuksel, "Developing a fuzzy analytic hierarchy process (AHP) model for behavior-based safety management," *Information Sciences*, vol. 178, pp. 1717–1733, 2008.

[70] W. Karwowski and W. Zhang, *The Discipline of Human Factors and Ergonomics*, Wiley Online Library, New York, NY, USA, 2021.

[71] L. Gualtieri, E. Rauch, and R. Vidoni, "Emerging research fields in safety and ergonomics in industrial collaborative robotics: a systematic literature review," *Robotics and Computer-Integrated Manufacturing*, vol. 67, Article ID 101998, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 8338903, 8 pages
https://doi.org/10.1155/2021/8338903

Research Article

Providing an Adaptive Routing along with a Hybrid Selection Strategy to Increase Efficiency in NoC-Based Neuromorphic Systems

Mohammad Trik,[1] **Saadat Pour Mozaffari**,[2] **and Amir Massoud Bidgoli**[1]

[1]*Department of Computer Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran*
[2]*Computer Engineering and Information Technology Department, Amirkabir University of Technology, Tehran, Iran*

Correspondence should be addressed to Amir Massoud Bidgoli; am_bidgoli@iau-tnb.ac.ir

Received 3 May 2021; Accepted 21 July 2021; Published 16 September 2021

Academic Editor: Navid Razmjooy

Effective and efficient routing is one of the most important parts of routing in NoC-based neuromorphic systems. In fact, this communication structure connects different units through the packets routed by routers and switches embedded in the network on a chip. With the help of this capability, not only high scalability and high development can be created, but by decreasing the global wiring to the chip level, power consumption can be reduced. In this paper, an adaptive routing algorithm for NoC-based neuromorphic systems is proposed along with a hybrid selection strategy. Accordingly, a traffic analyzer is first used to determine the type of local or nonlocal traffic depending on the number of hops. Then, considering the type of traffic, the RCA and NoP selection strategies are used for the nonlocal and local strategies, respectively. Finally, using the experiments that performed in the simulator environment, it has been shown that this solution can well reduce the average delay time and power consumption.

1. Introduction

The Neuromorphic Computing, also known as "Neuromorphic engineering," operates using a model inspired by the mechanism of the human brain. This technology not only models theories of neuroscience but also solves machine learning problems. The term neuromorphic computing is a concept developed by Carver Mead in the late 1980s describing the use of very-large-scale integration (VLSI) systems containing electronic analog circuits to mimic the neural and biological architecture present in the nervous system. Currently, the term neuromorphic is used to describe analog systems, digital systems, analog/digital complex systems, and software that model neural systems [1–3]. Interprocessor communication is supported on an effective multicast foundation managed by neurobiology. It uses a packet-switched network to achieve the very high coupling of biological systems. The packets are source-routed, i.e., they move only information near the packet issuer; the network is answerable for liberating them to their destinations. The function of a router is to be able to specify several different destinations to reach the desired packet when faced with a routing problem in sending packets. Accordingly, routers have the ability to broadcast, i.e., they can send a packet to multiple path simultaneously.

The increase in the number of components in a system on chip (SoC) coupled with the growth of interference problems caused by the bus system led to the appearance of NoC. These networks were introduced to eliminate these problems and to increase the performance of the NoC-based neuromorphic systems. In this infrastructure, instead of using wiring or communication paths, packet routing techniques are used in the network [3]. In these networks, there are several different paths to move from one node to another; therefore, there should be an algorithm to obtain the route to reach the destination. Routing algorithms may merely use the address of the current and destination nodes to compute the route (definite routing), or may use the

collected traffic information of other nodes to calculate the route (adaptive routing) [4, 5]. In the adaptive algorithms, the calculated path is stored in the packet header and used in the middle nodes to hold the channels. Deadlock occurs when packets need a channel available to other packets to continue their path. Wandering also means that the packet does not arrive at the destination for long and unreasonable time. Namely, the adaptive routing algorithm measures a set of acceptable output channels regarding the paths that the packet can pass through to reach the destination [6, 7]. Afterwards, according to the network characteristics, including the congestion rate or the length of one of the routes of the output channel, the selection function will be utilized to choose the output channel from a set of permitted output channels. In this case, a traffic analyzer is used that determines the type of function according to the type of local or nonlocal traffic. As a result, in this case, according to the type of traffic, the appropriate strategy can be determined with it. The overall schematic of using the routing algorithm and selection function is presented in Figure 1 [8]. Accordingly, in this research, a hybrid function with a traffic analyzer for adaptive routing algorithms is presented, in which, in addition to increasing the efficiency of the NoC-based neuromorphic systems, power consumption can be reduced by creating the balance in this infrastructure [9, 10].

1.1. Contribution. In a previous work, various selection strategies have been proposed to improve routing algorithms, each of which poses challenges in research results. For example, in Ref. [11], by using a virtual circuit switch, routing is minimized and thus energy consumption is reduced. In [12–15], the selection function is presented based on the input and output choices and the NOP technique in which they have been able to reduce energy consumption.

In addition, in [16], a selection strategy is used in *XY* routing to achieve improvements in reducing latency. In the proposed method, we have first separated a number of calculations that can be done offline from the main processing steps. In this way, the processing overhead can be reduced each time it is run. In the next step, a selection strategy is presented according to the traffic situation. In this case, a traffic analyzer is used, which determines the type of selection function according to the type of local or nonlocal traffic. As a result, in this case, according to the type of traffic, the appropriate strategy can be determined with it. In this way, in addition to reducing energy consumption, other parameters such as latency and congestion can be reduced compared to other solutions.

1.2. Paper Organization. The study is organized as follows. In the next section, related works are stated for the previously used algorithms in NoC-based neuromorphic systems along with selection functions. In Section 3, the suggested combined method is stated to propose a combined selection function. In Section 4, the results of analyzing the suggested model in different scenarios are shown. Section 5 concludes this paper. Finally, Section 6 discusses future works.

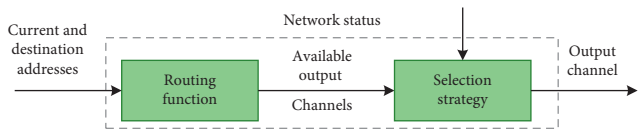


Figure 1: Routing structure and selection route blocks.

2. Related Works

Over recent years, numerous researchers have studied different utilized algorithms along with the selection functions for different fields in NoC-based neuromorphic systems, and we examine some of the performed studies in these subjects in the following sections. The neuromorphic model has scope in the development of VLSI systems, imitating the neurobiological networks of the nervous system—SNN. It is a large-scale parallel system consisting of a large number of computational units called neuromorphic nuclei interconnected by NoC. Communication management in the neuromorphic framework is the responsibility of NoC. In recent years, adaptive routing algorithms have been proposed that use local or nonlocal information for NoC. Due to the congestion of information in each router, regional congestion in NoC architectures can be divided into two categories: in some architectures, a router controls the status of the entire network, while in other architectures the routers are aware of the status of part of the network [17].

Regional congestion awareness (RCA) uses a lightweight network for accumulation and dissemination congestion information [18]. Other studies [11, 12, 19, 20] also provide information that global congestion has been reviewed and collected. Another type of architecture in neuromorphic systems that examines only a few nonlocal nodes instead of all nodes is the NoP architecture, which is based on destination-based adaptive routing. In general, the NoC architecture uses a routing strategy to avoid deadlocks. Based on the structure of NoC, two types of topology are commonly used in NoC: NoC tree and NoC mesh. Examples include NoC networks for TrueNorth and Loihi, NoC multistage networks for Dynapse [21], and NoC tree for CxQuad. SpiNNaker [13] can simulate the brain in real time by connecting 1 million ARM processors. By integrating eighteen ARM processors into a multichip processor (CMP) and 2ˆ16 CMPs, a system with a two-dimensional network structure is formed.

In [22], a method called PACMAN is presented to study the SNN mapping in SpiNNaker. PACMAN uses a simulated annealing algorithm to search for the best partitioning plan. A variety of previous studies have examined neural network accelerators in which parameters such as reduced power consumption [23, 24], increased throughput [25, 26], and the use of memory bandwidth for information processing are evaluated. A variety of previous studies have examined neural network accelerators in which parameters such as reduced power consumption [23, 24], increased throughput [25, 26], and the use of memory bandwidth for information processing are evaluated [27]. In [9], the behavior of various topologies under the mass communication traffic of neural networks

is investigated. The researchers concluded that mesh networks perform better than busses in point-to-point and tree links. Another large-scale neural network architecture is EMBRACE, which consists of a matrix of interconnected processing elements called a set of neural tiles [5]. In H-NoC, nodes are arranged in three layers, namely, module, tile, and cluster.

In the first layer, up to 10 nerve cells are connected to the router to form a neuron module. In addition, ten neuron modules are connected to the higher router to form a tile. In the higher layer, i.e., the third layer, four tiles are connected to a router, forming an EMBRACE cluster. Traffic in the neural network is multicast and all-broadcast; therefore, many previous studies that support mass communication in the NOC can be effective in managing interneuronal traffic [28]. Mass communication approaches are classified based on how the message is repeated. The main classes include single-broadcast, path-based, or tree-based multicast routing [29]. The development of processable systems is a major challenge for large amounts of multisensory data in the new age of cognitive computing. This type of intelligent computing has limitations such as real-time performance, low power consumption, and scalability. Structures and architectures imitating the brain hold great promise in this area. For this reason, in [30, 31], TrueNorth has developed a 65-megawatt real-time synaptic neural processor that uses a non-von Neumann architecture, uses low power, is highly parallel, is scalable, and is fault tolerant. NoC-enabled homogeneous CMP architectures focused on neuroscience programs that have already been explored. For example, a vastly parallelized CMP platform incorporating a custom-designed NoC architecture was used to put into effect spiking neural networks [32].

The study and analysis of neural networks in different topologies require very high configuration for these accelerators. Here are some examples of popular chips that are fast accelerators of the nervous system. Certain chips such as TrueNorth [33], Neurogrid [34], BrainScaleS [35], Loihi [32], and SpiNNaker [36] have used different features to inspire spiking neural networks. For further studies in this regard, we consider Furber solutions [16, 36] for these chips. SpiNNaker uses a processor architecture that connects to local memory on a chip. Compared to other architectures, it can be configured, but in terms of energy consumption, it is at a lower level compared to them. BrainScaleS are interconnected as several interconnected wafers, each of which consists of several HiCANN neurons. The purpose of simulating this architecture is to investigate biological neural behavior during rapid acceleration. Neurogrid is an SNN analyzer designed for analog electronics applications. This architecture operates in real time and follows several environmental methods. Finally, the architecture TrueNorth implemented in digital systems has a neuromorphic chip. These four architectures are a major step forward in the development of neural processors with the goal of mimicking the environment and reducing energy consumption in neuromorphic systems. Table 1, by imitating Furber [16], examines the characteristics of each architecture in comparison with each other from the perspective of different parameters.

3. The Proposed Method

The proposed strategy that focuses on the selection algorithm develops within the framework of the infrastructure sections of these systems. Initially, the main requirements of the proposed solution are examined, and the main core of the research, which is the expression of the traffic analyzer and hybrid selection function, is presented and its capabilities are expressed. Part of the calculations of the selected strategy is done before the simulation operation; the offline execution of the calculation allows the algorithm to reduce its delay due to reduced computational time.

3.1. Offline Computational Strategies. Parts of the selection of computational strategies in this research, such as the link connection and the equality of resistances, are computed offline and before the algorithm is implemented, which is explained as follows:

Link contention (CL): link contention is referred to as the traffic value, which can pass through a specific link based on the communications provided in the communication graph [29]. If p_i is a path for arbitrary communications, ρ_{comm} is the set of all possible paths, t_{comm} is the traffic generated by communications, and n_{comm} is the number of all communications that are specified in the program traffic and extracted from the communication graph, the link contention CL can be expressed as follows:

$$\text{CL} = \sum_{\text{comm}=1}^{n_{\text{comm}}} (\mu * t_{\text{comm}}), \quad \mu = 1\colon \exists \rho_i \varepsilon \rho_{\text{comm}}\colon L \varepsilon \rho_i, \text{ else: } 0. \tag{1}$$

Equivalent resistance (ER): by defining this concept (ER) for each given communication, and using the electrical concepts of the Kirchoff law, each node in the topology is considered as a circuit node and each link is deemed as a resistor with a volume equal to the contention of that link [37, 38].

3.2. Requirements of Online Function. During the execution of the algorithm, other additional available data are also used for routing, which are as follows:

(i) Free buffer rows (d) (B): the number of rows in the input buffer is the neighbors adjacent to path d in which d is one of the north, south, east, and west directions.

(ii) Instantaneous power (Δp): the instantaneous power is the difference between the power consumed by the router at t and $t - 1$, where t is the moment the final output channel is selected for the packet.

$$\Delta p = \text{power}(t) - \text{power}(t - 1). \tag{2}$$

Using this information, a better estimation can be obtained for the traffic load of the network and a more appropriate decision would be made at the moment of the next hop.

Table 1: Comparison of nervous system processors in terms of structural features.

Platform	Neuromorphic system models				
	BrainScaleS	TrueNorth	Loihi	Neurogrid	SpiNNaker
NoC	Hierarchical	2D mesh	2D mesh	Tree multicast	2D mesh
Run-time plasticity	Yes	No	Yes	Yes	Yes
Implementation	Analog	Digital	Digital	Analog	Digital
On-chip learning	Yes	No	Yes	No	Yes
Neuron model	Diverse	Diverse, fixed	Adaptive quad	LIF	Fixed
Energy efficiency	Yes	Yes	Yes	No	Yes
Time	Discretized	Discretized	Discretized	Real time	Discretized
Microchip	HiCANN	EMBRACE	—	Neurocore	HiCANN

3.3. Traffic Analyzer. In order to avoid deadlock in the routing algorithm and reduce the delay time, an analyzer and selection function are added to the routing algorithm, so that it could be used to select the best outlet based on the local or nonlocal nature of the packet. Accordingly, the analyzer first extracts the destination address of each packet that is routed through the router in each T cycle and examines its data. For this purpose, two 5-bit counters are used to determine the local or nonlocal nature of requests in the router [14]. If the intended destination of the packet is two hops away or higher from the current router, it is considered as a nonlocal packet; otherwise, it is considered local. The analyzer calculates packet hops periodically, and accordingly, it updates the counter of locality and (L) and nonlocality (N) of the packets. This information is sent to the switcher to decide on the selection strategy. The counter is cleared at the end of every T cycle. The pseudocode for determining the traffic type is shown in Figure 2.

In fact, with the help of the traffic analyzer, it would be possible to obtain appropriate information about the rate of traffic and its convergence to local or nonlocal traffic [11], and then, in the next step, the routing operations can be done accordingly. Figure 3 shows the schematic view of the solution.

Accordingly, at the end of every 32 cycles, the traffic pattern is determined by the analyzer's output, and the local traffic rate is calculated as nonlocal (x). If $x \geq 0.3$, then traffic is nonuniform and the RCA algorithm [20] must be used; otherwise, the NoP [39] selection strategy will be used. In other words, if the traffic pattern is oriented towards local traffic destination, the NoP-based selection strategy is activated; otherwise, the RCA-based strategy will be activated as a proposed strategy for nonlocal traffic. The general algorithm for switching operation based on the traffic analyzer is presented in Figure 4. The data input to this algorithm is of local or nonlocal data type, and the related output is also the best strategy. It should also be noted that since the analyzer and the switch only take the router data at any one time, there is no additional overhead in network communications.

3.4. Formulation of the Solution. When the routing function receives multiple outputs, by reviewing the reservation table, the selected algorithm for each of these outputs checks whether the channel is available to transfer the packet (header flit) or its reserved by the other header flits [40]. The channel must be available so that the selected score is

```
for every T clock cycles do
L and N value from analyzer;          Catch
Compute x = N/(L + N);
 if x = x < 0.3 then
;Switch to NoP
        else
;Switch to RCA
end
```

Figure 2: A pseudocode for determining the type of traffic.

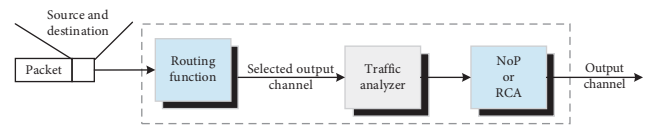


Figure 3: The schematic view of the solution.

```
Data: Packet hops (pkt_dst_hops), (Initializing: L = 0, N = 0)
Result: Local and Non-Local value (L = Local value, N = non_Local value)
 2 then = if pkt_dst_hops>
N++;
else
L++;
end
```

Figure 4: Pseudocode for performing switching operations.

calculated, and eventually, the channel with the highest score is selected. If more channels have the same score, the first one will be chosen. The calculation of the score is done by the following formula:

$$\text{Score}[d] = \alpha \times Psel[d] + \beta \times \left(\frac{B[d]}{\max\_\text{buffer\_size} + \gamma}\right) \times \left(\frac{\Delta p}{\max\_\text{power}}\right), \tag{3}$$

where α, β, and γ are the weight factors for the probability of selecting links, open buffers, and instantaneous power consumption. These coefficients result in the full-dynamic adaptability of the selected algorithm, and thus, the set values will be at their best state [41]. Since open buffers (B) and instantaneous power consumption ($p\Delta$) have different units, they are normalized using max-buffer-size and max-

power factors. In addition, since *Psel* is within the range (0 and 1), there is no need for normalization. Then, using the following formula, the score of the adaptive routing functions and all possible values of α, β, and γ are evaluated and the best coefficients are obtained for each of the routers:

$$\alpha + \beta + \gamma = 1, \quad \alpha = 0, 0.1, \ldots, 1, \beta = 0, 0.1, \ldots (1 - \alpha), \gamma = 1 - (\alpha + \beta). \tag{4}$$

For example, the best values of α, β, and γ in even-odd routing are 0.3, 0.4, and 0.3, respectively, under the MMS traffic scenario. Another important feature of this algorithm is its adaptability with any network topology [42].

4. Experiments and Simulation Environment

This section provides a platform for simulating the structure and framework of NoC-based neuromorphic systems. A Nirgam simulator is used to evaluate the suggested algorithm whose capabilities are listed in Table 2 [13, 19]. The main components in this simulator are routers, processing elements, links, and buffers [43, 44]. Moreover, the configuration parameters for the analysis and simulation of the suggested method are given in Table 3. In addition, the average delay, maximum delay, and power consumption are considered as the efficiency criteria. Delay is assumed as the time between entering the header flit to the network and the arrival of the tail flit to the destination node. In order to evaluate the proposed method, random, RCA, and NoP strategies have been compared [45, 46]. Here, the results of the study are shown in various traffic scenarios.

4.1. First Assessment: Load = 40%. In this assessment, the solutions are evaluated in the case of a traffic load of 40%. Through this assessment, the performance can be measured in low traffic and nontraffic situations. The average simulation results are achieved after 5000 runs. The evaluation results are shown in Figure 5–10. First, in Figure 5, the average overall delay in each channel is shown. As it can be observed, the proposed strategy in this case is more optimal than any other selected solutions.

Since the RCA selection strategy for a nonlocal packet has a better performance, it is, therefore, natural to lower performance in the channels; on the other hand, since the proposed strategy has different selection functions based on the length of the hop and locality and nonlocality of the channels, it has the highest efficiency. The total average delay is shown in Figure 6.

As can be seen, in total delay, the RCA efficiency increased due to the increase in nonlocal traffic rates, but given that the proposed solution core in addition to using RCA uses the NoP function for routing, naturally it has higher efficiency. In this case, the reason for this is the use of different strategies for local and nonlocal packets in which the packet will be sent with less delay and through better paths, which can lead to load balance in the network and reduce power consumption, which is clear from Figure 7.

Table 2: Main capabilities of Nirgam simulator.

Types of production traffic	Routing algorithm type	Switching mechanism	Topology type
Constant bit rate trace and bursty based	Odd-even, *XY*	Wormhole	Torus, mesh

Table 3: Simulation parameters.

Parameter	Configuration
Network size	8 ∗ 8 mesh
Schemes	Random, RCA [19], NoP [13], proposed
Packet size	8 flits
Reset_time cycles	5000
Simulation time	10

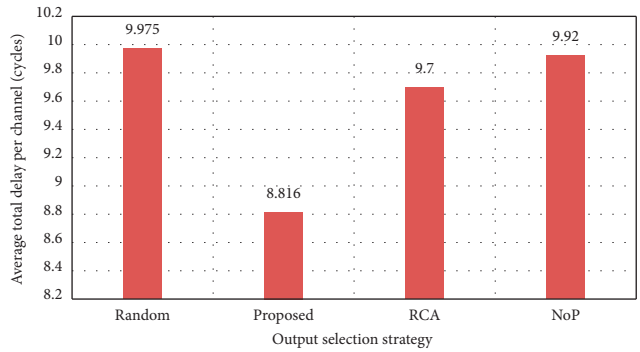


Figure 5: Average total delay per channel (load = 40%).

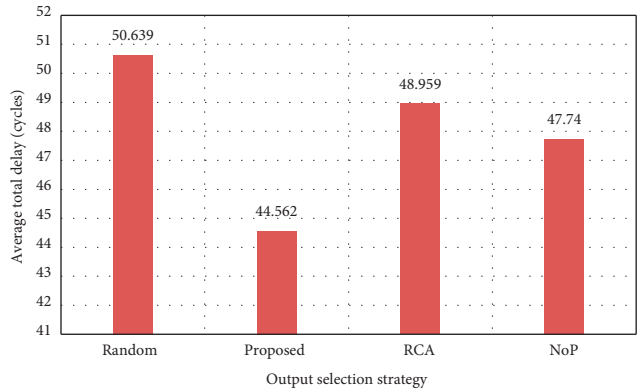


Figure 6: Average total delay (load = 40%).

As shown in Figure 7, the proposed solution has been able to reduce the amount of consumed power so that it is reduced by more than 50% compared to Random solution. In fact, one of the main reasons for this decrease in power consumption, in addition to offline calculations, is utilizing different strategies depending on the status of the packets, and, therefore, the best possible output path is always selected based on the network status. Accordingly, an optimal load balance is created, which reduces power consumption because in other methods such as Random, the output buffer status is not studied and, therefore, the packet may be led in a

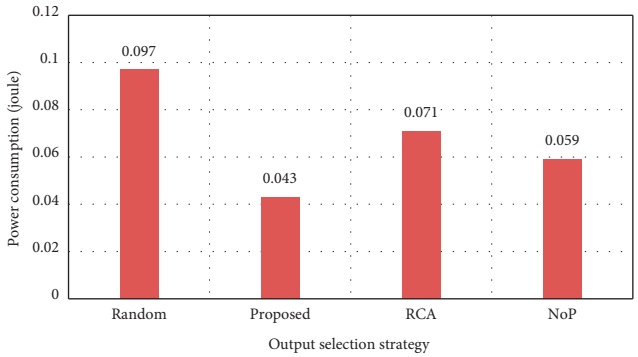


Figure 7: Power consumption (load = 40%).

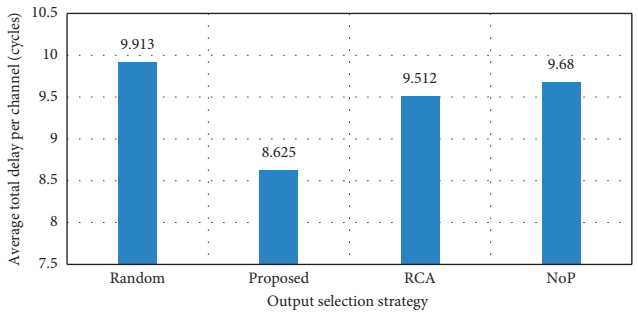


Figure 8: Average total delay per channel (load = 70%).

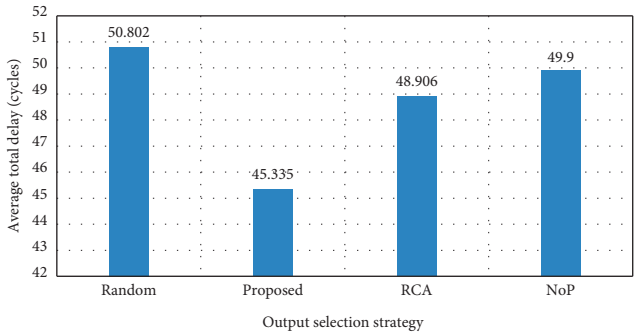


Figure 9: Average total delay (load = 70%).

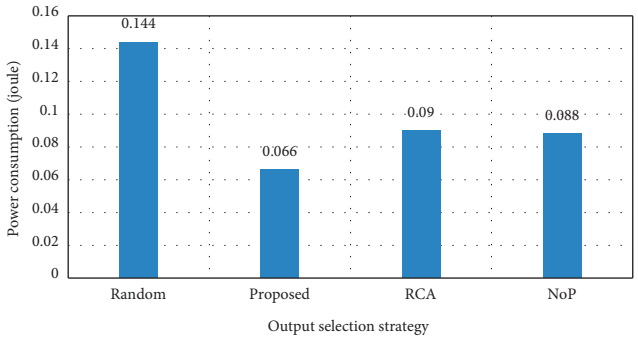


Figure 10: Power consumption (load = 70%).

path leading to increased traffic or the placement in the queue of the input buffers of other routers, which can cause bottlenecks and increased heat and power consumption.

4.2. Second Assessment: Load = 70%. In this assessment, the solutions are evaluated in the case of a traffic load of 70%. Through this assessment, one can examine the performance level in a high load case. The average simulation results have been obtained after 5000 runs. The assessment results are shown in Figures 8–10.

As shown in Figure 8, although the average delay has had a significant increase in the solutions, the best output channels have been selected in the proposed solution using the selection functions that detect the local and nonlocal traffic, which results in lower delay than other solutions. The average total delay is also shown in Figure 9, and again, the optimality of the solution can be verified. Finally, in Figure 10, the power consumption is tested in the solutions and shown with a loading rate of 70%. The optimality of the proposed power consumption strategy is also well presented in this figure.

5. Conclusion

There are usually several different paths to move from one node to another in NoC-based neuromorphic systems; accordingly, the selection functions are used along with the routing algorithms. The effect of any routing algorithm depends on the selection strategy. When the routing function returns a set of output channels, the selection function is used to select the output channel to which the packet is sent [19]. In this research, an adaptive routing algorithm with a hybrid selection strategy is presented, and by using it and the type of traffic, it would be possible to select the best outlet channel in terms of local and nonlocal nature of the packets. Finally, in the Nirgam simulation, it was shown that this proposed method increases the efficiency significantly. The conducted tests demonstrated that this method reduces the average and maximum delay significantly compared to RCA, NoP, and Random strategies, and when some of the calculations are done offline, power consumption is reduced significantly.

6. Future Works

Based on the results that we have obtained in this paper, we will improve the efficiency of our method to make it more suitable for the real-world physical environment. Besides, more criteria will be taken into consideration to make our method satisfy more parallel processing in NoC-based neuromorphic systems.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. Javed, J. Harkin, L. McDaid, and J. Liu, "Exploring spiking neural networks for prediction of traffic congestion in networks-on-chip," in *Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–5, IEEE, Seville, Spain, October 2020.

[2] M. Bouvier, A. Valentian, T. Mesquida et al., "Spiking neural networks hardware implementations and challenges," *ACM Journal on Emerging Technologies in Computing Systems*, vol. 15, no. 2, pp. 1–35, 2019.

[3] N. Jafarzadeh, M. Palesi, A. Khademzadeh, and A. Afzali-Kusha, "Data encoding techniques for reducing energy consumption in network-on-chip," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, no. 3, pp. 675–685, 2014.

[4] X. Zhou, L. Liu, Z. Zhu, and D. Zhou, "A routing aggregation for load balancing network-on-chip," *Journal of Circuits, Systems and Computers*, vol. 24, no. 9, Article ID 1550137, 2015.

[5] E.-J. Chang, H.-K. Hsin, C.-H. Chao, S.-Y. Lin, and A.-Y. Wu, "Regional ACO-based cascaded adaptive routing for traffic balancing in mesh-based network-on-chip systems," *IEEE Transactions on Computers*, vol. 64, no. 3, pp. 868–875, 2015.

[6] G. Dimitrakopoulos, A. Psarras, and I. Seitanidis, *Microarchitecture of Network-On-Chip Routers*, Springer, Berlin, Germany, 2015.

[7] T. S. Arulananth, M. Baskar, U. S. SM et al., "Evaluation of low power consumption network on chip routing architecture," *Microprocessors and Microsystems*, vol. 82, Article ID 103809, 2021.

[8] G. Ascia, V. Catania, M. Palesi, and D. Patti, "Implementation and analysis of a new selection strategy for adaptive routing in networks-on-chip," *IEEE Transactions on Computers*, vol. 57, no. 6, pp. 809–820, 2008.

[9] A. Asghari, A. A. Zoraghchian, and M. Trik, "Presentation of an algorithm configuration for network-on-chip architecture with reconfiguration ability," *International Journal of Electronics Communication and Computer Engineering (IJECCE)*, vol. 5, no. 5, 2014.

[10] G. Ascia, V. Catania, S. Monteleone, M. Palesi, D. Patti, and J. Jose, "Analyzing networks-on-chip based deep neural networks," in *Proceedings of the 13th IEEE/ACM International Symposium on Networks-on-Chip*, pp. 1-2, New York, NY, USA, October 2019.

[11] A. Rezaei, M. Daneshtalab, F. Safaei, and D. Zhao, "Hierarchical approach for hybrid wireless network-on-chip in many-core era," *Computers & Electrical Engineering*, vol. 51, pp. 225–234, 2016.

[12] T. Mak, P. Y. Cheung, K. P. Lam, and W. Luk, "Adaptive routing in network-on-chips using a dynamic-programming network," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 8, pp. 3701–3716, 2010.

[13] S. B. Furber, D. R. Lester, L. A. Plana et al., "Overview of the spinnaker system architecture," *IEEE Transactions on Computers*, vol. 62, no. 12, pp. 2454–2467, 2012.

[14] S. Werner, J. Navaridas, and M. Luján, "A survey on network-on-chip architectures," *ACM Computing Surveys (CSUR)*, vol. 50, no. 6, p. 89, 2017.

[15] A. R. Young, M. E. Dean, J. S. Plank, and G. S. Rose, "A review of spiking neuromorphic hardware communication systems," *IEEE Access*, vol. 7, pp. 135606–135620, 2019.

[16] S. Furber, "Large-scale neuromorphic computing systems," *Journal of Neural Engineering*, vol. 13, no. 5, Article ID 051001, 2016.

[17] T. S. Das, P. Ghosal, and N. Chatterjee, "VCS: a method of in-order packet delivery for adaptive NoC routing," *Nano Communication Networks*, vol. 28, Article ID 100333, 2020.

[18] P. Gratz, B. Grot, and S. W. Keckler, "Regional congestion awareness for load balance in networks-on-chip," in *Proceedings of the 2008 IEEE 14th International Symposium on High Performance Computer Architecture*, pp. 203–214, IEEE, Salt Lake City, UT, USA, February 2008.

[19] R. Akbar and F. Safaei, "A novel heterogeneous congestion criterion for mesh-based networks-on-chip," *Microprocessors and Microsystems*, vol. 84, Article ID 104056, 2021.

[20] M. Ramakrishna, V. K. Kodati, P. V. Gratz, and A. Sprintson, "GCA: global congestion awareness for load balance in networks-on-chip," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 7, pp. 2022–2035, 2016.

[21] S. Moradi, N. Qiao, F. Stefanini, and G. Indiveri, "A scalable multicore architecture with heterogeneous memory structures for dynamic neuromorphic asynchronous processors (DYNAPs)," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 1, pp. 106–122, 2017.

[22] F. Galluppi, S. Davies, A. Rast, T. Sharp, L. A. Plana, and S. Furber, "A hierachical configuration system for a massively parallel neural hardware platform," in *Proceedings of the 9th Conference on Computing Frontiers*, pp. 183–192, New York, NY, USA, May 2012.

[23] Q. Zhang, T. Wang, Y. Tian, F. Yuan, and Q. Xu, "ApproxANN: an approximate computing framework for artificial neural network," in *Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 701–706, IEEE, Grenoble, France, March 2015.

[24] A. Yasoubi, R. Hojabr, and M. Modarressi, "Power-efficient accelerator design for neural networks using computation reuse," *IEEE Computer Architecture Letters*, vol. 16, no. 1, pp. 72–75, 2016.

[25] E. Painkras, L. A. Plana, J. Garside et al., "SpiNNaker: a 1-W 18-core system-on-chip for massively-parallel neural network simulation," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 8, pp. 1943–1953, 2013.

[26] S. Carrillo, J. Harkin, L. J. McDaid et al., "Scalable hierarchical network-on-chip architecture for spiking neural network hardware implementations," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 12, pp. 2451–2461, 2012.

[27] Y.-H. Chen, J. Emer, and V. Sze, "Eyeriss: a spatial architecture for energy-efficient dataflow for convolutional neural networks," *ACM SIGARCH Computer Architecture News*, vol. 44, no. 3, pp. 367–379, 2016.

[28] H. Zhang, J. Thompson, M. Gu et al., "Efficient on-chip training of optical neural networks using genetic algorithm," *ACS Photonics*, vol. 8, no. 6, pp. 1662–1672, 2021.

[29] O. M. Ikechukwu, K. N. Dang, and A. B. Abdallah, "On the design of a fault-tolerant scalable three dimensional NoC-based digital neuromorphic system with on-chip learning," *IEEE Access*, vol. 9, pp. 64331–64345, 2021.

[30] F. Akopyan, J. Sawada, A. Cassidy et al., "Truenorth: design and tool flow of a 65 mw 1 million neuron programmable neurosynaptic chip," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 10, pp. 1537–1557, 2015.

[31] A. Amir, P. Datta, W. P. Risk et al., "Cognitive computing programming paradigm: a corelet language for composing networks of neurosynaptic cores," in *Proceedings of the 2013*

International Joint Conference on Neural Networks (IJCNN), pp. 1–10, IEEE, Dallas, TX, USA, August 2013.

[32] M. Davies, N. Srinivasa, T.-H. Lin et al., "Loihi: a neuromorphic manycore processor with on-chip learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, 2018.

[33] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza et al., "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, 2014.

[34] B. V. Benjamin, P. Gao, E. McQuinn et al., "Neurogrid: a mixed-analog-digital multichip system for large-scale neural simulations," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 699–716, 2014.

[35] J. Schemmel, D. Brüderle, A. Grübl, M. Hock, K. Meier, and S. Millner, "A wafer-scale neuromorphic hardware system for large-scale neural modeling," in *Proceedings of the 2010 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1947–1950, IEEE, Paris, France, May 2010.

[36] S. B. Furber, F. Galluppi, S. Temple, and L. A. Plana, "The spinnaker project," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 652–665, 2014.

[37] F. A. Samman, T. Hollstein, and M. Glesner, "Runtime contention and bandwidth-aware adaptive routing selection strategies for networks-on-chip," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1411–1421, 2012.

[38] H. C. Touati and F. Boutekkouk, "Reliable weighted globally congestion aware routing for network on chip," *International Journal of Embedded and Real-Time Communication Systems*, vol. 11, no. 3, pp. 48–66, 2020.

[39] F. Bahman, A. Reza, M. Reshadi, and S. Vazifedan, "CACBR: congestion aware cluster buffer base routing algorithm with minimal cost on NOC," *CCF Transactions on High Performance Computing*, vol. 2, no. 3, pp. 1–10, 2020.

[40] M. Tang, X. Lin, and M. Palesi, "Local congestion avoidance in network-on-chip," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 7, pp. 2062–2073, 2016.

[41] L. Wang, X. Wang, and T. Mak, "Adaptive routing algorithms for lifetime reliability optimization in network-on-chip," *IEEE Transactions on Computers*, vol. 65, no. 9, pp. 2896–2902, 2016.

[42] A. Ben Ahmed and A. Ben Abdallah, "Graceful deadlock-free fault-tolerant routing algorithm for 3D network-on-chip architectures," *Journal of Parallel and Distributed Computing*, vol. 74, no. 4, pp. 2229–2240, 2014.

[43] Y. Amar Babu, G. M. V. Prasad, and J. B. Solomon, "FPGA implementation of buffer-less NoC router for SDM-based network-on-chip," in *Progress in Advanced Computing and Intelligent Engineering*, K. Saeed, N. Chaki, B. Pati, S. Bakshi, and D. Mohapatra, Eds., Springer, Singapore, pp. 561–567, 2018.

[44] Y.-Y. Chen, E.-J. Chang, H.-K. Hsin, K.-C. Chen, and A.-Y. A. Wu, "Path-diversity-aware fault-tolerant routing algorithm for network-on-chip systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 3, pp. 838–849, 2017.

[45] V. Catania, A. Mineo, S. Monteleone, M. Palesi, and D. Patti, "Noxim: an open, extensible and cycle-accurate network on chip simulator," in *Proceedings of the 2015 IEEE 26th International Conference on Application-Specific Systems, Architectures and Processors (ASAP)*, pp. 162-163, IEEE, Toronto, Canada, July 2015.

[46] A. B. Gabis, P. Bomel, and M. Sevaux, "Bi-objective cost function for adaptive routing in network-on-chip," *IEEE Transactions on Multi-Scale Computing Systems*, vol. 4, no. 2, pp. 177–187, 2018.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 3583275, 17 pages
https://doi.org/10.1155/2021/3583275

Research Article

Leveled Design of Cryptography Algorithms Using Cybernetic Methods for Using in Telemedicine Applications

Ali Mohammad Norouzzadeh Gil Molk,[1] Mohammad Reza Aref,[2] and Reza Ramazani Khorshiddoust[3]

[1]*Department of Computer Engineering, Islamic Azad University, North Tehran Branch, Tehran, Iran*
[2]*Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran*
[3]*Department of Industrial Engineering & Management Systems, Amirkabir University of Technology, Tehran, Iran*

Correspondence should be addressed to Mohammad Reza Aref; aref@sharif.edu

Received 28 May 2021; Revised 31 July 2021; Accepted 12 August 2021; Published 11 September 2021

Academic Editor: Navid Razmjooy

The technology world is developing fast with the developments made in the hardware and software areas. Considering that privacy and security of telemedicine applications are among the main necessities of this industry, as a result, there is a need to use lightweight and practical algorithms to be used in applications in the field of telemedicine, while security have the least negative impact. The distinct and contradicting components in the design and implementation of the cryptography algorithm, to achieve various objectives in medicine-based applications, have made it a complicated system. It is natural that, without identifying the components, indices, and properties of each system component, the hardware and software resources are lost and a proper algorithm cannot be designed. Accordingly, this paper presents a leveled model of cryptography algorithms using the cybernetic method. First, the main objectives and measures in the design of the cryptography algorithms are extracted using the measure reduction methods, and some of the excess and overlapping measures are eliminated. Then, three general classes of the cryptography algorithm design and implementation measures, applications of cryptography algorithms, and cryptography implementation techniques are extracted. Since the complexity of the cryptography algorithm design is relatively high, the cybernetic methodology is used to present a supermodel to make the cryptography algorithm design objective. Such design prevents examining unnecessary details and establishes a bidirectional relationship between the main design and implementation process and the support process. This relationship provides the support requirements of the main process by the support process at each step. Finally, the *Q*-analysis tools are used to analyse the proposed method, and the efficiency results are represented.

1. Introduction

Since telemedicine technology relies on data transmission, data security is critical in order to keep information transmission confidential and patients' privacy, and any potential threat or attack on telemedicine networks such as unauthorized access to data and alteration or destruction of patient data should be considered. In other words, any weakness in any part of the telemedicine network can affect the entire system. Therefore, in order to create security in the field of storage and exchange of information in the medical network, enforcement mechanisms using relevant standards should be considered. Accordingly, this study has focused on the surface design of cryptographic algorithms for use in telemedicine. Security of cryptography systems depends on "algorithm power" and "key size," [1] and the general cryptography levels are divided into three levels, including cryptography algorithms, security protocols, and applications [2, 3]. However, the cryptography algorithms are designed and implemented to achieve goals such as confidentiality, authentication, and integrity [4], but various components such as speed, resource consumption, application type, flexibility, scalability, and reliability should be considered for their design. Design of a cryptography

algorithm should be systematic, comprehensive, and staged. All required components of the information security should be considered in an excellence pattern in terms of technical, organization, procedural, and humanitarian aspects. Identifying new cryptographic challenges such as post-quantum cryptography and its agility, mobile applications, robustness of algorithms, and the role of implementation methods to achieve the above goals can be implemented in a comprehensive model [5–7]. Providing all these requirements simultaneously in the design of an algorithm is difficult and sometimes impossible. If contradictory objectives are considered for formulation of an objective/objectives of an algorithm, most algorithms might be broken, and if attacker has sufficient time, motivation, and resources, he can track the information [8]. Accordingly, presenting a model to make the cryptography algorithm design targeted is very important. Therefore, in this study, an approach is presented to design a leveled model of cryptography algorithms using the cybernetic method. The rest of this paper is structured as follows. In the second section, the literatures review is presented. In the third section, the cybernetic methodology is described, and a model is specified for design and implementation of cryptography algorithms based on extracted indices. In the fourth section, the proposed model is used to examine the design of the cryptography algorithm using the cybernetic supermodel. Finally, the proposed approach is evaluated using the *Q*-analysis method.

2. Literature Review

In [9], a security approach based on cryptography has been presented through examining the security issues in mobile devices and the available solutions. Also, it is mentioned that asymmetric cryptography is not a proper option for securing the resource-limited infrastructures such as IoT due to high complexity of the design and implementation. On the contrary, employing symmetric algorithms has other security issues. Accordingly, it has studied the design of cryptography algorithms based on position. To this end, an approach based on the AES algorithm and position of an efficient cryptography approach has been designed. In this approach, the application diagram is described and the user operation is studied. In the following, the operation flows of the system are described. Finally, it was evaluated that security can be increased through employing this approach. In [10], the design of stream cipher algorithms has been studied. It has been mentioned that stream cipher is one of the essential branches of symmetric cryptography, which requires limited hardware resources for execution. Therefore, considering the development of the communication technologies, the need to these algorithms is increasing. Accordingly, in the following, the design procedures and performance of various encryption algorithms, including NFSR, eStram, FCSR, and Panama, are presented through describing the main requirements of the cryptography algorithm design. The main purpose of this study is to present a perspective of stream cipher algorithm design and their performance. In 2016, NIST published a document called cipher standard and development instructions [11]. Transparency, openness, balance, accuracy, technical merit, global acceptability, usability, continuous improvement, and innovation and intellectual property (IIP) are the guidance principles of NIST cipher standards and development procedures. Also, NIST has started a procedure to request, evaluate, and standard of one or multiple public key cipher algorithms robust against quantum attacks [12]. In [11], lifecycle management processes and policies of cipher standard have been presented, where its main principles include: identifying and evaluating the needs, announcing the user's intention on a standard or instruction, considering the requirements and solutions, defining a specific program and procedure and design and development of a standard, and evaluating and maintaining the standard. In [13], an analytical framework has been presented to hardware and software implementation using cipher programs that verifies an integrated statistical framework which can implement the classified algorithms successfully based on a combination of heterogeneous hardware features and their software applications. The model presented in this paper includes six elements of goal, input, activities, output, outcomes, and performance. In [14], software engineering methodologies have been used to propose an adaptive approach for presenting a robust cipher key generation algorithm. The technique used in this method is based on self-checking procedures that can detect the system-level errors. Therefore, it can be used to check the security keys generated via employing random factors. These factors have been presented in the NIST evaluation results. In this software method, the values of the random factors are smaller than the acceptance values, and the key is generated when a valid value is detected. The generated keys are generated through shift register and SIGBA technique. The evaluation results indicate the efficiency of the presented approach in generating valid cipher keys. In [15, 16], security issues of mobile devices and processing infrastructure, including mobile computing and edge computing, have been studied. It has been concluded the importance of cipher algorithms and necessity of employing new models consider the complexity of these infrastructures. Bhowmik et al. [17] focused on security issues in telemedicine and introduced a double-tier (nDTCS) encryption system. Accordingly, this solution has proposed a modified logistic map and a congruence-based security model to secure telemedicine medical transactions. Two keys have been used for the encryption and decryption process, intermediate key and session key. The results of the evaluation indicate the effectiveness of the solution in order to secure the information through the proposing method [18]. In order to protect telemedicine communications, a key exchange solution is proposed by improving the Diffie–Hellman cryptographic algorithm. In this method, a randomized key generation is used to generate the key. The proposed solution is naturally safe and reliable due to the use of the Diffie–Hellman algorithm, so there is no need for recalculations or key reversal. The results of the evaluation also indicate that the proposed method is safe against guessing key attacks. In [19], an intelligent and secured transmission security solution for heart disease reports based on session key-based methods is presented. For this

purpose, matrix confusion operations are used. The innovation of this solution is in the process of matrix transfer, which is transmitted in the form of a number of cardiologists in particular. Finally, the efficiency of the proposed method is evaluated with regard to cryptographic engineering, transparency, and strength. The results indicate that this method provides more security in the medical data transmission process. Hosseinian et al. [20] examined the importance of information security needs in telemedicine technology in the field of information transmission. In this study, the data collection tool was a questionnaire that was designed based on the criteria of the Association of Information Management and Health Care Systems (HMISS) in the field of telemedicine network security and security standards of the American Telemedicine Association. This questionnaire has been calculated separately based on a score of 1 to 5. Table 1 summarizes the importance of each section.

3. Cybernetic Supermodel

Considering the complexity of the cipher context in terms of various aspects, designing a cipher algorithm should be systematic, comprehensive, and stage. To design cipher algorithms, different technologies, including mathematics, physics, biometric, biology, and social engineering, are used [21, 22]. Also, concepts and basic sciences such as theory of numbers, Boolean functions [23], and random functions [24, 25] are very essential. Depending of the application of cipher algorithms, various technical and nontechnical requirements should be considered for their design. Detecting new cryptography challenges such as postquantum cipher and its agility [26] and mobile applications [6, 7], making an algorithm robust, and the role of implementation methods to achieve the above goals are the issues that should be considered in a comprehensive model. Amidst, considering the large number of components in the design of cipher algorithms and their relationship and impact on each other, the design and implementation of these algorithms has become complicated. One of the best tools to design a complicated system is to present a model for that system. The steps associated with the design and implementation procedure of the cipher algorithm regardless of the triple classification of the hash, symmetric, and asymmetric functions at the highest level are shown in Figure 1.

Cryptography is one of the main information security components to transmit information from the sender to the receiver using the most secure method [27]. Design of the algorithms has different requirements depending on its application in the embedded or nonembedded system [28]. To design a robust algorithm, various technical and nontechnical factors should be considered so that the designed algorithm has sufficient robustness [29]. On the contrary, the effective factors should be in a coherent model with logical integration so that their impact on each other can be measured and evaluated; for instance, in [30–33], various algorithms have been evaluated in terms of some parameters. In fact, designing a conceptual model for the cipher algorithm requires considering all factors, components, and indices that affect the design and implementation of the cipher algorithms. Accordingly, in the design of the cipher algorithm, there should be a balance between "efficiency" and "resources" required for a specific security level [34]. Considering the design and generation process of the cipher algorithm and classification of factors, components, and indices, the conceptual cybernetic supermodel is used for design and implementation. Cybernetic is mainly focused on system performance and how they control their activities and communicate with their components. Therefore, the cybernetic pattern might be a scientific basis for making the cipher algorithms targeted. The cybernetic model of the cipher algorithms has four components of approach/strategy, main process, support process, and control process. The interactions of the main and support processes constitute the structure of the cipher system. These interactions result in a complicated diagram. To overcome this complexity, a leveled structure and mathematical facilities such as graph and matrix are used. Accordingly, the general cybernetic model for the design and implementation of the cipher algorithms is shown in Figure 2. This model is comprised of four sections: development approach/strategy process, main process, support process, and control process. The main process includes cryptography algorithms. The support process is divided into two general classes of hardware and software. The control process includes controlling the design, implementation, and controlling the outcome. As mentioned, this model is designed in the general level; and, its processes and components are studied in detail in Section 3.

Since the cipher algorithms have specific complexities, the component model is used to facilitate the processes. This type of design prevents spending time on unnecessary details. The strategic model for design of cipher algorithms should be presented at a level of detail that creates a tradeodd between "inclusion" and "applicability." "Inclusion" indicated including various cipher algorithms. Accordingly, considering the component extracted for the main, support, and control section, a cybernetic model can be used to design and implement cipher algorithms in three levels, as shown in Figure 3. According to this model, there is a bidirectional relationship between the main design and implementation process and the support process; at each step, as a result of this relationship, the support requirements are demanded by the main process and provided by the support process.

3.1. Data Matrix of the Design Model Components. Since there are a large number of extracted objectives or measures in the design of cipher algorithms and some of them overlap, or eliminating some of them causes no problem for achieving the main goals, the criterion reduction method can be used to eliminate some measures. Accordingly, in this study, the approach presented in [19] is used to reduce the number of measures. In the feature reduction process, if eliminating one measure does not change the effective set of the problem, it is unnecessary. Therefore, after feature reduction, the component extraction process is carried out. In this step, three general classes of measures are extracted,

Table 1: The importance of information security needs in telemedicine technology in the field of information transmission.

Data transfer	Very important	Important	No idea	Nonsignificant
Implement network protocols to ensure the transmission of information and check its integrity	2/65%	4/30%	4/3	0
Establish a communication protocol to share information between local health institutions	50%	37%	13	0
Encrypt important files and information	2/62%	6/35%	0	2/2
Investigation of encryption mechanism by technical team of security assessor	5/56%	3/28%	15/2	0
Use combinations of numbers and letters for encryption	2/68	25	6/8	0
Use uppercase and lowercase letters for encryption to access remote network networks	50%	24%	17/4	8/7
Methods for controlling the integrity of application information	3/53%	6/35%	8/9	2/2

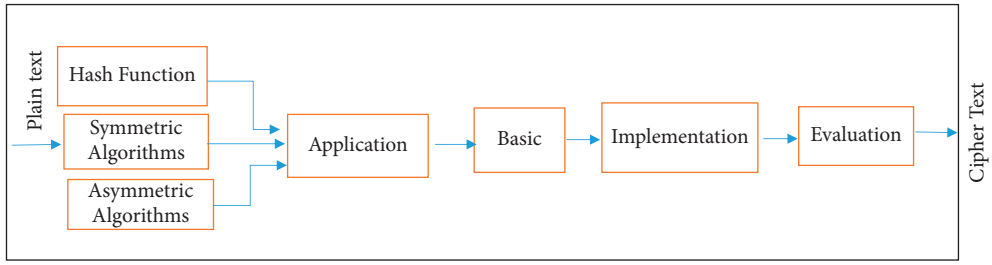


Figure 1: Main process of design and implementation of cipher algorithms at the highest level.

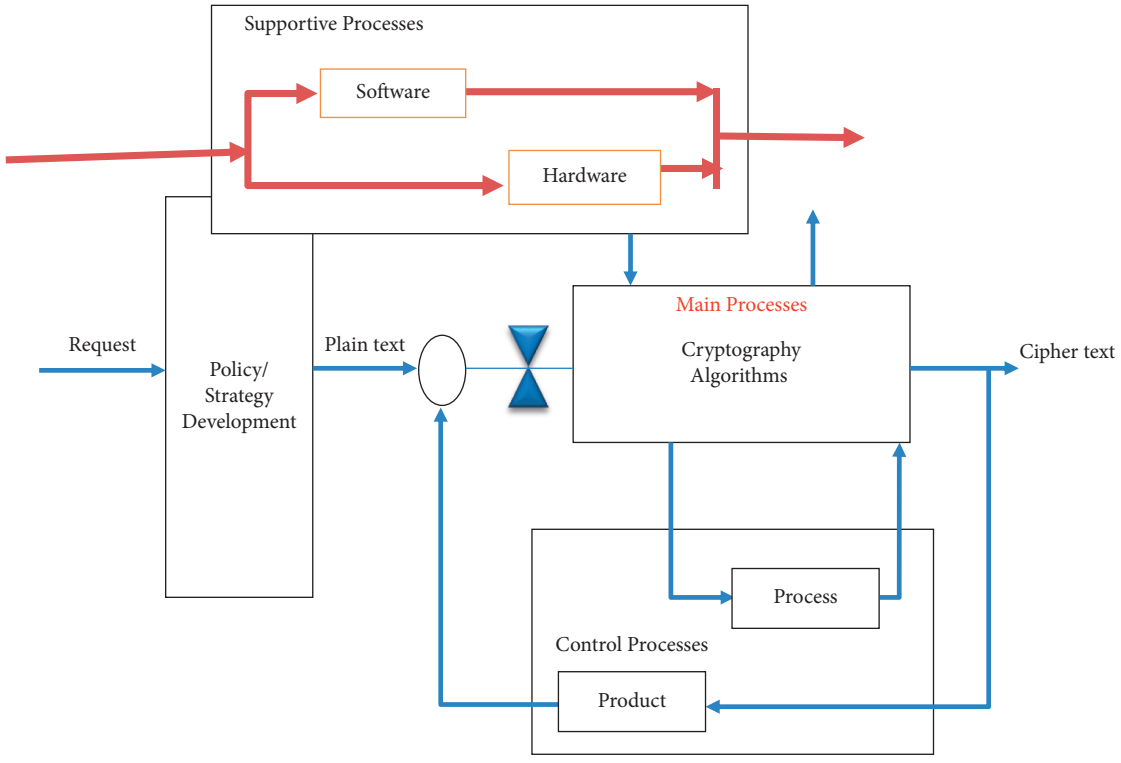


Figure 2: General schematic of the conceptual model of the cipher algorithm design.

including design and implementation objectives of cipher algorithms, applications of cipher algorithms, and implementation methods of the cipher algorithms. In the following, the reduced measures are classified into the above classes and modeling is carried out based on available measures of these three classes. Accordingly, based on the level-3 cybernetic model (Figure 3), the following three matrices are constituted to design and implement the cipher algorithms:

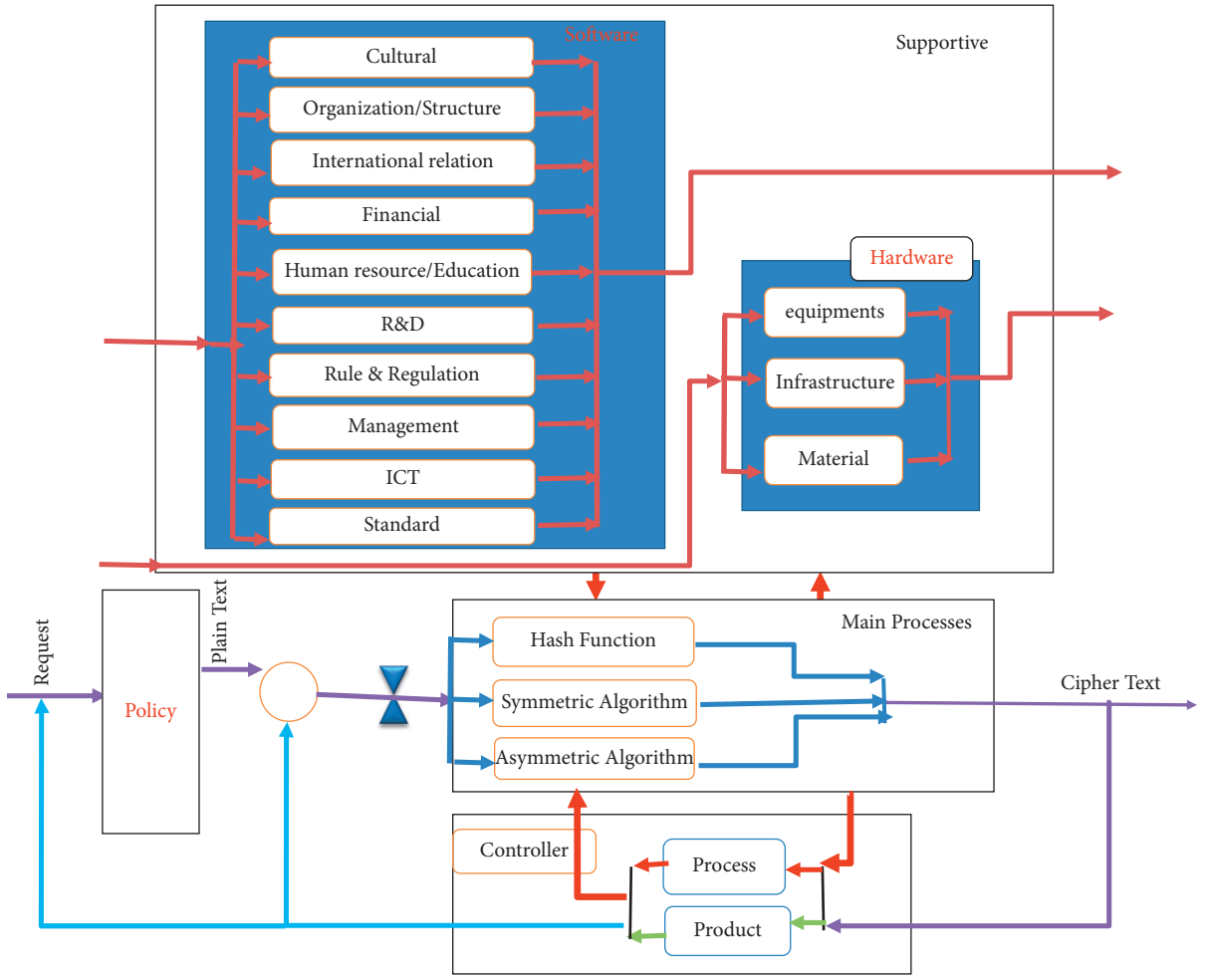


Figure 3: The cybernetic model of design and implementation of cipher algorithms at level 3.

(i) The relationship matrix of support indices with main design and implementation processes of the cipher algorithms: since there are 13 support indices in the level-3 model (Figure 3) and 4 steps in the level-1 model (Figure 2), a 13∗4 matrix is constituted to determine the relationship between the members of these two processes, where its rows are the elements of the support process and its columns are the quadruple elements of the design and implementation of cryptography algorithms. The elements of this matrix are between 0 and 10, as shown in Table 2 [35]. The value of each element represents the effectiveness of each support index on each design and implementation step. Analysis of this matrix and its modeling computation provides the possibility for the policy-makers and developers to manage the resources required for each step of design and implementation of the cryptography algorithms and use the available hardware and software resources optimally.

(ii) The relationship matrix of the objectives with the main design process of the cryptography algorithms: considering the seven objectives extracted in Section 2 and four steps in the main process, to determine the role of each step in achieving the seven objectives, a 7∗4 matrix is constituted. Since the number of final resources used to extract data is 814 papers, technical reports, and documents, the elements of this matrix are between 0 and 814, as given in Table 2. In this matrix, the value of each element represents the number of studies, documents, and reports indicating the relationship between two components. An interesting point in this matrix is that all algorithms implemented in the studied documents are evaluated.

(iii) The relationship matrix of the implementation techniques and the design objectives of the cryptography algorithms: considering the 29 extracted techniques for cryptography algorithm implementation and seven main objectives of the cryptography algorithms, a $29 * 7$ matrix is constituted to determine how much each technique is used for cryptography implementation to achieve each objective (Table 3). The elements of this matrix are between 0 and 814. A part of this matrix is shown in Table 4.

Table 2: The 13-component interaction matrix of the support and the main design and implementation processes of cryptography algorithms.

			Main process				Level 1
			Cryptography algorithms				Level 2
			Application	Theoretical basis	Implementation	Evaluation	Level 3
Support process	Software	Culture	5	6	5	2	.
		Organization/structure	8	7	7	7	.
		International and public relations	5	4	4	7	.
		Financial resources	7	6	7	6	.
		Human resources and education	8	10	10	9	.
		Research and development	10	8	8	10	.
		Rules	7	4	5	10	.
		Management	9	7	8	7	.
		FAVA	10	5	8	7	.
		Standard	10	3	8	10	.
	Hardware	Equipment	8	6	7	8	
		Infrastructure	9	5	5	7	
		Material	7	2	5	5	

Table 3: The relationship matrix of the objectives with the design process of the cryptography algorithms.

Main process objectives	Application	Theoretical basis	Implementation	Evaluation
Security	203	386	516	516
Simplicity	5	23	38	38
Resources	41	30	81	81
Flexibility	15	28	45	45
Scalability	6	11	22	22
Speed	72	146	247	247
Reliability	2	10	11	11

Table 4: A part of the interaction matrix between the cryptography algorithm implementation techniques and the seven objectives.

Level 3		Security	Simplicity	Using resources	Flexibility	Scalability	Speed	Reliability
Support process	Avalanche effect	**32**	**0**	**0**	**0**	**0**	**0**	**0**
	Digital signature	**1**	**0**	**0**	**0**	**0**	**0**	**0**
	Block size	**1**	**1**	**0**	**0**	**2**	**0**	**0**
	Image sharing	**0**	**0**	**0**	**1**	**0**	**0**	**0**
	Parallel processing	**0**	**0**	**0**	**1**	**12**	**0**	**0**
	Threshold technique	**3**	**1**	**0**	**0**	**0**	**0**	**0**
	Data mining	**1**	**0**	**0**	**0**	**0**	**0**	**0**
	Binary tree	**1**	**0**	**0**	**0**	**3**	**0**	**0**
	Cycle	**5**	**2**	**0**	**0**	**12**	**0**	**0**
	Multistage crypto	**1**	**0**	**0**	**0**	**0**	**0**	**0**
	Hybrid method	**30**	**3**	**0**	**2**	**9**	**0**	**2**
	Hardware	**0**	**0**	**7**	**0**	**0**	**0**	**0**

3.1.1. Design of Cryptography Algorithms Using the Cybernetic Supermodel. Considering the above discussion and presence of numerous indices and components in the design and implementation of the cryptography algorithm, which make it a complex system, identifying the relationship between these indices and ranking them is a necessity. Accordingly, in this section, the *Q*-analysis method is used and the output of the three matrices is analyzed. According to the level three of the proposed cybernetic model, the design and implementation support process of the cryptography algorithms includes 13 components. On the contrary, the design and implementation steps of the cryptography algorithms also include four steps, constituting a $13 * 4$ matrix.

3.2. Calculating the Incidence Matrix. First, the incidence matric is obtained based on the data matrix 4–1. This matrix represents the "impact of support indices on the main design process of the cryptography algorithms." The data matrix is comprised of two sets D, support indices, set C, and quadruple design and implementation steps (Table 5). The incidence matrix calculated using the data matrix for $\alpha = \%70$ is represented in Table 6. By assigning different values to the parameter $\boldsymbol{\alpha}$, difference incidence matrices are obtained. The results of the Q-analysis using C++ coding for $\alpha = \%70$ are given in Figure 4:

$$\begin{aligned} D &= \{d_1, d_2, \ldots, d_{13}\}, \\ C &= \{c_1, c_2, c_3, c_4\}. \end{aligned} \tag{1}$$

3.2.1. Geometric Representation. Multidimensional properties of the system are defined by a simple or complex set $K_D(C, \lambda)$, such that the entities of the set D represent the support indices and entities of the set C represent the quadruple design and implementation steps of the cryptography algorithms.

In the sample with $\alpha_{\%70} = 7$, $\mathrm{d_i}$s is as follows:

$$\begin{aligned} d_1 &= \{\}, \\ d_2 &= \{c_1, c_2, c_3, c_4\}, \\ d_3 &= \{c_4\}, \\ d_4 &= \{c_1, c_3\}, \\ d_5 &= \{c_1, c_2, c_3, c_4\}, \\ d_6 &= \{c_1, c_2, c_3, c_4\}, \\ d_7 &= \{c_1, c_4\}, \\ d_8 &= \{c_1, c_2, c_3, c_4\}, \\ d_9 &= \{c_1, c_3, c_4\}, \\ d_{10} &= \{c_1, c_3, c_4\}, \\ d_{11} &= \{c_1, c_3, c_4\}, \\ d_{12} &= \{c_1, c_4\}. \end{aligned} \tag{2}$$

Table 5: Sets of d_is and c_is; support indices based on the data matrix.

d_1	Culture
d_2	Structure/organization
d_3	International/public relations
d_4	Financial
d_5	Human resources/education
d_6	Research and development
d_7	Rules
d_8	Management
d_9	FAVA
d_{10}	Standard
d_{11}	Equipment
d_{12}	Infrastructure
d_{13}	Material
C_1	Application
C_2	Theoretical basis
C_3	Implementation
C_4	Evaluation

Table 6: The incidence matrix of the support indices' impact of the design steps of the cryptography algorithms with $\alpha = \%70$.

	C_1	C_2	C_3	C_4
d_1	0	0	0	0
d_2	1	1	1	1
d_3	0	0	0	1
d_4	1	0	1	0
d_5	1	1	1	1
d_6	1	1	1	1
d_7	1	0	0	1
d_8	1	1	1	1
d_9	1	0	1	1
d_{10}	1	0	1	1
d_{11}	1	0	1	1
d_{12}	1	0	0	1
d_{13}	1	0	0	0

The simplexes of $\sigma_{\mathbf{p}}(\mathbf{d_i})$ are also

$$\begin{matrix} \sigma_{-1}(d_1) & \sigma_3(d_2) & \sigma_0(d_3) & \sigma_1(d_4) & \sigma_3(d_5) & \sigma_0(d_3) & \sigma_1(d_7) \\ \sigma_3(d_8) & \sigma_2(d_9) & \sigma_2(d_{10}) & \sigma_2(d_{11}) & \sigma_1(d_{12}) & \sigma_1(d_{13}) & \end{matrix}. \tag{3}$$

Therefore, the complex dimension is 3. In other words, the diagnosis classes d_2 (structure/organization), d_5 (human resources/education), d_6 (research and development), and d_8 (management) have the largest dimension.

3.3. Calculating Dimensions and Q-Link. Q-link is defined as the link between a subset with smallest interface between two subsequent d_is in the chain of d_1 to d_n. Q-link between two subsequent d_is with $\boldsymbol{\alpha}_{\%70} = 7$ is

-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	3	0	1	3	3	1	3	2	2	2	1	0
		0	-1	0	0	0	0	0	0	0	0	-1
			1	1	1	0	1	1	1	1	0	0
				3	3	1	3	2	2	2	1	0
					3	1	3	2	2	2	1	0
						1	1	1	1	1	1	0
							3	2	2	2	1	0
								2	2	2	1	0
									2	2	1	0
										2	1	0
											1	0
												0

q	Q	SETS
3	1	{X2, X5, X6, X8}
2	1	{X2, X5, X6, X8, X9, X10, X11}
1	1	{X2, X4, X5, X6, X7, X8, X9, X10, X11, X12}
0	1	{X2, X3, X4, X5, X6, X7, X8, X9, X10, X11, X12, X13}

Figure 4: Implementation results of the model for support components of the cryptography algorithms' design with α = %70.

$$\begin{aligned}
&\sigma_{-1}(d_1), \sigma_3(d_2) \longrightarrow 1,\ \sigma_3(d_2), \sigma_0(d_3) \longrightarrow 0, \sigma_0(d_3), \sigma_1(d_4) \longrightarrow -1,\\
&\sigma_1(d_4), \sigma_3(d_5) \longrightarrow 1,\ \sigma_3(d_5), \sigma_3(d_6) \longrightarrow 3, \sigma_3(d_6), \sigma_1(d_7) \longrightarrow 1,\\
&\sigma_1(d_7), \sigma_3(d_8) \longrightarrow 1,\ \sigma_3(d_8), \sigma_2(d_9) \longrightarrow 2, \sigma_2(d_9), \sigma_2(d_{10}) \longrightarrow 2,\\
&\sigma_2(d_{10}), \sigma_2(d_{11}) \longrightarrow 2,\ \sigma_3(d_{11}), \sigma_1(d_{12}) \longrightarrow 1, \sigma_1(d_{12}), \sigma_0(d_{13}) \longrightarrow 0.
\end{aligned} \tag{4}$$

The maximum link dimension is 3, indicating the relationship between the diagnosis classes.

3.3.1. Calculating the Structure Vectors. As mentioned in the definitions, the vector Q_q is a simplification basis, created to eliminate the additional impacts in the equivalent simplex sets. The maximum complex dimension with $\boldsymbol{\alpha}_{\%70} = 7$ is 3. Therefore, the first structure vector based on the output is

$$\text{Dimension} \quad 3 \quad 2 \quad 1 \quad \{0\}. \tag{5}$$

The second structure vector P is

$$\begin{aligned}
&\text{dimensions} \quad 3 \quad 2 \quad 1 \quad 0,\\
&P = \left(P_{\dim 3} \ \ P_{\dim 2} \ \ P_{\dim 1} \ \ P_{\dim 0}\right),\\
&P = \left(4 \ \ 7 \ \ 10 \ \ 12\right),
\end{aligned} \tag{6}$$

where P_q is the number simplexes greater than or equal to q in the set K in which P is the number of simplex link repetitions (support indices) in the quadruple design and implementation steps of the cryptography algorithms. Based on the values of these two vectors, it is seen that the relationship between the support indices and the quadruple design and implementation steps of the cryptography algorithms is high. This issue indicates the role of support components in the design and implementation of the cryptography algorithms, which should be considered seriously.

3.3.2. Calculating the Obstruction or Flexibility Vector. $\mathbf{Q}^*\mathbf{K}$ represents the number of structural obstructions for simplex interactions in dimension k:

$$\begin{aligned}
Q^* &= Q - I \longrightarrow Q^* = \left[1 \ \ 1 \ \ 1 \ \ 1\right] - \left[1 \ \ 1 \ \ 1 \ \ 1\right] \longrightarrow Q^*\\
&= \left[0 \ \ 0 \ \ 0 \ \ 0\right].
\end{aligned} \tag{7}$$

As can be seen, there is no obstruction in any of the communication levels, indicating that there is a significant relationship between the support components at each equivalence class.

3.3.3. Calculating Irregularity. The value of $(\mathrm{ecc}'(\sigma))$ is calculated using the Chinese method. The results for $\boldsymbol{\alpha}_{\%70} = 7$ are shown in Table 7. The calculated irregularity value shows that the indices d_2 (structure/organization), d_5 (human resources/education), d_6 (research and development), and d_8 (management) affect other indices.

3.3.4. Calculating Complexity. Also, the results of Q-analysis can be used to describe structure complexity. According to equations (4)–(9) and for $\boldsymbol{\alpha} = 7$, the complexity measure is

$$Q = (1,\ 1,\ 1,\ 1),$$
$$\psi(K) = 2\left[\frac{(1+2+3+4)}{(4*5)}\right] = 1. \tag{8}$$

Since, in the above model, there is no obstruction among the component of the equivalent class, it was expected that the complexity of the support components is not high, and the obtained complexity index of 1 verifies this expectation. The system complexity for different alpha-cuts is shown in Figure 5.

3.3.5. Ranking the Support Components of the Design and Implementation of the Cryptography Algorithms. The results of using *A*-analysis are shown in Table 7. The connection strength of the factors in one group is specified with alpha-cut. Therefore, the support components are grouped in 5 levels. Each level describes the priority and importance of the group in developing the cryptography algorithms. In Figure 6 the ranking pyramid of the support components using *Q*-analysis is shown. To allocate proper resources, the components existing in higher levels of the pyramid (Figure 6) are of higher priority.

3.3.6. Validation of the Results. In this section, the results of the cybernetic model and *Q*-analysis for support components' ranking are compared with the results reported in the global cybersecurity index (GCI) in 2015, 2017, and 2018 presented by ITU [36–39]. The GCI reports are focused on five indices, including "legal cases, organization necessities, technical issues, capacity building, and cooperation," and the subindices include legal, technical, organization, capacity building, and cooperation. According to the presented indices and subindices, it is clear that the "rules," "standard," "research and development," "education," and "management" are of higher priority in security establishment. Although our research is more skilled and detailed compared to the GCI reports, but the results verify our findings. Also, Table 8 shows the ranking of the support components obtained using *Q*-analysis.

3.3.7. Executing the Model and Analyzing the Results of the Objectives' Impact on Design Steps of the Cryptography Algorithms' Matrix. In this section, the role of the seven components on the quadruple design and implementation steps of the cryptography algorithm is analyzed with *Q*-analysis. Using the *Q*-analysis method and the 7∗4 matrix obtained from the relationship of the cryptography algorithms' design objectives on their quadruple steps, their indices are ranked.

3.3.8. Calculating the Incidence Matrix and Executing the Model. First, the incidence matrix is obtained for the data matrix shown in Figure 2. It can be seen in Table 9 that each element of the two sets represents which indice. The incidence matrix calculated from the data matrix for $\alpha_{\%5} = 40$ is given in Table 10. The results of implementing the model for $\alpha_{\%5} = 40$ are given in Figure 7.

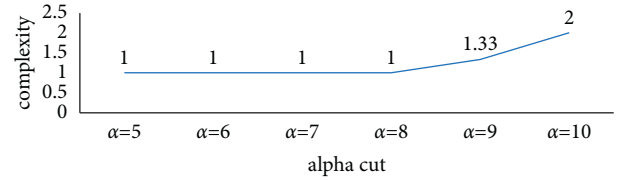


Figure 5: System complexity of the support indices of design and implementation of the cryptography algorithms for different alpha-cuts.

4. Geometric Representation

In the sample with $\alpha_{\%5} = 40$, $d_i s$ are

$$\begin{aligned} d_1 &= \{c_1, c_2, c_3, c_4\}, \\ d_2 &= \{\}, \\ d_3 &= \{c_1, c_3, c_4\}, \\ d_4 &= \{c_3, c_4\}, \\ d_5 &= \{\}, \\ d_6 &= \{c_1, c_2, c_3, c_4\}, \\ d_7 &= \{\}. \end{aligned} \tag{9}$$

The simplexes of $\boldsymbol{\sigma}_{\mathbf{p}}(\mathbf{d}_{\mathbf{i}})$ are also

$$\sigma_3(d_1)\sigma_{-1}(d_2)\sigma_2(d_3)\sigma_1(d_4)\sigma_{-1}(d_5)\sigma_3(d_6)\sigma_{-1}(d_7). \tag{10}$$

Therefore, the complex dimension is 3. In other words, the discriminant classes d_1 (security) and d_6 (speed) have the largest dimension.

4.1. Calculating the Structure Vectors. As mentioned, the vector Q_q is a simplification basis to eliminate the additional effects in the set of equivalent simplexes. The maximum complex dimension for $\alpha_{\%5} = 40$ is 3. Therefore, the first structure vector based on the output is

$$\begin{aligned} &\text{Dimension} \quad 3 \quad 2 \quad 1 \quad 1, \\ &Q = (1 \ \ 1 \ \ 1 \ \ 1), \\ &\text{Dimension} \quad 3 \quad 2 \quad 1 \quad 0, \\ &P = (2 \ \ 3 \ \ 4 \ \ 7). \end{aligned} \tag{11}$$

The second structure vector P is

$$P = (2 \ \ 3 \ \ 4 \ \ 7). \tag{12}$$

4.2. Calculating the Obstruction Vector or Inflexibility. $\mathbf{Q}^*_{\mathbf{K}}$ represents the number of structural obstructions for simplex interactions in dimension k, which is as follows for $\alpha_{\%5} = 40$:

Table 7: The irregularity of the data matrix parameters for $\boldsymbol{\alpha = 7}$.

$\boldsymbol{\sigma}$	$\mathbf{q_i}$	$\sum \mathbf{q_i/\sigma_i}$	$\mathbf{q_{max}}$	$\mathbf{ecc'(\sigma) = 2\sum q_i/\sigma_i/q_{max}(q_{max}+1)}$
$\mathbf{d}_2, \mathbf{d}_5, \mathbf{d}_6, \mathbf{d}_8$	$q_i = 3, 2, 1$	3.64	3	0.61
$\mathbf{d}_9, \mathbf{d}_{10}, \mathbf{d}_{11}$	$q_i = 2, 1$	0.64	3	0.11
$\mathbf{d}_4, \mathbf{d}_7, \mathbf{d}_{12}$	$q_i = 1$	0.14	3	0.02

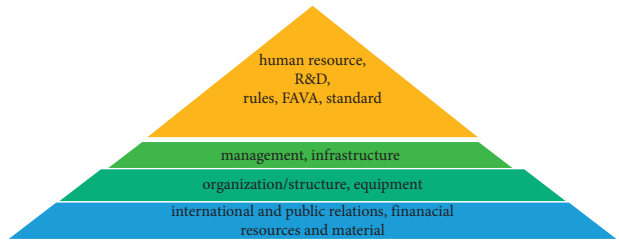


Figure 6: The ranking pyramid of the support components using Q-analysis.

Table 8: Ranking of the support components using Q-analysis.

Relationship of the support components with the design and implementation of the cryptography algorithms ($q = 0$) No relationship: α = 0%; complete relationship: α = 100%	
Human resource, R&D, rules, FAVA, and standard	$\alpha_{\%100} = 10$
Management and infrastructure	$\alpha_{\%90} = 9$
Organization/structure and equipment	$\alpha_{\%80} = 8$
International and public relations, financial resources, and material	$\alpha_{\%70} = 7$
Culture	$\alpha_{\%60} = 6$
All components	$\alpha_{\%60} = 5$

Table 9: Set of d_is and c_is; the objectives' indices.

Theoretical basis	C_2	Reliability	d_7	Scalability	d_5	Resources	d_3	Security	d_1
Implementation	C_3	Application	C_1	Speed	d_6	Flexibility	d_4	Simplicity	d_2
Evaluation	C_4								

Table 10: The incidence matrix of the design objectives of the cryptography algorithms with $\boldsymbol{\alpha = 40}$.

	c_1	c_2	c_3	c_4
d_1	1	1	1	1
d_2	0	0	0	0
d_3	1	0	1	1
d_4	0	0	1	1
d_5	0	0	0	0
d_6	1	1	1	1
d_7	0	0	0	0

$$Q^* = Q - I \longrightarrow Q^* = [2\ 1\ 1\ 3] - [1\ 1\ 1\ 1] \longrightarrow Q^* = [0\ 0\ 0\ 0]. \tag{13}$$

According to the obtained values, it can be concluded that there are no structural obstructions among the main design indices of the cryptography algorithms. That is, multiple objectives are considered simultaneously by the cryptography algorithm designers.

```
 3  -1  2  1  -1  3  -1
    -1  -1 -1  -1  -1  -1
        2  1  -1  2  -1
           1  -1  1  -1
              -1  -1  -1
                  3  -1
                     -1
```

q	Q	SETS
3	1	{X1,X6}
2	1	{X1,X3,X6}
1	1	{X1,X3,X4,X6}
0	1	{X1,X3,X4,X6}

Figure 7: The results of executing the model for design objectives' indices with $\alpha = 40$.

4.3. Calculating Irregularity. The results of applying the Chinese method (ecc'(σ)) for calculating irregularity for $\alpha_{\%5} = 40$ are shown in Table 11. The calculated irregularity value shows that indices d_1 (security) and d_6 (speed) have received more attention compared to other indices.

4.4. Calculating Complexity. Structural complexity for a sample with $\alpha_{\%5} = 40$ is

$$\begin{aligned} &\text{dimensions} \quad 3 \quad 2 \quad 1 \quad 0, \\ &Q = \left(Q_{\text{dim3}} \;\; Q_{\text{dim2}} \;\; Q_{\text{dim1}} \;\; Q_{\text{dim0}} \right), \\ &Q = \left(1 \;\; 1 \;\; 1 \;\; 1 \right), \\ &\psi(K) = 2\left[\frac{(1+2+3+4)}{(4*5)} \right] = 1. \end{aligned} \tag{14}$$

Since, in the above model, there was no obstruction among the components of the equivalent classes in any of the communication levels, it was expected that there is not a high complexity among objectives of the cryptography algorithm design; and, the complexity index of 1 verified this expectation. The system complexity for different alpha-cuts is shown in Figure 8. As can be seen, the system complexity is 1 for all significant values of alpha.

4.5. Prioritizing the Parameters of the Main Objectives of the Cryptography Algorithm Design Using Q-Analysis. According to the analysis results, the parameters can be classified into 5 levels. These levels are shown in Table 12 and Figure 9. Each level indicates a priority and importance of the group in the development of cryptography algorithms. To allocate proper resources, the components at the higher levels of the pyramid are of higher priority. As can be seen, three objectives of security, speed, and optimal usage of resources have the highest priority for the design of cryptography algorithms.

4.6. Executing the Model and Analyzing the Results for the Cryptography Algorithm Implementation Techniques. The interaction matrix between the 29 extracted techniques and the seven main objectives is shown in Figure 10. The purpose of this section is to rank the cryptography algorithm implementation techniques to achieve the goals of interest.

4.7. Calculating the Incidence Matrix and Executing the Model. The data matrix A is comprised of two sets. The set D represents the employed techniques, and the set C represents the seven main design objectives (Table 13):

$$\begin{aligned} D &= \{d_1, d_2, \ldots, d_{29}\}, \\ C &= \{c_1, c_2, \ldots, c_7\}. \end{aligned} \tag{15}$$

Some parts of the incidence matrix calculated from the data matrix A for $\alpha_{\%3}$ are shown in Table 14, and the results of the Q-analysis model for $\alpha_{\%3} = 24$ are shown in Figure 11.

4.8. Geometric Representation. In the sample with $\boldsymbol{\alpha}_{\%3} = 24$, the dis are

$$\begin{aligned} d_1 &= \{c_1\}, \\ d_2 &= \{\}, \\ d_3 &= \{\}, \\ d_4 &= \{\}, \\ d_5 &= \{\}, \\ d_6 &= \{\}, \\ d_7 &= \{\}, \\ d_8 &= \{\}, \\ d_9 &= \{\}, \\ d_{10} &= \{\}, \\ d_{11} &= \{c_1\}, \\ d_{12} &= \{\}, \\ d_{13} &= \{c_5\}, \\ d_{14} &= \{\}, \\ d_{15} &= \{c_1, c_4, c_5\}, \\ d_{16} &= \{\}, \\ d_{17} &= \{\}, \\ d_{18} &= \{\}, \\ d_{19} &= \{c_1\}, \\ d_{20} &= \{\}, \\ d_{21} &= \{\}, \\ d_{22} &= \{c_3\}, \\ d_{23} &= \{\}, \\ d_{24} &= \{\}, \\ d_{25} &= \{\}, \\ d_{26} &= \{\}, \\ d_{27} &= \{c_1\}, \\ d_{28} &= \{\}, \\ d_{29} &= \{\}. \end{aligned} \tag{16}$$

The simplexes of $\boldsymbol{\sigma}_{\mathbf{p}}(\mathbf{d}_{\mathbf{i}})$ are

Table 11: Irregularity of the cryptography objectives' parameters in the data matrix A for $\alpha_{\%5}$.

σ	q_i	$\sum q_i/\sigma_i$	$q_{\max}$	$\mathrm{ecc}'(\sigma) = 2\sum q_i/\sigma_i/q_{\max}(q_{\max}+1)$
d_1 and d_6	$q_i = 3, 2,$ and 1	4.33	3	0.72
d_3	$q_i = 2$ and 1	1.33	3	0.26
d_4	$q_i = 1$	0.33	3	0.06

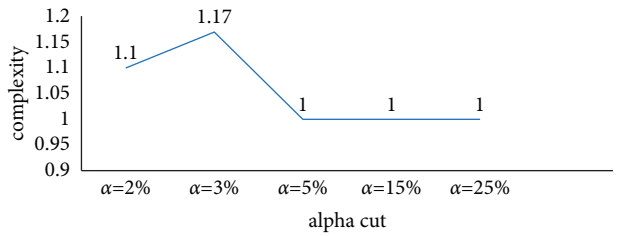


Figure 8: System complexity of the cryptography algorithm design objectives for different alpha-cuts.

$$
\begin{aligned}
&\sigma_0(d_1),\\
&\sigma_{-1}(d_2),\\
&\sigma_{-1}(d_3),\\
&\sigma_{-1}(d_4),\\
&\sigma_{-1}(d_5),\\
&\sigma_{-1}(d_6),\\
&\sigma_{-1}(d_7),\\
&\sigma_{-1}(d_8),\\
&\sigma_{-1}(d_9),\\
&\sigma_{-1}(d_{10}),\\
&\sigma_0(d_{11}),\\
&\sigma_{-1}(d_{12}),\\
&\sigma_0(d_{13}),\\
&\sigma_{-1}(d_{14}),\\
&\sigma_2(d_{15}),\\
&\sigma_{-1}(d_{16}),\\
&\sigma_{-1}(d_{17}),\\
&\sigma_{-1}(d_{18}),\\
&\sigma_0(d_{19}),\\
&\sigma_{-1}(d_{20}),\\
&\sigma_{-1}(d_{21}),\\
&\sigma_0(d_{22}),\\
&\sigma_{-1}(d_{23}),\\
&\sigma_{-1}(d_{24}),\\
&\sigma_{-1}(d_{25}),\\
&\sigma_{-1}(d_{26}),\\
&\sigma_0(d_{27}),\\
&\sigma_{-1}(d_{28}),\\
&\sigma_{-1}(d_{29}).
\end{aligned}
\tag{17}
$$

Therefore, the complex dimension is 2. In other words, the discriminant class d_{15} (basic sciences) has the largest dimension.

4.9. Calculating the Dimensions and the Q-Link. In the alpha defined for cryptography algorithm implementation techniques to achieve the defined goals, the obtained q-link shows that these techniques are relatively independent although there is a weak relationship between some techniques. Q-link in the samples with $\alpha_{\%3} = 24$ between each two subsequent d_i is as follows:

$$
\begin{aligned}
&\sigma_0(d_1), \sigma_{-1}(d_2) \longrightarrow -1,\\
&\sigma_{-1}(d_2), \sigma_{-1}(d_3) \longrightarrow -1,\\
&\sigma_0(d_3), \sigma_{-1}(d_4) \longrightarrow -1,\\
&\sigma_{-1}(d_4), \sigma_{-1}(d_5) \longrightarrow -1,\\
&\sigma_{-1}(d_5), \sigma_{-1}(d_6) \longrightarrow -1,\\
&\sigma_{-1}(d_6), \sigma_{-1}(d_7) \longrightarrow -1,\\
&\sigma_{-1}(d_7), \sigma_{-1}(d_8) \longrightarrow -1,\\
&\sigma_{-1}(d_8), \sigma_{-1}(d_9) \longrightarrow -1,\\
&\sigma_{-1}(d_9), \sigma_{-1}(d_{10}) \longrightarrow -1,\\
&\sigma_{-1}(d_{10}), \sigma_0(d_{11}) \longrightarrow -1,\\
&\sigma_0(d_{11}), \sigma_{-1}(d_{12}) \longrightarrow -1,\\
&\sigma_{-1}(d_{12}), \sigma_0(d_{13}) \longrightarrow -1,\\
&\sigma_0(d_{13}), \sigma_{-1}(d_{14}) \longrightarrow -1,\\
&\sigma_{-1}(d_{14}), \sigma_2(d_{15}) \longrightarrow -1,\\
&\sigma_2(d_{15}), \sigma_{-1}(d_{16}) \longrightarrow -1,\\
&\sigma_{-1}(d_{16}), \sigma_{-1}(d_{17}) \longrightarrow -1,\\
&\sigma_{-1}(d_{17}), \sigma_{-1}(d_{18}) \longrightarrow -1,\\
&\sigma_{-1}(d_{18}), \sigma_0(d_{19}) \longrightarrow -1,\\
&\sigma_0(d_{19}), \sigma_{-1}(d_{20}) \longrightarrow -1,\\
&\sigma_{-1}(d_{20}), \sigma_{-1}(d_{21}) \longrightarrow -1,\\
&\sigma_{-1}(d_{21}), \sigma_0(d_{22}) \longrightarrow -1,\\
&\sigma_0(d_{22}), \sigma_{-1}(d_{23}) \longrightarrow -1,\\
&\sigma_{-1}(d_{23}), \sigma_{-1}(d_{24}) \longrightarrow -1,\\
&\sigma_{-1}(d_{24}), \sigma_{-1}(d_{25}) \longrightarrow -1,\\
&\sigma_{-1}(d_{25}), \sigma_{-1}(d_{26}) \longrightarrow -1,\\
&\sigma_{-1}(d_{26}), \sigma_0(d_{27}) \longrightarrow -1,\\
&\sigma_0(d_{27}), \sigma_{-1}(d_{28}) \longrightarrow -1,\\
&\sigma_{-1}(d_{28}), \sigma_{-1}(d_{29}) \longrightarrow -1.
\end{aligned}
\tag{18}
$$

Table 12: Ranking the objectives' parameters considering different cuts in Q-analysis.

The equivalence class for the parameters with minimum relation level ($q = 0$)	α
Security	$\alpha_{\%50} = 244$
Speed	$\alpha_{\%30} = 244$
Resources	$\alpha_{\%15} = 122$
Flexibility	$\alpha_{\%5} = 40$
Simplicity-scalability	$\alpha_{\%2} = 16$
(All parameters)	?

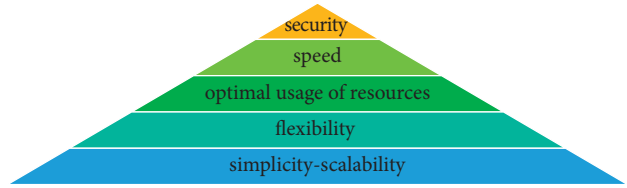


Figure 9: Prioritizing the design and implementation objectives of the cryptography algorithms using Q-analysis.

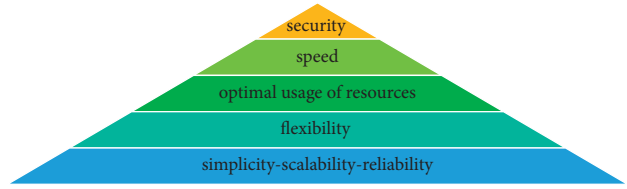


Figure 10: Prioritizing the design and implementation objectives of the cryptography algorithms using Q-analysis.

Table 13: Set of dis and cis; the techniques employed for implementing cryptography algorithms.

Flexibility	C_4	Fuzzy logic	d_{25}	Storage space	d_{17}	Cycle	d_9	Avalanche effect	d_1
Scalability	C_5	Software	d_{26}	Clustering	d_{18}	Multiple step	d_{10}	Digital signature	d_2
Speed	C_6	Steganography	d_{27}	Key	d_{19}	Hybrid method	d_{11}	Block size	d_3
Reliability	C_7	Music harmony	d_{28}	Graph	d_{20}	Hardware	d_{12}	Image sharing	d_4
		Artificial intelligence	d_{29}	Characteristic oriented	d_{21}	Hardware	d_{13}	Parallel processing	d_5
		Security	C_1	Energy consumption	d_{22}	Occupation area	d_{14}	Threshold technique	d_6
		Simplicity	C_2	Bandwidth consumption	d_{23}	Basic science	d_{15}	Data mining	d_7
		Resources	C_3	Memory consumption	d_{24}	Compression	d_{16}	Binary tree	d_8

Table 14: The incidence matrix of the cryptography algorithm implementation techniques with $\alpha = 24$.

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
d_1	1	0	0	0	0	0	0
d_2	0	0	0	0	0	0	0
d_3	0	0	0	0	0	0	0
d_4	0	0	0	0	0	0	0
d_5	0	0	0	0	0	0	0
d_6	0	0	0	0	0	0	0
d_7	0	0	0	0	0	0	0
d_8	0	0	0	0	0	0	0
d_9	0	0	0	0	0	0	0
d_{10}	0	0	0	0	0	0	0
d_{11}	1	0	0	0	0	0	0
d_{12}	0	0	0	0	0	0	0

4.10. Calculating the Structure Vectors. Therefore, the first structure vector based on the software output is

$$Q = (1 \quad 1 \quad 2). \tag{19}$$

The second structure vector P is

$$P = (1 \quad 1 \quad 29). \tag{20}$$

Since the large values of P of higher dimensions demonstrate more links, the second structure vector calculated for the alpha of interest shows that the relationship between the cryptography algorithm implementation techniques is minimum.

4.11. Calculating the Obstruction or Inflexibility Vector. The obstruction vector ($\mathbf{Q}^*$) for a sample with $\alpha_{\%3} = 24$ is

q	Q	SETS
2	1	{X15}
1	1	{X15}
o	2	{X1,X11,X15,X19,X27,X13} & {X22}

Figure 11: Results of executing the model for cryptography algorithm implementation techniques with $\alpha_{\%3} = 24$.

$$Q^* = Q - I \longrightarrow Q^* = [1\ \ 1\ \ 2] - [1\ \ 1\ \ 1] \longrightarrow Q^* = [0\ \ 0\ \ 1]. \tag{21}$$

The value of $\mathbf{Q}^*_{\mathbf{K}}$ represents the number of structural limitations or obstructions for the cryptography techniques' interaction at dimension k. As can be seen from the calculated obstruction vector, there is a significant relationship at some levels.

4.12. Calculating Irregularity. Irregularity is the integration degree of a cryptography method in the total complex. Measuring irregularity (ecc') for $\alpha_{\%3} = 24$ is shown in Table 15. As can be seen, irregularity for the discriminant class d_{15} indicates that this technique is isolated from other techniques.

4.13. Calculating Complexity. Structure complexity for a sample with $\alpha_{\%3} = 24$ is

$$Q = (1\ \ 1\ \ 2),$$
$$\psi(K) = 2\left[\frac{(2+2+3)}{(3*4)}\right] = 1.17. \tag{22}$$

The system complexity for different values of alpha is shown in Figure 12. As can be seen, the total complexity of the system for large values alpha tends to stability.

4.14. Ranking the Cryptography Algorithm Implementation Techniques. These techniques can be classified into 5 levels using the analysis of the obtained results. These levels can be seen in Table 16 and Figure 13. The link power of the factors in one group is specified with alpha. Each level describes priority and importance of the group in the development of cryptography algorithms. Accordingly, six methods of basic science, key management, using hardware methods, using steganography, Avalanche effect, and using hybrid method

Table 15: Irregularity of the cryptography algorithm implementation techniques for $\alpha_{\%3} = 24$.

σ	q_i	$\sum q_i/\sigma_i$	q_{max}	$ecc'(\sigma) = 2\sum q_i/\sigma_i/q_{max}(q_{max}+1)$
d_{15}	$q_i = 2, 1$	3	2	1
-	$q_i = 1$	1	2	0.33

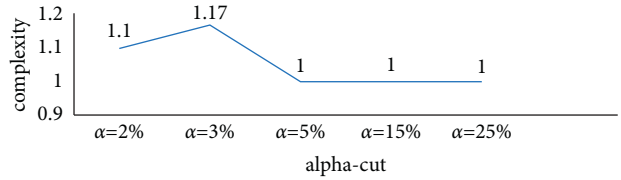


Figure 12: System complexity of cryptography algorithm implementation techniques for different alpha-cuts.

Table 16: Ranking the implementation techniques of the cryptography algorithms considering various alpha-cuts using Q-analysis.

Equivalence classes for the parameters with minimum relationship level ($q = 0$)	α
Basic sciences	$\alpha_{\%20} = 163$
Key	$\alpha_{\%10} = 81$
Hardware and steganography	$\alpha_{\%5} = 40$
Avalanche effect and hybrid method	$\alpha_{\%3} = 24$
{Artificial intelligence, software}, {occupation area, energy consumption, memory consumption}	$\alpha_{\%2} = 16$

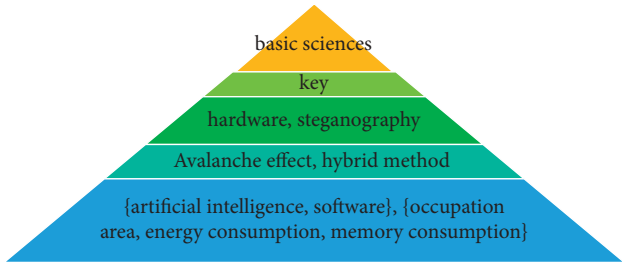


Figure 13: Prioritizing the implementation techniques of the cryptography algorithms using Q-analysis.

are the most important cryptography algorithm implementation techniques to achieve the main objectives.

5. Conclusion

In this study, the design and implementation model of the cryptography algorithms is designed in three levels and its effective components are extracted. To organize the components, in addition to elimination of the additional components through the measure reduction algorithm, a proper classification is applied to examine the mutual effects. After classification, three $13 * 4$, $7 * 4$, and $29 * 7$ matrices are constituted, and the model is implemented on these three matrices. To implement the designed model, Q-analysis is used. Accordingly, for support indices, five indices of high priority include human resources, research and development, management, organization/structure, and equipment. For the algorithm design objectives index, five high priority indices include security, speed, optimal usage of resources, and simplicity. For the indices related to implementation techniques of the cryptography algorithms, the most applied techniques for achieving the determined objectives include using basic science, hardware and software methods, using key management, hybrid method, steganography, Avalanche effect, and using artificial intelligence. In the future work, we will consider the following cases.

A plan should be formulated according to the priorities. According to the outputs of this study and the presented priorities, the following topics can be investigated in future studies:

(1) Presenting a comprehensive model for generating various cryptography algorithms based on the priorities of interest

(2) Examining and formulating a model about the role of basic science for design and implementation of cryptography algorithms considering the significant role of "basic science" in implementation of the cryptography algorithms and about the role of schools and universities

(3) Presenting a model for design and implementation of cryptography algorithms in the IoT infrastructure with optimal resource usage

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was partially supported by Iran National Science Foundation (INSF) under contract No. 96/53979.

References

[1] D. Nilesh and M. Nagle, "The new cryptography algorithm with high throughput," in *Proceedings of the 2014 International Conference on Computer Communication and Informatics*, pp. 3–7, Coimbatore, India, January 2014.

[2] S. Ravi, P. Kocher, R. Lee, G. McGraw, and A. Raghunathan, "Security as a new dimension in embedded system design," in *Proceedings of the 41st annual conference on Design automation-DAC '04*, pp. 753–760, San Diego, CA, USA, March 2004.

[3] P. Kuppuswamy and S. Q. Y. A. Khalidi, "Hybrid encryption/decryption technique using new public key and symmetric key algorithm," *International Journal of Information and Computer Security*, vol. 6, no. 4, pp. 372–382, 2014.

[4] S. Bhat and V. Kapoor, "Secure and efficient data privacy, authentication and integrity schemes using hybrid cryptography," *International Conference on Advanced Computing Networking and Informatics*, vol. 870, pp. 279–285, 2019.

[5] D. Ott and C. Peikert, "And other workshop participants, "identifying research challenges in post quantum cryptography migration and cryptographic agility," 2019, http://arxiv.org/abs/1909.07353.

[6] P. Manickam, K. Shankar, E. Perumal, M. Ilayaraja, and K. Sathesh Kumar, "Secure data transmission through reliable vehicles in VANET using optimal lightweight cryptography," in *Advanced Sciences and Technologies for Security Applications*, pp. 193–204, Springer, New York, NY, USA, 2019.

[7] S. Prakash and A. Rajput, "Hybrid cryptography for secure data communication in wireless sensor networks," *Advances in Intelligent Systems and Computing*, vol. 696, pp. 589–599, 2018.

[8] A. A. Soofi, I. Riaz, and U. Rasheed, "An enhanced vigenere cipher for data security," *International Journal of Scientific & Technology Research*, vol. 4, no. 8, pp. 141–145, 2015.

[9] N. Shamsuddin, M. Syafiqah, and S. Ali Pitchay, "Implementing location-based cryptography on mobile application design to secure data in cloud storage," *Journal of Physics: Conference Series*, IOP Publishing, vol. 1551, no. 1, 2020.

[10] L. Jiao, H. Yonglin, and D. Feng, "Stream cipher designs: a review," *Science China Information Sciences*, vol. 63, no. 3, pp. 1–25, 2020.

[11] NIST, "NIST cryptographic standards and guidelines development process," *Inside NIST*, vol. 27, 2016.

[12] G. Alagic, J. Alperin-Sheriff, D. Apon et al., *Status Report on the First Round of the NIST Post-Quantum Cryptography Standardization Process*, National Institute of Standards and Technology, Gaithersburg, MD, USA, pp. 1–27, 2019.

[13] I. Damaj and S. Kasbah, "An analysis framework for hardware and software implementations with applications from cryptography," *Computers & Electrical Engineering*, vol. 69, pp. 572–584, 2018.

[14] M. S. Croock, Z. A. Hassan, and S. D. Khudhur, "Adaptive key generation algorithm based on software engineering methodology," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 1, p. 589, 2021.

[15] K. Sha, T. A. Yang, W. Wei, and S. Davari, "A survey of edge computing-based designs for iot security," *Digital Communications and Networks*, vol. 6, no. 2, pp. 195–202, 2020.

[16] M. Yahuza, M. Y. I. B. Idris, A. W. B. A. Wahab et al., "Systematic review on security and privacy requirements in edge computing: state of the art and future research opportunities," *IEEE Access*, vol. 8, pp. 76541–76567, 2020.

[17] A. Bhowmik, J. Dey, and S. Karforma, *A Novel Double Tier Cryptographic System (nDTCS) to Reinforce Patients' Privacy in Contemporary COVID-19 Telemedicine*, Springer, New York, NY, USA, 2021.

[18] N. Gupta, *User Integrity Protection Security Model for Enhanced Telemedicine for Healthcare Networks*, 2020, https://osf.io/69jdq/.

[19] J. Dey, A. Bhowmik, A. Sarkar, S. Karforma, and B. Chowdhury, *Cryptographic Engineering on COVID-19 Telemedicine: An Intelligent Transmission through Recurrent Relation Based Session Key*, Springer, New York, NY, USA, 2021.

[20] V. Hosseinian, H. Ayatollahi, H. Haghani, and E. Mehraeen, "Requirements of information security in a telemedicine network: review of IT managers' opinion," *Journal of Paramedical Sciences & Rehabilitation*, vol. 4, no. 2, pp. 31–40, 2015.

[21] D. Karaoğlan Altop, A. Levi, and V. Tuzcu, "Deriving cryptographic keys from physiological signals," *Pervasive and Mobile Computing*, vol. 39, pp. 65–79, 2017.

[22] O. Uzunkol and M. S. Kiraz, "Still wrong use of pairings in cryptography," *Applied Mathematics and Computation*, vol. 333, pp. 467–479, 2018.

[23] S. Picek, D. Jakobovic, J. F. Miller, L. Batina, and M. Cupic, "Cryptographic Boolean functions: one output, many design criteria," *Applied Soft Computing*, vol. 40, pp. 635–653, 2016.

[24] R. Saha and G. Geetha, "Symmetric random function generator (SRFG): a novel cryptographic primitive for designing fast and robust algorithms," *Chaos, Solitons & Fractals*, vol. 104, pp. 371–377, 2017.

[25] A. Achuthshankar and A. Achuthshankar, "A novel symmetric cryptography algorithm for fast and secure encryption," in *Proceedings of the 2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO)*, pp. 4–9, Coimbatore, India, January 2015.

[26] D. Ott and C. Peikert, "Identifying research challenges in post quantum cryptography migration and cryptographic agility," 2019, http://arxiv.org/abs/1909.07353.

[27] W. Liu, B. Ying, H. Yang, and H. Wang, "Accurate modeling for predicting cryptography overheads on wireless sensor nodes,"vol. 2, pp. 997–1001, in *Proceedings of the 2019 11th International Conference on Advanced Communication Technology*, vol. 2, pp. 997–1001, IEEE, Bangalore, India, February 2019.

[28] S. Feizi, A. Ahmadi, and A. Nemati, "A hardware implementation of simon cryptography algorithm," in *Proceedings of the 2014 4th International Conference on Computer and Knowledge Engineering (ICCKE)*, pp. 245–250, IEEE, Mashhad, Iran, 2014.

[29] E. Thambiraja, G. Ramesh, and R. Umarani, "A survey on various most common encryption techniques," 2012, http://www.ijarcsse.com.

[30] A. Gupta and N. K. Walia, "Cryptography Algorithms : a review," *International Journal of Engineering Research and Development*, vol. 2, no. 2, pp. 1667–1672, 2014.

[31] R. Bhanot and R. Hans, "A review and comparative analysis of various encryption algorithms," *International Journal of Security and Its Applications*, vol. 9, no. 4, pp. 289–306, 2015.

[32] S. Chandra, S. Paira, S. S. Alam, and G. Sanyal, "A comparative survey of symmetric and asymmetric key

cryptography," in *Proceedings of the 2014 International Conference on Electronics, Communication and Computational Engineering (ICECCE)*, pp. 83–93, Hosur, India, November 2014.

[33] F. Maqsood, M. Ahmed, M. Mumtaz, and M. Ali, "Cryptography: a comparative analysis for modern techniques," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 6, p. 6, 2017.

[34] S. B. Sadkhan and A. O. Salman, "A survey on lightweight-cryptography status and future challenges," in *Proceedings of the International Conference on Advances in Sustainable Engineering and Applications*, pp. 105–108, Wasit, Iraq, June 2018.

[35] N. V. Thoai, "Criteria and dimension reduction of linear multiple criteria optimization problems," *Journal of Global Optimization*, vol. 52, no. 3, pp. 499–508, 2012.

[36] C. E. Shannon, "Communication theory of secrecy systems. 1945," *MD Computing*, vol. 15, no. 1, pp. 57–64, 1998.

[37] L. J. Fennelly, M. Beaudry, and M. A. Perry: Security in 2025.

[38] ITU, Global Cybersecurity Index 2018, 2019.

[39] ITU, Global Cybersecurity Index (GCI) 2017, 2017.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 7552185, 16 pages
https://doi.org/10.1155/2021/7552185

Research Article

Dual-Path Attention Compensation U-Net for Stroke Lesion Segmentation

Haisheng Hui, Xueying Zhang, Zelin Wu, and Fenlian Li

College of Information and Computer, Taiyuan University of Technology, Taiyuan 030024, China

Correspondence should be addressed to Xueying Zhang; tyzhangxy@163.com

Received 9 July 2021; Accepted 19 August 2021; Published 31 August 2021

Academic Editor: Suresh Manic

For the segmentation task of stroke lesions, using the attention U-Net model based on the self-attention mechanism can suppress irrelevant regions in an input image while highlighting salient features useful for specific tasks. However, when the lesion is small and the lesion contour is blurred, attention U-Net may generate wrong attention coefficient maps, leading to incorrect segmentation results. To cope with this issue, we propose a dual-path attention compensation U-Net (DPAC-UNet) network, which consists of a primary network and auxiliary path network. Both networks are attention U-Net models and identical in structure. The primary path network is the core network that performs accurate lesion segmentation and outputting of the final segmentation result. The auxiliary path network generates auxiliary attention compensation coefficients and sends them to the primary path network to compensate for and correct possible attention coefficient errors. To realize the compensation mechanism of DPAC-UNet, we propose a weighted binary cross-entropy Tversky (WBCE-Tversky) loss to train the primary path network to achieve accurate segmentation and propose another compound loss function called tolerance loss to train the auxiliary path network to generate auxiliary compensation attention coefficient maps with expanded coverage area to perform compensate operations. We conducted segmentation experiments using the 239 MRI scans of the anatomical tracings of lesions after stroke (ATLAS) dataset to evaluate the performance and effectiveness of our method. The experimental results show that the DSC score of the proposed DPAC-UNet network is 6% higher than the single-path attention U-Net. It is also higher than the existing segmentation methods of the related literature. Therefore, our method demonstrates powerful abilities in the application of stroke lesion segmentation.

1. Introduction

Recent global statistics on the incidence of stroke cases demonstrate that there are up to 10.3 million new cases annually [1]. Stroke has become one of the top three lethal diseases, besides chronic diseases. When a stroke occurs, accurate diagnosis of the severity of the stroke and timely thrombolytic therapy can effectively improve blood supply in the ischemic area and significantly reduce the risk of disability or even death. Therefore, it is clinically significant to quickly and accurately locate and segment the stroke lesions [2]. Since manual segmentation relies on the doctor's professional experience and medical skills, individual subjectivity can reduce segmentation accuracy. Furthermore, manual segmentation of the stroke lesion is time-consuming. It may take a skilled tracer several hours to complete accurate labeling and rechecking of a single large complex lesion on magnetic resonance imaging (MRI) [3].

This situation has changed after the advent of convolutional neural network (CNN) [4] and its continuously evolving network structures, such as fully convolutional network (FCN) [5] and SegNet [6], which have achieved success in the field of image segmentation, especially medical image segmentation [7]. However, CNN-based segmentation networks require a large amount of labeled medical data for training, which is limited by the high cost of acquiring and accurate labeling [8]. The multilevel U-shaped network (U-Net) [9] based on CNN, consisting of the contraction and expansion paths, mitigates the problem of requiring huge amounts of labeled data. The U-Net network structure and its improved network structure, such as the attention U-Net

[10], U-Net++ [11], and R2U-Net [12], have been applied successfully in medical segmentation tasks, such as skin cancer [13], brain tumor [14], colorectal tumor [15], liver [16], colon histology [17], kidney [18], and vascular borders [19]. The U-Net network has thousands of feature channels, especially the standard U-Net model with a five-level structure with enormous parameters to be trained. During the training process, the contraction path (encoder) and expansion path (decoder) need to repeatedly extract deep-scale features. The deep-scale features of standard U-Net are considered abstract and low-resolution features, which increase the training difficulty and make the training unstable and inadequate.

To reduce the training difficulty caused by repeated extraction of deep-scale features and improve segmentation accuracy, many researchers employed a two-step method to locate the lesion and segment the target area [20, 21]. However, these methods introduce additional positioning operations and cannot achieve end-to-end training. Schlemper et al. introduced a self-attention mechanism and proposed an attention U-Net with an attention gate (AG) [10] to avoid additional operations. The self-attention mechanism reduces the dependence on external information obtained from additional steps by utilizing the correlation coefficient of feature signals from different scales. This mechanism captures the internal correlation of features and focuses attention on the target area. The attention U-Net uses AG to generate a 2D attention coefficient map to suppress irrelevant regions in an input image while highlighting salient features useful for specific tasks. The AG module can be integrated into the standard U-Net model for end-to-end learning without additional pretraining steps. Compared with the standard U-Net training parameters, the number of training parameters slightly increased with additional computation of AG operations. The use of the built-in self-attention module eliminates the use of additional target location operations. It achieves the goal of reducing training difficulty, improving training efficiency, and improving model segmentation performance.

However, the self-attention mechanism based on correlation operation has some deficiencies. The attention coefficient α for constraining the area of interest is generated by the current-scale feature signal x and the rougher-scale feature signal g derived from x, leading to a potential risk of the segmentation network using the self-attention mechanism. It implies that a small lesion with a nondistinct lesion feature may cause the current level feature signal x to learn the lesion feature inadequately. Consequently, the deviation of the attention area from the lesion area due to the wrong or insufficient attention coefficient learning leads to incorrect segmentation results.

To solve the problem of the attention area deviating from the lesion area, we proposed a dual-path attention compensation U-Net (DPAC-UNet) network, which is composed of the primary path network (primary network) and auxiliary path network (auxiliary network). Both networks are all attention U-Net segmentation models based on the self-attention mechanism with an identical structure. The primary network is the core part of DPAC-UNet, which performs lesion segmentation and outputs the final segmentation result. The auxiliary network is used to generate an auxiliary attention compensation coefficient map sent to the primary network to compensate for possible attention coefficient learning errors. The auxiliary network realizes its compensation ability by focusing on a larger area than the actual lesion area, which increases the coverage of the attention coefficient map generated by the auxiliary network. The attention coefficient map with a larger attention area is defined as a tolerant attention coefficient map, which is used as an auxiliary compensation attention coefficient to compensate for possible errors in the primary network attention coefficient map. To study our lesion segmentation network, we use the ATLAS dataset [3], consisting of 239 T1-weighted subacute and chromic stroke MRI scans released in 2018.

The main contributions of this article are summarized as follows:

(1) We proposed a DPAC-UNet that uses the auxiliary network to generate an attention coefficient map with a larger area to compensate for the possible defect of the primary network's attention coefficient map.

(2) We proposed the WBCE-Tversky loss and tolerance loss to train the primary and auxiliary networks of the DPAC-UNet to realize their effects on the entire network, respectively, and explore the optimal hyperparameter configurations of the two proposed loss functions.

The remainder of this work is organized as follows: In Section 2.1, we describe the network structure of the DPAC-UNet and how to use the auxiliary network to compensate for attention in the primary network. Section 2.2 proposes two compound loss functions, the WBCE-Tversky loss and the tolerance loss. In this section, we also conducted experiments to discuss the effect of different hyperparameter values of the loss functions on the performance of the segmentation task. Finally, the steps to select the optimal hyperparameter configuration of the two proposed loss functions are listed. In Section 3, we train the DPAC-UNet by the WBCE-Tversky and the tolerance loss functions with the optimal hyperparameter configurations. In this section, a visualization example is also presented to demonstrate the effectiveness of the DPAC-UNet network further. We also discussed the time consumption of the primary and auxiliary networks of the DPAC-UNet, and we also tried to execute the auxiliary network's compensation mechanism for other segmentation models with self-attention mechanisms.

2. Materials and Methods

2.1. DPAC-UNet. The attention U-Net introduces several attention gates (AG) to generate attention coefficient maps that suppress irrelevant regions in an input image while highlighting salient features useful to improve segmentation performance without introducing additional positioning operations. However, it sometimes makes mistakes. A small lesion with indistinct lesion features is difficult to distinguish from the surrounding healthy tissues, leading to the current

scale feature signal x of a certain layer not learning the lesion feature well. As a result, the attention coefficient generated using x and its derived rougher feature g will deviate from the lesion area. Therefore, the wrong attention coefficient results in the AG outputting the wrong feature signal, which affects the segmentation results. Thus, if the attention U-Net finds the correct lesion in the AG module, it will emphasize the relevant area and suppress the unrelated area to improve the segmentation performance. Conversely, if the lesion location is not found in the AG or is wrong, it will result in diametrically opposite effects and degrade the segmentation performance. To cope with the previously mentioned issues, using the attention U-Net as the basic segmentation model, we propose the DPAC-UNet network.

2.1.1. Overview of the Structure. The schematic of DPAC-UNet is presented in Figure 1. We used two identical attention U-Net models as the primary and auxiliary network segmentation models, which correspond to the upper and lower half of Figure 1, respectively. The WBCE-Tversky loss function trains the primary network for accurate segmentation. The auxiliary network is trained by the tolerance loss to generate a tolerant auxiliary compensation attention coefficient that compensates for the defect of the attention coefficient map of the primary network. The details of the two loss functions are described in Section 2.2. As presented in Figure 1, the auxiliary network compensates for the auxiliary compensation attention coefficient to the primary network through the vertical dark red arrow line from the AG marked (II) to the AG marked (I), in order to perform the compensation operation. We just selected the second-level AG of the primary and auxiliary networks for additive compensation operation. This is because the resolution of the attention coefficient maps generated by the two bottom AGs (13×11 and 26×22) is too low. The difference between the attention maps of the two networks on this resolution scale is larger due to the difference of one pixel. When the level is deeper, the receptive field affected by a single pixel is very large. Consequently, the compensation operation at this scale by the auxiliary network has a significant impact on the primary network, and the compensating operation generates a significant attention fluctuation. Furthermore, the first-level AG, which is close to the uppermost layer's output, does not perform auxiliary attention compensation operation because the feature map here is too close to the output and affects the segmentation result. In summary, we only selected the second-level AG to implement the compensation operation in order to effectively compensate for the defective attention coefficient map of the primary network and ensure that it does not directly affect the accuracy of lesion segmentation of the primary network.

Figure 2 presents the AG schematic of the primary and auxiliary networks at the second level. The AG of the first, third, and fourth levels are shown in Figure 1, which are not involved in auxiliary attention coefficient compensation operation and are identical in structure to the AG in the literature [10]. The AG marked as (II) in the lower half of Figure 1 is the second-level AG in the auxiliary network's attention U-Net, and its detailed structure is shown in Figure 2(a). In Figure 2(a), ① and ② are the input of the auxiliary network AG, ④ is the output of the current level for skip connection (SC), where l is the level number of current AG (in this case $l = 2$), and feature signals x_i^l and g_i^l correspond to the inputs labeled ① and ②. The feature signals $g_i^l \in R^{F_g}$ and $x_i^l \in R^{F_x}$ are sent to the AG block to generate the attention coefficient α^l using the additive attention generation operation in order to determine the area to focus, where i is the pixel number,F^x is the number of feature channels of input feature signal x^l at the current level, andF^g is the number of feature channels of input feature signal g^l at the rougher level. When the additive attention coefficient map α^l is generated using x_i^l and g_i^l, the feature signal x_i^l is multiplied by α^l and used as the output of the AG gate and sent to the decoding path through the SC at the current level. The additive attention coefficient α^l marked as ③ is the auxiliary compensation attention coefficient map and sent to the AG marked as (I) at the same level and in the same position of the primary network in the upper half of Figure 1. The equations for generating the attention coefficient of the auxiliary network are as follows:

$$q_{att}^l = W_\psi^T\left(\sigma_1\left(W_x^T x_i^l + W_g^T g_i^l + b_g\right)\right) + b_\psi, \tag{1}$$

$$\alpha_i^l = \sigma_2\left(q_{\text{att}}^l\left(x_i^l, g_i^l; \Theta_{\text{att}}\right)\right), \tag{2}$$

$$\left(\alpha_i^l\right)_{rs} = \text{resample}\left(\alpha_i^l\right), \tag{3}$$

$$\widehat{x_i^l} = x_i^l \cdot \left(\alpha_i^l\right)_{rs}. \tag{4}$$

As presented in Figure 2(a), considering the inconsistent spatial resolution and feature channel dimensions of feature g_i^l and x_i^l, we also need to use the upsampling operation to change the spatial resolution of the signal g_i^l to make it consistent with x_i^l. Moreover, we need to use the linear transformation $W_g \in R^{F_g \times F_{\text{int}}}$ and $W_x \in R^{F_x \times F_{\text{int}}}$ to make the number of feature channels of these two signals the same, where $b_g \in R^{F_{\text{int}}}$ and $b_\psi \in R$ denote the biases of the two linear transformations. In (1), σ_1 is the ReLU activation function, and the output of this activation function is linearly transformed by $W_\psi^T \in R^{1 \times F_{\text{int}}}$ that forms an attention coefficient matrix with only one feature channel. In (2), the sigmoid activation function σ_2 converts the attention coefficient matrix into a gridded attention coefficient map α_i^l to act on x_i^l. Resample α_i^l, and then, multiply the resampled result by x_i^l to generate the AG output feature signal $\widehat{x}_i^l$. Figure 2(b) presents the block diagram of the AG marked as (I) in the upper half of Figure 1, where the auxiliary compensation attention coefficient map compensates for the primary network. The structure and equations of the signal operation process are almost consistent with the auxiliary network, as presented in Figure 2(a). The difference is that when generating the final additive fused attention coefficient map, the auxiliary compensation attention coefficient map generated by the auxiliary network AG is marked as ③, and perform additive fusion together with the original attention coefficient map generated by the primary network AG

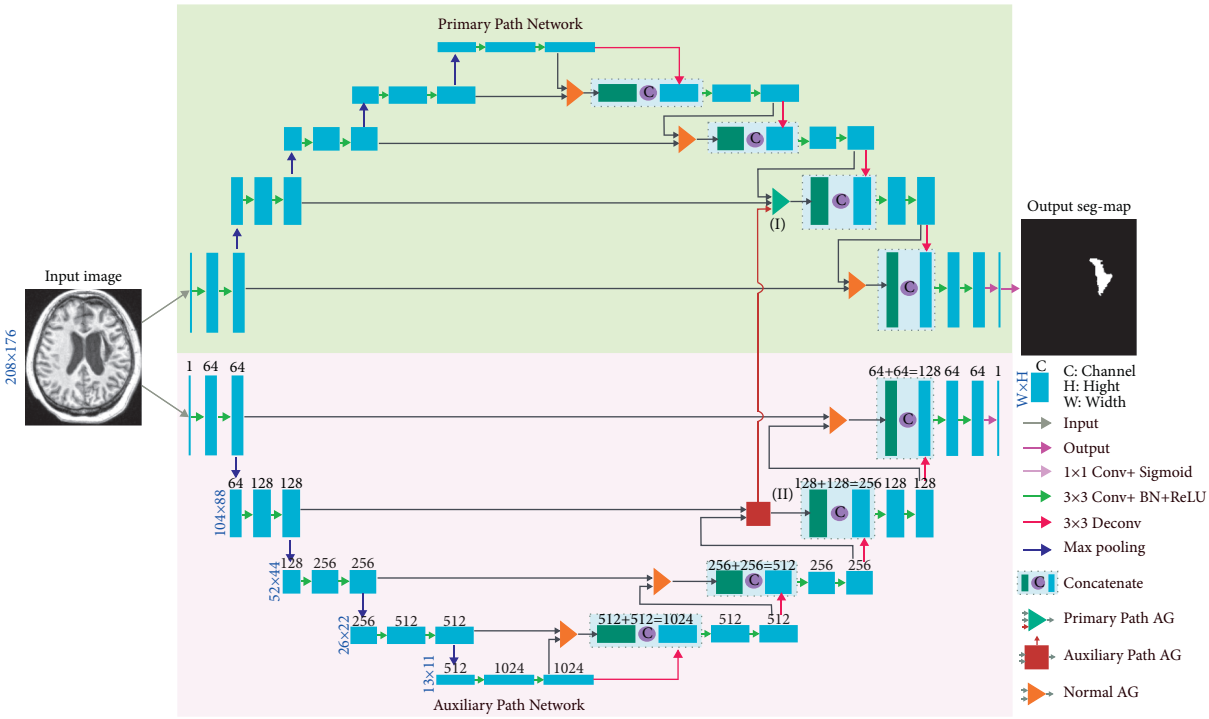
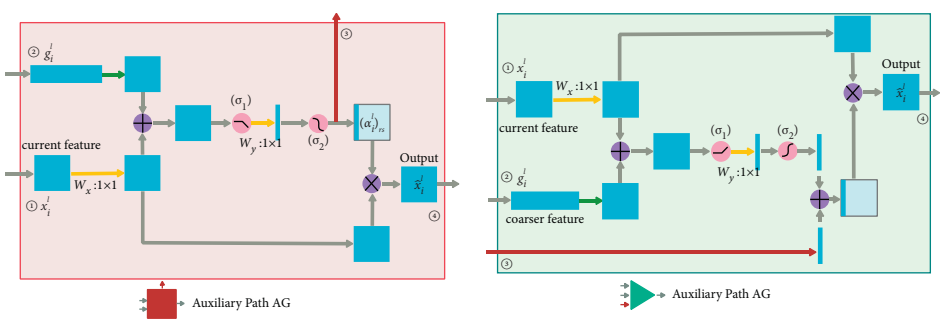


Figure 1: Schematic of DPAC-UNet.



(a) (b)

Figure	Description	Input	Output
	Conv 1×1 (W :1×1) and up-sampling	(26, 22, 512)	(52, 44, 256)
	Auxiliary path AG signal	(52, 44, 1)	(52, 44, 1)
	Current feature Conv 1×1	(52, 44, 256)	(52, 44, 256)
	Attn ceofficient Conv 1×1	(52, 44, 128)	(52, 44, 1)
	Element-wise addition	(52, 44, 256) (52, 44, 256)	(52, 44, 256)
	Element-wise production	(52, 44, 256) (52, 44, 256)	(52, 44, 256)
	ReLU activation	(52, 44, 256)	(52, 44, 256)
	Sigmoid activation	(52, 44, 1)	(52, 44, 1)
	Element-wise repeat to get multi-channel	(52, 44, 1)	(52, 44, 256)

(c)

Figure 2: (a) Schematic of the AG structure of the auxiliary network, (b) schematic of the AG structure of the primary network, and (c) the definition of various operation symbols and dimensional changes of input and output feature signals.

generated by inputs ① and ②. According to (3) and (4), the output feature signal ④ of the primary network AG is generated. Figure 2(c) presents the definition of various operation symbols and dimension changes of input and output feature signals in Figures 2(a) and 2(b).

2.1.2. Compensation Mechanism of the Auxiliary Network. The traditional single-path self-attention model generates a spatial attention coefficient map by the AG to cover the lesion area of features to pay more attention to the lesion area to improve the segmentation performance. Our proposed method builds an auxiliary network to generate an auxiliary attention coefficient map with a larger coverage area to compensate the segmentation network (primary network) to improve its hit rate of complete coverage of the lesion by spatial attention coefficient map. It should be noted that the attention compensation map will not deviate from the original attention area of the primary network but will be constrained to increase the attention area around it. This compensation mechanism is especially effective when the lesion feature is indistinct, the lesion's outline is unclear, or the segmentation model cannot generate the correct region of interest.

The qualitative analysis and comparison of using the primary network individually or combined with an auxiliary network are stated as follows. When DPAC-UNet uses the auxiliary network to compensate for the primary network, there are three situations:

Situation 1. (1) Use the primary network individually: when the focus area of the attention coefficient map of the single-path network is partially correct (Figure 3(a), ①), which will lead to reduced segmentation performance. (2) Combined with an auxiliary network: after the auxiliary network compensates the primary network's attention coefficient map with a larger focus area through additive compensation, the compensated attention coefficient map may be correct (Figure 3(a), ②) or remain unchanged (Figure 3(a), ③), which will eventually improve the segmentation performance or maintain the segmentation performance.

Situation 2. (1) Use the primary network individually: when the focus area of the attention coefficient of the primary network is completely correct, which will generate correct segmentation results (Figure 3(b), ①). (2) Combined with an auxiliary network: although the auxiliary network compensates it for a larger attention coefficient map, after the addition compensation operation, the value of the original correct focus area becomes larger, and the values of other areas are still smaller than the value of the correct area (Figure 3(b), ②). Therefore, the primary network of DPAC-UNet can still pay higher attention value in the correct area and keep the segmentation performance unchanged.

Situation 3. (1) Use the primary network individually: when the focus area of the primary network attention coefficient is completely wrong (Figure 3(c), ①), which will lead to reduced segmentation performance. (2) Combined with an auxiliary network: the larger auxiliary attention coefficient compensation map generated by the auxiliary network covers a larger area, and the compensated attention coefficient map may be still wrong (Figure 3(c), ②), or correct partially (Figure 3(c), ③), or correct completely (Figure 3(c), ④). At this time, correspondingly, the segmentation performance will remain unchanged, or improve to some extent, or improve significantly.

Therefore, by combining the previously mentioned three situations, the overall average segmentation performance of the whole dataset will be improved. It can also be seen from Figure 3 that the attention coefficient map generated by the auxiliary network does not deviate from the attention coefficient map area generated by the primary network.

2.2. Loss Functions of DPAC-UNet. We proposed two different compound loss functions to train the primary and auxiliary networks. First, we proposed the WBCE-Tversky loss for the primary network to generate an attention coefficient map focused on the target area and an accurate segmentation result. Second, we proposed the tolerance loss for the auxiliary network to generate an auxiliary compensation attention coefficient map with a larger coverage area to compensate for the primary network. It is called a tolerance loss because it can generate an attention coefficient map that covers a larger area and does not deviate from the lesion area, which means a higher fault tolerance for attention errors.

2.2.1. WBCE-Tversky Loss. The Tversky loss [22], which was proposed to address data imbalance in medical image segmentation, is introduced as a component of our WBCE-Tversky. The Tversky loss is as follows:

$$T_{\text{loss}}(\alpha,\beta) = 1 - \frac{\sum_{i=1}^{N} p_1 i \cdot g_1 i}{\sum_{i=1}^{N} p_1 i \cdot g_1 i + \alpha \sum_{i=1}^{N} p_1 i \cdot g_0 i + \beta \sum_{i=1}^{N} p_0 i \cdot g_1 i}, \tag{5}$$

where $p_{1,i}$ denotes the probability that a voxel is a lesion and $p_{0,i}$ denotes the opposite, and $g_{1,i}$ denotes the probability of whether a voxel is a lesion and $g_{0,i}$ denotes the opposite. The Tversky loss achieves a trade-off between false positives (FP) and false negatives (FN) by configuring the value of its hyperparameter β and α, where $\alpha + \beta = 1$. A higher β value implies that the trained model's *recall* is given greater weight than the precision, and the network pays more attention to FN. Often, the volume of the lesion is significantly smaller than that of healthy tissue. For example, in the 239 MRI scans of the ATLAS dataset, the voxel number ratio of the lesion to the background is about 3 : 1000. The high ratio of the nonlesion to lesion makes the segmentation network prone to focusing on the nonlesion area, therefore, predicting the lesions as nonlesions and increasing FN in the predicted results. To solve this problem, we increased the value of the hyperparameter β of Tversky loss. Larger β gives greater weight to *recall* than *precision* by placing more emphasis on FN. We assume that using higher β in our

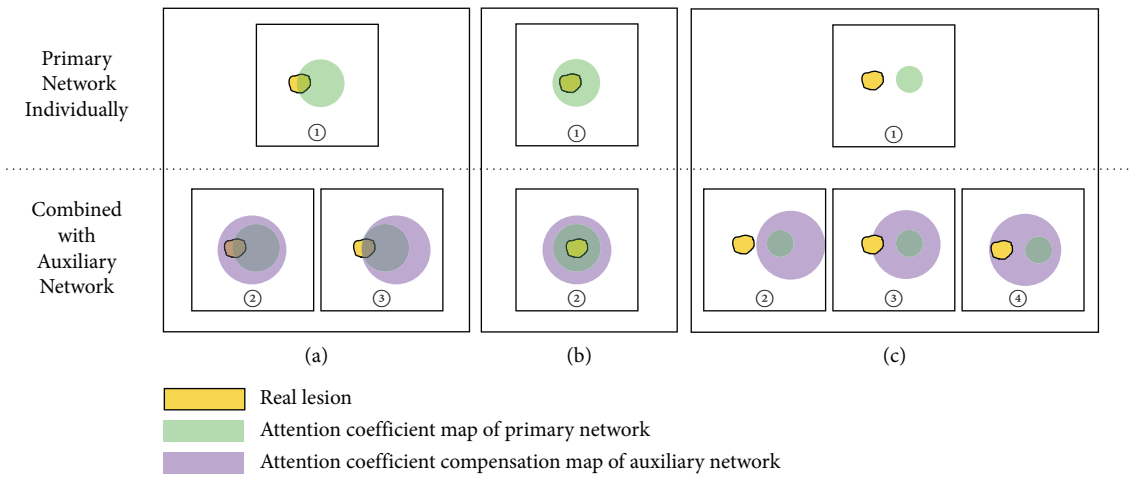


Figure 3: Qualitative analysis of compensation mechanism of the auxiliary network.

generalized loss function in training will lead to higher generalization and improved performance for the imbalanced dataset. So, we use the Tversky loss with higher β as a part of the WBCE-Tversky loss for training the primary network of DPAC-UNet. Meanwhile, in the tolerance loss, we also need to use a Tversky loss function to constrain the growth of the attention coefficient map to ensure that the larger and more tolerant focus area will not deviate from the lesion area. To compare the segmentation performance of the Tversky loss with the different hyperparameter values of β and select the appropriate hyperparameter β for the WBCE-Tversky loss and tolerance loss, we used the Tversky loss for the training the basic segmentation model, attention U-Net. The hyperparameter β of the Tversky loss ranges from 0.5 to 0.95, using 0.5 as the value interval. We conducted an experiment using the sixfold cross-validation, which is often used to train a model in which hyperparameters need to be optimized. We split the 239 stroke MRI scans into training, validation, and test sets by sixfold cross-validation according to Figure 4.

First, in each fold, we divided the data into training and test sets using a ratio of about 5 : 1 (199 : 40), and we ensured that all MRI scans of all test sets are not repeated. Second, we further split the training set in each fold into the inner training and validation sets using a ratio of about 4 : 1 (160 : 39). The validation set is used to select the best-performing model trained by the training set. Moreover, we also ensured that the training, validation, and test sets of each fold have the same lesion volume distribution for the accuracy of the experiment results. The lesion size distribution of fold 1 is presented in Figure 5.

The experimental configuration and results of training the attention U-Net using Tversky are presented in Table 1. We used 10 different β values to perform sixfold cross-validation and computed the average metric scores of all test sets' results. We used the dice similarity coefficient (DSC), F2 score (F2), precision (PRE), and recall (RE) as the metrics for the model evaluation. DSC is a widely used metric for evaluating the performance of the models; F2 score is often used to evaluate the performance of the models for imbalanced data; PRE quantifies the number of positive class predictions that belong to the positive class; RE quantifies the number of positive class predictions made out of all positive examples in the dataset. The experimental results of training the attention U-Net with different hyperparameter β values for the Tversky loss are presented in Table 1.

As presented in Table 1, the maximum RE value is obtained when β takes a large value of 0.95, and the maximum PRE value is obtained when the minimum value of 0.05 is taken. DSC and F2 scores reached the maximum when $\beta = 0.80$. Simultaneously, a trade-off between PRE and RE has been made, indicating that, for the imbalanced ATLAS dataset, training a model using the Tversky loss with hyperparameter $\beta = 0.80$ improves the segmentation accuracy. We need a loss function that can train the primary network of the DPAC-UNet to achieve an accurate segmentation. To improve the segmentation performance, we can handle the imbalanced dataset by selecting the hyperparameter β value of the Tversky loss to train the basic segmentation model in order to reduce the tendency of the lesion to be classified as nonlesion. As presented in Table 1, the use of the Tversky loss with hyperparameter $\beta = 0.80$ to train the attention U-Net on the ATLAS dataset achieves the highest segmentation performance. However, as presented in (5), if the denominator of the Tversky loss is a small value, it causes instability in backpropagation and derivation. To solve this problem, we introduced the WBCE loss [23] as another part of the WBCE-Tversky loss. On the one hand, it avoids the problems of backpropagation and gradient calculation instability caused by the Tversky loss for small denominators. On the other hand, using the WBCE loss and giving greater weight to the minority class in the equation adapts to the imbalance of dataset and further improves the overall segmentation performance. The WBCE loss function has differentiable properties, which simplifies the optimization process. The equation of the proposed WBCE-Tversky loss is presented in (8). The compound loss function is composed of the Tversky loss ($\beta = 0.80$) and WBCE loss, and their respective equations are presented as

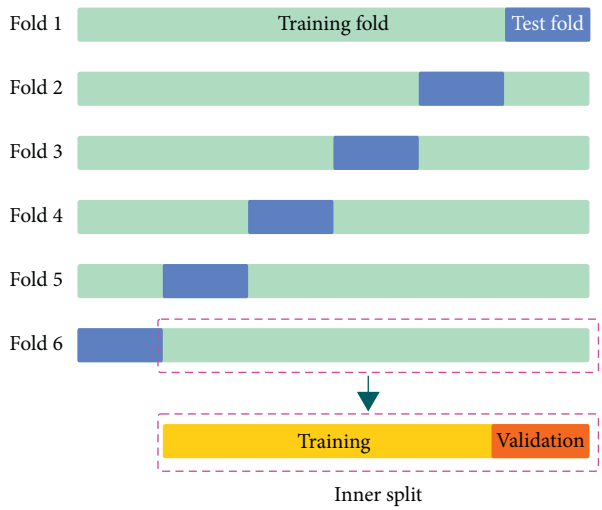


Figure 4: Schematic of sixfold cross-validation.

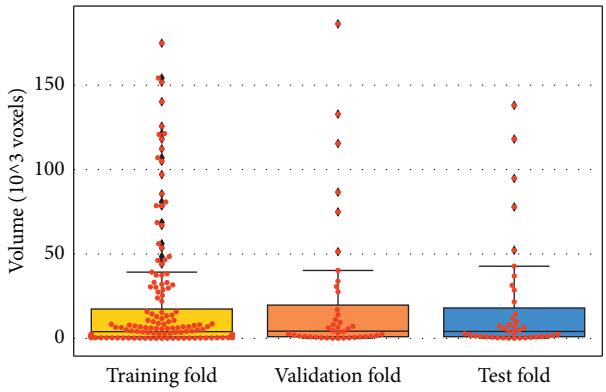


Figure 5: Distribution of lesion volume in the training, test, and validation sets.

$$\mathrm{WBCE}_{\mathrm{loss}} = -\frac{1}{N\sum_{i=1}^{N} wg_n}\log(p_n) + (1-g_n)\log(1-p_n), \tag{6}$$

$$w = \frac{N}{\mathrm{smooth.} + \sum_n g_n}, \tag{7}$$

$$\mathrm{WBCE-Tversky} = \mathrm{WBCE}_{\mathrm{loss}} + T_{\mathrm{loss}}(\beta = 0.8). \tag{8}$$

The WBCE loss adds weight w to the standard BCE loss to give the pixels more importance, and a higher training weight when the area of the lesion is small, thereby improving the segmentation performance for unbalanced datasets. As presented in (6), the main part of the WBCE loss is the same as the BCE loss [23]. The only difference is that we modified the calculation method of the weight w as presented in (7) and took the reciprocal of the proportion of lesion pixels to all pixels as the weight w, where N denotes the number of pixels in the entire image to be segmented and $\sum_n g_n$ is the number of lesion pixels to be segmented, andsmooth = 1 is used to prevent division by zero error.

Table 1: Experimental results when using the Tversky loss with different β values to train the attention U-Net.

Weights	Metrics (%)			
	DSC	F2	PRE	RE
$\alpha = 0.50, \beta = 0.50$	49.9	46.4	64.3	45.0
$\alpha = 0.45, \beta = 0.55$	50.8	48.6	62.8	47.5
$\alpha = 0.40, \beta = 0.60$	51.1	52.5	58.0	51.1
$\alpha = 0.35, \beta = 0.65$	50.9	52.1	57.8	53.7
$\alpha = 0.30, \beta = 0.70$	51.5	52.6	59.5	54.8
$\alpha = 0.25, \beta = 0.75$	52.0	51.5	61.3	52.5
$\alpha = 0.20, \beta = 0.80$	**52.7**	**55.4**	56.7	58.3
$\alpha = 0.15, \beta = 0.85$	50.5	52.5	53.4	55.5
$\alpha = 0.10, \beta = 0.90$	50.2	52.7	53.2	56.5
$\alpha = 0.05, \beta = 0.95$	51.6	55.0	53.5	59.4

To test and verify the proposed WBCE-Tversky loss, we conducted a series of comparative experiments using the WBCE loss, Tversky loss with different hyperparameter β, and WBCE-Tversky loss with different β. The model used in the experiment, the experiment datasets, and the experiment conditions are the same as the experiments corresponding to Table 1. The experiment parameter configuration and results are presented in Table 2. As can be seen from Table 2, for the same hyperparameter β, the DSC and F2 scores of the WBCE-Tversky loss are better than that of the Tversky loss. The WBCE-Tversky loss also performs best at $\beta = 0.80$. Compared with the WBCE loss, the segmentation accuracy improved significantly, the DSC score improved by 6.5%, and the F2 score increased by 12.5%. In summary, on the imbalanced ATLAS dataset, using the WBCE-Tversky loss with $\beta = 0.80$ to train the attention U-Net model achieves the best segmentation performance. Therefore, we used WBCE-Tversky loss with $\beta = 0.80$ as the loss function of the DPAC-UNet's primary network for accurate lesion segmentation.

2.2.2. Tolerance Loss. When the focus area is larger than the actual lesion area, the FP of the model segmentation result increases. The FP and FPR are proportional, implying that we can indirectly measure the tolerant degree of the lesion area using FPR. To indirectly measure the tolerant degree of the auxiliary compensation attention coefficient map, we used the FPR value as an indicator to determine the tolerant degree of attention coefficient generated by the auxiliary network. To provide the primary network with a more tolerant auxiliary compensation attention coefficient map and a much larger coverage area, we proposed the tolerance loss by introducing a *specificity reducing item* combined with the Tversky loss. It is called tolerance loss because the compound loss function's training goal is to obtain an attention coefficient map with high tolerance. The proposed tolerance loss is presented in (11), where $S_{\mathrm{loss}}(\lambda, \delta)$ denotes the *specificity reducing item* presented in (10). The concept of *specificity reducing item* is based on the adjustment of *specificity*, which measures the proportion of negatives that are correctly identified, and s is presented in

Table 2: Comparing the segmentation performance of the WBCE-Tversky loss under different hyperparameter configurations.

Loss functions	Weights	Metrics (%)				
		DSC	F2	PRE	RE	FPR
WBCE only	None	46.7	43.1	62.3	41.6	0.08
Tversky only	$\alpha = 0.50, \beta = 0.50$	49.9	46.4	64.3	45.0	0.06
WBCE-Tversky		51.5	49.5	63.2	49.5	0.10
Tversky only	$\alpha = 0.40, \beta = 0.60$	51.1	52.5	58.0	51.1	0.14
WBCE-Tversky		52.1	51.5	59.6	52.0	0.10
Tversky only	$\alpha = 0.30, \beta = 0.70$	51.5	52.6	59.5	54.8	0.14
WBCE-Tversky		51.9	50.4	62.2	50.3	0.10
Tversky only	$\alpha = 0.20, \beta = 0.80$	52.7	**55.4**	56.7	58.3	0.16
WBCE-Tversky		**53.2**	55.6	62.6	56.2	0.12
Tversky only	$\alpha = 0.10, \beta = 0.90$	50.2	52.7	53.2	56.5	0.20
WBCE-Tversky		51.5	51.6	57.7	53.1	0.14

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{9}$$

$$S_{\text{loss}}(\lambda, \delta) = \lambda\left(\frac{\sum_{i=1}^{N} p_0 i \cdot g_0 i}{\sum_{i=1}^{N} p_0 i \cdot g_0 i + \sum_{i=1}^{N} p_1 i \cdot g_0 i} - \delta\right)^2, \tag{10}$$

$$T_{\text{loss}} = S_{\text{loss}}(\lambda, \delta) + T^2_{\text{loss}}(\beta = 0.8). \tag{11}$$

Generally, the nonlesions in the imbalanced dataset occupy a large part of the total area. Using the ATLAS dataset as an example, the *specificity* of the segmentation results is reached as high as 95%. Since FPR = 1 −Specificity, it implies that the larger the proportion of nonlesions identified as nonlesions, the smaller the FPR, and the less tolerant the auxiliary compensation attention coefficient map. Therefore, we introduce a *specificity reducing item* to reduce the *specificity* of segmentation results, increase the FPR of the auxiliary network's training results, and increase the size of the coverage area of the attention coefficient map. As presented in (10) and (11), we used the hyperparameters λ and δ to control the weight of the *specificity reducing item* in the tolerance loss. We squared the *specificity reducing item* and the Tversky loss to balance the equation to make the backward derivation and backpropagation easier.

In (10), the *specificity reducing item* is the square of the difference between the *specificity* equation and δ. Since the training goal of any loss function is to make the value as small as possible, the training goal of (10) is to make value 0, which means that the value of *specificity* is close to the value of hyperparameter δ. Therefore, setting a reasonable δ can control the specificity value to the desired degree. The smaller the δ, the smaller the *specificity* obtained by the network training. As mentioned earlier, since FPR = 1 − Specificity, the smaller the *specificity*, the larger the obtained FPR value, and the resulting attention coefficient map is more tolerant with a larger coverage area. We set the hyperparameter δ value of our tolerance loss to 0.6, 0.7, 0.8, or 0.9. The other hyperparameter λ is set to 1, 2, 3, 4, or 5 to adjust the contribution of the *specificity reducing item* of the tolerance loss. The value of the hyperparameter β is set to 0.8 according to the conclusion discussed in Section 2.2.1. The experiment results are presented in Table 3.

Table 3: FPR values of the tolerance loss using different hyperparameter configurations.

Loss functions	Weights		Metrics (%)				
			DSC	F2	PRE	RE	FPR
Tolerance loss $\beta = 0.8$	$\lambda = 1$	$\delta = 0.9$	45.9	55.2	38.5	67.7	0.44
		$\delta = 0.8$	44.6	55.7	36.7	71.6	0.57
		$\delta = 0.7$	40.7	51.2	33.1	67.2	0.61
		$\delta = 0.6$	30.2	44.1	21.0	77.1	1.27
	$\lambda = 2$	$\delta = 0.9$	45.2	55.4	36.3	70.6	0.51
		$\delta = 0.8$	32.2	45.3	23.0	72.6	1.09
		$\delta = 0.7$	30.4	44.6	20.6	70.9	1.34
		$\delta = 0.6$	14.8	26.0	8.9	83.5	4.44
	$\lambda = 3$	$\delta = 0.9$	36.1	48.0	27.4	70.2	0.74
		$\delta = 0.8$	22.1	35.1	14.1	74.1	2014
		$\delta = 0.7$	22.8	36.1	14.8	79.4	2.01
		$\delta = 0.6$	11.8	18.8	7.6	83.5	4.57
	$\lambda = 4$	$\delta = 0.9$	39.4	50.9	31.4	68.9	0.69
		$\delta = 0.8$	23.9	37.6	15.6	74.2	1.89
		$\delta = 0.7$	14.9	25.4	9.2	80.6	4.09
		$\delta = 0.6$	7.9	11.2	5.7	82.8	4.99
	$\lambda = 5$	$\delta = 0.9$	34.4	47.5	24.9	72.4	0.90
		$\delta = 0.8$	20.7	33.5	13.3	82.3	2.74
		$\delta = 0.7$	13.2	24.1	7.8	84.1	5.63
		$\delta = 0.6$	5.7	11.8	3.1	92.8	**18.97**

As presented in Table 3, the different FPR values generated by the tolerance loss with different hyperparameters λ and $\delta$are compared. Based on (10), when $\lambda = 5$, the tolerance loss gives the most significant weight to the *specificity reducing item*. Increasing λ and keeping δ constant produce higher FPR. Furthermore, the smaller the value of δ, the smaller the value of *specificity*, and the higher the FPR. In Table 3, the largest FPR value was obtained when $\lambda = 5, \delta = 0.6$, and the FPR reaches as high as 18.97%. We also introduce a Tversky loss part to constrain the spatial position and contour shape of the lesion and restrict the growth of the attention coverage area with a high FPR value, rather than randomly increasing the FPR of the results.

As visual examples, we export the attention coefficient heatmaps of four MRI slices of different lesion sizes, which were segmented by the attention U-Net trained by tolerance loss with 10 varying configurations of hyperparameter. The attention coefficient heatmaps are generated by the AG (marked as II) in the auxiliary network in Figure 1. Note that, in the tolerance loss, the hyperparameter $\beta = 0.8$ is fixed, because we used the other two parameters to adjust the FPR value. Considering the FPR of some values may be caused by a smaller λ and a larger δ or by a larger λ and a smaller δ, to draw the heatmaps, we sorted the FPR values in Table 3 and evenly selected 10 hyperparameter configurations of the tolerance loss according to the different FPR values. The attention coefficient heatmaps from the selected 10 hyperparameter configurations from Table 3 are also presented in Figure 6. It can be seen that as the FPR value increases, the coverage area of the attention coefficient map gradually increases. Due to the restriction of the Tversky loss part in the tolerance loss, although the focus area increased gradually, it did not deviate from the lesion area. Therefore, when tolerance loss is used in the auxiliary network of the

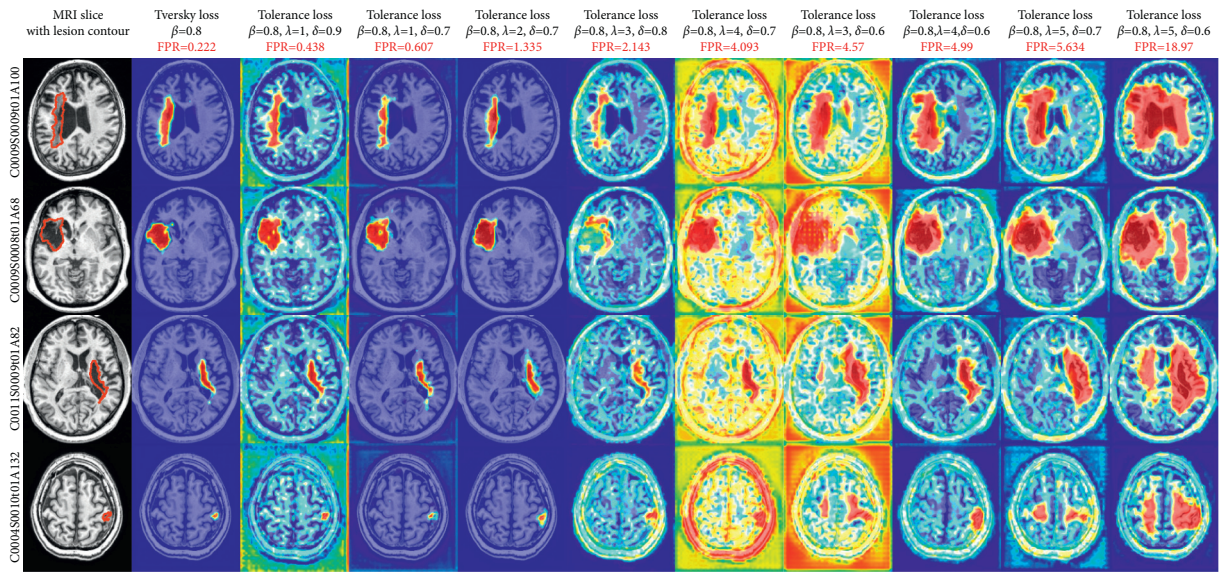


Figure 6: Attention coefficient heatmaps generated by the attention U-Net with different hyperparameters of the tolerance loss.

DPAC-UNet, the primary network gets a compensation attention coefficient with the correct region irrespective of the increase of the FPR value and the coverage area. However, for the coverage area of the auxiliary compensation attention coefficient map, the case is not the larger the better, indicating that FPR is not as high as possible. We need to set a moderate value of hyperparameters λ and δ to provide the best segmentation performance for DPAC-UNet. Therefore, in Session 3, the optimal λ and δ hyperparameters will be selected based on the DPAC-UNet model depending on the experiment performance.

2.2.3. Hyperparameter Selection. In order for the auxiliary network to generate a larger proper attention coefficient map, it needs to be trained by the tolerance loss proposed. Only when the hyperparameter configuration of the tolerance loss function is selected appropriately, the auxiliary network can provide moderate compensation to the attention module of the primary network to improve the segmentation performance. The selection process of loss function hyperparameter configuration of the primary and auxiliary network follows the following two steps:

Step 1. With 0.05 as the interval, from 0.5 to 0.95, using 10 different β values of Tversky loss to train the single-path Attention U-Net model, take the β value with the best segmentation performance as the selected β value of the proposed WBCE-Tversky loss and Tolerance loss.

Step 2. To select appropriate δ and λ values for the tolerance loss, so that the auxiliary network can provide appropriate attention coefficient map compensation and achieve the best segmentation performance of the entire DPAC-UNet, we use the WBCE-Tversky loss function (fix the β value that has been selected in the first step) to train the primary network. We set the tolerance loss δ value to 0.6, 0.7, 0.8, or 0.9, and set λ value to 1, 2, 3, 4, or 5; that is, we use a total of 20 different parameter pairs of tolerance loss to train the auxiliary network, and take the δ and λ pair with the best segmentation performance as the selected values of proposed tolerance loss.

When our method is applied to other different types of datasets of medical segmentation tasks or different segmentation models, the hyperparameter configurations of loss functions are different, and the hyperparameter values need to be redetermined. This is because the hyperparameter selection of the loss function needs to consider the imbalance of different datasets and the individual differences of attention maps generated by different models.

3. Experimental Results and Analysis

3.1. Dataset and Training. The ATLAS dataset has a high 3D resolution that can meet the requirements of rotation slicing operations, which contains 239 MRI data and focuses on the subacute and chronic stages of stroke disease. The operations of MNI-152 [24] image registration, intensity normalization [25], bias field correction [26], and changing the resolution of MRI scans to $176 \times 208 \times 176$ through cropping and interpolation operation to fit our method have been performed. We use the sixfold cross-validation to ensure that the test sets can cover the entire dataset. We also divide the training set of each fold into the inner loop training set and the inner loop validation set for best model selection. It should be noted that since the distribution of the number of MRIs of different sizes is extremely imbalanced in the dataset, it is necessary to ensure that the training, validation, and test sets have similar lesions sizes' distribution.

We use the deep learning framework PyTorch to conduct our experiments on three NVIDIA Tesla T4 GPUs. We train the models 100 epochs at most and save the best model when the validation set loss is the smallest. We used the lookahead optimizer [27] for model training. The optimizer improves the stability of the optimization process while considering

the dynamic adjustment of the learning rate and the acceleration of the gradient descent. We set the initial learning rate to 1×10^{-4}. The same experiment conditions and environment, used in the previous experiments in Section 2, are used for reproducing the single-path segmentation models, such as U-Net and attention U-Net. We applied the WBCE-Tversky loss for accurate segmentation to train these single-path models and use their results to compare the results of our DPAC-UNet method.

3.2. Experiment and Results. In Section 2.1, we elaborated on the principle of the proposed DPAC network structure. Using the attention U-Net as the basic segmentation model of the primary and auxiliary networks of the DPAC method, we proposed a specific segmentation model, DPAC-UNet. In Section 2.2, we also proposed the WBCE-Tversky loss and tolerance loss to train the primary and auxiliary networks, respectively. Moreover, we explored and verified the value of hyperparameter β of the WBCE-Tversky loss through the experiments presented in Tables 1 and 2 and found that when $\beta = 0.8$, the primary network based on the attention U-Net achieves the best segmentation performance trained by the WBCE-Tversky loss.

We also explained the relationship between the values of different hyperparameters δ and λ and the coverage area of the auxiliary compensation attention coefficient map in Section 2.2. The coverage area of the auxiliary attention coefficient map is proportional to the FPR value, and the FPR value is proportional to λ and inversely proportional to δ. We need to select a suitable set of λ and δ values to obtain an auxiliary attention coefficient map with a suitable coverage area in order to enable the DPAC-UNet to achieve the best segmentation performance. Therefore, based on the experiment results, as presented in Table 3, we explored the optimal hyperparameter configuration of λ and δ to train the best DPAC-UNet model. We used the tolerance loss ($\beta = 0.8$) configured with different hyperparameters λ and δ to train the auxiliary network of DPAC-UNet and the WBCE-Tversky loss ($\beta = 0.8$) to train the primary network of the DPAC-UNet.

Table 4 presents the experiment results corresponding to the experiment of DPAC-UNet trained by the tolerance loss function with different hyperparameters. In Table 4, the FPR* represents the FPR results of single-path attention U-Net trained by tolerance loss functions with different hyperparameter configurations from Table 3. We sort FPR* in ascending order and identified the corresponding tolerance loss functions and hyperparameter configurations. We use tolerance loss functions with these sorted configurations to train the auxiliary network of the DPAC-UNet and the WBCE-Tversky loss ($\beta = 0.8$) to train the primary network. Then, we got the experiment results of the different configurations of DPAC-UNet to select the best hyperparameter configuration.

By observing the relationship between FPR* and segmentation metrics, as presented in Table 4, it is evident that as the coverage area of the attention coefficient generated by the auxiliary network increases (indicated by FPR*), the DSC and F2 scores of the DPAC-UNet gradually increase. When the values of the hyperparameters are $\lambda = 4$ and $\delta = 0.7$, the DSC and F2 scores get the maximum value. As the FPR* further increases, the segmentation performance gradually declines. When the coverage area significantly increases with the FPR* value, it negatively affects the primary network. As presented in Figure 6, when $\lambda = 5$ and $\delta = 0.6$, the FPR* value reaches the maximum, as well as the coverage area of the auxiliary compensation attention, which occupies a quarter of the brain slice. At this time, the coverage area is too large to constrain the primary network to focus on the correct lesion area effectively. Its attention coefficient map generated by this hyperparameter configuration even interferes with the primary network, so its DSC and F2 scores are negatively affected as presented in Table 4. The change of FPR* is determined by the hyperparameters λ and δ together. FPR* is proportional to λ and inversely proportional to δ. Therefore, the smallest λ and the largest δ will generate the smallest FPR*, and the largest λ and smallest δwill lead to the largest FPR*. Figure 7 presents a line chart of the segmentation accuracy changing with FPR*. The line chart indicates that the DPAC-UNet segmentation accuracy changes as the FPR* increases. As the FPR* increases, the DSC and F2 scores increase and then decrease. It shows that when the FPR* is small, the coverage area of the corresponding auxiliary attention compensation coefficient map is also small. It cannot compensate for the primary network adequately and effectively. When the FPR* value is too large, it tends to overcompensate. Only when the hyperparameter values are moderate and its corresponding FPR* value is moderate can the DPAC-UNet achieve the best segmentation performance.

Simultaneously, it can be seen from Table 4 that the FPR values generated by the DPAC-UNet's primary network are all small, irrespective of the loss function of the auxiliary network used and the corresponding FPR* value. This is because the compensation operation of the auxiliary compensation attention coefficient map generated by the auxiliary network does not directly affect the segmentation result of the primary network. It is an additive compensation operation from the auxiliary network to the primary network during the training process; therefore, it does not participate in the gradient operation and backpropagation of the primary network. However, it partially modified the size of the coverage area of the primary network's attention coefficient map. The primary network still considers accurate segmentation as its training purpose. It does not generate FP as high as the auxiliary network due to the increased attention area after compensation.

In summary, when the primary network uses the WCBE-Tversky loss function with hyperparameter configuration of $\beta = 0.8$, and the auxiliary network uses tolerance loss function with hyperparameter configuration of $\beta = 0.8$, $\lambda = 4$, and $\delta = 0.7$, our DPAC-UNet can achieve the highest segmentation accuracy.

3.3. Visualization Examples. To show the principle of the DPAC-UNet, we give the attention coefficient heatmaps and segmentation results of using attention U-Net (primary

Table 4: The segmentation performance of the DPAC-UNet using different hyperparameter configurations.

Loss functions		Weights	Metrics (%)					
			FPR*	DSC	F2	PRE	RE	FPR
Tolerance loss, $\beta = 0.8$	1.	$\lambda = 1, \delta = 0.9$	0.438	54.8	54.1	63.6	55.1	0.111
	2.	$\lambda = 2, \delta = 0.9$	0.508	53.0	52.4	61.5	53.3	0.101
	3.	$\lambda = 1, \delta = 0.8$	0.573	55.2	54.7	64.4	55.7	0.120
	4.	$\lambda = 1, \delta = 0.7$	0.607	54.1	54	62.2	55.1	0.117
	5.	$\lambda = 4, \delta = 0.9$	0.689	55.9	56.6	63	57.4	0.124
	6.	$\lambda = 3, \delta = 0.9$	0.743	55.3	56	61.1	58.1	0.173
	7.	$\lambda = 5, \delta = 0.9$	0.898	53.8	54.1	62.8	55.7	0.142
	8.	$\lambda = 2, \delta = 0.8$	1.091	54.9	55.4	61.2	56.9	0.140
	9.	$\lambda = 1, \delta = 0.6$	1.27	55.5	55.6	63.7	57	0.126
	10.	$\lambda = 2, \delta = 0.7$	1.335	55.8	55.8	64.8	57.2	0.133
	11.	$\lambda = 4, \delta = 0.8$	1.888	53.6	53	64.4	53.9	0.111
	12.	$\lambda = 3, \delta = 0.7$	2.006	56.9	57.7	61.6	59.6	0.149
	13.	$\lambda = 3, \delta = 0.8$	2.143	56.7	57.3	61.9	59.1	0.157
	14.	$\lambda = 5, \delta = 0.8$	2.744	56.7	56	65.8	56.8	0.103
	15.	$\lambda = 4, \delta = 0.7$	**4.093**	**59.3**	**59.8**	65.6	59.9	0.106
	16.	$\lambda = 2, \delta = 0.6$	4.44	58.2	58.6	62.6	60.3	0.151
	17.	$\lambda = 3, \delta = 0.6$	4.57	57.5	57.5	64	58.8	0.137
	18.	$\lambda = 4, \delta = 0.6$	4.99	56.5	56.9	62.5	61.6	0.153
	19.	$\lambda = 5, \delta = 0.7$	5.634	56.2	57.5	63	59.3	0.132
	20.	$\lambda = 5, \delta = 0.6$	18.97	52.8	51.5	65.9	52.1	0.196

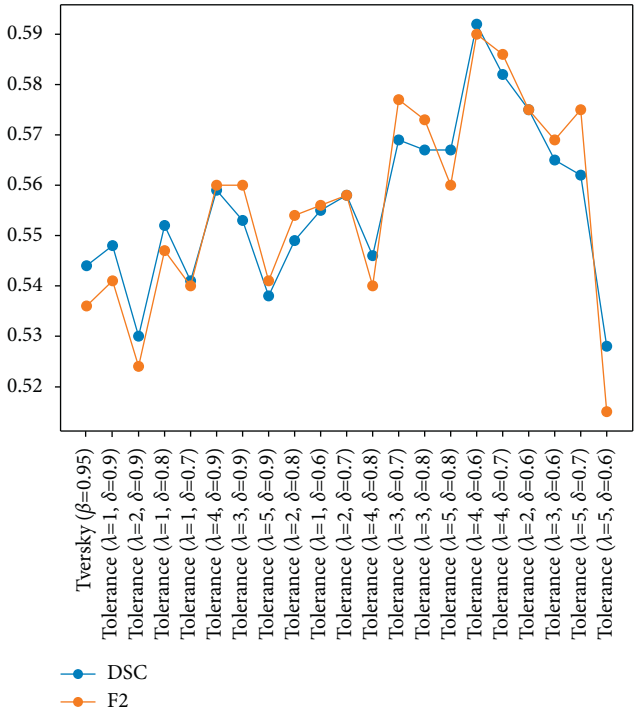
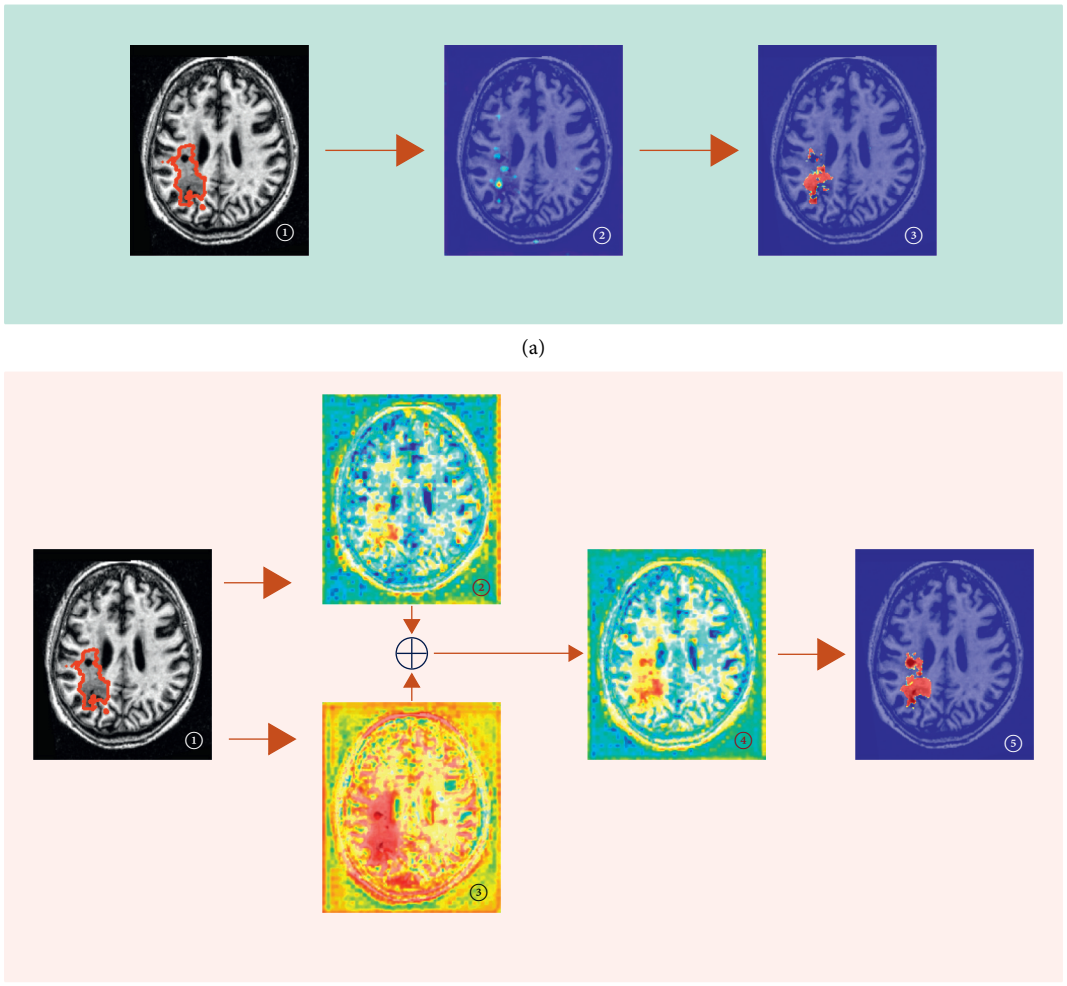


Figure 7: Segmentation performance of DPAC-UNet with the change in FPR*.

network) individually and using DPAC-UNet with the auxiliary network when segmenting an MRI slice, as presented in Figure 8.

Using the primary network individually as presented in Figure 8(a), ② is the attention coefficient heatmap generated by the second-level AG of classic Attention U-Net; it can be observed that its attention coefficient map has obvious defects. Although the lesion's location is correct, the coverage area of the lesion is too small to perform accurate segmentation. ③ is the segmentation result; comparing ③ with the truth label of ①, it can be seen that there is a big difference between the segmentation result and the ground truth. When using the DPAC-UNet to segment the slice, as presented in Figure 8(b), ② is the attention coefficient heatmap generated by the primary network at the location marked as (I) in Figure 1. It is evident from the figure that the attention coefficient heatmap has obvious defects that are consistent with ②, as presented in Figure 8(a), which is also a defective attention heatmap with a smaller coverage area than the actual lesion. Notably, the attention coefficient heatmap ②, as presented in Figure 8(b), introduces a certain amount of noise. As presented in Figure 8(b), ③ is the auxiliary compensation attention coefficient generated by the DPAC-UNet's auxiliary network at the location marked as (II) in Figure 1. It is evident that the coverage area is moderately larger than the actual lesion, and covering the correct lesion region. After compensating the auxiliary compensation attention coefficient map of ③ to the primary network's attention coefficient map of ② through an additive compensation operation, a new attention coefficient map after compensation is obtained, as shown in ④. Comparing ④ and ②, as presented in Figure 8(b), the insufficient coverage area of attention coefficient in ② has been compensated, and the noise has also been significantly reduced. ⑤ is the final segmentation result of the DPAC-UNet. After using the DPAC-UNet, the segmentation result has been significantly improved in terms of both lesion contour and area. One thing to note here is when we compare the heatmap ② of Figure 8(a) generated by single-path attention U-Net and the heatmap ② of Figure 8(b) generated by DPAC-UNet's primary network, the attention heatmaps of Figures 8(a)



(a)

(b)

Figure 8: Visualization examples of the attention coefficient maps of different methods: (a) single-path primary network individually; \(b) DPAC-UNet.

and 8(b) are slightly different in noise level because they are two independent trained models, but the respective heatmap ② has the defects of the same pattern.

3.4. Comparison of Different Methods. Many lesion segmentation methods have been studied recently using the ATLAS dataset. Zhou et al. proposed a new architecture called dimension-fusion-UNet (D-UNet) [28], which combines 2D and 3D convolution in the encoding stage. Yang et al. proposed a CLCI-Net using cross-level fusion and a context inference network [29]. The previously mentioned existing segmentation results serve as a comparison for our experiments.

Using the same conditions as the previous experiments, we conducted a comparison experiment of the following models and loss functions:

(1) the U-Net [9] model trained by the WBCE-Tversky loss ($\beta = 0.8$)

(2) the attention U-Net [10] trained by the WBCE-Tversky loss ($\beta = 0.8$)

(3) the DPAC-UNet model proposed in this paper, trained by the WBCE-Tversky loss and tolerance loss ($\beta = 0.8, \delta = 0.7, \lambda = 4$)

Cases (2) and (3) are, respectively, using the primary network individually and combined with the auxiliary network.

The final experiment comparison results are presented in Table 5 that the DPAC-UNet achieved the highest DSC and F2 scores. Comparing the single-path model attention U-Net with our DPAC-UNet, from using primary network individually to the introduction of the auxiliary attention compensation mechanism, the DSC score improved by 6%. Comparing the classic U-Net with attention U-Net, from no attention to the introduction of self-attention mechanism, the DSC score only improved by 2.1%. The previously mentioned comparison shows that our DPAC-UNet has a very significant performance improvement compared to the single-path self-attention segmentation model. Compared with the methods in the existing literature, it is 5.7% higher than the D-UNet and 1.1% higher than the CLCI-NET. This suggests that our DPAC-UNet achieved improved

Table 5: Comparison of segmentation performance of different methods.

Models	Loss functions	Metrics (%)			
		DSC	F2	PRE	RE
D-UNet	Enhance mixing loss	53.5	—	63.3	52.4
CLCI-NET	Dice loss	58.1	—	64.9	58.1
U-Net	WBCE-Tversky($\beta = 0.8$)	51.1	49.2	59.3	48.7
Attention U-Net	WBCE-Tversky($\beta = 0.8$)	53.2	55.6	62.6	56.2
DPAC-UNet	WBCE-Tversky($\beta = 0.8$), tolerance($\delta = 0.7, \lambda = 4$)	**59.2**	**59.0**	**65.6**	**59.9**

segmentation performance than the existing methods. As shown in Figure 9, we present a group of boxplots of the segmentation performance distribution of all 239 MRI scans to evaluate the performance of the different models. The 239 segmentation results are generated from the six nonrepeated test sets split by sixfold nested cross-validation. From the boxplots, we can state the following: first, comparing our DPAC-UNet model with the other two models, the overall segmentation accuracy increases significantly, and also, the minimum value of the boxplot of DSC and F2 scores and its lower quartile value increase significantly. This proves that our method significantly improves the data with poor performance using the other two methods. Second, when comparing the middle value and upper quartile of boxplots, we can see that, for the data with better segmentation performance segmented by the other two models, the DPAC-UNet has a slight improvement. For data with distinct lesion characteristics that are easy to segment, the primary network can generate a correct attention coefficient map with a high probability. At this time, using the auxiliary network to compensate the primary network will not reduce the segmentation accuracy or even slightly improve it. By observing the boxplots of the FPR results, it is evident that the FPR values of the three models are consistently small. This proves that although the auxiliary compensation attention coefficient map generated by the DPAC-UNet's auxiliary network has a high FPR, after compensating it to the primary network, the segmentation result of the primary network maintains a small FPR.

3.5. Time Consumption. The parameter amount, training, and testing computation time for each part of DPAC-UNet are listed in Table 6 to understand which part of the network needs more time for executing. Since the primary and auxiliary networks are trained in parallel as a whole, the computation time of each part cannot be measured separately at the same time. Therefore, we compared the computation complexity and time consumption of the primary and auxiliary networks of DPAC-UNet by training them independently.

The amount of our DPAC-UNet's training parameters is double compared with the single-path attention U-Net (primary network or auxiliary network). The training time of the DPAC-UNet (5.11 hours on average) is about 1.7 times that of each subnetwork (3.06 hours on average). The testing time of the DPAC-UNet (17 secs on average) is about 1.7 times that of each subnetwork (10 secs on average). Although DPAC-UNet has significantly increased the total number of model parameters and training time consumption after the introduction of the auxiliary network compensation mechanism, the significant improvement in segmentation performance makes up for the shortcoming of model complexity.

3.6. DPAC Structure of Other Models. The DPAC structure proposed in this paper that uses the auxiliary network to compensate the primary network can be applied to most segmentation models with spatial self-attention. We implemented our method on two other segmentation models with self-attention mechanism, RA-UNet [30] and AGResU-Net [31], and compared the experimental results of single-path with dual-path networks with auxiliary networks. The experimental results are shown in Table 7. The previously mentioned two single-path segmentation models can effectively improve the segmentation performance after using the auxiliary network for attention compensation. It shows that our method can be applied to other segmentation networks with the self-attention mechanism. It should be noted that, in accordance with the hyperparameter selection steps in Section 2.2.3, when the dataset and segmentation model change, the hyperparameters of the tolerance loss function need to be redetermined. As shown in Table 7, when the δ value of AGResU-Net is 0.6, the DPAC structure can achieve the best segmentation performance.

4. Discussion and Conclusions

In this paper, we proposed the DPAC-UNet using the classic self-attention model, attention U-Net, as the basic segmentation model. To realize the functions of the DPAC-UNet's primary and secondary networks, we proposed the WBCE-Tversky and tolerance losses as the training loss functions, respectively. We explored the hyperparameter configuration of the loss functions by applying sixfold cross-validation on the 239 MRI data of the ATLAS stroke segmentation dataset. We discovered that the WBCE-Tversky loss achieves the most accurate segmentation performance for the primary network when $\beta = 0.8$. The tolerance loss generates a tolerant auxiliary compensation attention coefficient map with a moderate coverage area to compensate for the primary network's defective attention coefficient map. It achieves the best segmentation performance when $\beta = 0.8$, $\lambda = 4$, and $\delta = 0.7$. The experiment results indicate that the DSC score of the proposed DPAC-UNet with the auxiliary network is 6% higher than that without the auxiliary network. Compared with the methods in the existing

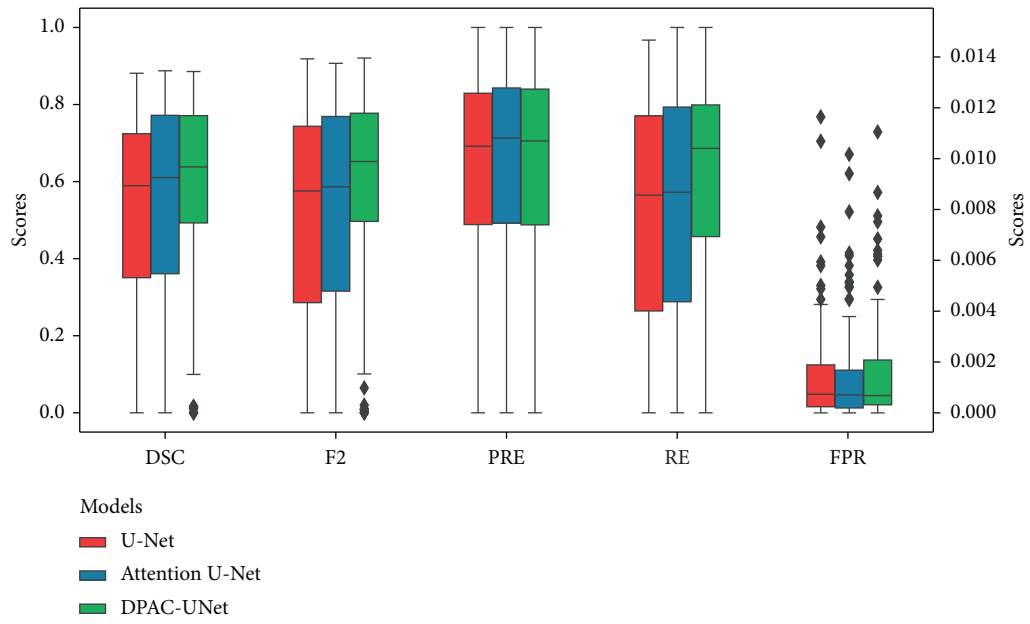


Figure 9: Boxplots of metric results for different models.

Table 6: Time consumption of DPAC-UNet.

Networks	Parameters (M)	Training (hours)	Testing (seconds)
Primary network	40.4	3.07	10
Auxiliary network	40.4	3.05	10
DPAC-UNet	80.8	5.11	17

Table 7: Experimental results of DPAC structure based on other models.

No.	Networks	Auxiliary	Loss functions	Metrics (%)			
				DSC	F2	PRE	RE
1	RA-UNet	Without	WBCE-Tversky ($\beta = 0.8$)	54.1	56.5	63.8	58.1
		With	WBCE-Tversky ($\beta = 0.8$), tolerance ($\delta = 0.7, \lambda = 4$)	**60.3**	**59.9**	**67.1**	**60**
2	AGResU-Net	Without	WBCE-Tversky ($\beta = 0.8$)	55.2	59.7	61.4	57.5
		With	WBCE-Tversky ($\beta = 0.8$), tolerance ($\delta = 0.6, \lambda = 4$)	**60.5**	**62.2**	**66.6**	**61.1**

literature, the DSC score of the proposed DPAC-UNet is 5.7% higher than the D-UNet and 1.1% higher than the CLCI-NET. The results indicate that the proposed method achieved an improved segmentation performance and verified the effectiveness of the proposed method.

It should be noted that although we used the same dataset in the proposed method as D-UNet and CLCI-NET, the version varied. We used the version without defacing that contains 239 MR images, and D-UNet and CLCI-NET used the version with defacing that contains 229 MR images. Furthermore, considering that the cross-validation dataset splitting methods do not generate the same training, validation, and testing sets, and also considering that the loss functions used are also different, achieving the best segmentation performance does not directly prove that the proposed method is the best. It proves that we have reached a higher level of segmentation performance in the current methods.

The purpose and focus of our work are to improve the performance of the single-path attention mechanism segmentation model by using our DPAC method. As shown in Table 6, although our method obviously requires more computing resources and takes more training time, the improvement in the segmentation performance of our method balances out the shortcomings in increased model complexity. The five-hour training time is currently at a lower or average level in some of the latest existing network models, which are currently used for stroke lesion segmentation. Moreover, we will implement our DPAC network structure on other basic segmentation models with a self-attention mechanism to verify our method's versatility. We also proved that if our DPAC structure is applied to other models based on the self-attention mechanism, it can also effectively improve the segmentation performance. In future work, we plan to use other stroke segmentation datasets to compare the effectiveness of our method across various datasets.

Data Availability

The ATLAS dataset is publicly available at http://fcon_1000.projects.nitrc.org/indi/retro/atlas_download.html.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (NSFC) under grant no. 62171307, Key Research and Development Project of Shanxi Province under grant no. 201803D31045 (China), Natural Science Foundation of Shanxi Province under grant no. 201801D121138 (China), research project supported by Shanxi Scholarship Council under grant no. 201925 (China), and Graduate Education Innovation Project of Shanxi Province under grant no. 2018BY051 (China).

References

[1] A. G. Thrift, D. A. Cadilhac, T. Thayabaranathan et al., "Global stroke statistics," *International Journal of Stroke*, vol. 9, no. 1, pp. 6–18, 2014.

[2] R. Zhang, L. Zhao, W. Lou et al., "Automatic segmentation of acute ischemic stroke from DWI using 3-D fully convolutional DenseNets," *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2149–2160, 2018.

[3] S. L. Liew, J. M. Anglin, N. W. Banks et al., "A large, open source dataset of stroke anatomical brain images and manual lesion segmentations," *Scientific data*, vol. 5, no. 1, Article ID 180011, 2018.

[4] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[5] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2017.

[6] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: a deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[7] K. Suzuki, "Overview of deep learning in medical imaging," *Radiological Physics and Technology*, vol. 10, no. 3, pp. 257–273, 2017.

[8] G. Litjens, T. Kooi, B. E. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[9] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Munich, Germany, October 2015.

[10] J. Schlemper, O. Oktay, M. Schaap et al., "Attention gated networks: learning to leverage salient regions in medical images," *Medical Image Analysis*, vol. 53, no. 12, pp. 197–207, 2019.

[11] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: a nested U-net architecture for medical image segmentation," in *Proceedings of the 4th International Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3–11, Granada, Spain, September 2018.

[12] M. Z. Alom, C. Yakopcic, T. M. Taha, and V. K. Asari, "Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net)," in *Proceedings of the NAECON 2018-IEEE National Aerospace and Electronics Conference*, pp. 228–233, Dayton, OH, USA, July 2018.

[13] B. S. Lin, K. Michael, S. Kalra, and H. R. Tizhoosh, "Skin Lesion Segmentation: U-Nets versus Clustering," in *Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1–7, Honolulu, HI, USA, November 2017.

[14] M. Noori, A. Bahri, and K. Mohammadi, "Attention-guided version of 2D UNet for automatic brain tumor segmentation," in *Proceedings of the 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE)*, Mashhad, Iran, October 2019.

[15] Y. J. Huang, Q. Dou, Z. X. Wang et al., "3-D RoI-aware U-net for accurate and efficient colorectal tumor segmentation," *IEEE Transactions on Cybernetics*, 2020, Early Access.

[16] P. F. Christ, M. E. A. Elshaer, F. Ettlinger et al., "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random elds," in *Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 415–423, Athens, Greece, October 2016.

[17] K. Sirinukunwattana, J. P. W. Pluim, H. Chen et al., "Gland segmentation in colon histology images: the glas challenge contest," *Medical Image Analysis*, vol. 35, pp. 489–502, 2017.

[18] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-net: learning dense volumetric segmentation from sparse annotation," in *Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 424–432, Athens, Greece, October 2016.

[19] J. Merkow, A. Marsden, D. Kriegman, and Z. Tu, "Dense volume-to-volume vascular boundary detection," in *Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 371–379, Athens, Greece, October 2016.

[20] M. Khened, V. A. Kollerathu, and G. Krishnamurthi, "Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers," *Medical Image Analysis*, vol. 51, pp. 21–45, 2019.

[21] Y. Li and L. Shen, "Deep learning based multimodal brain tumor diagnosis," in *Proceedings of the 3rd International MICCAI Brainlesion Workshop*, pp. 149–158, Quebec City, Canada, September 2017.

[22] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3D fully convolutional deep networks," in *Proceedings of the International Workshop on Machine Learning in Medical Imaging*, pp. 379–387, Quebec City, Canada, September 2017.

[23] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 240–248, Springer, Berlin, Germany, 2017.

[24] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans, "Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space," *Journal of Computer Assisted Tomography*, vol. 18, no. 2, pp. 192–205, 1994.

[25] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Transactions on Medical Imaging*, vol. 17, no. 1, pp. 87–97, 1998.

[26] N. J. Tustison, B. B. Avants, P. A. Cook et al., "N4ITK: improved N3 bias correction," *IEEE Transactions on Medical Imaging*, vol. 29, no. 6, pp. 1310–1320, 2010.

[27] M. Zhang, J. Lucas, J. Ba, and G. E. Hinton, "Lookahead Optimizer: k steps forward, 1 step back," in *Proceedings of the 33rd Conference on Neural Information Processing Systems*, pp. 9593–9604, Vancouver, Canada, December 2019.

[28] Y. Zhou, W. Huang, P. Dong, Y. Xia, and S. Wang, "D-UNet: a dimension-fusion U shape network for chronic stroke lesion segmentation," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 3, pp. 940–950, 2021.

[29] H. Yang, W. Huang, K. Qi et al., "CLCI-net: cross-level fusion and context inference networks for lesion segmentation of chronic stroke," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 266–274, Cham, Germany, October 2019.

[30] Q. Jin, Z. Meng, C. Sun, H. Cui, and R. Su, "RA-UNet: A hybrid deep attention-aware network to extract liver and tumor in CT scans," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 1471, 2020.

[31] J. Zhang, Z. Jiang, J. Dong, Y. Hou, and B. Liu, "Attention gate resU-Net for automatic MRI brain tumor segmentation," *IEEE Access*, vol. 8, pp. 58533–58545, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 7567870, 9 pages
https://doi.org/10.1155/2021/7567870

Research Article

An Improved CNN Architecture to Diagnose Skin Cancer in Dermoscopic Images Based on Wildebeest Herd Optimization Algorithm

Biying Zhou[1] **and Behdad Arandian**[2]

[1]*School of Computer, Weinan Normal University, Weinan 714099, Shaanxi, China*
[2]*Department of Electrical Engineering, Dolatabad Branch, Islamic Azad University, Isfahan, Iran*

Correspondence should be addressed to Behdad Arandian; b.arandian@iauda.ac.ir

Received 14 July 2021; Revised 7 August 2021; Accepted 16 August 2021; Published 29 August 2021

Academic Editor: V. Rajinikanth

Skin cancer is one of the most common types of cancers that is sometimes difficult for doctors and experts to diagnose. The noninvasive dermatoscopic method is a popular method for observing and diagnosing skin cancer. Because this method is based on ocular inference, the skin cancer diagnosis by the dermatologists is difficult, especially in the early stages of the disease. Artificial intelligence is a proper complementary tool that can be used alongside the experts to increase the accuracy of the diagnosis. In the present study, a new computer-aided method has been introduced for the diagnosis of the skin cancer. The method is designed based on combination of deep learning and a newly introduced metaheuristic algorithm, namely, Wildebeest Herd Optimization (WHO) Algorithm. The method uses an Inception convolutional neural network for the initial features' extraction. Afterward, the WHO algorithm has been employed for selecting the useful features to decrease the analysis time complexity. The method is then performed to an ISIC-2008 skin cancer dataset. Final results of the feature selection based on the proposed WHO are compared with three other algorithms, and the results have indicated good results for the system. Finally, the total diagnosis system has been compared with five other methods to indicate its effectiveness against the studied methods. Final results showed that the proposed method has the best results than the comparative methods.

1. Introduction

Skin cancer is the most common malignant cancer of the body [1]. This cancer is caused by the uncontrolled and abnormal growth of skin cells. Ultraviolet rays from sunlight and tanning lamps cause genetic changes in the nucleus of skin cells [2]. If the body's immune system fails to repair the damage, skin cells begin to grow rapidly and unrestrained, initially appearing as fast-growing moles with bleeding, tumors, or wounds that do not heal such that if not treated, it spread to other areas (metastasis) [3].

One of the most dangerous types of skin cancer is melanoma. Melanoma is the deadliest type of skin cancer. Melanoma commonly begins with changes in an old or normal mole or suddenly appears as a black or dark brown area on the skin. Early detection is very important [4]. Here are the important points in early diagnosis. There is usually no symmetry in the skin lesion, and one half of the lesion is different from the other. The margin of the lesion is not clear and is not well diagnosable from the surrounding healthy skin. The lesion may be seen in several different colors. Brown and black are more common, and red, white, and blue are less common [5]. The diameter of the lesions is usually more than six millimeters. The last important symptom is the difference between a fleshy spot or a brown spot and other spots on the body, which usually begin to change shape and color or cause itching and bleeding [6].

Melanoma, as one of the skin cancers, is responsible for 50% of deaths associated with skin cancers. The source of this disease is in the epidermal and dermal layers of the skin [7]. This disease is formed by the accumulation of melanin granules, and it is spread to the outermost layer of the skin.

Since determining the extent of the lesion and somehow extracting the exact boundary between the lesion and the background is one of the most influential parameters in clinical methods of diagnosing cancerous lesion, therefore, providing lesion diagnosis methods is of great importance [8].

Goldman first invented the dermatoscope, which evaluates changes in skin pigmentation under disease conditions. Dermatologists then used this device as a suitable and noninvasive tool to observe skin lesions. With the development of science in recent years, the digital dermatoscope has been replaced by conventional dermatoscopes with the ability to capture and store skin images. Therefore, it is possible to provide commercial software packages to help diagnose some skin lesions, create databases, and create a medical resume for each patient [9].

Melanoma diagnosis by modern smart devices, in the early stages of its formation, is a very important issue that has involved many researchers [10]. Meanwhile, melanoma due to its high prevalence has attracted a large amount of research. For example, Razmjooy et al. presented a new well-organized technique to diagnose the melanoma from dermoscopy images. Firstly, the additional scales were removed using edge detection and smoothing. Then, the data were segmented, and mathematical morphology was used as postprocessing section. They also used an optimized multilayer perceptron (MLP) for the final classification. The optimization was based on the World Cup Optimization (WCO) Algorithm. Simulation results indicated the high accuracy of the system based on the authors explanations [11].

Pathan et al. proposed an automatic system for the diagnosis of the melanoma. The method contained a set of sequential stages. 2D Gabor filters were used for detection of the pigment network masks. A machine learning based on the rules was then produced from the pigment network masks to detection of the typical and atypical pigment network patterns. The method was implemented on PH2 dataset. The method was also compared with some state-of-the-art techniques from the literature to show the method effectiveness [12].

Xu et al. presented another automatic CAD-based technique for early diagnosis of melanoma. In the method, the noise of the input images was first reduced by a median filter. Then, a new optimized method for segmentation of the region of interest was proposed based on the convolutional neural network (CNN). To improve the CNN efficiency, it is optimized based on the satin bowerbird optimization (SBO) algorithm. Afterward, the main features of the region of interest were extracted for the next step. Because of the large volume of the features, an optimized feature selection based on the SBO algorithm was also utilized. At last, the images were classified based on a support vector machine (SVM). Final results were implemented on the American Cancer Society database, and the results were compared with some state-of-the-art methods to show the method dominance [13].

As can be observed from the literature, several and different techniques were proposed for skin cancer diagnosis. However, providing an optimal technique for the diagnosis system is not too simple, and each of them has its advantages and shortages. Moreover, using deep learning for this purpose is also too useful. However, using the combination of these works is not provided or may be too rare. Therefore, in this paper, a new diagnosis method was proposed for melanoma detection. The method extracted features of the images by the help of a CNN that is done from the melanoma dermoscopic images. Afterward, the Water Strider Algorithm (WSA) is for optimal selection of the only relevant features. Here, we utilized WSA as a feature selector from deep features which are generated by the CNN with large redundancy. The results showed that this process can develop the diagnosis accuracy of the melanoma. The next sections of this paper are organized as follows. In Section 2, the main materials and methods are explained. This section contains the structure of the CNN and how the CNN can be used for the features extraction. Section 3 describes about the definition of the Wildebeest Herd Optimization Algorithm. In Section 4, the utilized dataset has been described. Section 5 includes the simulation results and the discussions about the results. The paper has been concluded in Section 6.

2. The Convolutional Neural Network

2.1. CNN-Based Features Extraction. The present paper uses a convolutional neural network (CNN) for establishing the feature extraction of the melanoma dermoscopic images. A pretrained model of CNN has been trained on the dataset that includes several of images on different classes. So, the transfer learning has been implemented by transferring weights which were already trained and reserved into the structure of the pretrained model, like Inception, in this paper. Three main parameters are considered for the CNN: Filter size, Stride, and Max pooling. The present study uses size 2 filters, and also size 2 stride parameter and 2×2 Max pooling filter are employed. The main target of convolutional layers here is to provide a proper features extraction from the input images. Here, we used different convolutional layers that have been implemented to extract various features like texture, edges, high-lighted patterns, and colors from the images.

Three fully connected layers are then used for classification. We also used Softmax as activation function to provide a binary output (positive melanoma and negative melanoma). The last concatenation layer that includes the extracted features has been kept, and the top layers such as the Flatten, Drop out, and the Dense layers which the later performs classification have been removed. The weights have been updated by RMSprop optimizer, and cross entropy loss function and learning rate are set as $1e-4$. The input shape for the feature extractor (Inception) is [225, 225, 3]. Figure 1 shows an overview of Inception network [14].

A significant disadvantage of the pretrained models like Inception is that their construction needs a big memory as well as storage capacity, that makes a high time complexity for the system. Therefore, some statistical operations are added to ignore immaterial features. In the following, these features have been briefly given [14].

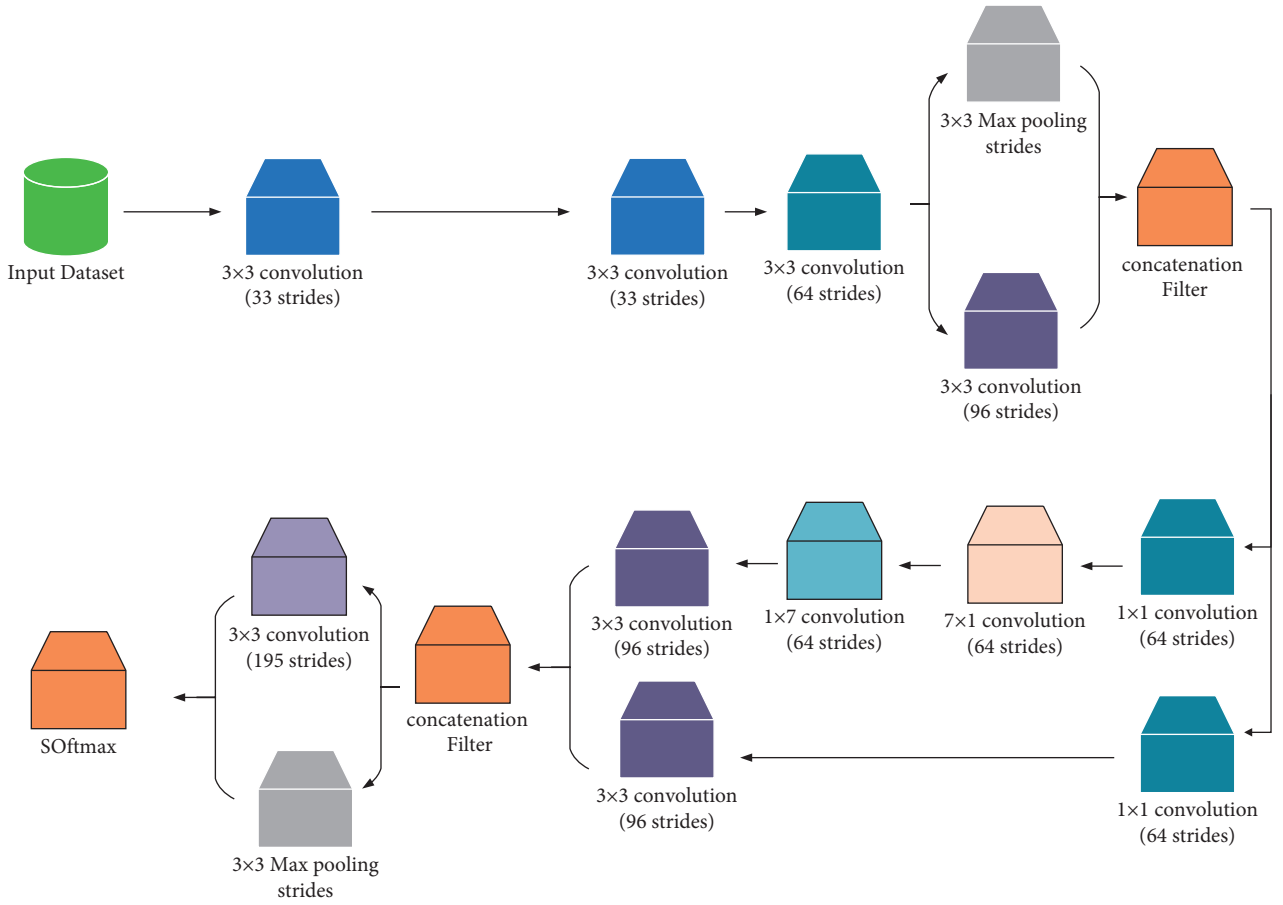


Figure 1: An overview of Inception network [14].

2.1.1. Tree-Based Classifier. This feature is one of the most popular methods to advance the classification ratio. In addition to high accuracy, they are simple and robust. For all decision trees, node importance has been evaluated by Gini importance, and this is given in the following formula:

$$n_k = w_k C_k - w_k^{\text{left}} C_k^{\text{left}} - w_k^{\text{right}} C_k^{\text{right}}, \tag{1}$$

where w_k signifies the weighted number of samples that reaches the node k, n_k describes the importance of node l, C_k defines the contamination value of the k^{th} node, and left and right represent the child nodes from the left and the right splits on the k^{th} node, respectively. The importance of the features is then evaluated.

$$f_i = \frac{\sum_{j:\text{node } k \text{ splits on feature } i} n_k}{\sum_{j \in \text{all nodes}} n_j}, \tag{2}$$

where n_k describes the importance of the k^{th} node and f_i describes the importance of feature i.

Normalize the values in the range [0, 1] by dividing the sum of all feature importance values, i.e.,

$$\overline{f_i} = \frac{f_i}{\sum_{k \in \text{all nodes}} f_k}. \tag{3}$$

Finally, the sum of the importance value in the features for the trees is evaluated and then divided by the total number of trees, i.e.,

$$Rf_i = \frac{\sum_{k \in \text{all trees}} \overline{f_i}}{T}, \tag{4}$$

where Rf_i describes the importance of the i^{th} feature that are evaluated from all trees, where $\overline{f_i}$ describes the normalized feature importance for the i^{th} feature in the j^{th} tree and T defines the total number of trees.

2.1.2. Chi-Square. This feature has been implanted to eliminate the features with high correlation values by evaluating their dependency. This is evaluated on all classes between each feature:

$$\psi^2 = \sum_{j=1}^{n} \frac{\left(Y_j - E_j\right)^2}{E_j}, \tag{5}$$

where Y_j and E_j represent the real and the expected feature values, respectively.

Here, after establishing Chi-square, the feature vector has been minimized for the dataset.

3. Wildebeest Herd Optimization Algorithm

The main target of optimization is to provide the most desirable solution to the problem by considering its limitations and other features. Numerous solutions have been introduced for solving an optimization problem. Formerly, classical methods like Pontryagin maximum principle [15] and distributed newton method [16] were used for this purpose. However, by increasing the complexity of the problems, the classical methods have been weakened such that they failed in solving some kinds of NP-hard problems. This made the researchers to move toward and offer newer methods to resolve the classical optimization methods' issues. Metaheuristic algorithms are some kinds of new optimization algorithms that recently are used extensively for this purpose. In recent years, lots of different versions for metaheuristic algorithms have been introduced, for example, Manta-Ray Foraging Optimization (MRFO) [17], World Cup Optimization (WCO) Algorithm [18], Locust Swarm Optimization (LSO) Algorithm [19], and Wildebeest Herd Optimization (WHO) Algorithm [20].

In this study, the Wildebeest Herd Optimization (WHO) Algorithm [20] has been utilized for optimizing of our deep network. The main reason for selecting this algorithm is that it is too recent among different types of the metaheuristic algorithms. Also, its results on the benchmark functions based on the paper provide too better results. This makes us to use this metaheuristic technique to improve the efficiency of the proposed CNN. The WHO algorithm has been inspired from the behavior of food searching by the Wildebeests. Wildebeests are social mammal animals which travel to find food sources. The male sex challenges with other competitors to get females for mating.

The Wildebeest Herd Optimization (WHO) Algorithm starts by random initializing a number of population (wildebeests) as candidates. The population is limited between the lower ($X_{\min}$), and the higher ($X_{\max}$) boundaries, i.e.,

$$X_i \in [X_{\min}, X_{\max}], \tag{6}$$

where $i = 1, 2, \ldots, N$.

The milling action is then used for local movement of the wildebeest individuals. The phase is modeled by considering a consistent number (n) as the minor random movement based on the solution spaces and continuing to find the optimal position. A random phase Z_n has been employed by the candidates in position X that should frequently search for the small random phase positions. A tunable length is accomplished by a random step size to the candidates. Consequently, the local experimental phase Z_n is obtained in the following formula:

$$Z_n = X_i + \varepsilon \times \theta \times \nu, \tag{7}$$

where ε signifies the learning rate variable, X_i describes the candidate number i, θ represents a randomly uniform value ranged between 0 and 1, and ν defines a random unit vector.

Then, a constant number (n) of minor random candidates is evaluated, and the wildebeest updates its position to get an optimal random location, i.e.,

$$X_i = \alpha_1 \times Z_n^* + \beta_1 \times (X_i - Z_n^*), \tag{8}$$

where α_1 and β_1 describe the leader variables to lead the local movement of the candidates.

The last phase is to model the swarm instinct of the wildebeests. This is simulated once the other candidates are positioned in a location with proper food source, i.e.,

$$X_i = \alpha_2 \times X_i + \beta_2 \times X_h, \tag{9}$$

where X_h defines a random candidate and α_2 and β_2 represent the leader variables to lead the local movement of the crew.

In the Wildebeest Herd optimization algorithm, there is another term to avoid the candidates from moving to spaces with insufficient food source. This term is mathematically modeled as follows:

$$X_i \coloneqq X_i + \theta \times (X_{\max} - X_{\min}) \times \overline{\nu}, \tag{10}$$

where $\overline{\nu}$ describes a random unit vector.

Another term in the algorithm is to simulate the crowded spaces. The crowd happens when there is a wide fertility for the grass. This term is named "individual pressure." Based on this term, a challenge is accomplished and the strongest candidate demolishes the others using the following equation:

$$\begin{aligned} &\text{if}\left(\left\|X^* - X_i\right\|\right) < \eta, \quad \left(\left\|X^* - X_i\right\|\right) > 1 \\ &\text{then: } X_i = X^* + \varepsilon \times \widehat{n}, \end{aligned} \tag{11}$$

where η signifies a threshold to avoid crowding in the spaces and $\widehat{n}$ describes exploitable parts number close to the optimum solution point.

And the final phase is to simulate the swarm social memory to offer better positions that is called swarm social memory and is obtained by the following equation:

$$X = X^* + 0.1 \times \widehat{\nu}. \tag{12}$$

By combination of the concept of the Wildebeest Herd Optimization Algorithm with the Inception network, the total method of the proposed system has been determined as follows:

(1) Extracting deep features from Inception.

(2) Initializing solutions for the Wildebeest Herd. The Herd follows the leader during searching the fertile grass source (solution space).

(3) Attempting the wildebeest to catch the best fertile food source based on exploration and exploitation terms in the algorithm.

(4) All stages are repeated until the termination criteria have been reached.

4. The Database

The present study uses International Skin Imaging Collaboration (ISIC-2008) for validation. This archive defines an

international attempt to progress the research studies for the melanoma diagnosis which is supported by the International Society for Digital Imaging of the Skin (ISDIS). This dataset includes the largest widely available collection of dermoscopic images for skin lesions. The dataset includes over 13,000 dermoscopic images that have been gathered from international clinical centers and collected from different devices within the centers. This international collaboration is to guarantee to collect a proper dataset. Also, a subset of the images has been labeled by the experts. Figure 2 shows some examples of the ISIC-2008 dataset [21].

5. Simulation Results

Because the WHO algorithm is a metaheuristic and has stochastic nature, the achieved results of this algorithm may change during different executions. Consequently, in this study, the training method by the WHO algorithm has been repeated for 30 times and the best network has been used for the analysis. The WHO algorithms and the Inception network have been programmed in MATLAB 2019b 64 bit version and executed on computation environment of Intel Core i7Intel processor CPU 2.00 GHz, 2.5 GHz, and 32 GB RAM with and two SLI GeForce Titan GPUs and 64 bit operating system. The convolutional neural networks were under Google lab network.

The section confirms and discusses about the results of the proposed skin cancer diagnosis system. The method is performed to the ISIC-2008 dataset. For the dataset, the segmented cancer has been assessed by a comparative study with the expert's ground truth image.

To assess the efficiency of the suggested method, the average value of both best values and the worst values (Max) along with standard deviation (Std) and computational time has been evaluated. The mathematical definitions of all indicators are given as follows:

$$\begin{aligned} \mathrm{ACC} &= \frac{(\mathrm{TP}+\mathrm{TN})}{(\mathrm{FP}+\mathrm{TP}+\mathrm{TP}+\mathrm{FN})}, \\ \mathrm{SPC} &= \frac{\mathrm{TN}}{\mathrm{FP}+\mathrm{TN}}, \\ \mathrm{SNS} &= \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \\ F_{\mathrm{score}} &= 2\times\frac{\mathrm{SPC}\times\mathrm{SNS}}{\mathrm{SPC}+\mathrm{SNS}}, \end{aligned} \tag{13}$$

where TP is the true positive, TN is the true negative, FP is the false positive, and FN is the false negative, and ACC, SPC, and SNS represent the accuracy, specificity, and sensitivity, respectively.

And the optimization is based on the following equation:

$$\begin{aligned} \mathrm{Best}_{\mathrm{ACC}} &= \mathrm{Max}\,\mathrm{ACC}, \\ \mathrm{fit}_{\mathrm{Best}} &= \mathrm{Min}\,\mathrm{Fit}_i, \\ \mathrm{fit}_{\mathrm{worst}} &= \mathrm{Max}\,\mathrm{Fit}_i, \\ \mathrm{fit}_{\mathrm{avg}} &= \frac{1}{m}\sum_{i=1}^{N}\mathrm{Fit}_i, \\ \mathrm{Std} &= \sqrt{\frac{1}{m-1}\times\sum_{i=1}^{m}\left(\mathrm{Fit}_i-\mathrm{fit}_{\mathrm{avg}}\right)^2}, \\ & \qquad i=1,2,\ldots,m, \end{aligned} \tag{14}$$

where m describes the number of evaluations and Fit_i signifies the fitness function value.

Once Inception trains the skin cancer images continuously in the epochs, these rapid ups and downs are gradually minimized in the future training part.

After feature extraction based on Inception, the WHO algorithm has been performed to optimal selection of the features and to eliminate the useless features. Here, the results of the proposed WHO-Inception have been compared with some state-of-the-art optimization algorithms including Chimp optimization algorithm (ChOA) [22], Biogeography-Based Optimizer (BBO) [23], and Locust Swarm Optimization (LS) [24] for feature selection. The parameter setting of the utilized optimization algorithms is reported in Table 1.

Here, the feature selection methods have been employed for optimal selection of the generated feature vector from Inception to provide just the useful relevant features. Some of the parameters are selected similar to fair comparison. All of the algorithms have 15 iterations and 30 number of candidates. All of the algorithms are run for 35 independent runs to offer consistent results. Table 2 reports the results of four measurement indicators applied by the before mentioned optimization algorithms of the feature selection for the dataset.

As can be observed from Table 2, the proposed WHO algorithm has the higher efficiency toward the other comparative algorithms in all four metrics followed by LS, ChOA, and BBO, respectively. The best results of $\mathrm{fit}_{\mathrm{worst}}$ indicated that the proposed WHO algorithm provides the best value of the fitness function against the others. Also, LS is placed on the second rank. Figure 3 shows the running time of the analyzed algorithms.

As can be observed from Figure 3, the proposed WHO algorithm considers the fastest results among all of the compared algorithms followed by ChOA and BBO, respectively, while LS has the slowest results. Figure 4 shows the simulation results for number of selected features based on the studied comparative algorithms.

As can be observed from Figure 4, the proposed WHO algorithm with 210 features provided less features than other algorithms. BBO and LS are placed in the second and the third ranks with 314 and 352 selected features, respectively. The largest features have been selected by ChOA.

In the following part, the image similarity and statistical measures like dice coefficient (DC), accuracy (ACC), specificity (SPC), PPV, NPV, and sensitivity (SNS) are evaluated. The mathematical model of the parameters is given as follows:

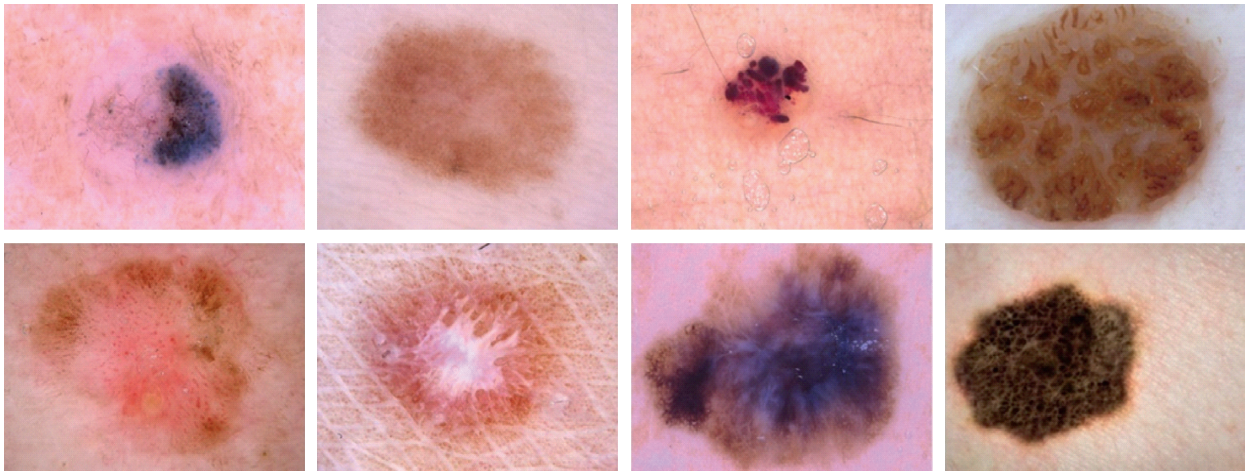


Figure 2: Some samples of the ISIC-2008 dataset [21].

Table 1: The parameter setting of the utilized optimization algorithms.

Algorithm	Parameter	Value
ChOA [22]	r_1 and r_2	Random
	m	Chaotic
BBO [23]	Habitat modification probability	1
	Immigration probability bounds per gene	[0, 1]
	Step size for numerical integration of probabilities	1
	Max immigration (I) and Max emigration (E)	1
	Mutation probability	0.005
LS [24]	F	0.6
	L	1
	g	20
WHO [20]	Inertia	0.2
	P_{Cross}	0.7
	P_{Mut}	0.15
	$\text{Tour}_{\text{size}}$	0.7

Table 2: Results of feature selection by the optimization algorithms.

Algorithm	Measurement indicator			
	fit_{avg}	Std	fit_{Best}	$\text{fit}_{\text{worst}}$
ChOA [22]	0.215	0.009	0.194	0.308
BBO [23]	0.597	0.034	0.351	0.640
LS [24]	0.085	0.009	0.041	0.062
WHO	0.037	0.005	0.039	0.051

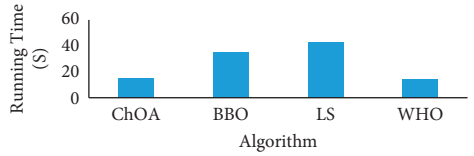


Figure 3: The running time of the analyzed algorithms.

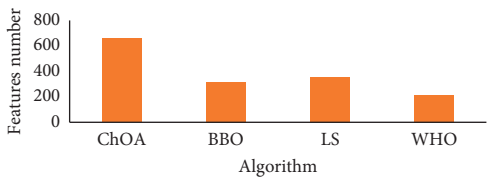


Figure 4: The simulation results for number of selected features based on the studied comparative algorithms.

Table 3: The simulation results of the suggested method compared with other state-of-the-art methods.

Method	Performance metric					
	DS	ACC	SNS	SPC	NPV	PPV
Dorj's [25]	0.91	0.90	0.93	0.91	0.95	0.85
Linsangan's [26]	0.89	0.84	0.19	0.64	0.87	0.69
Thanh's [6]	0.87	0.85	0.87	0.88	0.88	0.81
Khan's [2]	0.88	0.73	0.86	0.63	0.85	0.62
Angurana's [27]	0.91	0.76	0.84	0.79	0.88	0.78
Proposed method	0.94	0.96	0.96	0.95	0.97	0.89

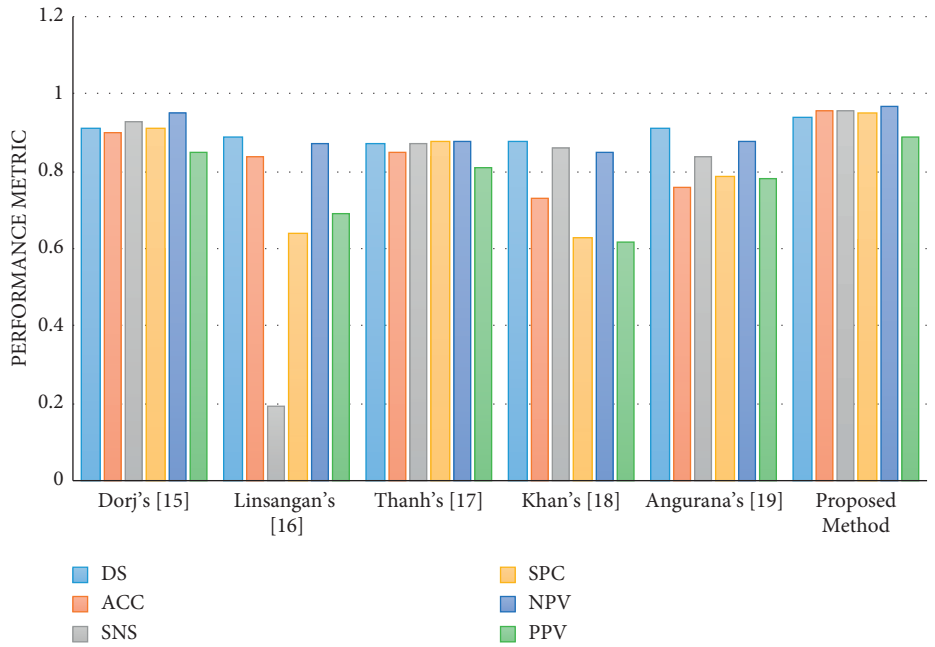


Figure 5: The simulation results of the suggested method compared with other state-of-the-art methods.

$$\text{Ds} = \frac{2 \times \text{TP}}{((\text{FP} + \text{TP}) + (\text{TP} + \text{FN}))},$$

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{15}$$

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}},$$

where TP is the true positive, TN is the true negative, FP is the false positive, and FN is the false negative.

The final results are compared with 5 state-of-the-art methods from the literature including Dorj's [25], Linsangan's [26], Thanh's [6], Khan's [2], and Angurana's [27]. Table 3 reports the simulation results of the suggested method compared with other state-of-the-art methods.

Figure 5 shows the bar plot of the classification rate. It can be observed from the results that the presented methodology offers the highest accuracy compared with the other methods from the literature.

As can be observed from Table 1, the suggested technique with 96% accuracy offers the uppermost efficiency among the other comparative methodologies. Similarly, Dorj's method and Angurana's method with 91% have the second rank in the comparison. Likewise, Linsangan's, Khan's, and Thanh's methods with 89%, 88%, and 87%, respectively, are in the next ranks. Also, the suggested method with 96% sensitivity as the highest value against the others indicates the method's higher consistency than the others. Furthermore, higher value of PPV and NPV for the suggested method against the others designates its higher condition occurrence for handling the likelihood of a test identifying the cancer. In addition, higher results for the specificity of the suggested technique indicate its sophisticated occurrence-independent achievement.

6. Conclusions

Skin cancers are one of the most common types of cancer. At least 40% of all cancers worldwide are skin cancers. This cancer is on the rise with a terrible trend. Twenty percent of Americans have the disease, of which more than 36,000 are women and have been diagnosed with melanoma. Therefore, early detection of this cancer can play a significant role incurring this cancer and survive from its dangers. Recently, for reducing the experts' errors, image processing techniques have been employed for these purposes. These techniques

can play a complementary role alongside of the experts. Therefore, in the present study, an optimal computer-aided diagnosis system based on deep learning was proposed for early detection of the skin cancer. The idea was to propose an optimized version of Inception convolutional neural network for features extraction from the skin cancer images. The results features were then pruned by the Wildebeest Herd Optimization (WHO) Algorithm as a feature selection technique. Using the WHO algorithm is to select just the useful features and to neglect the irrelevant ones. After designing the method, it was applied to a standard dataset, namely, International Skin Imaging Collaboration (ISIC-2008) dataset with over 13,000 dermoscopic images. The simulation results of the proposed method were compared with Chimp optimization algorithm (ChOA), Biogeography-Based Optimizer (BBO), and Locust Swarm Optimization (LS) for defining the feature selection efficiency. The results indicated that the proposed method provides better effectiveness in terms of accuracy, minimizing the number of extracted features and the speed. The final diagnosis system were also compared with some well-known methods including Dorj's, Linsangan's, Thanh's, Khan's, and Angurana's methods to show its dominance toward the studied methods. According to the results, the suggested method based on the EHO algorithm and Inception CNN provides the best results among all studied methods. However, the results showed the best results for the proposed WHO algorithm, and some other new metaheuristics can be testified to solve the problems like monarch butterfly optimization (MBO), earthworm optimization algorithm (EWA), elephant herding optimization (EHO), moth search (MS) algorithm, Slime mould algorithm (SMA), and Harris hawks optimization (HHO) which are a good motivation for the future work.

Data Availability

The data for the dataset can be obtained in https://www.isic-archive.com/.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] S. Sugiarti, Y. Yuhandri, J. Na'am, D. Indra, and J. Santony, "An artificial neural network approach for detecting skin cancer," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 17, no. 2, pp. 788–793, 2019.

[2] M. Q. Khan, A. Hussain, S. U. Rehman et al., "Classification of melanoma and nevus in digital images for diagnosis of skin cancer," *IEEE Access*, vol. 7, pp. 90132–90144, 2019.

[3] M. Kumar, M. Alshehri, R. AlGhamdi, P. Sharma, and V. Deep, "A DE-ANN inspired skin cancer detection approach using fuzzy c-means clustering," *Mobile Networks and Applications*, vol. 25, no. 4, pp. 1319–1329, 2020.

[4] A. Arshaghi, M. Ashourian, and L. Ghabeli, "Detection of skin cancer image by feature selection methods using new Buzzard optimization (BUZO) algorithm," *Traitement du Signal*, vol. 37, no. 2, pp. 181–194, 2020.

[5] Q. U. Ain, H. Al-Sahaf, B. Xue, and M. Zhang, "A genetic programming approach to feature construction for ensemble learning in skin cancer detection," in *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, pp. 1186–1194, Cancún, Mexico, July 2020.

[6] D. N. Thanh, V. S. Prasath, and N. N. Hien, "Melanoma skin cancer detection method based on adaptive principal curvature, colour normalisation and feature extraction with the ABCD rule," *Journal of Digital Imaging*, vol. 33, no. 3, pp. 574–585, 2019.

[7] K. Manasa and D. G. V. Murthy, "Skin cancer detection using VGG-16," *European Journal of Molecular and Clinical Medicine*, vol. 8, no. 1, pp. 1419–1426, 2021.

[8] S. Jain, V. Jagtap, and N. Pise, "Computer aided melanoma skin cancer detection using image processing," *Procedia Computer Science*, vol. 48, pp. 735–740, 2015.

[9] A. Hekler, J. S. Utikal, A. H. Enk et al., "Superior skin cancer classification by the combination of human and artificial intelligence," *European Journal of Cancer*, vol. 120, pp. 114–121, 2019.

[10] M. Anas, K. Gupta, and S. Ahmad, "Skin cancer classification using K-means clustering," *International Journal of Technical Research and Applications*, vol. 5, no. 1, pp. 62–65, 2017.

[11] N. Razmjooy, F. R. Sheykhahmad, and N. Ghadimi, "A hybrid neural network-world cup optimization algorithm for melanoma detection," *Open Medicine*, vol. 13, pp. 9–16, 2018.

[12] S. Pathan, K. G. Prabhu, and P. C. Siddalingaswamy, "A methodological approach to classify typical and atypical pigment network patterns for melanoma diagnosis," *Biomedical Signal Processing and Control*, vol. 44, pp. 25–37, 2018.

[13] Z. Xu, F. R. Sheykhahmad, N. Ghadimi, and N. Razmjooy, "Computer-aided diagnosis of skin cancer based on soft computing techniques," *Open Medicine*, vol. 15, no. 1, pp. 860–871, 2020.

[14] A. T. Sahlol, D. Yousri, A. A. Ewees, M. A. A. Al-Qaness, R. Damasevicius, and M. A. Elaziz, "COVID-19 image classification using deep features and fractional-order marine predators algorithm," *Scientific Reports*, vol. 10, no. 1, pp. 15364–15415, 2020.

[15] R. E. Kopp, "Pontryagin maximum principle," *Mathematics in Science and Engineering*, vol. 5, pp. 255–279, 1962.

[16] A. Jadbabaie, A. Ozdaglar, and M. Zargham, "A distributed Newton method for network optimization," in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) Held Jointly with 2009 28th Chinese Control Conference*, pp. 2736–2741, IEEE, Shanghai, China, 2009.

[17] W. Zhao, Z. Zhang, and L. Wang, "Manta ray foraging optimization: an effective bio-inspired optimizer for engineering applications," *Engineering Applications of Artificial Intelligence*, vol. 87, Article ID 103300, 2020.

[18] R. Navid, "A new meta-heuristic optimization algorithm inspired by FIFA world cup competitions: theory and its application in PID designing for AVR system," *Journal of Control, Automation and Electrical Systems*, vol. 27, no. 4, pp. 419–440, 2016.

[19] I. Benmessahel, K. Xie, M. Chellal, and T. Semong, "A new evolutionary neural networks based on intrusion detection systems using locust swarm optimization," *Evolutionary Intelligence*, vol. 12, no. 2, pp. 131–146, 2019.

[20] D. Amali and M. Dinakaran, "Wildebeest herd optimization: a new global optimization algorithm inspired by wildebeest herding behaviour," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 6, pp. 8063–8076, 2019.

[21] ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection, 2008, https://challenge2018.isic-archive.com.
[22] M. Khishe and M. R. Mosavi, "Chimp optimization algorithm," *Expert Systems with Applications*, vol. 149, Article ID 113338, 2020.
[23] D. Simon, "Biogeography-based optimization," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 6, pp. 702–713, 2008.
[24] E. Cuevas, F. Fausto, and A. González, "The locust swarm optimization algorithm," in *New Advancements in Swarm Algorithms: Operators and Applications,* Springer, Cham, Switzerland, 2020.
[25] U.-O. Dorj, K.-K. Lee, J.-Y. Choi, and M. Lee, "The skin cancer classification using deep convolutional neural network," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 9909–9924, 2018.
[26] N. B. Linsangan, J. J. Adtoon, and J. L. Torres, "Geometric analysis of skin lesion for skin cancer using image processing," in *Proceedings of the 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, pp. 1–5, Baguio City, Philippines, 2018.
[27] N. Angurana, A. P. Rajan, and I. Srivastava, "Skin cancer detection and classification," *International Journal of Engineering and Management Research*, vol. 9, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 4423407, 14 pages
https://doi.org/10.1155/2021/4423407

Research Article

Machine Learning-Based Facial Beauty Prediction and Analysis of Frontal Facial Images Using Facial Landmarks and Traditional Image Descriptors

Tharun J. Iyer,[1] **Rahul K.**,[2] **Ruban Nersisson**,[1] **Zhemin Zhuang**,[2] **Alex Noel Joseph Raj**,[2] **and Imthiaz Refayee**[3]

[1]*School of Electrical Engineering, Vellore Institute of Technology, Vellore 632014, India*
[2]*Department of Electronic Engineering, Shantou University, Shantou 515063, China*
[3]*Chairman NB Multispecialty Dental Hospital, Chennai, India*

Correspondence should be addressed to Zhemin Zhuang; zmzhuang@stu.edu.cn

Received 8 May 2021; Accepted 13 August 2021; Published 26 August 2021

Academic Editor: Ezequiel López Rubio

The beauty industry has seen rapid growth in multiple countries and due to its applications in entertainment, the analysis and assessment of facial attractiveness have received attention from scientists, physicians, and artists because of digital media, plastic surgery, and cosmetics. An analysis of techniques is used in the assessment of facial beauty that considers facial ratios and facial qualities as elements to predict facial beauty. Here, the facial landmarks are extracted to calculate facial ratios according to Golden Ratios and Symmetry Ratios, and an ablation study is performed to find the best performing feature set from extracted ratios. Subsequently, Gray Level Covariance Matrix (GLCM), Hu's Moments, and Color Histograms in the HSV space are extracted as texture, shape, and color features, respectively. Another ablation study is performed to find out which feature performs the best when concatenated with the facial landmarks. Experimental results show that the concatenation of primary facial characteristics with facial landmarks improved the prediction score of facial beauty. Four models are trained, K-Nearest Neighbors (KNN), Linear Regression (LR), Random Forest (RF), and Artificial Neural Network (ANN) on a dataset of 5500 frontal facial images, and amongst them, KNN performs the best for the concatenated features achieving a Pearson's Correlation Coefficient of 0.7836 and a Mean Squared Error of 0.0963. Our analysis also provides us with insights into how different machine learning models can understand the concept of facial beauty.

1. Introduction

Facial beauty has long been a topic of intellectual discussion and its various attributes have been researched upon and studied. In medieval times, renaissance painters used unique ratios named "The Golden Ratios" to represent through paintings what the perfectly shaped human face would look like [1]. The Golden Ratios are ratios based on the value of 1.6, which was considered by the Greeks to be a perfect number. Many examples exist in architecture, too, from the Greek Empire, where the golden ratio was used in buildings and pantheons. This golden ratio was applied to facial beauty, where different facial ratios were calculated and compared against the value of 1.6. Although facial attractiveness can be debated on its objectivity or subjectivity, recent empirical results support the idea that attractiveness is objective and quantifiable, which is achieved by measuring cross-cultural differences [2], brain activity [3], and cognitive psychology [4]. Studies from medicine and psychology have also concluded that aesthetic features like facial averageness [5] and symmetry [6] are important when assessing attractiveness. Facial skin colors and texture also significantly contribute towards facial attractiveness [7, 8] and have been included in assessments of facial attractiveness [9, 10]. Moreover, research has shown that attractive faces follow defined ratios of facial proportions, such as

neoclassical canons [11] and the golden ratio [12], which is considered the golden standard for beautiful faces since ancient times by artists, physicians, and cosmetic surgeons [13]. In machine learning, several methods have been proposed to assess facial attractiveness by using these facial features. However, due to less efficient feature extraction methods and the inability to combine various features together, this method has not been worked upon, even though this approach may close the gap between human and machine performance. Hence, there is a need for an efficient technique of facial beauty assessment from a machine's perspective. In this paper, a technique for assessing facial beauty based on facial proportion factors is developed, which is widely believed to be the gold standard for facial beauty.

Research conducted in Psychology and Biology settles the problem by making a hypothesis of which facial features contribute to attractiveness. Various features like sexual dimorphism, averageness, and symmetry influence the perception of beauty. Jones and Jaeger [10] showed that women appear more attractive based on these three features. Thornhill and Gangestad [14] proved that there is a significant correlation between average faces and facial beauty, but the most attractive faces are not average. Facial symmetry, although, does increase with facial averageness, as stated in Grammer and Thornhill [15], which also supports the same idea from evolutionary biology studies. Moreover, sexual dimorphism is shown to play a significant role in assessing a person's facial beauty, as stated by Perrett et al. [16]. Sexual dimorphism is the presence of secondary sexual characteristics which appear during beauty. These secondary characteristics make people appear more masculine or feminine. Many studies provide evidence that masculinity and femininity provide more depth to a person's beauty than symmetry [10, 17, 18]. Intrinsic features of the face, such as facial texture and skin color, can also affect the perception of beauty. Many researchers have proposed a connection between facial beauty and healthy skin, which consequently proves that the health of facial skin might be a surface-level feature that influences beauty assessments. Fink et al. [7] assessed facial beauty through the use of human ratings from facial textures and skin color. Fink et al. [19] show that the health of facial skin is positively correlated with the attractiveness index of the face. Also, since facial attractiveness is affected by various factors, facial shape features and appearance have also been considered for facial beauty assessments. Kagian et al. [9] analyzed facial beauty depending on the shape and facial geometry. Rhazi et al. [20] proposed a method to predict facial beauty based on golden ratios calculated from the extracted feature corners. Schmid et al. [21] have proposed a model to calculate facial beauty based on golden ratios, symmetry, and neoclassical canons. The neoclassical canons are ratios used by medieval painters in their paintings to represent their understanding of human beauty. Dornaika et al. [22] developed a semisupervised face beauty prediction technique using a graph-based method with a public dataset. Though semisupervised techniques require fewer training images, their model accuracy entirely depends on the graph density, which affects the prediction results. Lin et al. [23] used Attribute aware CNN to predict the facial beauty, with SCUT-FPB5500 dataset and trained with powerful GPU support. Xiao et al. [24] developed Beauty3DFaceNet, which is comprised a deep CNN to predict the attractiveness of 3D faces. They collected the 3D point cloud and facial texture of the image to train their network which will output attractiveness score. Although the approach is promising, they are limited by the data available to collect the 3D point cloud and also the training requires extensive computing. Wei et al. [25] assessed facial symmetry and attractiveness based on SVM and linear regression using a predefined dataset. Also, they have developed a mobile app based on their concluded features which are useful for plastic surgeons to plan reconstructive facial surgeries. Tong et al. [26] investigated facial attractiveness using facial putative ratios and DNN by training with 4512 face images. The DNN model was trained using NVIDIA Tesla K40 GPU. Recent research takes advantage of CNN and DNN to predict facial attractiveness, which is computationally expensive and requires a large number of training and testing datasets. Since facial beauty prediction mainly depends on how effectively the facial features are analyzed, it is required to determine which features influence the most. Thus, by determining the effective features, facial beauty prediction could be achieved with less computationally expensive state-of-the-art machine learning techniques.

This paper aimed to predict the facial beauty of frontal images using machine learning techniques and traditional feature extraction methods. The proposed technique improves the performance of machine learning models on facial images through the fusion of facial landmarks and Texture, Color, and Shape Features. Nineteen facial ratios based on Golden Ratios and Symmetry are used in this paper. These facial ratios are fed as input to four regression-based machine learning models (Linear Regression, Random Forest, K-Nearest Neighbors, and an Artificial Neural Network) and trained to predict the beauty of the facial image. An Ablation Study is performed on the nineteen ratios to find out the best performing combination of ratios called the "Feature Set." The Feature Set is concatenated with texture, color, and shape features (TCS Features). Also, another ablation study is performed to check the performance of each model in accurately predicting facial beauty. The ablation study is used to find out which of the TCS Features contributes to the prediction and how the model performance varies among each combination. The rest of the paper is organized as follows: At first, the steps to predict the facial beauty (extraction of the features from the dataset) are provided. Then, the discussion of various models used to predict facial beauty has been provided. Finally, the results of the implementation of various models and the best performing model are discussed.

2. Materials and Methods

The following steps were used to predict the facial beauty score (Figure 1 represents the block diagram of the entire process):

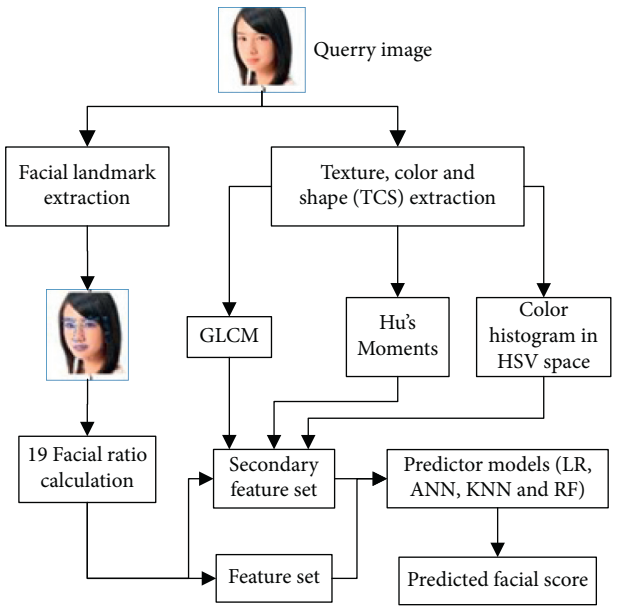


Figure 1: Block diagram of the analysis performed.

(i) Extraction of facial landmarks, texture, color, and shape (TCS) features.

(ii) Extraction of best performing facial landmark feature set.

(iii) Ablation study of TCS Features.

2.1. Dataset Considered. The dataset used in this study is SCUT-FBP5500, a dataset that consists of 5500 images of Asian and Caucasian males and females [27], where the dimension of each image is 350×350 pixels. The dataset contains 5500 frontal, unoccluded faces aged from 15 to 60 with a neutral expression. It can be divided into four subsets with different races and genders, including 2000 Asian females, 2000 Asian males, 750 Caucasian females, and 750 Caucasian males. All the images are labeled with beauty scores ranging from 1–5 by a total of 60 volunteers aged from 18–27 (average 21.6), where the beauty score 5 means most attractive and a score of 1 or less means least attractive. The metrics used in this study to measure performance are Pearson's Correlation Coefficient (PC), Mean Absolute Error (MAE), Mean Squared Error (MSE), and $R2$ Score. All the models are trained on the same dataset using Python 3.7 software.

2.1.1. Facial Landmark Localization. Facial ratios are calculated by measuring the distance between certain points on the face image. These points are called Facial landmarks, which are coordinates on the face image. The SCUT-FBP dataset contains predefined facial landmarks for all the 5500 images. Each image had a total of 86 landmarks that covered the most important points in the face. These facial landmarks were used to calculate the nineteen facial ratios used in this analysis. Figure 2 illustrates the facial landmarks (Figure 2(b)) for the input sample image (Figure 2(a)).

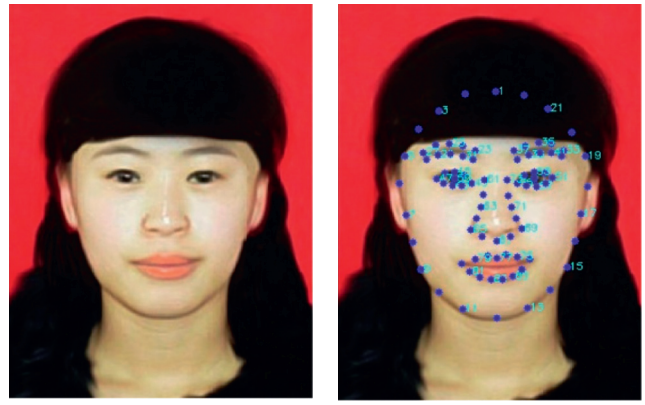
2.1.2. Facial Feature Set Extraction. The facial beauty rating for all the 5500 images available with the dataset was used as a label for the nineteen ratios. These ratios are used as input to the model for the prediction of facial beauty. The basic premise of the feature set is that certain proportions of the face should follow defined ratios. Here, 14 golden ratios and five symmetry values have been used. A total of 19 values are referred to as the feature set (FS). To assess facial beauty based on facial proportion features, 19 ratios in the FS are analyzed. A detailed description of the various ratios in the FS is given in Tables 1 and 2. In Tables 1 and 2, $d(m, n)$ refers to the Euclidean distance between landmarks m and n. The ratio values in the FS were different so normalization had to be performed on the ratios.

In the Golden Ratios, attractive faces should have a ratio of 1.618 and in Symmetry Ratios, attractive faces should have a ratio of 1. So, z-score normalization and linear scaling are used to normalize the ratio values into an interval of [0, 1]. The normalization formula is given by the following:

$$\begin{aligned} \mathrm{zi} &= \frac{\mathrm{si} - \mathrm{mean}(S)}{\mathrm{std}(S)}, \\ z_i' &= \mathrm{lb} + \frac{\mathrm{zi} - \min(Z)}{\max(Z) - \min(Z)} \times (ub - lb), \end{aligned} \tag{1}$$

where si and zi denote the i^{th} original and normalized score values, respectively, mean() and std() denote the mean and standard deviation of the FS, lb and ub denote the lower bound (zero) and upper bound (1.618) of a target score range, and min() and max() denote the minimum and maximum values of a given score set, respectively. Figure 3 shows the score distribution for each category in the dataset, namely Asian Male/Female and Caucasian Male/Female. The X-axis represents the score and the Y-axis represents the number of images.

2.1.3. Secondary Feature Set Extraction. From previous literature, it is obvious that only facial landmarks and facial ratios cannot be used to provide good results while predicting facial beauty. Facial landmarks can only provide limited information regarding facial beauty. Also, it is known that humans decide beauty based on other characteristics like face color, shape, texture, etc. AL Jones [28] analyzed the effect that facial color had on the perception of facial beauty. Their study also concluded that better facial health, i.e., clear skin, reduced abnormalities, etc., positively correlated to higher attractiveness. Face shape also corresponds to attractiveness [23, 29], as shown in Jones and Zhao et al. The studies show that a more narrow face shape with sharp features corresponds to higher beauty as compared to a round face. Facial textures are also shown to correlate to higher beauty standards, as shown in Tan et al., [30]. The combination of facial textures and color provides more information about facial beauty than facial landmarks. This study aims to use facial shape, textures, and colors to predict facial beauty and infer the performance and contribution of features towards the performance of the model. These



(a) (b)

Figure 2: Sample image and its facial landmarks. (a) Sample image. (b) Facial landmarks of the sample image.

Table 1: Description of facial ratios used.

S. No.	Description	Ratio vector
1.	Under eyes/Interocular	$d(49,57)/d(43,55)$
2.	Under eyes/Nose width	$d(49,57)/d(65,59)$
3.	Mouth width/Interocular	$d(80,87)/d(43,55)$
4.	Upper lip-jaw/Interocular	$d(77,12)/d(43,55)$
5.	Upper lip-jaw/Nose width	$d(77,12)/d(65,59)$
6.	Interocular/Lip height	$d(43,55)/d(77,83)$
7.	Nose width/Interocular	$d(65,69)/d(43,55)$
8.	Nose width/Upper lip height	$d(65,69)/d(77,84)/2$
9.	Interocular/Nose mouth height	$d(43,55)/d(67,77)$
10.	Face top-eyebrows/Eyebrows-Nose	$d(1,d(23,37)/2)/d(d(23,37)/2,67)$
11.	Eyebrows-nose/Nose-jaw	$d(d(23,37)/2,67)/d(67,12)$
12.	Face top-eyebrows/Nose-Jaw	$d(1,d(23,37)/2)/d(67,12)$
13.	Interocular/Nose width	$d(43,55)/d(65,69)$
14.	Face height/Face width	$d(1,12)/d(7,17)$

Table 2: Description of Symmetry ratios used.

S. No.	Description	Ratio vector
1.	Lower eyebrow length	$d(27,32)/d(38,42)$
2.	Lower lip length	$d(80,84)/d(74,84)$
3.	Upper eyebrow	$d(23,d(23,37)/2)/d(35,d(23,37)/2)$
4.	Upper lip	$d(80,77)/d(74,77)$
5.	Nose	$d(65,67)/d(69,67)$

features are called as Secondary Feature Set or (SFS) in this study, and the types of features extracted are as follows:

(i) Texture Features GLCM Features (Correlation, Contrast, Energy, and Homogeneity)

(ii) Shape Features Hu's Seven Invariant Moments

(iii) Color Features Color Histograms in HSV Color Space

GLCM or Gray Level Covariance Matrix is also known as the Gray Level Spatial Dependence Matrix, which learns about the texture of an image by calculating the frequency of pixel pairs with certain values in a spatial relationship that occurs in an image. Various statistical measures are then extracted from this matrix which provides textural information of the image. GLCM features are extracted for this study as they provide good information regarding the spatial relationships in the image. The statistical descriptors and their description are provided in Table 3. Each statistic

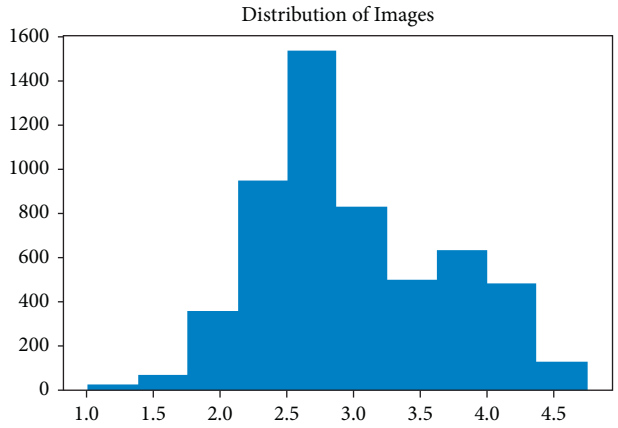


Figure 3: Distribution of beauty scores in the dataset.

Table 3: Statistical descriptors for GLCM.

Statistic	Description	Justification	Formula
Contrast	This measure returns the intensity contrast between a pixel and its neighbor over the image.	This feature can help in measuring the number of local variations in the image.	$\sum_{i,j}\lvert i-j\rvert^2 p(i,j)$
Homogeneity	This measure measures the similarity between the diagonal elements of the GLCM and the distribution of elements in the GLCM	This measure helps in finding out the degree to which each pixel differs from the other.	$\sum_{i,j} p(i,j)/1+\lvert i-j\rvert$
Correlation	This measure returns the correlation between a pixel and its neighbor over the whole image.	This feature can help in measuring how much a pixel relates to the whole image.	$\sum_{i,j}(i-\mu i)(j-\mu j)p(i,j)/\sigma_i\sigma_j$
Energy	This measure measures the sum of squared elements in GLCM. It is also known as uniformity.	This measure helps in finding out the disorders in the texture of an image.	$\sum_{i,j} p(i,j)^2$

returns a single feature value for an image and the four features of Correlation, Contrast, Energy, and Homogeneity make up a feature vector that is concatenated with the best performing FS's given in the facial feature set extraction section.

Hu's Moments [31] or Hu's invariant moments are a set of 7 numbers calculated using central moments that are invariant to image transformations. The first 6 moments have been proved to be invariant to translation, scale and rotation, and reflection. While the 7th moment's sign changes for image reflection. The 7 moments are calculated by the below equations:

$$
\begin{aligned}
h_0 &= \eta_{20}+\eta_{02},\\
h_1 &= (\eta_{20}-\eta_{02})^2+4\eta_{11}^2,\\
h_2 &= (\eta_{30}-3\eta_{12})^2+(3\eta_{21}-\eta_{02})^2,\\
h_3 &= (\eta_{30}+\eta_{12})^2+(\eta_{21}+\eta_{03})^2,\\
h_4 &= (\eta_{30}-3\eta_{12})(\eta_{30}+\eta_{12})\left[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2\right]+(3\eta_{21}-\eta_{03})\left[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right],\\
h_5 &= (\eta_{20}-3\eta_{03})\left[(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2+4\eta_{11}(\eta_{30}+3\eta_{12})(\eta_{21}+\eta_{03})\right],\\
h_6 &= (3\eta_{21}-\eta_{03})(\eta_{30}+\eta_{12})\left[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2\right]-(\eta_{30}-3\eta_{12})(\eta_{21}+\eta_{03})\left[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})\right]^2.
\end{aligned}
\tag{2}
$$

To calculate the facial color of the image, it is required to think about how an average person views a face. The HSV color space is more intuitive to how people experience color than the RGB color space [32]. As hue (H) varies from 0 to 1.0, the corresponding colors vary from red, through yellow, green, cyan, blue, and magenta, back to red. As saturation (S) varies from 0 to 1.0, the corresponding colors (hues) vary from unsaturated (shades of gray) to fully saturated (no white component). As value (V), or brightness, varies from 0 to 1.0, the corresponding colors become increasingly brighter. The sample image (Figure 4(a)) in Hue plane (Figure 4(b)), saturation plane (Figure 4(c)), and value plane (Figure 4(d)) are shown in Figure 4. With RGB, the color will have values like (0.5, 0.5, 0.25), whereas for HSV, it will be (30°, $\sqrt{3}/4$, 0.5). HSV is best used when a user is selecting a color interactively. It is usually much easier for a user to get the desired color as compared to using RGB [33].

2.1.4. Models Used for Prediction. In this study, the FS is used as input to the models and the corresponding scores as the labels. Also, four well-known regression models, Linear Regression (LR), K-Nearest Neighbors (KNN), Random Forest (RF), and Artificial Neural Network (ANN), were used for prediction. A small description for each model is given below.

2.2. Linear Regression. Linear Regression is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. It sets up a relationship between input variables and target variables which is represented by the following equation:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon, \tag{3}$$

where i is the number of observations, y_i is the target variable, x_i is the input variable, β_0 is the y-intercept, β_p is the coefficient for each input variable, and ε is the error term.

2.2.1. Random Forest. Random forest is a Supervised Learning algorithm that uses ensemble learning methods for classification and regression. Random forest is a bagging technique and not a boosting technique. The trees in random forests are run in parallel. There is no interaction between these trees while building the trees. It operates by

(a) (b) (c) (d)

Figure 4: Sample image and its facial landmarks in various planes. (a) Original image. (b) Hue plane. (c) Saturation plane. (d) Value plane.

constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. A random forest is a meta-estimator (i.e., it combines the result of multiple predictions) which aggregates many decision trees, with some helpful modifications:

(i) The number of features that can be split at each node is limited to some hyperparameter. This ensures that the model does not rely too heavily on any individual feature and makes fair use of all potentially predictive features.

(ii) Each tree draws a random sample from the original data set when generating its splits, adding a further element of randomness that prevents overfitting.

The above modifications help prevent the trees from being too highly correlated.

2.3. K-Nearest Neighbor (K-NN). kNN falls is a lazy learning method, which means that there is no explicit training phase before classification. The Euclidean distance formula and probability formula of the kNN method are given in the following equations, respectively:

$$d(x, x') = \sqrt{(x_1 - x_1')^2 + (x_2 - x_2')^2 + \cdots + (x_n - x_n')^2}, \tag{4}$$

$$P(y = j \mid X = x) = \frac{1}{K} \sum_{i \in A} I\left(y^{(i)} = j\right), \tag{5}$$

where A is the particular class/set, unseen observation x and a similarity metric d and K is an arbitrary integer. A weighted average of the K-nearest neighbors were used, where the weight was decided by the Euclidean distance of the K closest training samples. The number of clusters is set to [12, 22].

2.4. Artificial Neural Network. In ANN regression, a multilayer perceptron (MLP) is applied which is composed of an input layer, hidden layers, and an output layer. Each layer has one or more neurons directionally linked with the neurons from the previous and next layers. A sigmoid function was applied as the activation function to compute the output of the hidden layer in each neuron. Artificial Neural Networks have been used to predict facial beauty with much success, as shown in [34].

2.5. Metrics Used for Prediction. To measure the model performance and calculate the error generated by the models, four metrics are used. Mean Absolute Error (MAE), Mean Squared Error (MSE), R^2 Score, and Pearson's Correlation Coefficient (PC) [35] are provided in Table 4.

2.6. Experimental Setup. The dataset was split into training data (80%) and testing data (20%). All the experiments were run on an Intel i3 Processor with 12 GB of RAM using Python programming language. Two ablation studies were performed. One was to find the best performing FS out of all the primary features containing facial landmarks. The other ablation study was done to find out the performance variations amongst the concatenated SFS. The experimental results, along with the corresponding graphs, are explained in the results and discussion section.

3. Results and Discussion

In our study on facial beauty prediction, we used facial landmarks as a base feature set for extracting facial features. From the facial landmarks, we calculate 19 facial ratios that are used for predicting the beauty of the facial image (explained in Appendix A). It is observed that extracting and employing more features does not improve the performance of the proposed model. Further, these 19 facial ratios holistically describe the facial landmarks of the face, which indeed predict the symmetry and quantify the associated beauty [36, 37].

The analysis compared the performance of computer models against human ratings and examined the performance of each model. Table 5 shows the performance of each model concerning the correlation and error metrics.

Table 4: Various matrices used for prediction.

S. No	Metrics	Formula
1.	MAE	$\text{MAE} = 1/n \sum_{i=1}^{n} \lvert x_i - x \rvert$
2.	MSE	$\text{MSE} = 1/n \sum_{i=1}^{n} \lvert x_i - x \rvert^2$
3.	R^2	$R^2 = 1 - \sum_i (y_i - y_i')^2 / \sum_i (y_i - \overline{y_i})^2$
4.	PC	$\text{PC} = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$

Table 5: Best Performing FS for each model.

S. No	Model name	PC	R^2 score	MAE	MSE
1.	LR	0.5811	0.3310	0.3570	0.2031
2.	KNN	0.7115	0.4162	0.3504	0.1773
3.	RF	0.6201	0.3344	0.3777	0.2019
4.	ANN	0.5838	0.3096	0.3610	0.2096

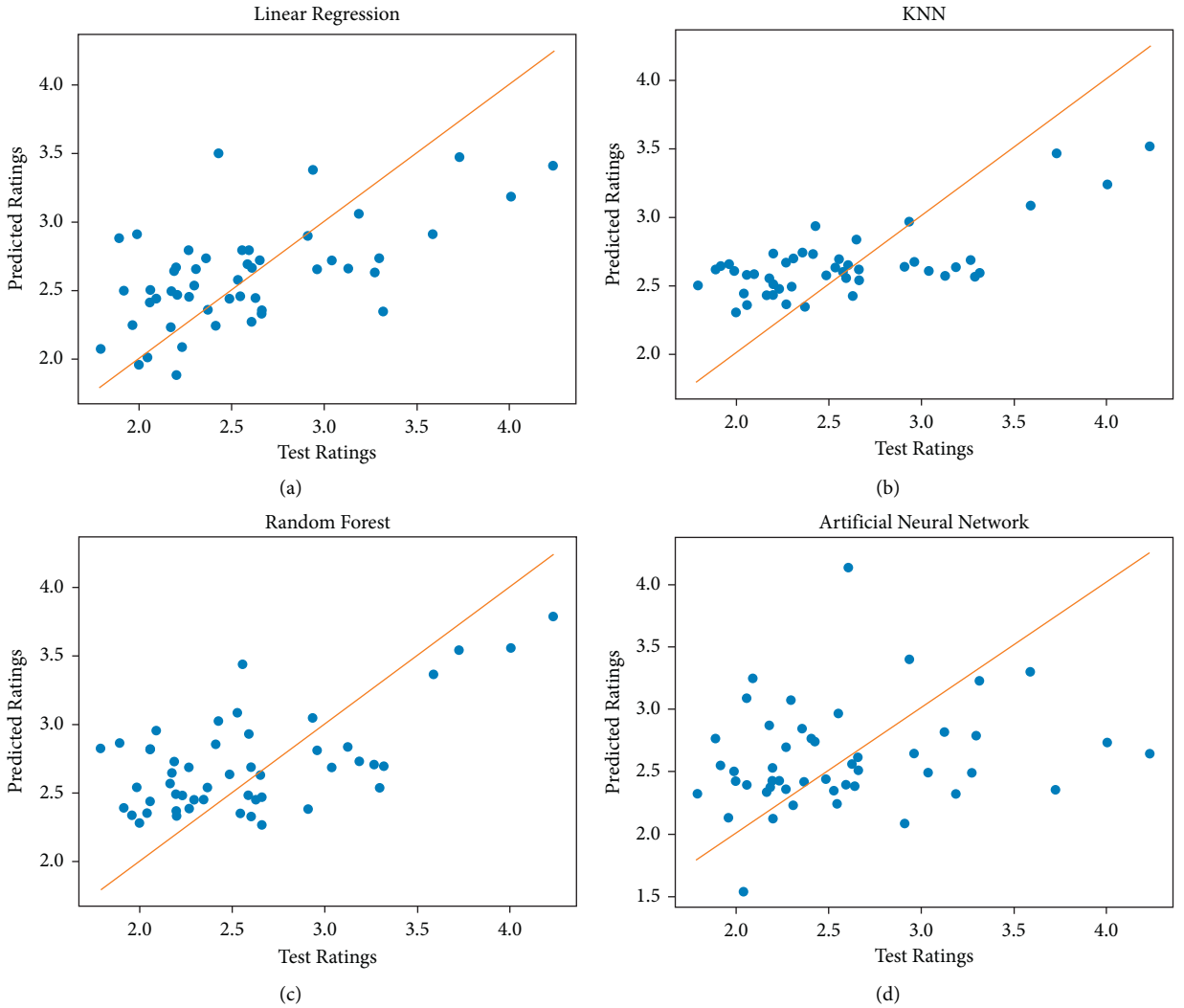


Figure 5: Regression curves for the best performing FS of each model. (a) Regression Curve for LR. (b) Regression curve for KNN. (c) Regression curve for RF. (d) Regression curve for ANN.

Table 6: Ablation study of PC and R^2 score for various models

S. No.	Features used	PC of SFS for				R^2 score of SFS for			
		LR	KNN	RF	ANN	LR	KNN	RF	ANN
1.	Without	0.58	0.71	0.62	0.58	0.33	0.41	0.33	0.30
2.	**T + C + S**	**0.65**	**0.78**	**0.70**	**0.62**	**0.40**	**0.48**	**0.40**	**0.36**
3.	$T + C$	0.64	0.76	0.68	0.60	0.39	0.46	0.38	0.35
4.	$T + S$	0.60	0.75	0.67	0.59	0.37	0.45	0.36	0.33
5.	$C + S$	0.62	0.74	0.67	0.61	0.38	0.45	0.37	0.34
6.	T	0.59	0.74	0.65	0.58	0.38	0.47	0.34	0.34
7.	C	0.58	0.72	0.63	0.59	0.35	0.42	0.37	0.32
8.	S	0.57	0.73	0.64	0.60	0.34	0.43	0.35	0.31

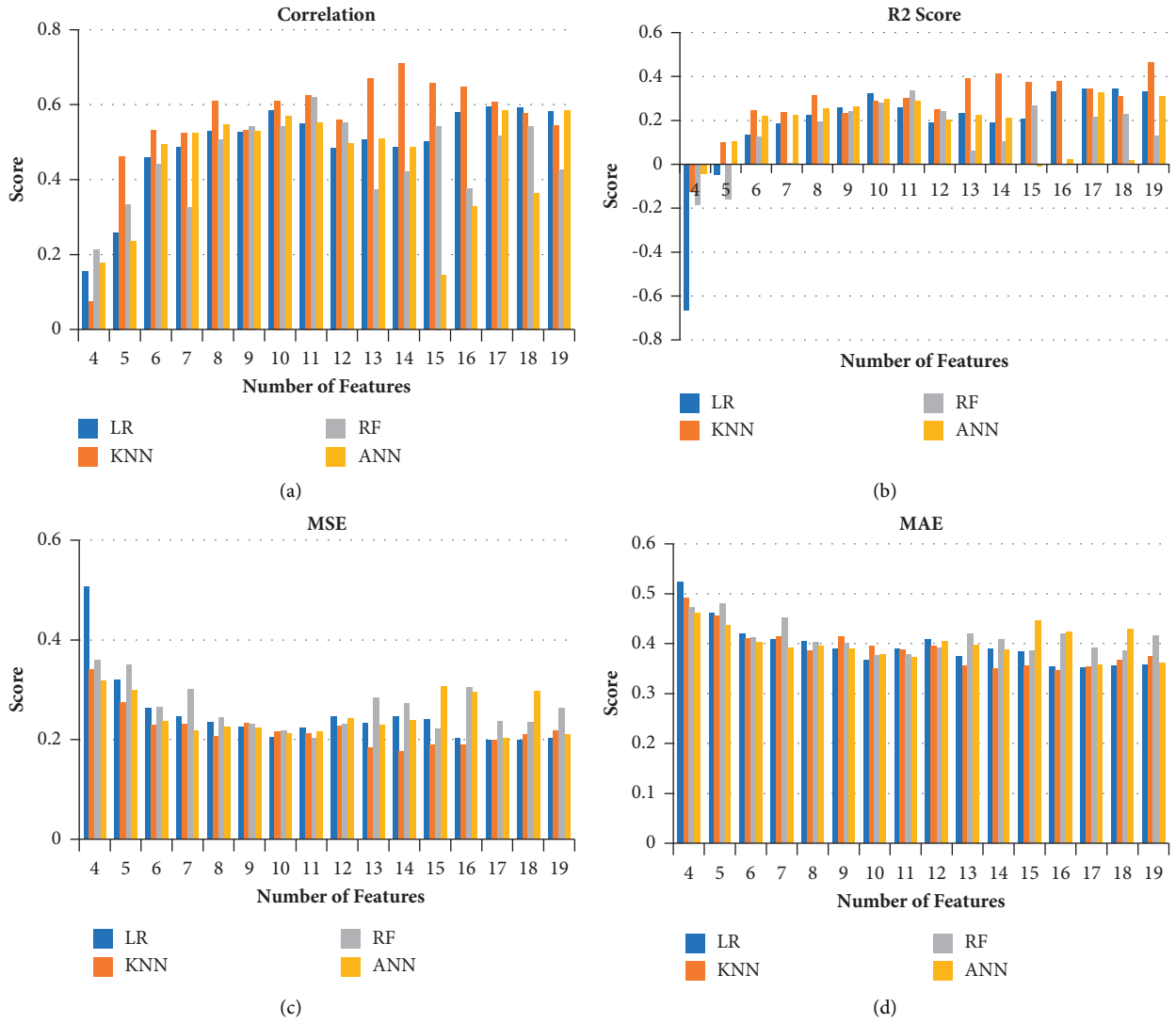


Figure 6: Ablation study of the metrics of FS for the various models. (a) Ablation Study of Correlation for LR, KNN, RF, and ANN. (b) Ablation studies of R2 Score for LR, KNN, RF, and ANN. (c) Ablation studies of MSE for LR, KNN, RF, and ANN. (d) Ablation studies of MAE for LR, KNN, RF, and ANN.

In Table 5, each model has a relatively high correlation with human ratings and low errors for each of their corresponding best performing FS.

In Figure 5, the x-axis represents human scores on a scale of 0–5 and the y-axis represents the computer-predicted scores in the same range. KNN has the highest correlation values with the lowest error. ANN has a higher correlation than LR, but LR has comparatively lower error than ANN. Also, ANN is much more scattered than LR. The red line in the graphs shows the regression fit of the data. ANN is most

Table 7: Ablation study of MAE and MSE score for various models.

S. No.	Features used	MAE of SFS for				MSE score of SFS for			
		LR	KNN	RF	ANN	LR	KNN	RF	ANN
1.	Without	0.35	0.35	0.37	0.36	0.20	0.17	0.20	0.20
2.	**T + C + S**	**0.30**	**0.28**	**0.29**	**0.31**	**0.15**	**0.09**	**0.11**	**0.15**
3.	$T + C$	0.31	0.30	0.30	0.33	0.16	0.12	0.12	0.16
4.	$T + S$	0.32	0.32	0.31	0.35	0.18	0.11	0.12	0.17
5.	$C + S$	0.31	0.31	0.30	0.33	0.17	0.11	0.11	0.15
6.	T	0.33	0.31	0.35	0.35	0.16	0.15	0.15	0.18
7.	C	0.32	0.34	0.32	0.34	0.15	0.16	0.13	0.17
8.	S	0.34	0.33	0.334	0.33	0.19	0.14	0.17	0.18

Table 8: Results obtained from the study through concatenation of TCS Features with SFS.

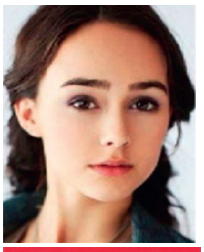
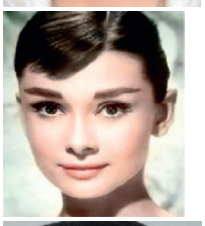
Actual value	Predicted value	Image
4.7	4.405	
4.63	4.65	
4.7	4.43	
2.7	3.8	
1.43	2.8	

Table 9: Performance of each model according to correlation for 11–25 features.

S. No.	Number of features	LR	KNN	RF	ANN
1.	11	0.5492	0.625	**0.6201**	0.5527
2.	12	0.4842	0.5603	0.5535	0.4972
3.	13	0.5076	0.6694	0.3729	0.509
4.	14	0.4859	**0.7115**	0.4211	0.486
5.	15	0.5019	0.6587	0.542	0.1443
6.	16	0.5792	0.6458	0.3761	0.3275
7.	17	**0.5952**	0.6067	0.5158	0.5837
8.	18	0.5921	0.5773	0.5424	0.3648
9.	19	0.5811	0.5449	0.4248	**0.5838**
10.	20	0.5954	0.5321	0.4156	0.5801
11.	21	0.5806	0.5495	0.4187	0.5796
12.	22	0.5854	0.5369	0.4258	0.58
13.	23	0.5897	0.5485	0.4296	0.5732
14.	24	0.5801	0.5331	0.4235	0.5721
15.	25	0.58	0.541	0.4113	0.5821

Table 10: Performance of each model according to R^2 Score for 11–25 features.

S. No.	Number of features	LR	KNN	RF	ANN
1.	11	0.2604	0.302	**0.3344**	0.2888
2.	12	0.1904	0.2509	0.2402	0.203
3.	13	0.2331	0.3945	0.0604	0.2264
4.	14	0.1892	**0.4162**	0.1029	0.213
5.	15	0.2082	0.3753	0.2676	-0.0114
6.	16	0.3303	0.3787	-0.003	0.0246
7.	17	**0.3459**	0.3462	0.2165	0.3288
8.	18	0.3457	0.309	0.2298	0.0178
9.	19	0.331	0.3952	0.1308	**0.3096**
10.	20	0.3369	0.3924	0.1395	0.3169
11.	21	0.3256	0.4021	0.1456	0.3178
12.	22	0.3248	0.4089	0.1324	0.316
13.	23	0.331	0.4097	0.1437	0.3099
14.	24	0.3358	0.3924	0.1587	0.3184
15.	25	0.3294	0.3910	0.1308	0.3269

similar to the ideal case. The second similar measure is LR, with the least predictive measures being KNN and RF. Therefore, while KNN and RF can correlate much more with the human values, ANN and LR are better models for fitting the data.

Even though Table 6 provides information about the best performing FS, Figure 6 shows us how each model learns from each feature in the FS and how each feature contributes to the learning process. The X-axis in Figure 6 represents the number of features used in the ablation study, and the Y-axis represents the Score for each metric. The X-axis starts from 4 because the metrics for features 3 and below were very poor, and hence it was decided not to be added in the graph as it would be negligible in comparing the performance of the models. It is trivial from Figure 6(b) when the number of features is less (for 5 and 6 in X-axis), the R^2 is in a negative trend which indicates that the fit with regression curve is worse. Meanwhile, by increasing the number of features, the R^2 values are improved significantly.

The nineteen features in the FS are ratios from Symmetry and Golden Ratios. From previous studies, it is known that Golden Ratios perform better than Symmetry Ratios. This phenomenon is proved when looking at the above tables. The best performing FS is obtained from the initial removal of Symmetry Ratios and the metrics decrease once the Golden Ratios are removed. Next, the ablation studies for the SFS have been performed. Tables 6 and 7 show the ablation studies performed for each model.

It is observed from Tables 6 and 7, the fusion of Texture, Color, and Shape features as the SFS along with the ratio features improves the overall performance of the model. The highest correlation is achieved by KNN, followed by RF. The least correlation is obtained through LR and ANN. In every model, each feature contributes differently to the performance of the model. Since the KNN model has achieved the highest correlation values and lowest errors, it is considered as the ideal model for facial feature analysis, and the obtained results are given in Table 8 (images and scores of few best and moderate prediction results).

From the above-obtained results, it is clear that our technique has avoided overfitting in our study through the following ways:

Table 11: Performance of each model according to MSE for 11–25 features.

S. No.	Number of features	LR	KNN	RF	ANN
1.	11	0.2246	0.2119	**0.2019**	0.2159
2.	12	0.2458	0.2274	0.2307	0.242
3.	13	0.2329	0.1839	0.2853	0.2295
4.	14	0.2462	**0.1773**	0.2724	0.239
5.	15	0.2404	0.1897	0.2224	0.3072
6.	16	0.2033	0.1886	0.3046	0.2963
7.	17	**0.1986**	0.1985	0.2379	0.2038
8.	18	0.1987	0.2098	0.2339	0.2983
9.	19	0.2031	0.2171	0.2639	**0.2096**
10.	20	0.2056	0.2069	0.2658	0.2165
11.	21	0.2098	0.2154	0.2645	0.2148
12.	22	0.2154	0.2167	0.2594	0.2259
13.	23	0.2106	0.2098	0.2561	0.226
14.	24	0.2198	0.2192	0.2632	0.2197
15.	25	0.2046	0.2367	0.2674	0.213

Table 12: Performance of each model according to MAE for 11–25 features.

S. No.	Number of features	LR	KNN	RF	ANN
1.	11	0.3903	0.3876	**0.3777**	0.3727
2.	12	0.4083	0.3957	0.3914	0.4034
3.	13	0.3753	0.3558	0.4188	0.3964
4.	14	0.3897	**0.3504**	0.4079	0.3876
5.	15	0.3832	0.3546	0.3844	0.446
6.	16	0.3543	0.3468	0.4193	0.4236
7.	17	**0.3521**	0.3542	0.3913	0.3572
8.	18	0.3561	0.3667	0.3858	0.4279
9.	19	0.357	0.3749	0.4153	**0.361**
10.	20	0.3589	0.3725	0.4188	0.3654
11.	21	0.3621	0.3763	0.4103	0.3789
12.	22	0.3652	0.3792	0.4069	0.3625
13.	23	0.3624	0.3824	0.4037	0.3764
14.	24	0.3527	0.3761	0.4152	0.371
15.	25	0.361	0.3601	0.4092	0.3762

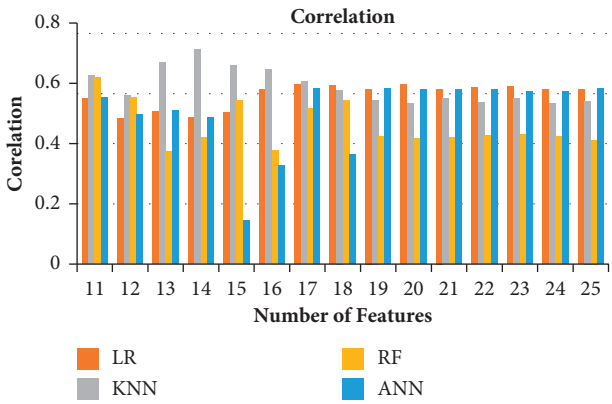


Figure 7: Ablation study of correlation for 11–25 features.

(i) The dataset used is from SCUT-FBP5500, which is a collection of 5500 frontal facial images. This dataset has been extensively used by many researchers, and there is no class imbalance, as shown from the histogram (Figure 3). The dataset has an equal number of images of each class in the training and testing set, and the target labels are normally distributed. The data used is enough, and since it is already balanced, no preprocessing or data augmentation was required.

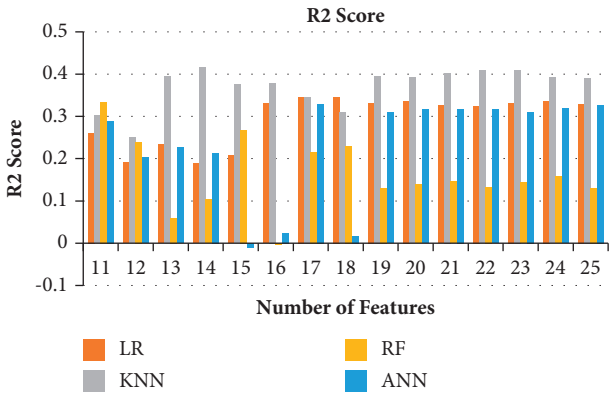


Figure 8: Ablation study of R2 Score for 11–25 features.

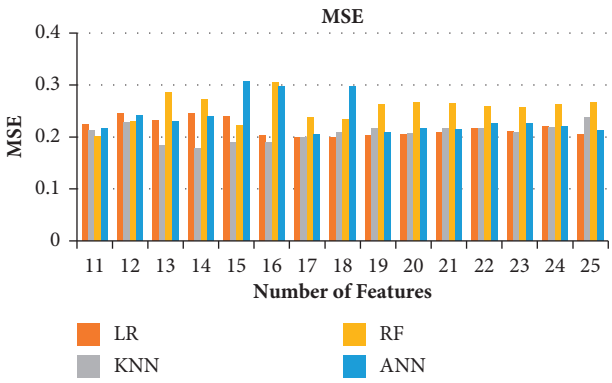


Figure 9: Ablation study of MSE for 11–25 features.

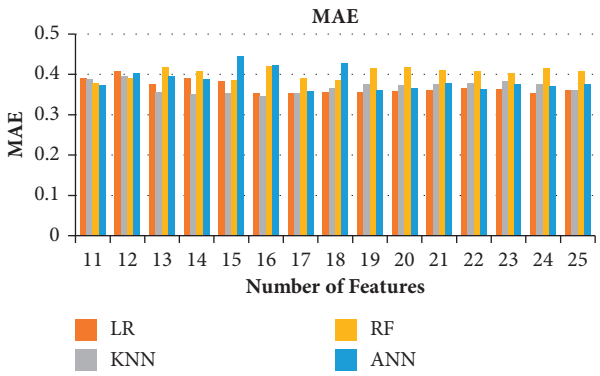


Figure 10: Ablation study of MAE for 11–25 features.

(ii) An Ablation study was conducted to find the best performing feature set and to find the optimal number of features which can be used to achieve the best performance. We avoided overfitting by removing features and limiting the maximum number of features to 19.

4. Conclusion

In this study, different computer models are analyzed to predict facial beauty using facial features like golden ratios, texture, shape, and color through Machine Learning. First, the models are analyzed and the performance on facial ratios is derived from Golden Ratios and Symmetry. Nineteen facial ratios were selected to represent the Feature Set (FS) from which the best performing features were derived by ablation study on each feature. From the best performing FS for each model, Texture, Color, and Shape Features were extracted from the dataset. This feature vector from secondary features (SFS) is fed to each model through an ablation study to measure the performance of each feature and observed the variation on performance metrics of the model. Experimental results showed that the beauty score obtained from KNN achieved the best metrics, followed by RF, LR, and ANN. The fusion of TCS Features with the FS to form the SFS performs the best, with the highest correlation being 78%. Therefore, a fusion of multiple feature types rather than a single feature type provided better performance than using one feature type. Also, heterogeneity of the feature vector increased the performance as compared to a homogenous feature vector. In addition, our analysis observed that Golden Ratios provided better information than Symmetry Ratios which was consistent with previous literature.

Our analysis has shown that human judgments regarding facial beauty are consistent with the facial ratios derived from painters, architects, etc. These ratios from Symmetry, Golden Ratios, Neoclassical canons, etc., correlate highly with human ratings. But, there does exist the element of variation, which exists as secondary features like sexual features, facial health, etc. Our analysis has shown that combining secondary features like facial health, texture, and shape with facial proportions allows a computer model to learn better and correlate much more with its human counterparts.

Overall, analysis has been limited with the facial beauty prediction of Asian and Caucasian males and females. Our future work will involve an analysis of facial beauty for different types of people. Alongside, our analysis will use more powerful algorithms to extract deep features and to use the same fusion technique on those features to achieve an even higher correlation and lower error. Also, it is aimed to increase the size of the dataset and variation to allow for better generalization. Another work is aimed to make a modular system that can rate beauty in real-time to allow for various beauty evaluation applications.

Appendix

A

The results from the Ablation study for each model to predict the facial beauty are shown as Tables and Figures that experimentally verify that 19 facial ratios are indeed present the best performance.

As can be seen from Tables 9–Table 12 and Figures 7–Figure 10 , the performance of each model after 19 features does not improve in any significant manner. Due to this reason, model performance was calculated until 19 features. As mentioned in the study, the below 4 features were not considered as the performance was poor, which can be seen from negative performance metrics.

Data Availability

The data set used for this study is openly available and the details are mentioned in the article. Data sets used in this study can be found on the website https://github.com/HCIILAB/SCUT-FBP5500-Database-Release/, and the codes are available openly in the repository https://github.com/IyerOnFyer/Facial-Beauty-Prediction.git.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors want to express their sincere thanks and gratitude to the Vellore Institute of Technology, Vellore, India, and Shantou University, Shantou, China, for their support and for the resources provided to carry out this research. This work is supported by the National Natural Science Foundation of China (grant no. 82071992); Basic and Applied Basic Research Foundation of Guangdong Province (grant no. 2020B1515120061); the Guangdong Province University Priority Field (Artificial Intelligence) Project (grant no. 2019KZDZX1013); National Key R&D Program of China (grant no. 2020YFC0122103); the Key Project of Guangdong Province Science and Technology Plan (grant no. 2015B020233018); and the Scientific Research Grant of Shantou University, China, (Grant No: NTF17016).

References

[1] E. P. Prokopakis, I. M. Vlastos, V. A. Picavet et al., "The golden ratio in facial symmetry," *Rhinology journal*, vol. 51, no. 1, pp. 18–21, 2013.

[2] D. I. Perrett, K. A. May, and S. Yoshikawa, "Facial shape and judgements of female attractiveness," *Nature*, vol. 368, no. 6468, pp. 239–242, 1994.

[3] J. S. Winston, J. O'Doherty, J. M. Kilner et al., "Brain systems for assessing facial attractiveness," *Neuropsychologia*, vol. 45, no. 1, pp. 195–206, 2007.

[4] A. J. Rubenstein, L. Kalakanis, and J. H. Langlois, "Infant preferences for attractive faces: a cognitive explanation," *Developmental Psychology*, vol. 35, no. 3, pp. 848–855, 1999.

[5] J. H. Langlois and L. A. Roggman, "Attractive faces are only average," *Psychological Science*, vol. 1, no. 2, pp. 115–121, 1990.

[6] G. Rhodes, F. Proffitt, J. M. Grady et al., "Facial symmetry and the perception of beauty," *Psychonomic Bulletin & Review*, vol. 5, no. 4, pp. 659–669, 1998.

[7] B. Fink, K. Grammer, and R. Thornhill, "Human (Homo sapiens) facial attractiveness in relation to skin texture and color," *Journal of Comparative Psychology*, vol. 115, no. 1, pp. 92–99, 2001.

[8] A. L. Jones, R. Russell, and R. Ward., "Cosmetics alter biologically-based factors of beauty: evidence from facial contrast," *Evolutionary Psychology : An International Journal of Evolutionary Approaches to Psychology and Behavior*, vol. 13, no. 1, pp. 210–29, 2015.

[9] A. Kagian, G. Dror, T. Leyvand et al., "A humanlike predictor of facial attractiveness," in *Advances in Neural Information Processing Systems*, pp. 649–656, MIT Press, Massachusetts, MA, USA, 2007.

[10] A. Jones and B. Jaeger, "Biological bases of beauty revisited: the effect of symmetry, averageness, and sexual dimorphism on female facial attractiveness," *Symmetry*, vol. 11, no. 2, p. 279, 2019.

[11] L. G. Farkas, T. A. Hreczko, J. C. Kolar et al., "Vertical and horizontal proportions of the face in young adult north American caucasians," *Plastic and Reconstructive Surgery*, vol. 75, no. 3, pp. 328–337, 1985.

[12] D. L. Narain, "The golden ratio," Availabel online: http://cuip.uchicago.edu/%7Edlnarain/golden/(accessed on, 2017.

[13] Y. Jefferson, "Facial beauty--establishing a universal standard," *International Journal of Orthodontics*, vol. 15, no. 1, pp. 9–22, 2004.

[14] R. Thornhill and S. W. Gangestad, "Human facial beauty," *Human Nature*, vol. 4, no. 3, pp. 237–269, 1993.

[15] K. Grammer and R. Thornhill, "Human (Homo sapiens) facial attractiveness and sexual selection: the role of symmetry and averageness," *Journal of Comparative Psychology*, vol. 108, no. 3, pp. 233–242, 1994.

[16] D. I. Perrett, K. J. Lee, I Penton-Voak et al., "Effects of sexual dimorphism on facial attractiveness," *Nature*, vol. 394, pp. 884–887, 1998.

[17] A. C. Little, B. C. Jones, C. Waitt et al., "Symmetry is related to sexual dimorphism in faces: data across culture and species," *PLoS One*, vol. 3, no. 5, Article ID e2106, 2008.

[18] N. Koehler, L. W. Simmons, G. Rhodes et al., "The relationship between sexual dimorphism in human faces and fluctuating asymmetry," *Proceedings of the Royal Society of London Series B Biological Sciences*, vol. 271, no. suppl_4, pp. S233–S236, 2004.

[19] B. Fink, K. Grammer, and P. Matts, "Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces☆," *Evolution and Human Behavior*, vol. 27, no. 6, pp. 433–442, 2006.

[20] M. E. Rhazi, A. Zarghili, A. Majda et al., "Facial beauty analysis by age and gender," *International Journal of Intelligent Systems Technologies and Applications*, vol. 18, no. 1, pp. 179–203, 2019.

[21] K. Schmid, D. Marx, and A. Samal, "Computation of a face attractiveness index based on neoclassical canons, symmetry, and golden ratios," *Pattern Recognition*, vol. 41, no. 8, pp. 2710–2717, 2008.

[22] F. Dornaika, K. Wang, I. Arganda-Carreras et al., "Toward graph-based semi-supervised face beauty prediction," *Expert Systems with Applications*, vol. 142, Article ID 112990, 2020.

[23] L. Lin, L. Liang, L. Jin et al., "Attribute-aware convolutional neural networks for facial beauty prediction," in *Proceedings of the 2019 International Joint Conference on Artificial Intelligence*, pp. 847–853, Macao, China, August 2019.

[24] Q. Xiao, Y. Wu, D. Wang et al., "Beauty3DFaceNet: deep geometry and texture fusion for 3D facial attractiveness prediction," *Computers & Graphics*, vol. 98, pp. 11–18, 2021.

[25] W. Wei, E. S. Ho, K. D. McCay et al., "Assessing facial symmetry and attractiveness using augmented reality," *Pattern Analysis and Applications*, vol. 28, pp. 1–7, 2021.

[26] S. Tong, X. Liang, T. Kumada et al., "Putative ratios of facial attractiveness in a deep neural network," *Vision Research*, vol. 178, pp. 86–99, 2021.

[27] L. Liang, L. Lin, L. Jin et al., "SCUT-FBP5500: a diverse benchmark dataset for multi-paradigm facial beauty prediction," in *Proceedings of the 2018 24th International Conference*

on Pattern Recognition (ICPR), pp. 1598–1603, IEEE, Beijing China, August 2018.

[28] A. L. Jones, "The influence of shape and colour cue classes on facial health perception," *Evolution and Human Behavior*, vol. 39, no. 1, pp. 19–29, 2018.

[29] J. Zhao, M. Zhang, C. He et al., "A novel facial attractiveness evaluation system based on face shape, facial structure features and skin," *Cognitive Neurodynamics*, vol. 14, no. 5, pp. 643–656, 2020.

[30] K. W. Tan, B. Tiddeman, and I. D. Stephen, "Skin texture and colour predict perceived health in Asian faces," *Evolution and Human Behavior*, vol. 39, no. 3, pp. 320–335, 2018.

[31] Z. Huang and J. Leng, "Analysis of Hu's moment invariants on image scaling and rotation," in *Proceedings of the 2010 2nd International Conference on Computer Engineering and Technology*, vol. 7, IEEE, Bali Island, Indonesia, March 2010.

[32] L. Liu, J. Xing, S. Liu et al., "Wow! You are so beautiful today!" *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 11, no. 1s, pp. 1–22, 2014.

[33] J. Gan, L. Li, Y. Zhai et al., "Deep self-taught learning for facial beauty prediction," *Neurocomputing*, vol. 144, pp. 295–303, 2014.

[34] L. Zhang, D. Zhang, M.-M. Sun et al., "Facial beauty analysis based on geometric feature: toward attractiveness assessment application," *Expert Systems with Applications*, vol. 82, pp. 252–265, 2017.

[35] M. M. Mukaka, "A guide to appropriate use of correlation coefficient in medical research," *Malawi Medical Journal*, vol. 24, pp. 69–71, 2012.

[36] P. M. Pallett, S. Link, and K. Lee, "New "golden" ratios for facial beauty," *Vision Research*, vol. 50, no. 2, pp. 149–154, 2010.

[37] J. Fan, K. P. Chau, X. Wan et al., "Prediction of facial attractiveness from facial proportions," *Pattern Recognition*, vol. 45, no. 6, pp. 2326–2334, 2012.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 7788491, 9 pages
https://doi.org/10.1155/2021/7788491

Research Article

A New Optimal Diagnosis System for Coronavirus (COVID-19) Diagnosis Based on Archimedes Optimization Algorithm on Chest X-Ray Images

Liping Chen[1] **and Tahereh Rezaei**[2]

[1]*College of Computer, Weinan Normal University, Weinan, Shaanxi, China*
[2]*Neuroscience Research Center, Shiraz, Iran*

Correspondence should be addressed to Tahereh Rezaei; dr.tahereh.rz@gmail.com

Received 28 May 2021; Revised 29 July 2021; Accepted 11 August 2021; Published 24 August 2021

Academic Editor: Navid Razmjooy

The new coronavirus, COVID-19, has affected people all over the world. Coronaviruses are a large group of viruses that can infect animals and humans and cause respiratory distress; these discomforts may be as mild as a cold or as severe as pneumonia. Correct detection of this disease can help to avoid its spreading increasingly. In this paper, a new CAD-based approach is suggested for the optimal diagnosis of this disease from chest X-ray images. The proposed method starts with a min-max normalization to scale all data into a normal scale, and then, histogram equalization is performed to improve the quality of the image before main processing. Afterward, 18 different features are extracted from the image. To decrease the method difficulty, the minimum features are selected based on a metaheuristic called Archimedes optimization algorithm (AOA). The model is then implemented on three datasets, and its results are compared with four other state-of-the-art methods. The final results indicated that the proposed method with 86% accuracy and 96% precision has the highest balance between accuracy and reliability with the compared methods as a diagnostic system for COVID-19.

1. Introduction

Coronaviruses are a large family of viruses that can infect animals and humans. Many of the known coronaviruses have caused a range of respiratory infections in humans, ranging from the common cold to more severe illnesses such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). The newly discovered coronavirus is the cause of COVID-19 disease. The emerging virus and its disease were unknown before the recent outbreak.

COVID-19 is a new infectious disease that first broke out in December 2019 in Wuhan Province, China.

The disease later spread to other countries in Asia, Europe (mainly Italy, Spain, France, and the United Kingdom), Africa, and the United States (mainly the United States). The COVID-19 disease is highly transmissible. According to the World Health Organization (WHO), as of Apr 14, 2021, there are 13,804,929 cases of infection and 2,972,662 deaths. However, these statistics are exponentially increasing. From the outset, the World Emergency Committee stressed the need for prompt diagnosis, quarantine, and prompt treatment.

Efforts to curb the spread of COVID-19, in which more than 170 countries are currently fighting and controlling the disease, are ongoing, with daily statistics on the number of deaths, which is based on a kind of global alliance among all countries of the world to prevent and stop the spread of this virus. With the spread of this virus and its epidemic worldwide, the diagnosis of this disease has become a priority for the medical system of the world and international organizations such as WHO. There are currently several ways to diagnose this disease. One of the important parameters in this regard is the use of advanced equipment and

the help of technology to control this virus. Artificial intelligence and machine learning techniques play a significant role in carrying out this mission.

Recently, the analysis of disease-related data, data preparation, prevention, and control of infectious diseases, such as COVID-19, has become one of the goals of artificial intelligence. Artificial intelligence is one of the technologies that can easily track the spread of the virus, identify high-risk patients, and help control infection in a timely manner. It can also predict patient mortality risk by analyzing previous data. Artificial intelligence can help us fight the virus with population screening, medical help, information, and suggestions for infection control. This technology has the potential to improve the planning, treatment, and reported outcomes of the COVID-19 patient and is an evidence-based medical tool.

Artificial intelligence can quickly analyze the symptoms and alert patients and treatment staff. This technology helps make decisions faster, which is cost-effective. Artificial intelligence helps create a new COVID-19 patient diagnosis and management system through efficient algorithms. Artificial intelligence uses medical imaging technologies such as X-ray imaging, Computed tomography (CT), and magnetic resonance imaging (MRI) to diagnose infected cases. At the same time, in the field of medical imaging, machine learning helps to identify patterns in images and enhances the ability of radiologists to diagnose and detect disease early.

The diagnosis of COVID-19 is problematic due to the appearance of different types of coronaviruses. However, sampling offers well diagnosis, it is possible only by surgery, which gives an unpleasant experience to the patient. For avoiding redundant sampling, researchers have reviewed numerous noninvasive techniques for diagnosing COVID-19. Almost, all diagnosis systems based on image processing include three main parts of image preprocessing, boundary identification, feature extraction, and the final classification [1]. However, COVID-19 is a new disease, because of its highly dangerous and pandemic status, different works have been done in the diagnosis and purpose of this cancer from medical images.

Song et al. [2] proposed diagnosis method based on CT images. They collected chest CT scans from hospitals of two provinces in China. Deep learning was used for the diagnosis of COVID-19. The final results indicated that the proposed method with 0.96 precision has good accuracy for detecting the COVID-19 patients. The method was also provided a model to extract significant lesion features from the CT images.

Zhang et al. [3] proposed deep anomaly detection system to provide a fast and reliable screening. For providing the model efficiency, Github, repository1, and Chest X-ray14 datasets were analyzed. Simulation results showed that the proposed method provides about 96.00% accuracy in the diagnosis of the COVID-19 cases with 96.00% sensitivity which are proper results for this purpose.

Jin et al. [4] presented another method based on medical imaging for a quick diagnosis of COVID-19 patients from CT scans. The method was based on a difficult multiclass diagnosis system based on Convolutional Neural Network (CNN) to reach the defective region. Moreover, the diagnosis accuracy of chest X-ray (CXR) was compared with that of CT to show its higher accuracy.

As mentioned above, X-ray imaging provides better results for diagnosing the COVID-19 disease. Therefore, in this study, we propose a medical imaging system based on X-ray to provide better results of the diagnosis. However, the results of the researches in the literature are well, but there are also lots of works that can be done on the system consistency and accuracy. Therefore, in this study, the following contributions are considered:

(i) A new CAD-based approach for optimal diagnosis of COVID-19 from chest X-ray images

(ii) Using a new algorithm to decrease the complexity of feature extraction

(iii) Using a new algorithm for the final classification of the images

(iv) The new algorithm is based on a new metaheuristic called Archimedes optimization algorithm (AOA)

(v) Validating the method by three different datasets and comparing with four other state-of-the-art methods

The next sections of this paper are organized as follows. In Section 2, the main preprocessing stages of the paper, including min-max normalization and histogram equalization, are explained. Section 3 explains about feature extraction and the features that are utilized for the classification. In Section 4, the theory of the Archimedes optimization algorithm has been explained. Section 5 describes the method of classification of the data based on optimization and multilayer perceptron. In Section 6, the simulation results have been completely performed and discussion proved the method capability. Finally, the paper has been concluded in Section 7.

2. Preprocessing of the CXR Images

2.1. Min-Max Normalization. Data normalization is one of the most important stages of preprocessing in data mining science. In explaining the importance of normalization, imagine that the intervals corresponding to the values of the two properties are significantly different from each other. For example, suppose the interval for one of the attributes is the interval [1, 0] and the interval for another attribute of the same data set is [500, 1], in which case it is clear that when using criteria that are based on distance A feature with a smaller range will have virtually no significant effect on the calculations. According to these explanations, to achieve more accurate results, it is necessary that the intervals related to the different features be identical or close to each other. Normalization methods are used for this purpose. There are various methods for normalization that are used in different studies. In this study, the Min-Max method has been used.

In this simple method, each set of data is mapped to arbitrary intervals, the minimum and maximum values of which are already known. In this method, any desired

interval can be mapped to a new interval with just a simple conversion. Suppose attribute A is to be mapped from the data set between $\min_A$ to $\max_A$ and the new range new_{Min} to new_{Max}. For this purpose, any initial value such as v in the initial interval will be converted to the new value v' in the new interval according to the following equation:

$$v' = (v - \min_A) \times \frac{\text{New}_{\text{Max}} - \text{New}_{\text{Min}}}{\max A - \min A} + (\text{New}_{\text{Min}}). \quad (1)$$

2.2. Histogram Equalization. Image histogram equalization is a process that involves changes in the pixel intensity of the input image, so that the output image looks better intuitively and perceptually. Therefore, the purpose of image enhancement is to improve the interpretation or perception of the information contained therein for the viewer, or to provide better input for automated processing systems. One of the important quality factors in processing different images is contrast. If the contrast of an image is too dark, too bright, or too focused, the image information is lost in areas that are over and evenly focused. Therefore, the contrast of the image should be enhanced to show all information in the input image. The general histogram equalizer in almost all types of images is due to its simplicity and relatively good performance. The main idea of modifying the histogram is to reprogram the gray surfaces of the input image based on the uniform expansion of its probability density function. This technique flattens the dynamic range of the image histogram, increasing the contrast of the input image as much as possible.

By considering a discrete grayscale image, x with n_i as the number of occurrences of the i^{th} gray level, the probability of an occurrence of a pixel of the level i in an image is as follows:

$$v' = (v - \min_A) \times \frac{\text{New}_{\text{Max}} - \text{New}_{\text{Min}}}{\max A - \min A} + (\text{New}_{\text{Min}}), \quad (2)$$

where L describes the total number of gray levels in an image (typically 256), n signifies the image total number of pixels, and $p_x(i)$ determines the histogram of the image for pixel value i, normalized to [0, 1]. By considering the cumulative distribution function based on p_x, we have

$$\text{CDF}_x(N) = \sum_{i=0}^{N} p_x, \quad x = i. \quad (3)$$

That shows the accumulated normalized histogram of the image.

The idea is to generate a transformation of the form $y = T(x)$ to provide a new image (y), with a flat histogram. The generated image is a linearized cumulative distribution function (CDF) across the value range for some constant K is as follows:

$$\text{CDF}_y(N) = N \times K. \quad (4)$$

The CDF properties let us establish such a transform as follows:

$$\text{CDF}_y(y') = \text{CDF}_y(T(k)) = \text{CDF}_x(k), \quad (5)$$

where k is limited in the range between 0 and L. Notice that T maps the levels into the range [0, 1]; subsequently, a normalized histogram of x has been employed. To map the values back into the original range, the following simple transformation should be performed:

$$y' = y \times (\max x - \min x) + \min x. \quad (6)$$

More details can be derived in [5]. Figure 1 shows some examples of the image preprocessing step.

3. Feature Extraction

The purpose of feature extraction is to make the raw data more usable for future statistical processing. Feature extraction is a widespread process in various types of data processing such as image processing, audio processing, etc. Feature extraction means selecting a feature that can describe the image with little information. These features must have properties so that a set of these features for each image is described uniquely. Therefore, feature extraction is a process in which data is mapped from a high-dimensional space to a lower-dimensional space. This mapping can be linear (like the principal component analysis method) or nonlinear. How to select these features requires the study of data properties, and to extract it, preprocessing operations and filters must be applied to the image to turn the image into the desired information.

Numerous features are introduced for image feature extraction. In this research, geometric features, statistical features, and texture features have been employed. In the following, the formulation of this feature is explained as follows:

$$\text{perimeter} = \sum_{I=1}^{M} \sum_{j=1}^{N} B_p(i, j), \quad (7a)$$

$$\text{correlation} = \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{p(i, j) - \mu_r \mu_c}{\sigma_r \sigma_c}, \quad (7b)$$

$$\text{area} = \sum_{i=1}^{M} \sum_{j=1}^{N} p(i, j), \quad (7c)$$

$$\text{solidity} = \frac{\text{area}}{\text{convex area}}, \quad (7d)$$

$$\text{elongation} = \frac{2\sqrt{\text{area}}}{a\sqrt{\pi}}, \quad (7e)$$

$$\text{rectangularity} = \frac{\text{area}}{b \times a}, \quad (7f)$$

$$\text{irregularity index} = 4\pi \times \frac{\text{area}}{\text{perimeter}^2}, \quad (7g)$$

$$\text{eccentricity} = 2a^{-1}\left(a^2 - b^2\right)^{0.5}, \quad (7h)$$

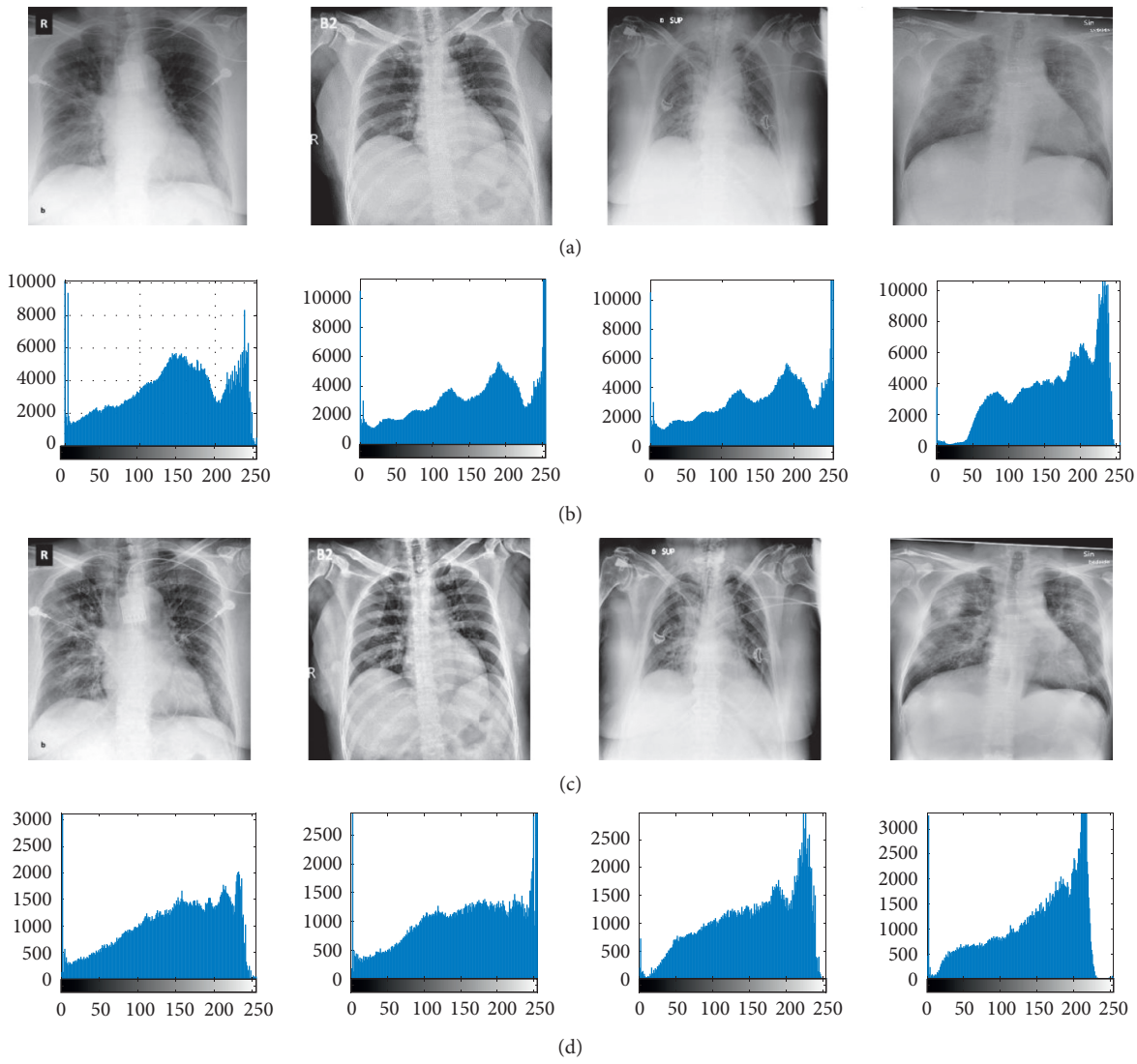


Figure 1: Some examples of preprocessing. (a) Original CXR images, (b) histogram of (a), (c) preprocessed images, and (d) histogram of (c).

$$\text{form factor} = \frac{\text{area}}{a^2}, \tag{7i}$$

$$\text{contrast} = \sum_{i=1}^{M}\sum_{j=1}^{N} p^2(i,j), \tag{7j}$$

$$\text{mean} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j), \tag{7k}$$

$$\text{entropy} = -\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j)\log\ p(i,j), \tag{7l}$$

$$\text{variance} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} (p(i,j)-\mu), \tag{7m}$$

$$\text{standard deviation} = \text{variance}^{(1/2)}, \tag{7n}$$

$$\text{energy} = \sum_{i=1}^{M}\sum_{j=1}^{N} p^2(i,j), \tag{7o}$$

$$\text{homogenity} = \sum_{i=1}^{M}\sum_{j=1}^{N} \frac{p(i,j)}{1+|i-j|}. \tag{7p}$$

Invariant moments:

$$\begin{aligned} \varphi_1 &= \eta_{20} + \eta_{02}, \\ \varphi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \\ \varphi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \mu_{03})^2, \end{aligned} \tag{7q}$$

where MN defines the image size, B_p signifies the length of the external side for the boundary pixel, $p(i,j)$ states the value of the intensity of the pixel at location (i, j), a defines the major axis, and b determines the minor axis, μ signifies the mean value σ describes the standard deviation.

Yet, some of the explained features have a low effect and some others have a high effect on the diagnosis of the image. In this study, an optimized procedure is utilized for a better selection of the useful features from the image. Therefore, in this study, we consider the important features selection as an optimization problem. The objective function is formulated as follows:

$$\text{fitness} = \frac{(\text{TP} \times \text{TN}) - (\text{FP} \times \text{FN})}{\sqrt{((\text{TN} + \text{FP}) \times (\text{TP} + \text{FP}) \times (\text{TP} + \text{FN}) \times (\text{TN} + \text{FN}))}}, \tag{8}$$

where FP, FN, TP, and TN represent false positive, false negative, true positive (TP), and true negative (TN).

To get the best value of this problem, we need a proper optimizer. Different kinds of optimization methods can be used for this, but due to the complexity and nonlinear behavior of this study, it is better to use a metaheuristic. The present study uses Archimedes optimization algorithm as one of the newest and efficient metaheuristics to solve this problem. More explanation about this algorithm is given in the next section.

4. Archimedes Optimization Algorithm

The goal of optimization is to find the best acceptable solution, given the limitations and needs of the problem. For a problem, there may be different solutions, and to compare them and to select the optimal solution, a function called the objective function is defined. The choice of this function depends on the nature of the problem. The purpose of optimization is to determine the design variables so that the objective function is minimized or maximized.

There are different types of optimization methods. In the meantime, heuristic algorithms provide good results rather than classical methods in solving optimization problems. The purpose of heuristic algorithms is to provide a solution within an acceptable time frame that is appropriate to solve the problem. The heuristic algorithm may not be the best real solution to the problem, but it can be close to the best solution.

The heuristic algorithms can be combined with optimization algorithms to improve the performance of the algorithm. A metaheuristic algorithm combines heuristic algorithms designed to find, generate, or select any exploration at any stage and provides a good solution to problems that have optimization problems. Metaheuristic algorithms take into account some of the optimization assumptions that need to be solved.

Metaheuristic algorithms are techniques that are inspired by nature, physics, and human beings and are used to solve many optimization problems. Usually, metaheuristic algorithms are used in combination with other algorithms to achieve the optimal solution or to exit the local optimal solution status. In recent years, one of the most important and promising researches has been "heuristic methods taken from nature." These methods have similarities with social or natural systems.

Their application is derived from continuous metaheuristic methods that have had very good results in solving problem problems (NP-Hard). Based on the no-free-lunch theorem, each metaheuristic algorithm has some advantages and disadvantages. Therefore, an algorithm may have the best results for a specific problem, while it gives the worst results on another problem.

This motivation makes the researchers work on different new metaheuristics to provide the better one continuously. For example, Variance Reduction of Gaussian Distribution (VRGD) [6, 7], Chimp Optimization Algorithm (COA) [8], World Cup Optimization (WCO) Algorithm [9], Red Fox Optimization (RFO) algorithm [10], Equilibrium optimizer [11], and Archimedes optimization algorithm [12] are some of the newly introduced metaheuristic algorithms.

Among the mentioned algorithms, the Archimedes optimization is the most newly introduced population-based metaheuristic algorithm, which is inspired by the Archimedes' principle and the buoyancy law. This principle simulates the time an object is immersed partially or fully in fluid, the fluid employs an upward force on the object, which is equal to the fluid discharged weight by the object.

The Archimedes optimizer is an algorithm based on a population with consideration of the immersed objects as candidates. AOA starts with several random numbers being distributed consistently with a possible limitation, similar to other metaheuristics based on the population. The searching method of objects includes random volumes (V), densities (D), and accelerations (A). The Archimedes optimizer can be introduced as a global optimization method by consideration of exploitation and exploration. This technique is described in the following. The initialization of the location of candidates is a phased one as defined below:

$$x(i) = x_l(i) + \text{rand} \times \left(x_u(i) - x_l(i)\right), \quad i = 1, 2, \ldots, N, \tag{9}$$

where $x(i)$ is the i^{th} candidate with N candidates in a population and $x_l(i)$ and $x_u(i)$ stand for the minimum and maximum ranges of the solution space.

As aforementioned, V, D, and A are initialized randomly for each *candidate* by this algorithm.

The update of V, D, and A for the i^{th} candidate for the iteration $t+$ is phase two as follows:

$$\begin{aligned} V(i) &= \text{rand}, \\ D(i) &= \text{rand}, \\ A(i) &= x_l(i) + \text{rand} \times \left(x_u(i) - x_l(i)\right). \end{aligned} \tag{10}$$

This algorithm computes the primary individuals and the candidates with optimum cost numbers are chosen here. These optimum amounts have been specified as V^{best}, D^{best}, A^{best}, and x^{best}.

Then, by the algorithm parameters, the objects have been updated. By the collision condition of the acceleration of the object with neighborhood objects, they are updated. Next, V, D, and A parameters are used to achieve the updated locations of the candidates. This is defined as follows:

$$V^{t+1}(i) = V^{t}(i) + \text{rand} \times \left(V^{\text{best}} - V^{t}(i)\right),$$
$$D^{t+1}(i) = D^{t}(i) + \text{rand} \times \left(D^{\text{best}} - D^{t}(i)\right), \tag{11}$$

where V^{best} and D^{best} signify the volume and the density of the optimum candidate, and rand defines a consistent randomly distributed number.

The application of a transfer operator (TF) and D factor is to transfer from exploration to exploitation. This is obtained as follows:

$$\text{TF} = \exp\left(\frac{t - t^{\max}}{t^{\max}}\right), \tag{12}$$

where TF is stepped up gradually in duration to reach 1, t is the number of iterations, and $t^{\max}$ is the highest number of iterations. Consequently, a reduction factor of D tries the algorithm to obtain a global to local search, which is defined as follows:

$$D^{t+1} = \exp\left(\frac{t^{\max} - t}{t^{\max}}\right) - \left(\frac{t}{t^{\max}}\right), \tag{13}$$

where D^{t+1} minimizes slowly to give a suitable convergence. Thus, the balance between exploitation and exploration is possible by this.

If $\text{TF} \leq 0.5$, candidates are considered without collision for exploitation, and the acceleration of the candidate is updated by the algorithm for $i + 1$ iteration as defined below:

$$D^{t+1} = \exp\left(\frac{t^{\max} - t}{t^{\max}}\right) - \left(\frac{t}{t^{\max}}\right), \tag{14}$$

where D^{t+1} decreases over time which can converge in the former found promising area.

This leads to a proper balance between exploration and exploitation in the AOA algorithm. Then, to update the candidate acceleration for iteration $t + 1$, a random material (mr) was chosen as follows:

$$A^{t+1} = \frac{D_{\text{mr}} + V_{\text{mr}} \times A_{\text{mr}}}{D^{t+1}(i) \times V^{t+1}(i)}, \tag{15}$$

where $A(i)$, $V(i)$, and $D(i)$ denote the acceleration, volume, and density of the i^{th} candidate, respectively, while A_{mr}, D_{mr}, and V_{mr} define the acceleration, volume, and density of random material (mr). Exploration-exploitation performance can be altered by a performance value different from 0.5.

The exploitation of the algorithm was modeled without any collision between the candidate if $\text{TF} > 0.5$. The update of the candidates was performed as below in this step:

$$A^{t+1}(i) = \frac{D^{\text{best}} + V^{\text{best}} \times A^{\text{best}}}{D^{t+1}(i) \times V^{t+1}(i)}, \tag{16}$$

where the optimum candidate acceleration is specified by A^{best}.

Therefore, the normalized acceleration can be obtained as follows:

$$A^{t+1}(\bar{i}) = u \times \frac{A^{t+1}(i) - \min(A)}{\max(A) - \min(A)} + l, \tag{17}$$

where $A^{t+1}(i)$ defines the step variation percentage of each population and l and u explain the normalization restriction between 0.9 and 0.1, respectively.

In the later phase, the location of the i^{th} updated candidate for iteration $t + 1$ if $\text{TF} \leq 0.5$ was obtained as follows:

$$x^{t+1}(i) = x^{t}(i) + c_1 \times \text{rand} \times A^{t+1}(\bar{i}) \times D \times \left(x^{\text{rand}} - x^{t}(i)\right), \tag{18}$$

where C_1 denotes a constant equal to 2.

Otherwise, the candidates renew their locations if $\text{TF} > 0.5$ as follows:

$$x^{t+1}(i) = x^{\text{best}^{t}} + F \times c_2 \times \text{rand} \times A^{t+1}(\bar{i}) \times D \times \left(T \times x^{\text{best}} - x^{t}(i)\right), \tag{19}$$

where T increments over time between $c_3 \times 0.3$ and 1 and considers a defined percentage of the optimum location and c_2 denotes a constant amount of 6. This amount increases over time to decrease the difference between the present and the optimum locations to give a proper balance between exploration and exploitation, and F states the updated model of the candidates by their locations, which is defined as follows:

$$F = \begin{cases} +1, & \text{if } P \leq 0.5, \\ +1, & \text{if } P > 0.5, \end{cases} \tag{20}$$

where

$$P = 2 \times \text{rand} - c_4. \tag{21}$$

The AOA evaluates each candidate by fitness function f and saves the optimum solution. The application of AOA here is to minimize Fitness function that is described in the previous section.

5. Image Classification

To provide a proper classification, the CXR images are grouped into two target classes where each contains 140 images including normal and COVID-19. Here, an optimized multilayer neural network (MLP) has been utilized. In this research, the proposed AOA is used again to design an optimal classifier based on MLP neural network. The main purpose here is to optimize the value of the weight to accomplish an MLP with global optimal.

Generally, artificial neural networks (ANNs) are computing systems inspired by biological neural networks. These systems learn activities by looking at examples (i.e., they improve their performance by performing activities over time), and this usually happens without any activity-specific programming. An ANN is based on a set of connected units or nodes, called artificial neurons (similar to biological neurons in the animal brain). Any connection (synapse)

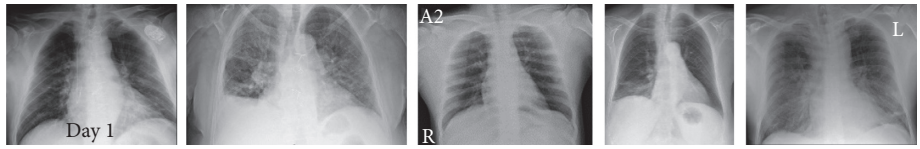


Figure 2: Some examples of the utilized dataset for COVID-19.

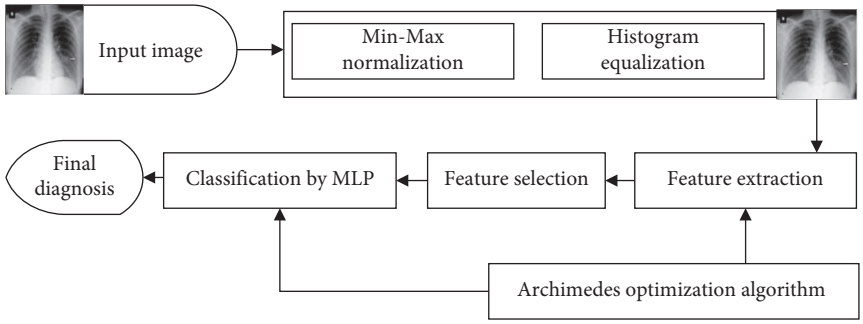


Figure 3: The graphical abstract of the proposed system.

between neurons can transmit a signal from one neuron to another. The receiving neuron (postsynaptic) can process the signal (s) and then the signal neurons attached to it. In typical ANN implementations, a synapse signal is a real number, and the output of each neuron is calculated by a nonlinear function of its inputs. Neurons and synapses usually have weights that are adjusted as learning progresses. This weight increases or decreases the signal strength it sends to the synapse. Neurons can have a threshold at which a signal can only be sent if the total signal is absorbed by that threshold.

One of the most basic neural models available is the multilayer perceptron (MLP) model, which simulates the transfer function of the human brain. In this type of neural network, most of the behaviors of human brain networks and signal propagation have been considered, and hence, they are sometimes referred to as feedforward networks. Each neuron in the human brain, called a neuron, processes the input (from another neuron or nonneuron) and transmits the result to another cell (neuron or nonneuron). This behavior continues until a definite result is reached, which is likely to eventually lead to a decision process, thought, or move. Widely used supervision in neural networks is backpropagation, which is based on the error correction law. Assume a neural network output as y. This function is reliant on the inputs, x_i, and their weights.

$$y(t) = f\left(\sum_i \omega_i x_i\right), \tag{22}$$

where α signifies the activation function (here the sigmoid function has been used).

Backpropagation is a classic method based on gradient descent to adjust the network weights such that network errors get minimized. This technique takes a lot of time.

Table 1: The main parameter setting for the network.

Parameter	Value
Batch size	2
Maximum epochs	120
Training percentage	70
Validation percentage	15
Testing	15
Dropout value	0.2
Learning rate	0.001

One big issue in using gradient descent algorithms is that they are sometimes stuck in the local minimum which is reliant on the initial weight. Therefore, they need a global minimizer for refining this issue. This leads us to use again AOA, this time for classification of the COVID-19 data.

The main purpose of AOA here is to adjust the network weights such that the network mean squared error (MSE) between the experimental data and the network value get minimized; that is,

$$\min \text{MSE} = \frac{1}{2}\sum_{k=1}^{n}\sum_{j=1}^{m}\left(Y_j - T_j\right)^2, \tag{23}$$

where m states the number of output nodes, n signifies the number of training samples, and Y_j and T_j represent the real output and the desired output, respectively.

6. Simulation Results

6.1. The Database. Recently, some different datasets have been introduced for COVID-19. This research adopts reliable resources provided by the Renmin Hospital of Wuhan University and two affiliated hospitals (the Third Affiliated Hospital and Sun Yat-sen Memorial Hospital) of the Sun Yat-sen University in Guangzhou with 76 and 12 patients,

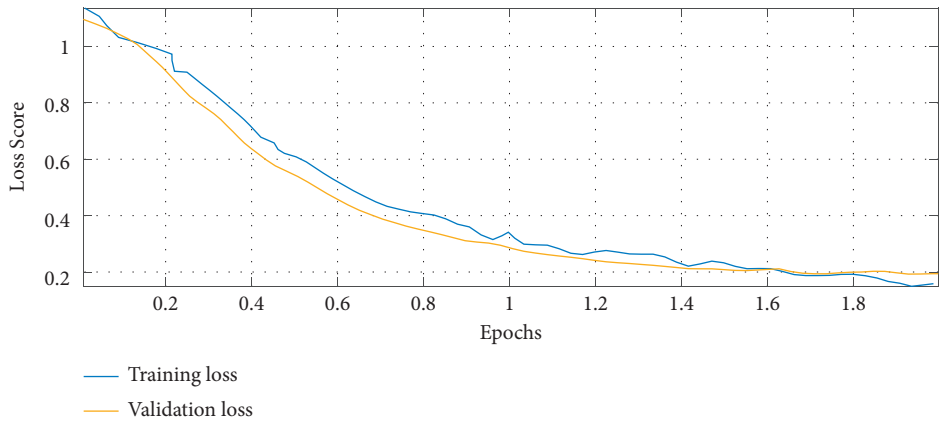


Figure 4: The loss graph of the model.

Table 2: Performance analysis of the proposed method compared with other state-of-the-art techniques.

Method	Performance metric				
	Accuracy (%)	Specificity (%)	Precision (%)	Recall (%)	*F*-measure (%)
Ismail's method [14]	84	80	93	88	90.5
Mangal's method [15]	83	75	93	93	93
Abbas's method [16]	84	78	94	92	93
Castiglioni's [17]	82	79	95	93	94
Proposed method	86	79	96	96	96

respectively [13]. Figure 2 shows some examples of the utilized dataset for the COVID-19.

6.2. System Configuration. In this study, we used an optimized X-ray diagnosis system for COVID-19 disease detection. Figure 3 shows the graphical abstract of the proposed system. As can be observed from Figure 3, the proposed CAD-based COVID-19 X-ray diagnosis system has four main steps. First, the input image has been preprocessed based on min-max normalization to scale all data into a normal scale, and the scaled images have been improved based on the histogram equalization. Then, geometric features, statistical features, and texture features are extracted from the images. In the next step, the useless or less informative features have been removed to increase the system speed, and the final step is the classification of the features based on an optimized MLP network.

As previously mentioned, Archimedes optimization algorithm is used to adjust the weights and train the proposed multilayer neural network with minimum error. The main parameter setting for the network is reported in Table 1.

The experiment has been established on MATLAB R2016b 64-bit version and executed on computation environment of Intel® Pentium® CPU 2.30 GHz, 2.3 GHz, 4 GB RAM, and 64-bit Windows 7 operating system. The training stage has been iterated 15 times independently to obtain reliable results. During the training of the proposed model, both training and validation sets were reached. Figure 4 shows the loss graph of the model.

As can be observed from the loss graph in Figure 4, there is a satisfying fit between the validation and the training curves that confirms the suggested model's efficiency and with no further overfitting or underfitting. The analysis has been accomplished based on five performance indexes including accuracy, specificity, precision, recall, and *F*1-score that are formulated follows:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}, \tag{24a}$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FN}}, \tag{24b}$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{24c}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{24d}$$

$$F_1 - \text{score} = \frac{\text{precision} + \text{recall}}{2}, \tag{24e}$$

where TP, FP, TN, and FN represent the numbers of true positives, false positives, true negatives, and false negatives, respectively.

The final results are compared with four other methods including Ismail's method [14], Mangal's method [15], Abbas's method [16], and Castiglioni's method [17]. Table 2 reports the performance analysis of the proposed method compared with other state-of-the-art techniques.

As can be observed from Table 2, the proposed method achieved the highest accuracy score among all studied methods, and Ismail's and Abbas's methods are close, and Mangal's method has the lowest score. By analyzing the *F*-

measurement, another balanced measurement, the proposed method has ranked and Mangal's and Abbas's methods with close values are ranked as the lowest scores, while Castiglioni's method has slightly higher than these two methods.

7. Conclusions

The new coronavirus, which causes COVID-19 disease, is one of the seven viruses in the coronavirus family that are transmitted from one human to another. This disease is spreading around the world and makes lots of cases and deaths every day. One of the methods for the detection of this disease is to use chest X-ray images. The present study proposed a CAD procedure for automatic diagnosis of COVID-19 from chest X-ray images. At first, the original X-ray images were preprocessed by min-max normalization to scale all data into a normal scale, and then histogram equalization was performed to improve the quality of the images before main processing. In the next step, geometric features, statistical features, and texture features were extracted from the X-ray images. Afterward, by performing an optimized method for feature selection, the useless features were removed. The optimization of feature selection and the MLP classifier was based on a newly introduced metaheuristic, called Archimedes optimization algorithm. Finally, an optimized MLP network was used to determine the suspected cases of the COVID-19. Simulations were performed in Renmin Hospital of Wuhan University and two affiliated hospitals (the Third Affiliated Hospital and Sun Yat-sen Memorial Hospital) of the Sun Yat-sen University in Guangzhou dataset, and the results were compared with four state-of-the-art methods to examine the model efficiency. Final results specified that with analyzing accuracy, specificity, precision, recall, and $F1$-score, the proposed method has the best efficiency toward the other methods. As can be inferred from the results, the proposed method provides a good optimal result for the diagnosis of COVID-19. However, in some cases, the method has weak results. This leads us to work more on developing the AOA in the future work to provide better results. The future work will be to propose a new method for improving the algorithm efficiency and to provide a method with higher speed to use in real-time applications.

Data Availability

This research adopts reliable resources provided by the Renmin Hospital of Wuhan University and two affiliated hospitals (the Third Affiliated Hospital and Sun Yat-sen Memorial Hospital) of the Sun Yat-sen University in Guangzhou with 76 and 12 patients. The dataset can be accessed by sending a request to these hospitals.

Conflicts of Interest

None of the authors have any conflicts of interest regarding the publication of this study.

References

[1] U. R. Acharya, S. L. Fernandes, J. E. WeiKoh et al., "Automated detection of Alzheimer's disease using brain MRI images–a study with various feature extraction techniques," *Journal of Medical Systems*, vol. 43, no. 9, p. 302, 2019.

[2] Y. Song, S. Zheng, L. Li et al., "Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021.

[3] J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "COVID-19 screening on chest x-ray images using deep learning based anomaly detection," 2020, https://arxiv.org/abs/2003.12338.

[4] C. Jin, W. Chen, Y. Cao et al., "Development and evaluation of an artificial intelligence system for COVID-19 diagnosis," *Nature Communications*, vol. 11, no. 1, pp. 1–14, 2020.

[5] K. G. Dhal, A. Das, S. Ray, J. Gálvez, and S. Das, "Histogram equalization variants as optimization problems: a review," *Archives of Computational Methods in Engineering*, vol. 28, no. 5, pp. 1–26, 2020.

[6] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67–82, 1997.

[7] A. Namadchian, M. Ramezani, and R. Navid, "A new metaheuristic algorithm for optimization based on variance reduction of guassian distribution," *Majlesi Journal of Electrical Engineering*, vol. 10, no. 4, p. 49, 2016.

[8] M. Khishe and M. R. Mosavi, "Chimp optimization algorithm," *Expert Systems with Applications*, vol. 149, Article ID 113338, 2020.

[9] N. Razmjooy, M. Khalilpour, and M. Ramezani, "A new metaheuristic optimization algorithm inspired by FIFA world cup competitions: theory and its application in PID designing for AVR system," *Journal of Control, Automation and Electrical Systems*, vol. 27, no. 4, pp. 419–440, 2016.

[10] D. Połap and M. Woźniak, "Red fox optimization algorithm," *Expert Systems with Applications*, vol. 166, Article ID 114107, 2021.

[11] A. Faramarzi, M. Heidarinejad, B. Stephens, and S. Mirjalili, "Equilibrium optimizer: a novel optimization algorithm," *Knowledge-Based Systems*, vol. 191, Article ID 105190, 2020.

[12] F. A. Hashim, K. Hussain, E. H. Houssein, M. S. Mabrouk, and W. Al-Atabany, "Archimedes optimization algorithm: a new metaheuristic algorithm for solving optimization problems," *Applied Intelligence*, vol. 51, no. 3, pp. 1531–1551, 2021.

[13] R. Kumar, R. Arora, V. Bansal et al., "Accurate prediction of COVID-19 using chest x-ray images through deep feature learning model with smote and machine learning classifiers," *medRxiv*, 2020.

[14] A. M. Ismael and A. Şengür, "Deep learning approaches for COVID-19 detection based on chest X-ray images," *Expert Systems with Applications*, vol. 164, Article ID 114054, 2021.

[15] A. Mangal, S. Kalia, H. Rajgopal et al., "CovidAID: COVID-19 detection using chest X-ray," 2020, https://arxiv.org/abs/2004.09803.

[16] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," *Applied Intelligence*, vol. 51, no. 2, pp. 854–864, 2021.

[17] I. Castiglioni, D. Ippolito, M. Interlenghi et al., "Artificial intelligence applied on chest X-ray can aid in the diagnosis of COVID-19 infection: a first experience from Lombardy, Italy," *medRxiv*, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 7529893, 12 pages
https://doi.org/10.1155/2021/7529893

Research Article

Method for Diagnosis of Acute Lymphoblastic Leukemia Based on ViT-CNN Ensemble Model

Zhencun Jiang,[1] **Zhengxin Dong**,[2] **Lingyang Wang**,[1] **and Wenping Jiang**[1]

[1]*School of Electrical and Electronic Engineering, Shanghai Institute of Technology, 100 Haiquan Road, Shanghai, China*
[2]*School of Electrical and Electronic Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai, China*

Correspondence should be addressed to Wenping Jiang; jiangwenping@sit.edu.cn

Received 26 May 2021; Accepted 7 August 2021; Published 23 August 2021

Academic Editor: Suresh Manic

Acute lymphocytic leukemia (ALL) is a deadly cancer that not only affects adults but also accounts for about 25% of childhood cancers. Timely and accurate diagnosis of the cancer is an important premise for effective treatment to improve survival rate. Since the image of leukemic B-lymphoblast cells (cancer cells) under the microscope is very similar in morphology to that of normal B-lymphoid precursors (normal cells), it is difficult to distinguish between cancer cells and normal cells. Therefore, we propose the ViT-CNN ensemble model to classify cancer cells images and normal cells images to assist in the diagnosis of acute lymphoblastic leukemia. The ViT-CNN ensemble model is an ensemble model that combines the vision transformer model and convolutional neural network (CNN) model. The vision transformer model is an image classification model based entirely on the transformer structure, which has completely different feature extraction method from the CNN model. The ViT-CNN ensemble model can extract the features of cells images in two completely different ways to achieve better classification results. In addition, the data set used in this article is an unbalanced data set and has a certain amount of noise, and we propose a difference enhancement-random sampling (DERS) data enhancement method, create a new balanced data set, and use the symmetric cross-entropy loss function to reduce the impact of noise in the data set. The classification accuracy of the ViT-CNN ensemble model on the test set has reached 99.03%, and it is proved through experimental comparison that the effect is better than other models. The proposed method can accurately distinguish between cancer cells and normal cells and can be used as an effective method for computer-aided diagnosis of acute lymphoblastic leukemia.

1. Introduction

Leukemia is cancer with an extremely high fatality rate. It is a hematological malevolent tumor caused by the malicious cloning of immature white blood cells in the bone marrow. Leukemia can be further divided into acute leukemia and chronic leukemia. Chronic leukemia normally has a long onset period. While quite the opposite, without special treatment, the average survival period for acute leukemia is only three months. As a type of acute leukemia, acute lymphocytic leukemia is the most important kind of childhood leukemia, and it accounts for 25% of childhood cancers [1]. In more than 50 years of development, the treatment of acute lymphoblastic leukemia has made great progress. With special treatment at the early stage, the initial complete response rate can reach more than 70% [2]. Therefore, it is particularly important to diagnose acute lymphoblastic leukemia in the early stage of its onset. In general, the main diagnostic method for acute lymphoblastic leukemia is through morphology. When there are a large number of B-lymphoblast cells (cancer cells) in the bone marrow, it can be diagnosed as acute lymphoblastic leukemia [3]. Accurately distinguishing B-lymphoid precursors (normal cells) from cancer cells is the key to the diagnosis of acute lymphoblastic leukemia. However, under a microscope, cancer cells are very similar to normal cells that it is hard to classify them.

Assisting in the diagnosis of diseases with computer vision technology is a promising research direction in recent years. In computer vision technology, image recognition

through deep learning is one of the important methods. As one of neural networks used most frequently in deep learning, convolutional neural network (CNN) has strong self-learning ability, adaptive ability, and generalization ability. Traditional image recognition methods require manual feature extraction and classification, while CNN only needs the image data as an input of the network, and the self-learning ability of the network can complete the image classification [4, 5]. Nahid et al. used a multichannel convolution neural network to identify chest radiographs and diagnose pneumonia. The classification accuracy rate of chest radiographs reached 97.92%, which provided a very reliable detection method [6]. Daoud et al. combined the methods of deep learning for extracting image features and manual extraction for processing breast ultrasound images to classify breast tumors. The average accuracy rate of classification reached 96.1%, which meant breast cancer could be accurately detected by breast ultrasound images [7]. Yang et al. applied deep learning to the recognition of bladder cancer, and the recognition accuracy rate in practical application reached 83.36%, which was the same as that of medical experts, proving the effectiveness of deep learning in the diagnosis of bladder cancer [8].

Similarly, there are some researchers using computer vision technology to diagnose leukemia. Ahmed et al. used machine learning algorithms and convolutional neural networks to classify four types of leukemia, with the highest accuracy of 88.25% [9]. Boldú et al. proposed a machine learning method for diagnosing acute leukemia based on peripheral blood images. Colour clustering and mathematical morphology were used to segment the images, and then machine learning algorithms were used to classify six types of cells. The accuracy rate of cell classification reached 85.8%, and the correct diagnosis rate for leukemia reached 94% [10]. Kasani et al. combined two models of NASNet-Large and VGG19 to classify leukemic B-lymphoblast cells and normal B-lymphoid precursor cells with a classification accuracy rate of 96.58%, which can accurately diagnose acute lymphoblastic leukemia and proved that the ensemble model was more effective than a single network [11].

The transfer learning method has also been widely used in medical image classification in recent years. Alshazly et al. used transfer learning to train chest CT images to diagnose COVID-19 patients, obtained 92.9% accuracy rate on the COVID-19-CT dataset, and then used visualization technology to explain the model predictions clearly [12]. El-Khatib et al. used deep learning methods to diagnose skin lesions. They trained a variety of image classification models based on convolutional neural network through training methods of transfer learning to distinguish different types of skin lesions. The experimental result showed that this method had a better effect in diagnosing skin lesions [13]. Brodzicki et al. trained a convolutional neural network model to classify *Clostridioides difficile* bacteria cytotoxicity using transfer learning methods and achieved a classification accuracy rate of 93.5% on 369 images, with excellent recognition [14].

The models used in these studies are all based on convolutional neural networks (CNNs). Unlike these CNN models, which rely too much on convolutional layers, vision transformer is based on the transformer structure, which is a deep neural network based on the self-attention mechanism. Transformer structure was first applied in the field of natural language processing (NLP), and the researchers extended it to the field of computer vision. Compared with the CNN model, the model based on transformer structure performs better in image classification. Bazi et al. applied the vision transformer to remote sensing image classification and used the CutMix data enhancement method to test multiple remote sensing image data sets. The experimental results showed that the vision transformer classification accuracy rate of remote sensing images exceeds the CNN model [15].

In summary, it is of great significance to assist in the diagnosis of acute lymphocytic leukemia by classifying leukemic B-lymphoblast cells (cancer cells) and B-lymphoid precursors (normal cells). In this article, we propose the ViT-CNN ensemble model to assist in the diagnosis of acute lymphoblastic leukemia. The main contributions are as follows:

(1) We propose the ViT-CNN ensemble model to distinguish cancer cells and normal cells. This is a model that uses two different methods to extract and combine features from cell images. This is the first time the vision transformer model and the CNN model have been combined to diagnose acute lymphocytic leukemia.

(2) We propose a data enhancement method of difference enhancement-random sampling (DERS), which solves the problem of data set imbalance.

(3) The ViT-CNN ensemble model has a classification accuracy rate of 99.03% for cancer cells and normal cells.

(4) We compare the ViT-CNN ensemble model with the ordinary CNN model and other ensemble models and prove that the ViT-CNN ensemble model proposed in this article performs better in classification accuracy.

The second section presents the data set used in this article, the method of processing the data, the model used in this article, the loss function, and the optimizer. The third section presents the experimental process and experimental results of the proposed method, as well as the comparison results with other models. The full article is summarized in the fourth section.

2. Materials and Methods

2.1. Data Set. The data set used to build the diagnostic model is the ISBI 2019 data set [16–20], and 10661 cell pictures of 73 subjects were selected, including 7272 pictures of leukemic B-lymphoblast cells (cancer cells) from 47 all patients and 3389 pictures of B-lymphoid precursors (normal cells) from 26 healthy persons. These cells have been segmented from the microscopic images, and each cell picture is a real image after collection. Some staining noise and illumination errors generated during the collection process have been repaired

to a large extent. As shown in Figure 1, the morphology of the two cells is very similar, so a professional oncologist will annotate the label of the image. The labels of normal cells images are positive samples, and the labels of cancer cells images are negative samples.

2.2. Difference Enhancement-Random Sampling. Directly training on an unbalanced data set can easily cause the model to fall into overfitting or cause the generalization ability of the model to be weak. In order to solve the problem of data imbalance, this article proposes a data enhancement method based on difference enhancement-random sampling (DERS). Suppose the unbalanced data set D, there are a images of category A and b images of category B, where $a < b$. N kinds of data enhancement are performed on category A, and M kinds of data enhancements are performed on category B so that the number of $a \times N$ and $b \times M$ are relatively close. Then L images are selected from $a \times N$ images of category A, and L images are selected from $b \times M$ images of category B, thus ensuring that the number of category A and the number of category B in the new data set are the same, so that the new data set becomes a balanced data set.

2.3. Data Processing. Figure 2 shows the number of normal cells and cancer cells in the data set. It can be seen that this data set is an unbalanced data set.

For the data set used in this article, there are two categories of cell images, which are images of normal cells and cancer cells. The number of images of cancer cells is more than twice the number of images of normal cells. We use the method of difference enhancement-random sampling to process the data set.

Three data enhancement methods of left and right flip, counterclockwise rotation 90°, and pixel matrix transpose are used to generate new cancer cells images. Six data enhancement methods of left and right, flipping up and down flip, counterclockwise rotating 90°, counterclockwise rotating 180°, counterclockwise rotating 270°, and pixel matrix transpose are used to generate new normal cells images. The examples of the generated images are shown in Figure 3.

The number of cancer cells images was 29,088, and the number of normal cells images was 23,702. A new data set is created with 20,000 images randomly drawn from two newly generated images. The new data set is a completely balanced data set. Figure 4 shows the number of normal cells and cancer cells in the new data set.

Before training the model, the images need to be preprocessed. The original size of the cells images is 450×450, the size of the cells images is adjusted to 224×224, and the image is normalized to prevent overfitting of the model.

2.4. The Overall Flow of the Method. The method proposed in this article diagnoses acute lymphoblastic leukemia by distinguishing leukemic B-lymphoblast cells (cancer cells) images and B-lymphoid precursors (normal cells). After the data is preprocessed, two image classification models are trained by transfer learning. The symmetric cross-entropy loss function is selected as the loss function, the RAdam optimizer is selected as the optimizer. We combine the two models to ViT-CNN ensemble model based on the weighted sum method, it helps doctors realize computer-aided diagnosis. The process is shown in Figure 5.

2.5. Model. This section presents the ViT model and the CNN model used in the ViT-CNN ensemble model proposed in this article.

2.5.1. Vision Transformer. Transformer structure is widely used in NLP (natural language processing) [21]. The vision transformer model is completely implemented based on the transformer structure without any CNN structure [22]. Transformer structure consists of a set of encoder components and a set of decoder components, whereas the vision transformer model is an image classification model and does not require a decoder. Therefore, there is only an encoder component in transformer structure of the vision transformer. The encoder component is composed of a stack of six identical encoders. Each encoder is composed of a multihead attention layer and a feed forward layer, and both layers contain the structure of residual connection and the structure of LayerNorm. The structure of an encoder component of the vision transformer is shown in Figure 6.

The multihead attention is a kind of self-attention structure, and it allows the model to pay attention to different aspects of information, as shown in formula (1)–formula (3) of multihead attention.

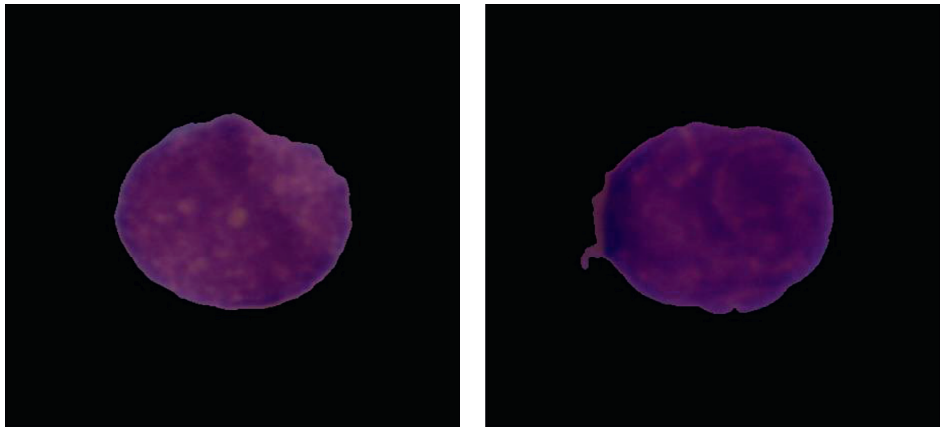
$$\begin{aligned} Q_i &= QW_i^Q, \\ K_i &= KW_i^K, \\ V_i &= VW_i^V, \\ &\quad i = 1, \ldots, 8, \end{aligned} \tag{1}$$

$$\begin{aligned} \text{head}_i &= \text{Attention}(Q_i, K_i, V_i), \\ &\quad i = 1, \ldots, 8, \end{aligned} \tag{2}$$

$$\text{MultiHead}(Q, K, V) = \text{Concact}(\text{head}_1, \ldots, \text{head}_8)W^o. \tag{3}$$

In these formulas, Q means the query vector, K means the key vector, V means the value vector, and W means the weight matrix.

The linear embedding layer is an important structure in the vision transformer model. The linear embedding layer divides the image into multiple patches and then flattens the patches into a one-dimensional tensor. After the patch embedding operation is completed, location embedding and class embedding are added and input into transformer encoder. After being output by the transformer encoder, it will go through an MLP head structure, which is composed of a fully connected layer and an activation function. The activation function used here is GELU (Gaussian error linear unit), and its formula is shown as follows:



(a) (b)

Figure 1: (a) Normal cells. (b) Cancer cells.

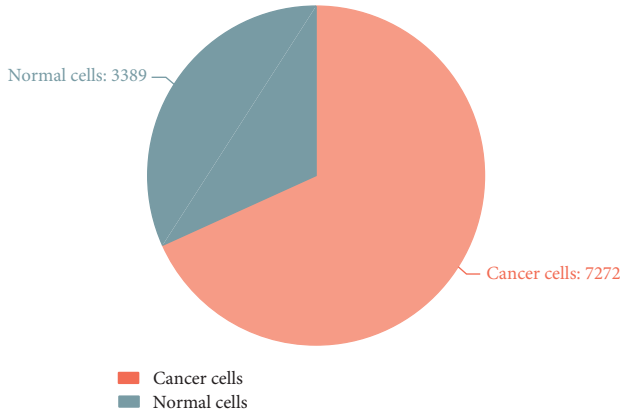


Figure 2: The number of normal cells and cancer cells.

$$\mathrm{GELU}(x) = 0.5x\left(1 + \tanh\left(\sqrt{\frac{2}{\pi}}\left(x + 0.044715x^3\right)\right)\right). \tag{4}$$

After outputting through the MLP head structure, a classification task will be performed. The problem studied in this article is a two-class classification problem, so the final output category of the vision transformer model is changed to two categories. The model overview of the vision transformer model is shown in Figure 7.

2.5.2. CNN Model. Convolutional neural network (CNN) is composed of convolutional layers, activation function, pooling layers, and fully connected layers. In the CNN classification models, the convolutional layer, the activation function, and the pooling layer constitute the feature extraction layer to extract the features, while the full connection layer forms a classification layer for classification [23, 24]. The pooling layer is a down-sampling operation to reduce the dimensionality of the extracted features while retaining important information of the features. The convolutional layer is the core structure of CNN, as shown in the following formula:

$$y(t) = \int_{-\infty}^{\infty} x(p)h(t-p)\mathrm{d}p = x(t) * h(t). \tag{5}$$

In this article, EfficientNet is selected as the CNN model. EfficientNet is an image classification model proposed by the Google team in 2019 [25]. It is known as the strongest image classification model today. The EfficientNet model balances resolution, depth, and width to optimize efficiency and accuracy. The main idea of the EfficientNet is that all convolutional layers of a convolutional neural network must be uniformly expanded by the same proportional constant. The EfficientNet model uses MBCConv in MobileNet V2 as the backbone network of the model and uses the squeeze and excitation method in SENet to optimize the network structure. EfficientNet has eight versions of models, and this article uses the EfficientNet-b0 model. The original model has 1000 classifications, and this article has improved the original model, redesigned the classification layer, and improved its final output category to two categories. The structure of the EfficientNet-b0 model is shown in Figure 8.

2.6. ViT-CNN Ensemble Model. The ensemble model is a method to improve the accuracy rate of the model [26, 27]. The greater the differences between the models, the greater the performance improvement after the ensemble. Two different feature extraction methods extract the features of cell images, which can more comprehensively distinguish the differences between images and obtain better classification results. The ViT-CNN ensemble model combines two completely different feature extraction methods, vision transformer model and CNN (EfficientNet) model, to extract the features of cells images and classify them. The ensemble model method used in this article is the weighted sum method. The output results of the vision transformer model are multiplied by a coefficient of 0.7, and the output results of the EfficientNet model are multiplied by a coefficient of 0.3, and then the two results are added up together as the final prediction result.

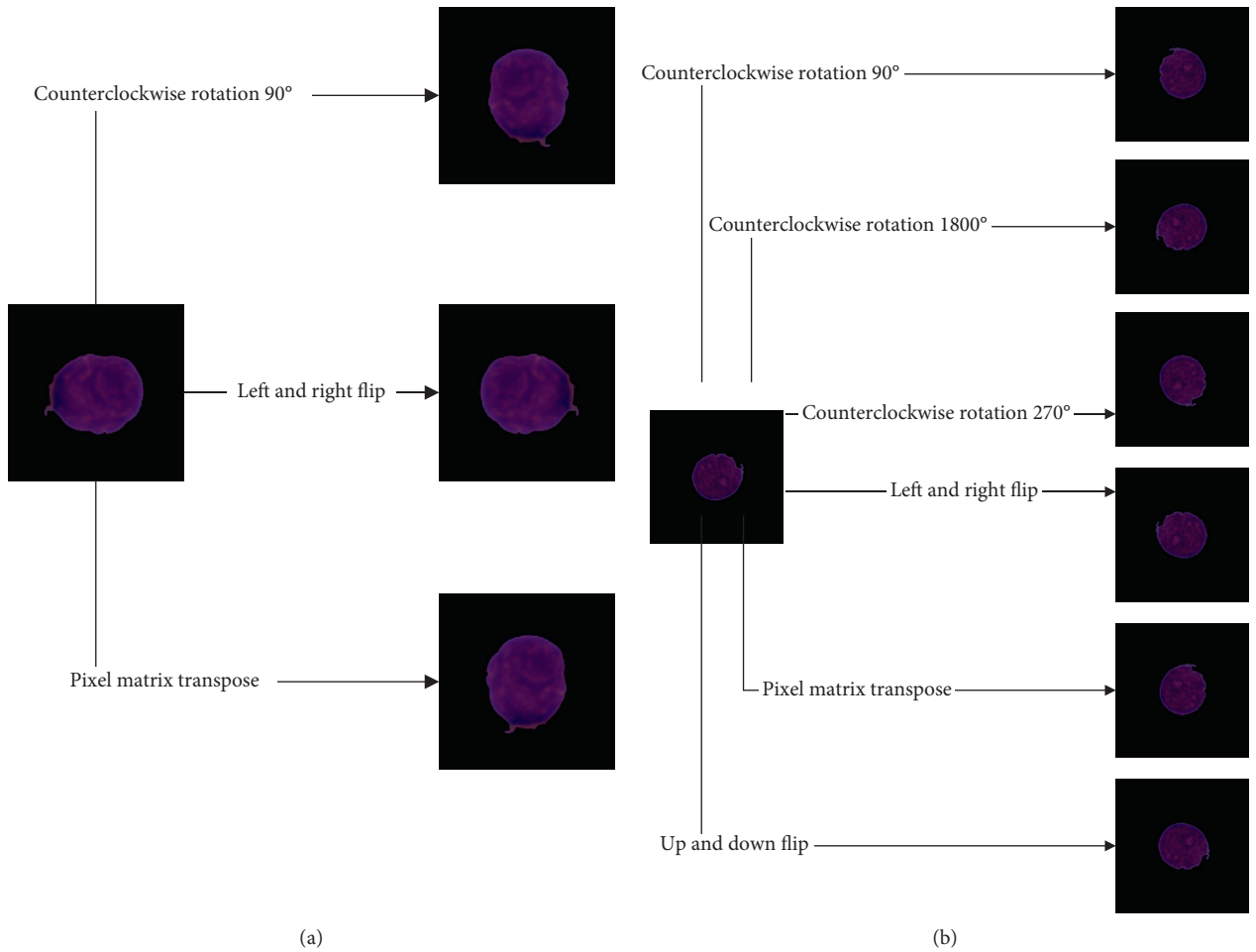


Figure 3: The examples of the generated images: (a) cancer cells; (b) normal cells.

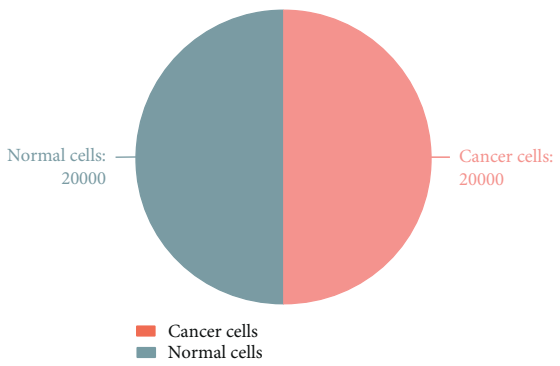


Figure 4: The number of normal cells and cancer cells in the new data set.

2.7. Transfer Learning. Transfer learning solves the shortcoming that deep learning needs numerous samples to train models. The pretraining model obtained by training on large data sets can be trained with a few data sets, and the training time required for deep learning is greatly shortened. Fine-tuning is a kind of transfer learning method, and it takes the weight of the pretraining model as the initial weight and trains on the basis of the initial weight without training the model from scratch. Fine-tuning not only increases the convergence speed and generalization ability of the model but also reduces the risk of overfitting [28–30]. In this article, both the ViT model and the CNN model are trained by the method of fine-tuning of transfer learning.

2.8. Symmetric Cross-Entropy Loss Function. For medical images data sets, a bit of noise is unavoidable. The morphological similarity of cancer cells and normal cells results in some noise on the label. The symmetric cross-entropy loss function can reduce the influence of noise and prevent overfitting [31]. The definition of the symmetric cross-entropy loss function is shown as follows:

$$l_{\text{sce}} = l_{\text{ce}} + l_{\text{rce}}. \tag{6}$$

Among them, l_{ce} is the cross-entropy loss function and l_{rce} is the reverse cross-entropy function. Its definition is shown as follows:

$$l_{\text{rce}} = -\sum_{k=1}^{K} p(k \mid x) \log q(k \mid x). \tag{7}$$

Therefore, the symmetric cross-entropy loss function is used in this article as the loss function of the model to reduce

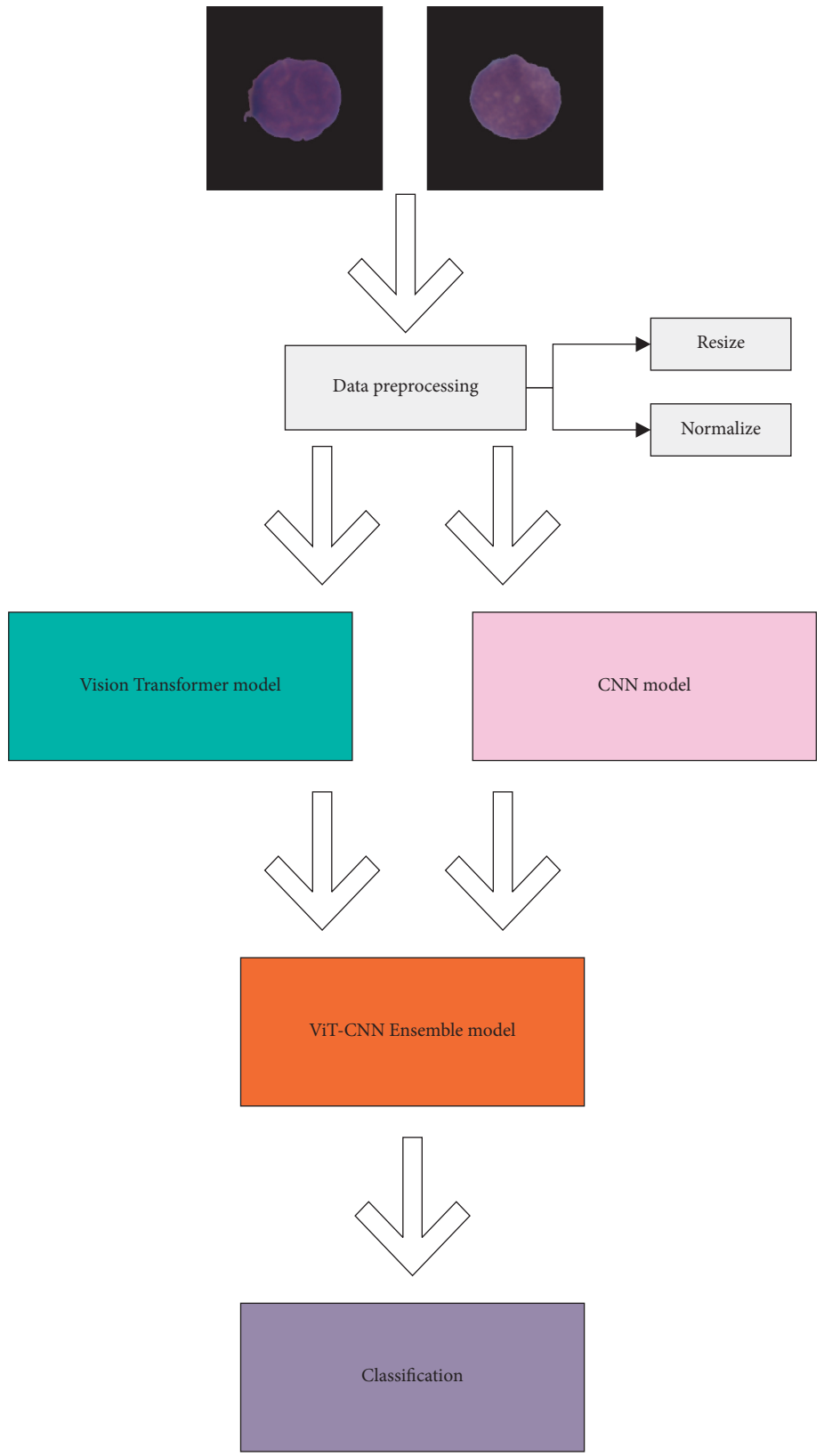


Figure 5: The process of the method.

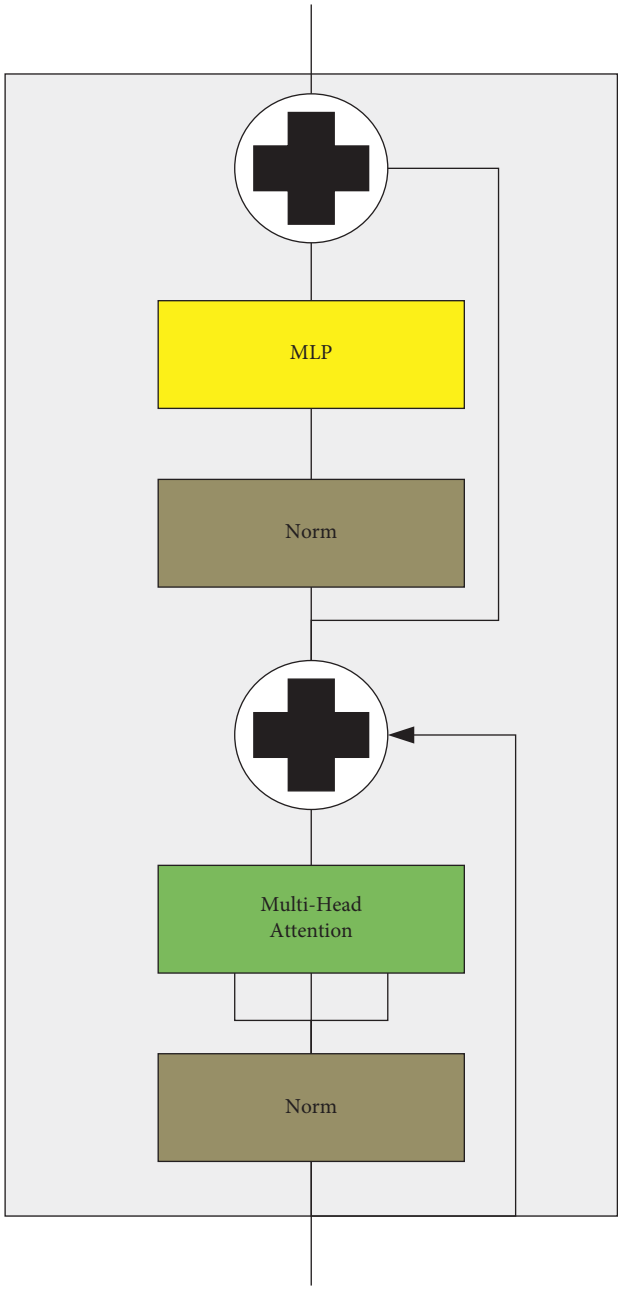


Figure 6: The structure of an encoder component of the vision transformer.

the influence of noise on the generalization ability of the model.

2.9. Optimizer. RAdam optimizer is a new optimizer based on the classic Adam optimizer. In the training process of the model, the initial learning rate is very important. The Adam optimizer has too large variance and more uncertain factors, which leads to its insufficient stability. The RAdam optimizer is based on the Adam optimizer, which aims to introduce a correction to the adaptive learning rate to correct the variance of the Adam optimizer [32]. Compared with SGD optimizer and Adam optimizer, SGD optimizer has better convergence effect but slow convergence speed and Adam optimizer has fast convergence speed but easily converges to the local optimal solution. RAdam optimizer has good performance in convergence effect and convergence speed and is relatively compared with Adam optimizer, and it can stably improve accuracy.

3. Results and Discussion

3.1. Experiment Platform. The experimental platform used in this article includes a hardware environment, which consisted of Intel Core i7-9700f processor, NVIDIA RTX2060s 8 GB graphics card, and 16 GB memory. The proposed model is implemented in python3.7 using PyTorch [33] framework.

3.2. Performance Metrics. There are many evaluation indicators for image classification, but for the problem of assisting diagnosis of cancer through image classification, accuracy and precision of the classification model play a significant role. Therefore, accuracy and precision are selected as the indicator for evaluating the performance of the proposed model. The accuracy rate is defined as the ratio of correctly recognized positive samples and negative samples to total samples, as shown in the following formula:

$$\mathrm{ACC} = \frac{(\mathrm{TP} + \mathrm{TN})}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}. \tag{8}$$

The precision rate is the ratio of correctly recognized positive samples to all positive samples, as shown in the following formula:

$$\mathrm{precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}. \tag{9}$$

In the above formulas, TP is the samples predicted to be positive samples in the positive samples, TN is the samples predicted to be negative samples in the negative samples, FP is the samples predicted to be positive samples in the negative samples, and FN is the samples predicted to be negative samples in the positive samples.

3.3. Experimental Comparison. The implementation process of the method proposed in this article is as follows:

Step 1. The new balanced data set is divided into training set, validation set, and test set at a ratio of 8 : 1 : 1

Step 2. The learning rate is set to 0.0001, the training batch size is set to 32, and the validation batch size is set to 32

Step 3. The vision transformer model is trained for 30 epochs, and then the parameter model with the highest accuracy is saved in the validation set

Step 4. The EfficientNet model is trained for 30 epochs, and then the parameter model with the highest accuracy is saved in the validation set

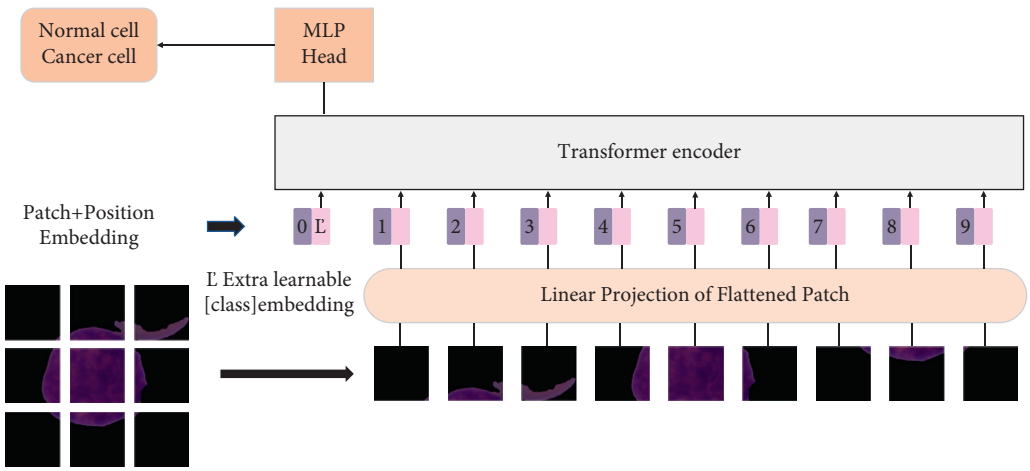


Figure 7: The model overview of the vision transformer model.

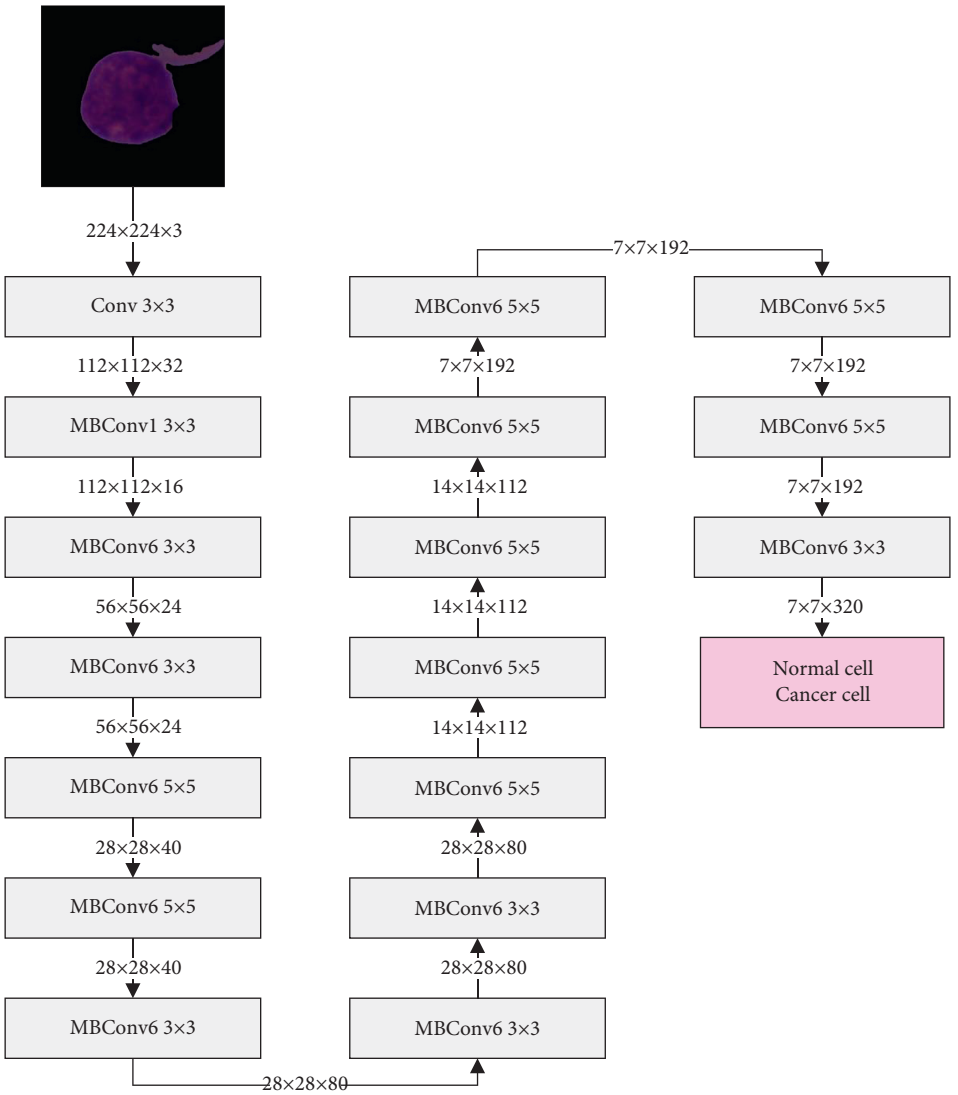


Figure 8: The structure of the EfficientNet-b0 model.

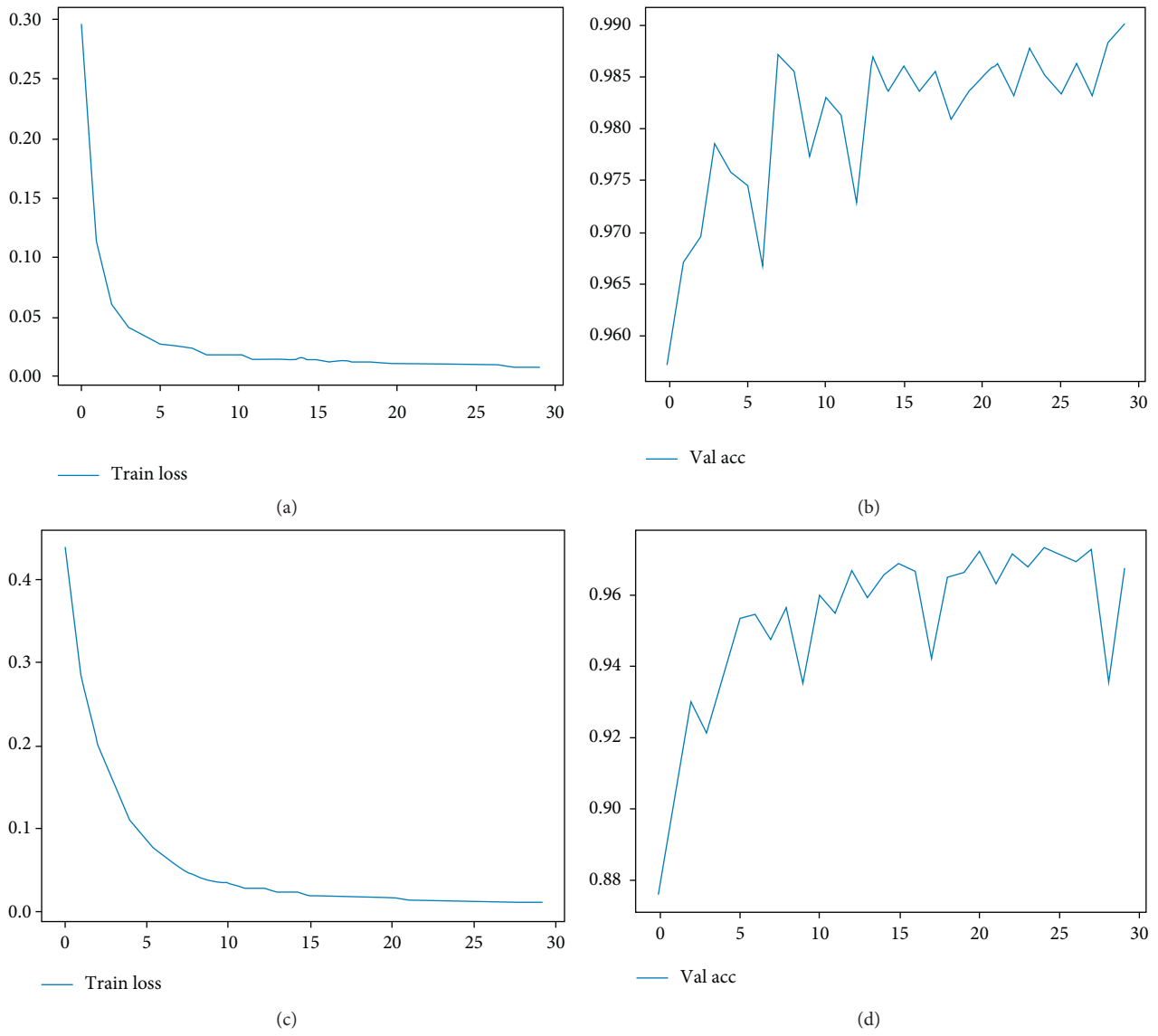


Figure 9: Loss change curve of training set and accuracy change curve of validation set. (a) Train loss of the vision transformer model. (b) Val accuracy of the vision transformer model. (c) Train loss of the EfficientNet model. (d) Val accuracy of the EfficientNet model.

Step 5. The vision transformer model and the EfficientNet model are integrated into the ViT-CNN ensemble model for testing

The experiment is carried out according to the above procedure. The loss change curve of the training set and the accuracy change curve of the validation set of the vision transformer model and the EfficientNet model are shown in Figure 9.

The accuracy rate and the precision rate of the vision transformer model, the EfficientNet model, and the ViT-CNN ensemble model on the test set are shown in Table 1.

As can be seen from Table 1, the accuracy rate of the vision transformer model is 3.72% higher than that of the EfficientNet model and the precision rate of the vision transformer model is 1.52% higher than that of the EfficientNet model, and those showed that the performance of the vision transformer model is better than the performance of the EfficientNet model. After being ensemble into the ViT-CNN ensemble model, the accuracy rate reached 99.03%, the precision rate reached 99.14%, and the performance is further improved.

Table 1: The accuracy and precision of three different models.

Model	Accuracy (%)	Precision (%)
Vision transformer	98.90	98.90
EfficientNet	95.18	97.38
ViT-CNN ensemble model	99.03	99.14

The accuracy rate can judge the classification ability of the model, but the specific details cannot be reflected. The confusion matrix is the comparison matrix between the predicted result and the actual value, which can

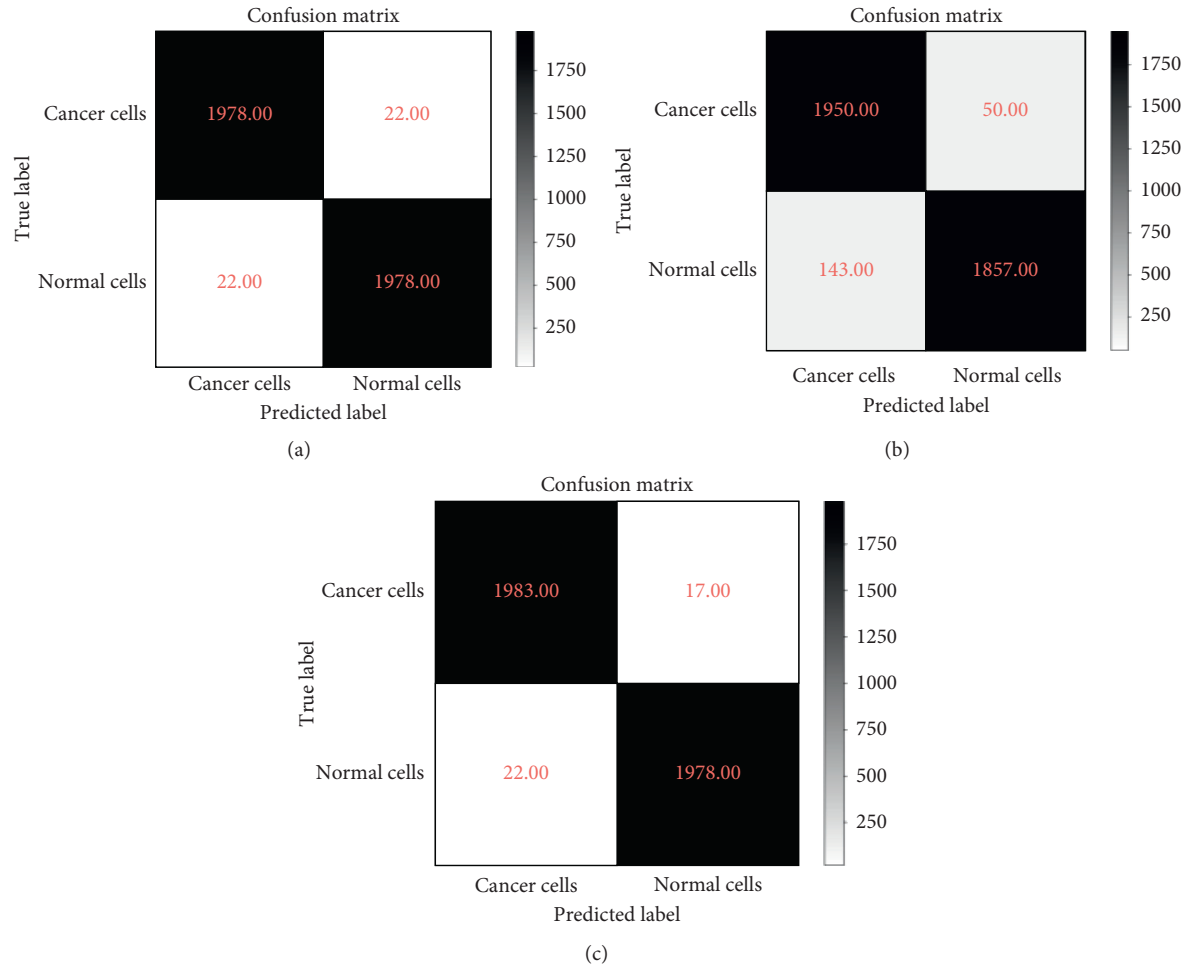


Figure 10: Confusion matrix of three models: (a) the vision transformer model; (b) the EfficientNet model; (c) the ViT-CNN ensemble model.

clearly indicate the prediction details of each category when the classification model is making predictions. The confusion matrix is used to further analyze the classification ability of the ensemble model proposed in this article. The confusion matrix of three models is shown in Figure 10.

It can be seen from the confusion matrix that the recognition ability of the vision transformer model is very balanced, with the recognition accuracy of cancer cells and normal cells as the same. The EfficientNet model has a difference in the ability to recognize cancer cells and normal cells, and the ability to recognize cancer cells is stronger. The ViT-CNN ensemble model has the same recognition accuracy for normal cells as the vision transformer model but has a stronger ability to recognize cancer cells. In practical applications, better identification of cancer cells can more accurately diagnose acute lymphocytic leukemia, so the ViT-CNN ensemble model can in a superior way assist in the diagnosis of acute lymphocytic leukemia.

In order to prove the effectiveness of the method, the proposed model in this article is compared with the following models:

Table 2: The comparison of accuracy of five different models.

Model	Accuracy (%)
Resnet50	94.95
Densenet121	93.65
VGG16	95.24
Model in literature [11]	96.58
ViT-CNN ensemble model	**99.03**

(1) *Other CNN Models*. This article compared the ViT-CNN ensemble model with Resnet50, Densenet121, and VGG16 three classic CNN models.

(2) *Model in Literature* [11]. Literature [11] has the best current research results on diagnosis of acute lymphoblastic leukemia. This article compared the accuracy of the model they proposed.

The specific comparison is shown in Table 2.

It can be seen from Table 2 that the accuracy rate of the ViT-CNN ensemble model is 4.08%, 5.38%, and 3.79% higher than the accuracy rate of the Resnet50, Densenet121, and VGG16. This shows that the ViT-CNN ensemble model

has a better ability to classify cancer cells and normal cells. Compared with the model in literature [11], the accuracy rate of the model proposed in literature [11] is 96.58%, while the accuracy rate of the ViT-CNN ensemble model is 99.03%; this accuracy rate is 2.45% higher than the model proposed in the literature. Obviously, the ViT-CNN ensemble model has better classification performance and can assist in the diagnosis of acute lymphocytic leukemia more accurately.

4. Conclusions

In this article, we proposed a diagnostic approach for acute lymphocytic leukemia, which could classify cancer cells and normal cells through an ensemble model to assist doctors in the diagnosis in reality. The ISBI 2019 data set was used in the article; we proposed the difference enhancement-random sampling (DERS) method to solve the problem of data imbalance. We designed an ensemble model that integrates the vision transformer model and the EfficientNet model into the ViT-CNN ensemble model. The accuracy of this model in the classification of B-lymphoblastic cells and normal B-lymphoid precursors was 99.03%. We compared the ViT-CNN ensemble model with Resnet50, Densenet121, and VGG16 three classic convolutional neural Nnetwork models. The ViT-CNN ensemble model significantly outperformed these previous models. The results showed that the model proposed in this article was superior to other models in accuracy and had a balanced classification ability, which could better assist in the diagnosis of acute lymphoblastic leukemia.

Data Availability

The data used in this work are from public datasets: ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging (https://competitions.codalab.org/competitions/20395). To apply for the access to dataset, a registration is required.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

Zhencun Jiang and Zhengxin Dong contributed equally to this work. J.Z.C. and D.Z.X. conceptualized the study. J.Z.C. and D.Z.X. developed methodology. J.Z.C. and D.Z.X. provided software. J.Z.C. and D.Z.X. validated the study. J.W.P. was responsible for formal analysis. J.Z.C. and D.Z.X. investigated the study. J.W.P. was responsible for resources. J.Z.C. and D.Z.X. were involved in data curation. J.Z.C. and W.L.Y. prepared the original draft. J.Z.C. and W.L.Y. reviewed and edited the article. J.Z.C. and D.Z.X. visualized the study. J.Z.C. and D.Z.X. supervised the study. J.Z.C. and D.Z.X. were involved in project administration. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The authors would like to thank SBILab for the data and the donations of 73 volunteers. This work was supported by the Natural National Science Foundation of China (81827807, 61675134, and 61307015) and Science and Technology Innovation Project of Shanghai Science and Technology Commission (19441905800).

References

[1] T. C. Fujita, N. Sousa-Pereira, M. K. Amarante, and M. A. E. Watanabe, "Acute lymphoid leukemia etiopathogenesis," *Molecular Biology Reports*, vol. 48, no. 1, pp. 817–822, 2021.

[2] L. Li and Y. Wang, "Recent updates for antibody therapy for acute lymphoblastic leukemia," *Experimental Hematology & Oncology*, vol. 9, no. 1, pp. 1–11, 2020.

[3] X. Peng, W. Yang, L. Jia et al., "Application of cell morphology-related technology in hematological tumors," *Chinese Journal of Biological Engineering*, vol. 39, no. 9, pp. 84–90, 2019.

[4] W. Zhao, F. Chen, H. Huang, D. Li, and W. Cheng, "A new steel defect detection algorithm based on deep learning," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5592878, 13 pages, 2021.

[5] A. Lavric and P. Valentin, "KeratoDetect: keratoconus detection algorithm using convolutional neural networks," *Computational Intelligence and Neuroscience*, vol. 2019, Article ID 8162567, 9 pages, 2019.

[6] A.-A. Nahid, N. Sikder, A. K. Bairagi et al., "A novel method to identify pneumonia through analyzing chest Radiographs employing a multichannel convolutional neural network," *Sensors*, vol. 20, no. 12, p. 3482, 2020.

[7] M. I. Daoud, S. Abdel-Rahman, T. M. Bdair, M. S. Al-Najar, F. H. Al-Hawari, and R. Alazrai, "Breast tumor classification in ultrasound images using combined deep and handcrafted features," *Sensors*, vol. 20, no. 23, p. 6838, 2020.

[8] R. Yang, Y. Du, X. Weng et al., "Automatic recognition of bladder tumours using deep learning technology and its clinical application," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 17, no. 2, p. e2194, 2020.

[9] N. Ahmed, A. Yigit, Z. Isik, and A. Alpkocak, "Identification of leukemia subtypes from microscopic images using convolutional neural network," *Diagnostics*, vol. 9, no. 3, p. 104, 2019.

[10] L. Boldú, A. Merino, S. Alférez, A. Molina, A. Acevedo, and J. Rodellar, "Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis," *Journal of Clinical Pathology*, vol. 72, no. 11, pp. 755–761, 2019.

[11] P. H. Kasani, S.-W. Park, and J.-W. Jang, "An aggregated-based deep learning method for leukemic B-lymphoblast classification," *Diagnostics*, vol. 10, no. 12, p. 1064, 2020.

[12] H. Alshazly, C. Linse, E. Barth, and T. Martinetz, "Explainable COVID-19 detection using chest CT scans and deep learning," *Sensors*, vol. 21, no. 2, p. 455, 2021.

[13] H. El-Khatib, D. Popescu, and L. Ichim, "Deep learning-based methods for automatic diagnosis of skin lesions," *Sensors*, vol. 20, no. 6, p. 1753, 2020.

[14] A. Brodzicki, J. Jaworek-Korjakowska, P. Kleczek, M. Garland, and M. Bogyo, "Pre-trained deep convolutional neural network for Clostridioides difficile Bacteria

cytotoxicity classification based on fluorescence images," *Sensors*, vol. 20, no. 23, p. 6713, 2020.

[15] Y. Bazi, L. Bashmal, M. M. A. Rahhal, R. A. Dayil, and N. A. Ajlan, "Vision transformers for remote sensing image classification," *Remote Sensing*, vol. 13, no. 3, p. 516, 2021.

[16] R. Gupta, P. Mallick, R. Duggal, A. Gupta, and O. Sharma, "Stain color normalization and segmentation of plasma cells in microscopic images as a prelude to development of computer assisted automated disease diagnostic tool in multiple myeloma," in *Proceedings of the 16th International Myeloma Workshop (IMW)*, New Delhi, India, March 2017.

[17] R. Duggal, A. Gupta, R. Gupta, M. Wadhwa, and C. Ahuja, "Overlapping cell nuclei segmentation in microscopic images using deep belief networks," in *Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)*, Guwahati, India, December 2016.

[18] R. Duggal, A. Gupta, and R. Gupta, "Segmentation of overlapping/touching white blood cell nuclei using artificial neural networks," in *Proceedings of the CME Series on Hemato-Oncopathology*, All India Institute of Medical Sciences (AIIMS), New Delhi, India, July 2016.

[19] R. Duggal, A. Gupta, R. Gupta, and P. Mallick, "SD-layer: stain deconvolutional layer for CNNs in medical microscopic imaging," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2017*, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. Collins, and S. Duchesne, Eds., pp. 435–443, Springer, Quebec City, Canada, September 2017.

[20] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, "Albumentations: fast and flexible image augmentations," *Information*, vol. 11, no. 2, p. 125, 2020.

[21] J. Gao, S. Qummar, J. Zhang, R. Yao, and F. G. Khan, "Ensemble framework of deep CNNs for diabetic retinopathy detection," *Computational Intelligence and Neuroscience*, vol. 2020, Article ID 8864698, 11 pages, 2020.

[22] A. Dosovitskiy, L. Beyer, A. Kolesnikov et al., "An image is worth 16 × 16 words: transformers for image recognition at scale," 2020, http://arxiv.org/abs/2010.11929.

[23] L. Ma, C. Ma, Y. Liu, and X. Wang, "Thyroid diagnosis from SPECT images using convolutional neural network with optimization," *Computational Intelligence and Neuroscience*, vol. 2019, Article ID 6212759, 11 pages, 2019.

[24] S. Kwadwo Asare, F. You, and O. T. Nartey, "A semi-supervised learning scheme with self-paced learning for classifying breast cancer histopathological images," *Computational Intelligence and Neuroscience*, vol. 2020, Article ID 8826568, 16 pages, 2020.

[25] M. Tan and Q. Le, "Efficientnet: rethinking model scaling for convolutional neural networks," in *Proceedings of the International Conference on Machine Learning*, pp. 6105–6114, Long Beach, CA, USA, June 2019.

[26] F. Ahmad, A. Farooq, and M. U. Ghani, "Deep ensemble model for classification of novel coronavirus in chest X-ray images," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8890226, 17 pages, 2021.

[27] L.-peng Jin and J. Dong, "Ensemble deep learning for biomedical time series classification," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 6212684, 13 pages, 2016.

[28] Z. Jiang, Z. Dong, W. Jiang, and Y. Yang, "Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning," *Computers and Electronics in Agriculture*, vol. 186, Article ID 106184, 2021.

[29] J. Zhen-Cun, W. E. N. Xiao-Jing, D. Zheng-Xin et al., "Research on fire detection of improved VGG16 image recognition based on deep learning," *Fire Science and Technology*, vol. 40, no. 3, p. 375, 2021.

[30] J. Wen-Ping and J. Zhen-Cun, "Research on early fire detection of Yolo V5 based on multiple transfer learning," *Fire Science and Technology*, vol. 40, no. 1, pp. 109–112, 2021.

[31] Y. Wang, X. Ma, Z. Chen et al., "Symmetric cross entropy for robust learning with noisy labels," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 322–330, Montreal, Canada, 2019.

[32] L. Liu, H. Jiang, P. He et al., "On the variance of the adaptive learning rate and beyond," 2019, http://arxiv.org/abs/1908.03265.

[33] A. Paszke, S. Gross, F. Massa et al., "Pytorch: an imperative style, high-performance deep learning library," 2019, http://arxiv.org/abs/912.01703.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 4454507, 11 pages
https://doi.org/10.1155/2021/4454507

Research Article

Optimal Diagnosis of COVID-19 Based on Convolutional Neural Network and Red Fox Optimization Algorithm

Ehsan Khorami,[1] **Fatemeh Mahdi Babaei**,[2] **and Aidin Azadeh**[3]

[1]*Department of Computer Engineering, Ravansar (Kermanshah) Branch, Islamic Azad University, Ravansar (Kermanshah), Iran*
[2]*Department of Electrical and Computer Engineering, Mahallat Branch, Islamic Azad University, Mahallat, Iran*
[3]*Shafaghkar Dena Technical Engineering Company, Ahvaz, Iran*

Correspondence should be addressed to Ehsan Khorami; ehsan.khorami@qiau.ac.ir

Received 24 April 2021; Revised 12 May 2021; Accepted 5 August 2021; Published 20 August 2021

Academic Editor: Navid Razmjooy

SARS-CoV-2 is a specific type of Coronavirus that was firstly reported in China in December 2019 and is the causative agent of coronavirus disease 2019 (COVID-19). In March 2020, this disease spread to different parts of the world causing a global pandemic. Although this disease is still increasing exponentially day by day, early diagnosis of this disease is very important to reduce the death rate and to reduce the prevalence of this pandemic. Since there are sometimes human errors by physicians in the diagnosis of this disease, using computer-aided diagnostic systems can be helpful to get more accurate results. In this paper, chest X-ray images have been examined using a new pipeline machine vision-based system to provide more accurate results. In the proposed method, after preprocessing the input X-ray images, the region of interest has been segmented. Then, a combined gray-level cooccurrence matrix (GLCM) and Discrete Wavelet Transform (DWT) features have been extracted from the processed images. Finally, an improved version of Convolutional Neural Network (CNN) based on the Red Fox Optimization algorithm is employed for the classification of the images based on the features. The proposed method is validated by performing to three datasets and its results are compared with some state-of-the-art methods. The final results show that the suggested method has proper efficiency toward the others for the diagnosis of COVID-19.

1. Introduction

Coronaviruses are a large family of viruses that can cause disease in animals or humans. In humans, several types of viruses cause respiratory infections, from the common cold to more serious illnesses such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). A recently discovered virus from this family is SARS-CoV-2. It causes an infectious disease that has recently been discovered. The virus and its new disease were unknown before the outbreak in Wuhan, China, in December 2019.

People can get COVID-19 from other infected people with the virus. The disease can be spread from person to person with small drops in the nose or mouth when coughing or exhaling. These droplets land on objects and surfaces around a person. Other people then catch COVID-19 by touching these objects or surfaces and then touch their eyes, nose, or mouth. People can also get COVID-19 if they breathe in drops from a person with COVID-19 who coughs. This is why there has been emphasis on staying more than 1 meter (3 feet) away from someone who is sick. Medical imaging is a definitive way to diagnose coronaviruses, for example, CT scan of the lungs and chest X-ray images. Chest X-ray imaging exams are better types due to their valuable medical tool for a wide range of investigations. Chest X-ray imaging exams are a noninvasive and painless test that supports surgical and medical treatment planning to help experts in the diagnosis of the disease. The number of COVID-19 patients in the world is growing day by day. Based on the World Health Organization (WHO), infected patients of COVID-19 for April 17, 2021, are 140,888,075, and total death numbers until this date is 3,016,755 [1, 2].

One way of reducing the COVID-19 death rate and for avoiding the spread of this disease is to diagnose it in the

initial stages. To provide a correct diagnosis system with high precision, chest X-ray imaging is an interesting method. The chest X-ray imaging as an early detection system can also increase the patients' survival chances.

Although X-ray imaging is typically one of the best imaging and nondestructive testing systems, some classes of the disease are not diagnosed by this method [3]. Therefore, offering a CAD-based computer-aided system can be used for efficient disease diagnosis [1]. Based on the literature, using computer-aided methods for the diagnosis of the disease is more efficient than that of physicians. Therefore, using these systems for helping physicians is a good technique to improve their accuracy.

Shorfuzzaman and Hossain [4] proposed a method based on artificial intelligence to better diagnose COVID-19 cases from chest X-ray (CXR) images. They utilized a synergistic methodology to participate in contrastive learning by a fine-adjusted ConvNet encoder for the detention of the unbiased feature representations to better diagnose COVID-19 cases. The performance of the suggested method was then validated based on two widely used datasets including normal, COVID-19, and pneumonia infected categories. The final results showed that the proposed model provides the highest accuracy against others in diagnosing COVID-19 from chest X-ray images.

Wu et al. [5] proposed a method based on Joint Classification and Segmentation (JCS) system to provide a real-time COVID-19 diagnosis system in chest CT scan images. The JCS system was trained by constructing a large-scale COVID-19 dataset. They also used lesion counts, opacification areas, and locations and thus benefit different aspects of diagnosis. Simulation results indicated that the suggested JCS diagnosis system has very well-organized results for COVID-19 detection.

Chen et al. [6] proposed a deep learning-based method for the automatic diagnosis of COVID-19. Contrastive learning of the study was utilized for encoder training to capture expressive feature representations on widely used lung datasets and the data is classified by the prototypical network. They authenticated the performance of the presented model compared with some other state-of-the-art methods. The results demonstrated the superiority of the proposed model toward the other comparative methods.

Zhang et al. [7] suggested a method based on deep learning for designing a detection system to determine COVID-19 from the chest X-ray concerning a consistent and fast diagnosis system. To provide the effectiveness of the model, ChestX-ray14 and GitHub repository 1 datasets were utilized. The final results indicated that the suggested technique gives better accurateness toward some state-of-the-art methods to show its superiority.

From the literature, it can be observed that several works were accomplished for the automatic diagnosis of COVID-19. However, literature shows that different kinds of methods have been proposed to develop a diagnosis system with higher efficiency. Also, using metaheuristic algorithms along with deep learning methods is scarce. This motivates us to work on this subject to assess how much this combination can be useful for the diagnosis of COVID-19. In this study, a new CAD-based system is utilized for the diagnosis of COVID-19 in chest X-ray images. Here, an optimized deep network based on a newly introduced metaheuristic, called Red Fox Optimization (RFO) algorithm, is employed for this purpose.

2. Image Preprocessing

Preprocessing is the process of improving the quality of an input image. The image preprocessing of medical images is vital to remove cases like noise or light. In medical images, some disturbances have been made due to their high-frequency reception, different brightness in the field, and problems with distant orientation. This can be corrected by artificial intelligent [8]. So, after medical image acquisition, the preprocessing operations should be accomplished to remove noise and develop the areas differentiation.

2.1. Contrast Enhancement. Contrast enhancement of an image is a process to facilitate the interpretation and understanding of images. During the image processing, it may not have enough resolution or the desired quality to display the information in the images. For this reason, each image may require special adjustments in terms of brightness and scatter [9]. Image processing and artificial intelligence are also used for medical imaging purposes [10]. This study uses contrast enhancement to highlight the lung areas on the images with no changes on the other areas. Here, 16-bit Lookup Table is used to develop the contrast of the image. This technique is accomplished by Lookup Table based on the following equation [11]:

$$y_{hist} = \frac{x_{\text{hist}} - \text{Min}_{\text{hist}}}{\text{Max}_{\text{hist}} - \text{Min}_{\text{hist}}}, \tag{1}$$

where Min_{hist} and Max_{hist} stand for the lowest and the highest levels of the gray magnitudes for the main image histogram, respectively, and x_{hist} and y_{hist} describe the input image before and after contrast enhancement, respectively.

2.2. Data Normalization. Data normalization is used to change the scale in most multivariate methods and techniques. Thus, in addition to the impact of the size of the variable criterion, all variables are important in terms of weights or coefficients created by the model. One method of resizing data is to use the min-max normalization method. In this way, in addition to unifying the data scale, the edges of their change will also be in the range [0, 1]. This calculation method is often used when we want to determine the degree of similarity between points. Suppose A as data that should be mapped from the dataset between $A_{\min}$ to $A_{\max}$. For this purpose, we have the following equation [12]:

$$A_{\text{normalized}} = \frac{A - A_{\min}}{A_{\max} - A_{\min}}. \tag{2}$$

The reason for using min-max normalization here is that it guarantees all features to be exacted in the same scale.

Figure 1 shows some examples of image processing on the input images.

3. Image Segmentation

The Otsu thresholding method along with mathematical morphology is used here to determine the region of interest (ROI). The Otsu method offers a low-cost segmentation to provide a method with less complexity.

The Otsu method works based on intergroup variance optimization and decreasing pixels' intragroup variance to automatic segmentation of the input images. Global thresholding has some problems during the times when the image background resolution is not enough (heterogeneity effect). For refining this effect, the local thresholding method is used [13]. Based on the Otsu thresholding method, the threshold value is examined that minimizes the class-in-between variance. This is done based on the following equation [14]:

$$\sigma_{\omega}^{2}(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t), \tag{3}$$

where ω_i represents the probability in two separated groups with a magnitude equal to t and σ_i^2 describes the value of the variance within the groups.

In other words, the Otsu method specifies that the class variance minimization is equal to variance maximization in class-within; in other words,

$$\sigma_b^2(t) = \sigma^2 - \sigma_{\omega}^2(t) = \omega_1(t)\omega_2(t)\left[\mu_1(t) - \mu_2(t)\right]^2, \tag{4}$$

where μ_i represents the mean value. The pseudocode of the Otsu algorithm is given in Algorithm 1.

Afterward, three postprocessing techniques based on mathematical morphology are used to improve the segmentation accuracy. The three steps are using mathematical filling, closing, and opening [15]. To do so, extra holes in the images have been first filled using a mathematical filling. The mathematical model of this operator is as follows:

$$X_k = \left(X_{k-1} \oplus B\right) \cap A^c, \quad k = 1, 2, 3 \ldots, \tag{5}$$

where A and B represent the processed area and the constructing element, respectively.

The next process is to use the mathematical opening to remove the useless low-cost information with no changes on the other parts of the image. The mathematical opening is modeled as follows:

$$A \circ B = (A \ominus B) \oplus B. \tag{6}$$

Finally, mathematical closing is utilized to connect the narrow parts by the following equation:

$$A \cdot B = (A \oplus B) \ominus B. \tag{7}$$

This study uses 3×3 identity matrix as a structuring element. Figure 2 shows some examples of total image segmentation to determine the ROI in chest X-ray images.

4. Feature Extraction

Although machine vision techniques are widely used to extract features, these techniques cannot be used in some images, such as medical images. In recent years, much attention has been paid to various areas of image processing and computer vision. To extract the properties of images, especially medical images, it is necessary to extract the features of color, extract the texture, determine the homogeneity of the image, and filter to improve the output. Before performing the recovery operation, all the required features for the image database must be extracted. For reducing the complexity of the processing, the image features are extracted instead of using pixel values. In this study, Discrete Wavelet Transform (DWT) and the gray-level cooccurrence matrix (GLCM) have been used for this purpose.

4.1. Discrete Wavelet Transform (DWT). Wavelet Transform is one of the most important mathematical transformations used in various fields of science. The main idea of wavelet transform is to overcome the weaknesses and limitations of the Fourier transform. Unlike Fourier transform, this conversion can also be used for nonstatic signals and dynamic systems.

Several kinds of DWT models have been proposed for different applications. One of the widely used methods for medical image feature extraction is to use multiresolution analysis (MRA). In MRA, the DWT implementation is accomplished using a series of uninterrupted operations where each part contains signal filtering and downsampling. On each step, the signal content has been decomposed into two orthogonal subspaces containing a low-pass filter (LL) and a high-pass (HH) filter [16], which is then divided into four segments: LH, LL, HH, and HL. To provide a high-resolution frequency, this decomposition has rendered successively, so that the approximation signal has been passed over a pair of high- and low-pass filters and decomposed into two new approximate and information signals. Then, the read speed was reduced to 50%. To offer more information, the HL subbands with more efficiency have been performed. This is mathematically expressed as follows [17]:

$$P_{\mathrm{dwt}}(s) = \begin{cases} d_{i,j} = \sum f(s) \times H * i(s - 2 \times i \times j), \\ d_{i,j} = \sum f(s) \times L * i(s - 2 \times i \times j), \end{cases} \tag{8}$$

where $d_{i,j}$ describes the feature of the component in signal $f(s)$, L and H represent the low-pass filter coefficient and the high-pass filter coefficient, respectively, and i and j provide the translation and the wavelet factors scales, respectively.

4.2. Gray-Level Cooccurrence Matrix. Gray-level cooccurrence matrix (GLCM) is one of the statistical methods of tissue extraction proposed by Haralick in 1973. This method is based on the distance and angle between two pairs of pixels located in a window with specific dimensions to calculate vectors features. This matrix consists of $G \times G$ arrays, each of

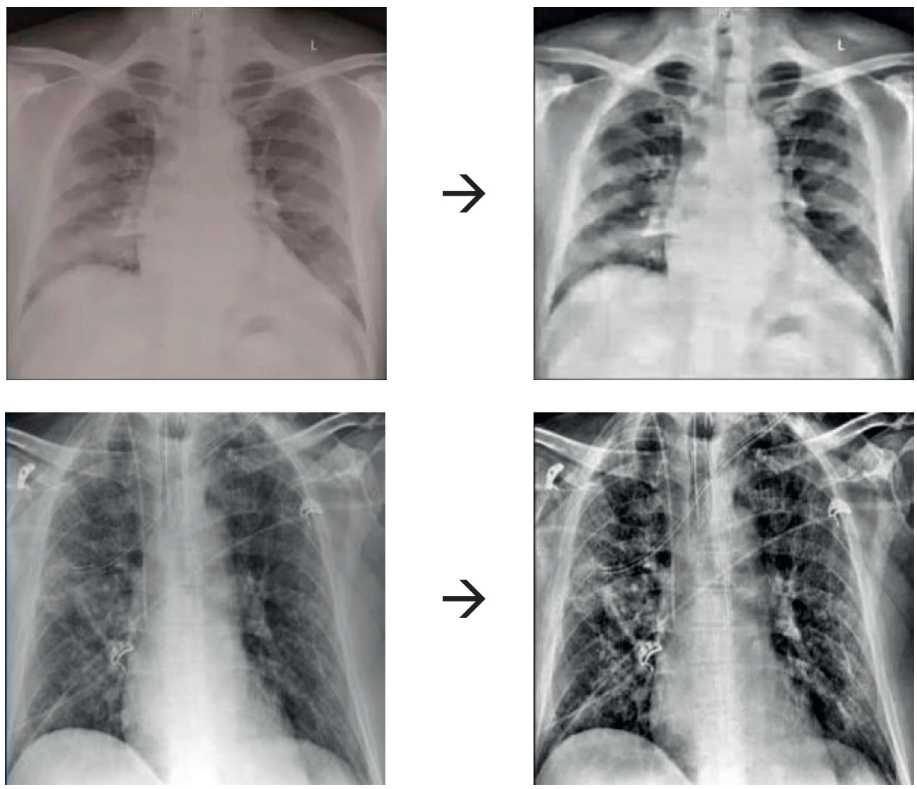


Figure 1: Some examples of image processing on the input images.

(i) Evaluate the histogram and the probability of the intensity levels
(ii) Initializing $\omega_i(0)$ and $\mu_i(0)$ for feasible threshold levels
(iii) Update ω_i and μ_i
(iv) Evaluate $\sigma_b^2(t)$
(v) Display optimal threshold as the maximum value of $\sigma_b^2(t)$

Algorithm 1: The pseudocode of the Otsu algorithm.

which is equal to Pd, where $d = (\mathrm{d}x, \mathrm{d}y)$ describes the displacement vector, i.e., [18]

$$P_d(i, j) = |[((r, s), (t, v)): I(r, s) = i, I(t, v) = j]|, \quad (9)$$

where

$$(r, s), (t, v) \in N \times N, \quad (t, v) = (r + \mathrm{d}x, s + \mathrm{d}y), \quad (10)$$

and $|.|$ describes the dimension of the set.

This matrix can be applied to the whole image or any area of it. To extract texture features, we must assign a number to each pixel of the image. So, we first put individual windows on each pixel, and in this window, we calculate the GLCM. Then, using this matrix, a large number of texture features can be extracted using statistical relationships, with each property assigning a number to each pixel. Among the most important properties from the GLCM, second angular momentum, contrast, inverse differential torque, irregularity, correlation, and autocorrelation have been considered.

This study uses a radial distance of 1 with four zero angles and several 256 gray surfaces. Among these features, homogeneity provides a local uniformity feature to generate multiple/single intervals for accusing the nontextured/textured characteristics; entropy determines the image selected interference; contrast determines the pixels intensity magnitude and their neighborhood; correlation states the spatial features dependence among the pixels; and energy defines the repetitive pixel pairs quantity. The miasmatical formulation of these methods is given below:

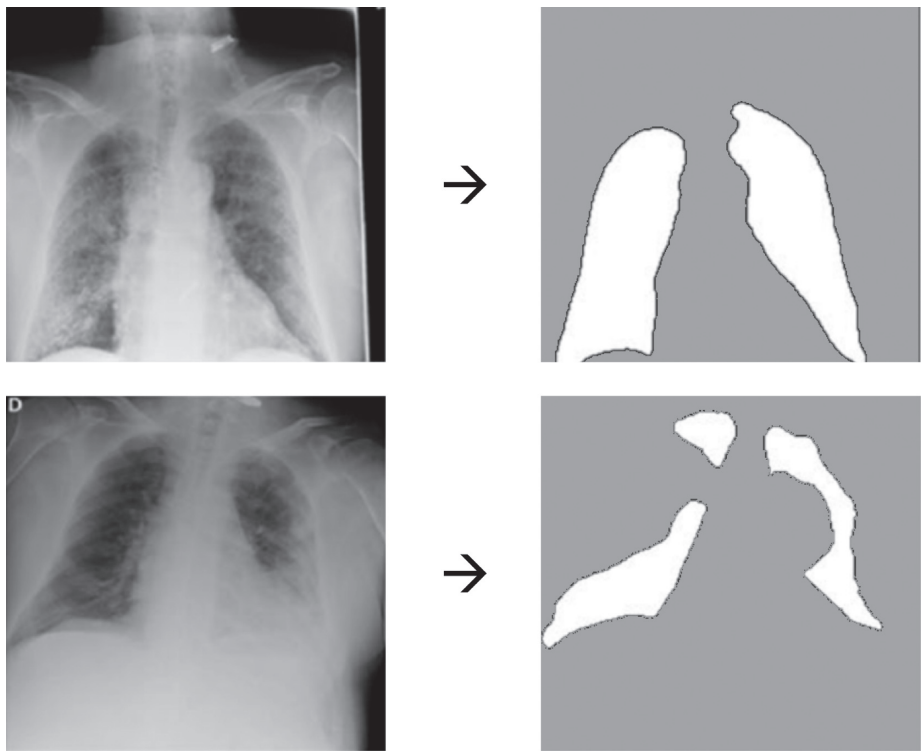


Figure 2: Some examples of image segmentation on the preprocessed images.

$$\begin{aligned}
\text{Homogeneity} &= \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\frac{1}{1+(i-j)^2}f(i,j),\\
\text{Entropy} &= -\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\log_2 f(i,j),\\
\text{Contrast} &= \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(i-j)^2 f(i,j),\\
\text{Correlation} &= \frac{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(i,j)f(i,j)-\mu_i\mu_j}{\sigma_i\sigma_j},\\
\text{Energy} &= \sqrt{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}f^2(i,j)}.
\end{aligned} \tag{11}$$

After extracting the features based on the GLCM and DWT, they have been injected into a classifier for the final diagnosis. The present study uses Convolutional Neural Network (CNN) as a second-order statistic to reflect the overall average for the degree of correlation between pairs of pixels in different aspects (in terms of homogeneity, uniformity, etc.).

5. Classification

5.1. Convolutional Neural Networks. Convolutional Neural Networks (CNNs) are deep types of neural networks that extended into three dimensions for increasing system efficiency. These types of neural networks are specially introduced for image processing applications.

Convolutional Neural Networks contain three main layers including the convolutional layer, pooling layer, and the fully connected layer. Each layer has a special duty. This is accomplished during two steps: feedforward and backward for preparation. At first, the features are injected into the network which is done by point multiplication between the input features and the variables of the neurons and the application of convolution operations in the layers.

Afterward, the network output has been evaluated. To accomplish the variables related to network training, network output results have been performed to evaluate the network error rate. To do this end, a comparison between the output of the network and the optimal solution has been performed. Then, by using the achieved error ratio, the postrelease phase has been started. The chain rule is used to achieve the gradient value of the variables and update them. The next step is feedforward. This phase continues until the network is prepared for application. The purpose of CNN in this research is to use it as a classifier of the features entered from the previous section in chest X-ray images for COVID-19 diagnosis. To give optimum weighting among network connections, the backpropagation technique is utilized and the activation mechanism is chosen rectified linear unit (ReLU). Size reduction of the output is based on max pooling and the cross-entropy loss reduction is based on the gradient descent method. The backpropagation process is modeled below:

$$L = \sum_{j=1}^{N}\sum_{i=1}^{M} -d_j^{(i)} \log z_j^{(i)}, \quad d_j = \left(0, \ldots, 0, \underbrace{1, \ldots, 1}_{k}, 0, \ldots, 0\right), \tag{12}$$

where d_j defines the desired output vector and z_j describes the obtained output vector for the m^{th} class. The Softmax function is mathematically modeled as follows:

$$z_j^{(i)} = \frac{e^{f_j}}{\sum_{i=1}^{M} e^{f_i}}, \tag{13}$$

where M represents the number of samples.

To provide a function with higher values, a weighting penalty (ρ) has been used as follows:

$$L = \sum_{j=1}^{N}\sum_{i=1}^{M} -d_j^{(i)} \log z_j^{(i)} + \frac{1}{2}\rho \sum_{K}\sum_{L} W_{k,l}^2, \tag{14}$$

where W_k describes the connections weight and L and K represent the total numbers of layers and the layer l connections, respectively.

The layouts of CNN are usually defined by trials and errors that present imprecise results. Several automated and optimized works are accomplished to address this issue [19]. Among different types of methods for solving this issue, metaheuristic algorithms are so popular. This study uses one of the newest metaheuristic techniques, called Red Fox Optimizer, to provide an effective CNN based on the previously described cases. The main advantage of CNN compared to its predecessors is that it automatically detects the important features without any human supervision.

5.2. Red Fox Optimizer (RFO). As mentioned before, for optimal diagnosis of COVID-19 based on CNN, an optimization tool is required. To obtain the optimum solution of a problem, it should be minimized or maximized (in this study, a minimization process is required) [20]. Different methods are introduced to achieve the optimal solutions [21]. The Karush–Kuhn–Tucker technique (KKT), Hamiltonian technique, and other classic techniques have an important benefit [22]. Also, these methods can achieve the exact solution for the problem, but there is a major deficiency: these methods often cannot find a solution for complex nonlinear problems. Metaheuristics are presented as new optimization techniques to solve these problems [23, 24]. These approaches can find a solution for all types of problems by obtaining a pseudooptimum solution. The metaheuristics are inspired by different phenomena such as behaviors of humans and animals contest to hunt [25]. Different kinds of these algorithms are introduced, for example, Grasshopper Optimizer (GO) [26], Cuckoo Optimizer (CO) [27], Harris Hawks Optimizer (HHO) [28], and Mayfly Optimizer [29].

Foxes are small to medium-sized, omnivorous mammals belonging to several genera of the family Canidae; due to their long, slender legs, pointed noses, thick tails, and slender limbs, foxes are distinguished from other family members or big dogs. The lifestyle of this animal is very social and thus it has flexible behavior. They can be seen all over the world with different diets. Their habitats are polar regions, deserts, and treeless plains. This animal is a type of populated fox species with great distribution. The foxes are small and fast animals that can hunt the prey while taking care about the other risks. Also, they are well-known because of their adaptability and higher intelligence. These types often chose one of two lifestyle types. In one lifestyle, they always live in a crew in a position with considering a territory [30–32]. In the second lifestyle, they live in a nomadic crew. The leader is an alpha couple in both ways of life. After enough growth of young individuals in each crew, they decide to leave or to stay with the crew.

The Red Fox Optimizer (RFO) is a new metaheuristic optimization algorithm that is inspired by the hunting lifestyle among the red foxes. When hunting, the red fox gets close to the prey gradually whereas it hides behind the bushes, and after, the prey is unexpectedly attacked. Similar to other metaheuristics, the RFO algorithm includes exploitation and exploration.

By the fox prey choice at faraway locations from the prey, the exploration term is defined, and the exploitation term is defined based on nearing the fox to the prey wherever possible to attack the prey. The RFO initialization is modeled by random individual generation as follows:

$$X = [x_0, x_1, \ldots, x_{n-1}], \tag{15}$$

where i denotes the number of population, $(X_j^i)^t$ describes the x_i in iteration t, and j is the problem dimension in the searching space.

With assuming f as condition function in R^n where n stands for the parameters in the range $[a, b]^n$,

$$(X)^i = \left[(x_0)^i, (x_1)^i, \ldots, (x_{n-1})^i\right], \tag{16}$$

where $a, b \in R$.

Thus, the optimum solution is achieved while $f((X)^i)$ suggests the global optimal. Individuals are assumed as a certain assignment to help the crew for exploration.

If the area does not include adequate prey, the individuals move to another region to get a better chance for prey. The location is shared with the others if there is a more proper region obtained. Therefore, the individuals were adapted based on the cost value. Squared Euclidean distance is applied in this regard, i.e.,

$$D\left(\left((X)^i\right)^t, \left(X_{\text{best}}\right)^t\right) = \sqrt{\left((X)^i\right)^t - \left(X_{\text{best}}\right)^t}. \tag{17}$$

Therefore, all candidates migrate by the optimum solution as follows:

$$\left((X)^i\right)^t = \left((X)^i\right)^t + \alpha \times \text{sgn}\left(\left(X_{\text{best}}\right)^t - \left((X)^i\right)^t\right), \tag{18}$$

where α is a random value.

Here, the new location of the candidates should suggest a proper solution. If not so, the previous location will have remained. When the red fox observes the prey, it gets close to the prey. It is defined as the exploitation of the Red Fox

Optimization algorithm, which is modeled by assuming a random value r in the range [0, 1]:

$$\begin{cases} \text{move closer} & \text{if } r > \frac{3}{4}, \\ \text{stay and hide} & \text{if } r \le \frac{3}{4}. \end{cases} \tag{19}$$

Afterward, an enhanced cochleoid formula is used to figure out the member's motion.

The next term is conditioned by a variable, defined as radius, which is based on two variables: a which is a random value in the range [0, 0.2] and the variable ϕ_0 which denotes a value in the range $[0, 2\pi]$ that defines the fox observation angle. Mathematically, this term can be modeled as follows:

$$r = \begin{cases} a \times \dfrac{\sin(\phi_0)}{\phi_0}, & \text{if } \phi_0 \neq 0, \\ \gamma, & \text{if } \phi_0 = 0, \end{cases} \tag{20}$$

where γ represents a random value between 0 and 1. Mathematically, the fox population nearing the prey is modeled as follows:

$$\begin{cases} x_0^{\text{New}} = a \times r \times \cos(\phi_1) + x_0^{\text{actual}} \\ x_1^{\text{New}} = a \times r \times \sin(\phi_1) + a \times r \times \cos(\phi_2) + x_1^{\text{actual}} \\ x_1^{\text{New}} = a \times r \times \sin(\phi_1) + a \times r \times \sin(\phi_2) + a \times r \times \cos(\phi_3) + x_2^{\text{actual}} \\ \qquad\qquad \vdots \\ x_{n-1}^{\text{New}} = a \times r \times \sum_{k=1}^{n-2} \sin(\phi_1) + a \times r \times \cos(\phi_{n-1}) + x_{n-2}^{\text{actual}} \\ x_{n-1}^{\text{New}} = a \times r \times \sin(\phi_1) + a \times r \times \sin(\phi_2) + \cdots + a \times r \times \sin(\phi_{n-1}) + a \times r \times \sin(\phi_{n-1}) + x_{n-a}^{\text{actual}}. \end{cases} \tag{21}$$

In the created population, 5% of the worst members are eliminated and several new members were included to the individuals to have fixed size of the population. Likewise, two optimum members were achieved as $(X(1))^t$ and $(X(2))^t$ as alpha couple in iteration t. Then, the territory center is obtained as given below:

$$H_c^t = \frac{1}{2}(X(1))^t - (\mathrm{X}(2))^t, \tag{22}$$

and the territory diameter by Euclidean distance is defined as follows:

$$H_d^t = \sqrt{(X(1))^t - (\mathrm{X}(2))^t}. \tag{23}$$

A random amount, σ, is chosen in this process that is between 0 and 1:

$$\begin{cases} \text{New nomadic candidate}, & \text{if } \sigma > 0.45, \\ \text{Reproduction of the alpha couple}, & \text{if } \sigma \le 0.45. \end{cases} \tag{24}$$

Random locations are obtained in the searching space. Then, the new members are set up by the alpha couple, i.e.,

$$(X^{\text{rep}})^t = \frac{\sigma}{2}(X(1))^t - (X(2))^t. \tag{25}$$

The utilized parameters for RFO in this study are as follows: $a = 0.2$ and $\phi_0 = 1$.

5.3. Hybrid RFO and CNN for Classification. The classification is one of the most important parts of each medical imaging. After features extraction, they should be injected into a classifier for diagnosing the disease. Based on the explanations above, CNN uses the backpropagation technique for learning. This study developed an RFO algorithm for the proper selection of network weights in CNN by minimizing the mean square error (MSE). The MSE can be mathematically defined as follows:

$$\text{MSE} = \frac{1}{T}\sum_{j=1}^{L}\sum_{i=1}^{M}\left(y_j^i - d_j^i\right)^2, \tag{26}$$

where M and L describe the output value of the layers and the data, respectively, and y_j^i and d_j^i represent the achieved and the proper magnitudes for j^{th} unit in the output layer of the network in time t, respectively.

6. Results and Discussions

The proposed method proposes an effective automated method for automatic diagnosis of the COVID-19 based on hybridizing the convolutional neural network and Red Fox Optimization algorithm. The method is a pipeline technique including preprocessing, segmentation, features extraction, and finally classification. Figure 3 shows the block diagram of the proposed system.

6.1. Dataset Description. To validate the effectiveness of the proposed method, it should be testified on a standard COVID-19 dataset. There are different types of datasets for COVID-19. In this study, consistent resources were collected by the Renmin Hospital of Wuhan University and two affiliated hospitals (the Third Affiliated Hospital and Sun Yat-Sen Memorial Hospital) of the Sun Yat-Sen University in Guangzhou with 76 and 12 patients employed [33].

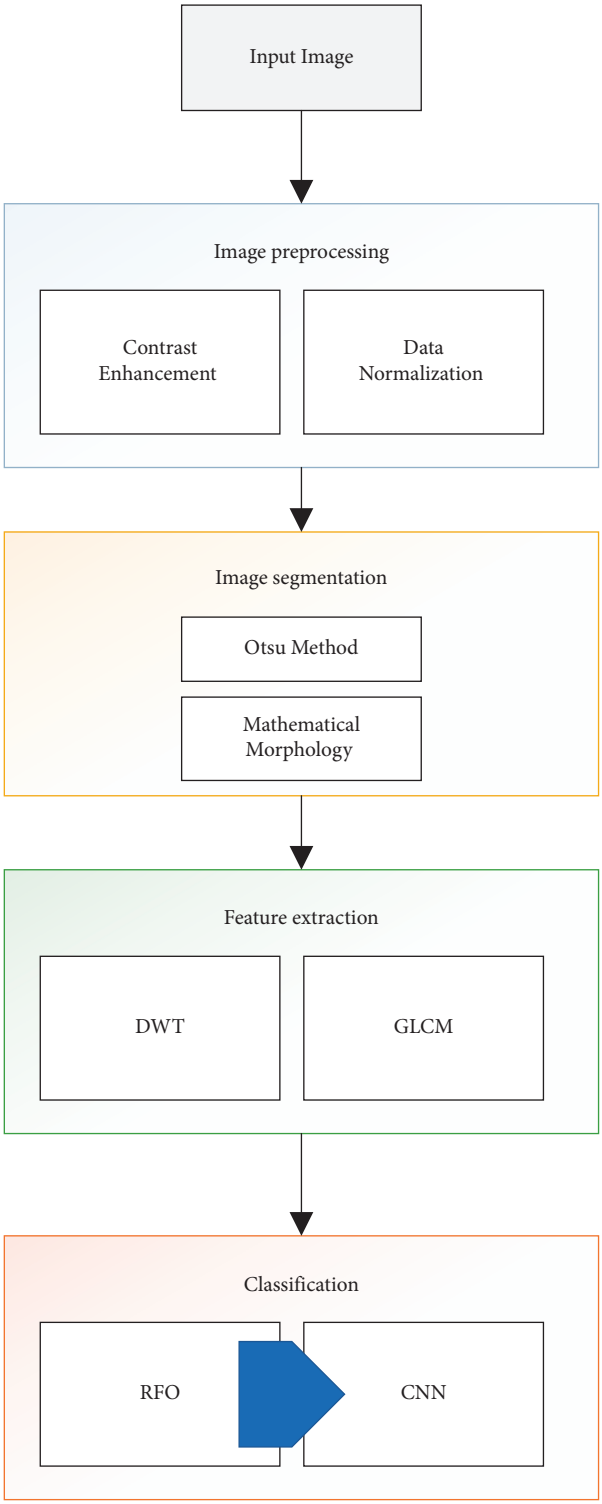


Figure 3: The block diagram of the proposed system.

Figure 4 shows some examples of the employed chest X-ray images for the analysis.

6.2. Simulations. After implementing the Discrete Wavelet Transform on the ROU in input chest X-ray images based on LL and HL decomposition characteristics, the GLCM features of these decompositions have been obtained. Afterward, the obtained features have been injected into the proposed CNN-AOA classifier for the final diagnosis. Here, five features including correlation (Cr), homogeneity (H, energy (E), entropy (ER), and contrast (CN) have been employed for applying into the LL and the HL subband levels on the image. Table 1 reports the feature extraction of the training data.

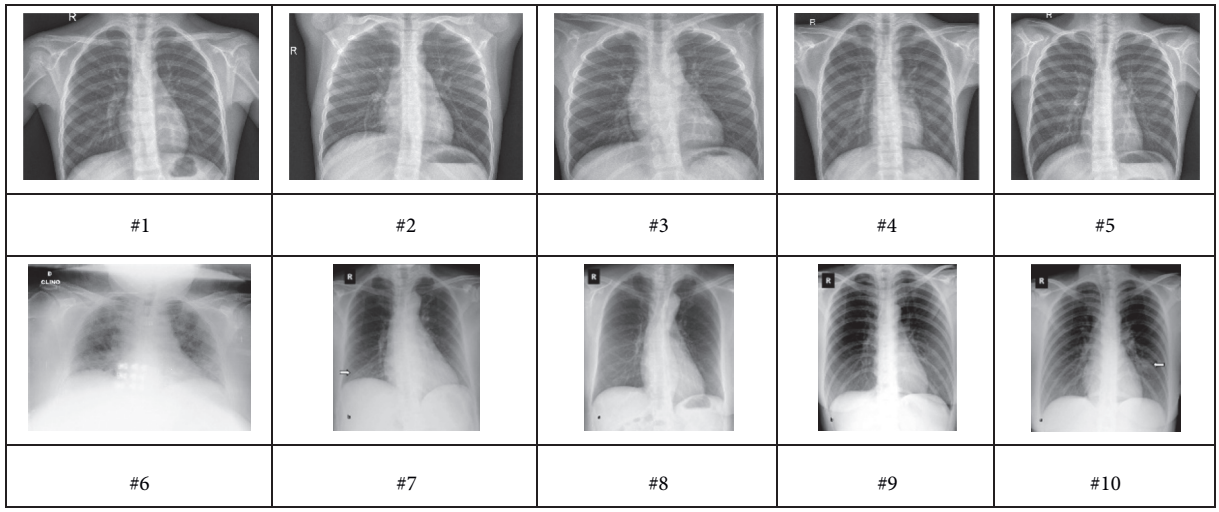


Figure 4: Some examples of COVID-19 chest X-ray images.

The feature extraction of the testing data is also reported in Table 2.

To provide proper analysis on the suggested methodology, it has been verified using three measurement indicators including sensitivity, specificity, and precision. The mathematical formulation of these three indicators is given below:

$$\begin{aligned} \text{specificity}\,(\%) &= \frac{\text{TN}}{\text{FP} + \text{TN}}, \\ \text{sensitivity}\ (\%) &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{accuracy}\,(\%) &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}, \end{aligned} \tag{27}$$

where TN and TP are true negative and true positive and FN and FP represent false negative and false positive, respectively. The suggested method has been compared with 3 state-of-the-art methods and the results are shown in Figure 5. The methods are as follows: the method of Aminu [34], Wu et al. method [5], and Singh et al. method [35].

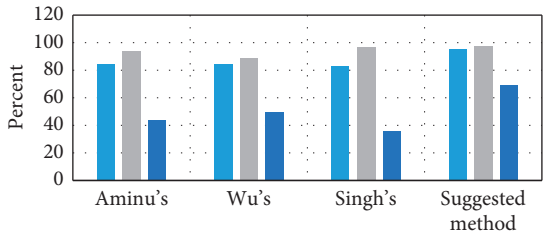
According to Figure 5, the proposed RFO-CNN classifier with 95.60% accuracy provides the highest accuracy among the comparative methods, and Aminu's, Wu's, and Singh's methods with 84.56%, 84.29%, and 83.34%, respectively, are in the next ranks. In addition to accuracy, the specificity of the suggested method with 69.14% provides the best achievements compared to the other methods. Specificity shows how well it can separate the disease from the healthy cases. Here, Wu's method with 49.38% specificity is in the next rank and Aminu's and Singh's methods with 44.08% and 36.09% are in the next ranks. At last, sensitivity is a fundamental characteristic of diagnostic imaging tests which shows how well it can be positive among all those with the condition. Here, the proposed method with 97.47% sensitivity has the best results, and Singh's, Aminu's, and Wu's methods with 96.58%, 93.76%, and 88.59%, respectively, are in the next ranks. Finally, the results showed the higher efficiency of the system for usage as a proper system for COVID-19 diagnosis.

Table 1: The feature extraction of the training data.

#	H	CR	E	CN	ER
1	0.720	0.056	0.788	0.311	0.311
2	0.840	0.174	0.973	0.052	0.275
3	0.832	0.039	0.889	0.029	0.295
4	0.807	0.038	0.905	0.027	0.407
5	0.764	0.035	0.974	0.116	0.413
6	0.836	0.064	0.897	0.010	0.228
7	0.613	0.068	0.984	0.043	0.306
8	0.788	0.006	0.941	0.028	0.386
9	0.612	0.073	0.986	0.030	0.319
10	0.715	0.037	0.974	0.053	0.427

Table 2: The feature extraction of the testing data.

#	H	CR	E	CN	ER
1	0.812	0.066	0.802	0.051	0.286
2	0.761	0.057	0.873	0.015	0.329
3	0.804	0.040	0.847	0.053	0.353
4	0.842	0.032	0.806	0.022	0.211
5	0.767	0.030	0.672	0.017	0.412
6	0.659	0.034	0.789	0.033	0.400
7	0.783	0.019	0.906	0.064	0.336
8	0.702	0.026	0.843	0.040	0.438
9	0.635	0.037	0.822	0.052	0.452
10	0.736	0.072	0.946	0.082	0.369



	Aminu's	Wu's	Singh's	Suggested method
Accuracy (%)	84.56	84.29	83.34	95.6
Sensitivity (%)	93.76	88.59	96.58	97.47
specificity (%)	44.08	49.38	36.09	69.14

Figure 5: The comparison analysis between the proposed method and the studied methods applied to the chest X-ray images dataset.

7. Conclusions

COVID-19 is a new member of coronaviruses named by the World Health Organization (WHO) on February 11, 2020, based on the year of its outbreak and its infectious agent. Recently, several types of research works were introduced on chest X-ray images to diagnose this disease with high accuracy based on image processing and computer vision. In the present research, a new pipeline methodology was proposed for the automatic diagnosis of COVID-19. At first, the input chest X-ray images were preprocessed based on contrast enhancement and data normalization to improve the input image quality and shift them into a standard interval. Then, simple image segmentation is followed by mathematical morphology. To provide the main features of the chest X-ray images, a combined gray-level cooccurrence matrix (GLCM) and Discrete Wavelet Transform (DWT) were established on the segmented images. Finally, the Red Fox Optimization algorithm was performed to a Convolutional Neural Network (CNN) for the final classification of the features. Experimental results were compared with three other state-of-the-art methods to indicate the method's efficacy. The results showed that although the proposed method is efficient for this purpose, it needs much time for the analysis (i.e., higher time complexity). Therefore, in future work, the main focus should be on determining a method for improving the system to decrease its time complexity.

Data Availability

In this study, consistent resources were collected by the Renmin Hospital of Wuhan University and two affiliated hospitals (the Third Affiliated Hospital and Sun Yat-Sen Memorial Hospital) of the Sun Yat-Sen University in Guangzhou with 76 and 12 patients employed that can be achieved by emailing to these places.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] R. Navid, F. R. Sheykhahmad, and N. Ghadimi, "A hybrid neural network—world cup optimization algorithm for melanoma detection," *Open Medicine*, vol. 13, pp. 9–16, 2018.

[2] A. Costa, Y. Kieffer, A. Scholer-Dahirel et al., "Fibroblast heterogeneity and immunosuppressive environment in human breast cancer," *Cancer Cell*, vol. 33, no. 3, pp. 463.e10–479.e10, 2018.

[3] A. Ulhaq, J. Born, A. Khan, D. P. S. Gomes, S. Chakraborty, and M. Paul, "Covid-19 control by computer vision approaches: a survey," *IEEE Access*, vol. 8, pp. 179437–179456, 2020.

[4] M. Shorfuzzaman and M. S. Hossain, "MetaCOVID: a siamese neural network framework with contrastive loss for n-shot diagnosis of COVID-19 patients," *Pattern Recognition*, vol. 113, Article ID 107700, 2021.

[5] Y.-H. Wu, S.-H. Gao, J. Mei et al., "Jcs: an explainable covid-19 diagnosis system by joint classification and segmentation," *IEEE Transactions on Image Processing*, vol. 30, pp. 3113–3126, 2021.

[6] X. Chen, L. Yao, T. Zhou, J. Dong, and Y. Zhang, "Momentum contrastive learning for few-shot COVID-19 diagnosis from chest CT images," *Pattern Recognition*, vol. 113, Article ID 107826, 2021.

[7] J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "Covid-19 screening on chest x-ray images using deep learning based anomaly detection," 2020, https://arxiv.org/abs/2003.12338.

[8] R. Navid, M. Ashourian, M. Karimifard et al., "Computer-aided diagnosis of skin cancer: a review," *Current Medical Imaging*, vol. 16, no. 7, pp. 781–793, 2020.

[9] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.

[10] A. Hu and R. Navid, "Brain tumor diagnosis based on metaheuristics and deep learning," *International Journal of Imaging Systems and Technology*, vol. 2020, no. 1, pp. 1–13, 2020.

[11] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE Transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1–8, 1997.

[12] S. Patro and K. K. Sahu, "Normalization: a preprocessing stage," 2015, https://arxiv.org/abs/1503.06462.

[13] M. Mir, M. Dayyani, T. Sutikno, M. Mohammadi Zanjireh, and N. Razmjooy, "Employing a gaussian particle swarm optimization method for tuning multi input multi output-fuzzy system as an integrated controller of a micro-grid with stability analysis," *Computational Intelligence*, vol. 36, no. 1, pp. 225–258, 2020.

[14] X. Xu, S. Xu, L. Jin, and E. Song, "Characteristic analysis of otsu threshold and its applications," *Pattern Recognition Letters*, vol. 32, no. 7, pp. 956–961, 2011.

[15] N. Razmjooy, B. S. Mousavi, B. Sadeghi, and M. Khalilpour, "Image thresholding optimization based on imperialist competitive algorithm," in *Proceedings of the 3rd Iranian Conference on Electrical and Electronics Engineering (ICEEE2011)*, Gonabad, Iran, January 2011.

[16] M. T. Hagh, H. Ebrahimian, and N. Ghadimi, "Hybrid intelligent water drop bundled wavelet neural network to solve the islanding detection by inverter-based DG," *Frontiers in Energy*, vol. 9, no. 1, pp. 75–90, 2015.

[17] P.-Y. Chen and H.-J. Lin, "A DWT based approach for image steganography," *International Journal of Applied Science and Engineering*, vol. 4, no. 3, pp. 275–290, 2006.

[18] P. Mohanaiah, P. Sathyanarayana, and L. GuruKumar, "Image texture feature extraction using GLCM approach," *International Journal of Scientific and Research Publications*, vol. 3, no. 5, pp. 1–5, 2013.

[19] R. Navid, V. V. Estrela, and H. J. Loschi, "A study on metaheuristic-based neural networks for image segmentation purposes," in *Data Science*, pp. 25–49, CRC Press, Boca Raton, FL, USA, 2019.

[20] Z. Wang, Q. Wang, Z. Zhang, and N. Razmjooy, "A new configuration of autonomous CHP system based on improved version of marine predators algorithm: a case study," *International Transactions on Electrical Energy Systems*, vol. 31, no. 1, Article ID e12806, 2021.

[21] X. Fei, R. Xuejun, and N. Razmjooy, "Optimal configuration and energy management for combined solar chimney, solid oxide electrolysis, and fuel cell: a case study in Iran," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, pp. 1–21, 2019.

[22] M. Eslami, M. A. Moghadam, H. Zayandehroodi, and N. Ghadimi, "A new formulation to reduce the number of variables and constraints to expedite SCUC in bulky power systems," *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, vol. 89, no. 2, pp. 311–321, 2018.

[23] Q. Meng, T. Liu, C. Su, H. Niu, Z. Hou, and N. Ghadimi, "A single-phase transformer-less grid-tied inverter based on switched capacitor for PV application," *Journal of Control, Automation and Electrical Systems*, vol. 31, no. 1, pp. 257–270, 2020.

[24] O. Abedinia, N. Amjady, and N. Ghadimi, "Solar energy forecasting based on hybrid neural network and improved metaheuristic algorithm," *Computational Intelligence*, vol. 34, no. 1, pp. 241–260, 2018.

[25] Z. Yuan, W. Wang, H. Wang, and N. Ghadimi, "Probabilistic decomposition-based security constrained transmission expansion planning incorporating distributed series reactor," *IET Generation, Transmission & Distribution*, vol. 14, no. 17, pp. 3478–3487, 2020.

[26] S. Arora and P. Anand, "Chaotic grasshopper optimization algorithm for global optimization," *Neural Computing and Applications*, vol. 31, no. 8, pp. 4385–4405, 2019.

[27] R. Rajabioun, "Cuckoo optimization algorithm," *Applied Soft Computing*, vol. 11, no. 8, pp. 5508–5518, 2011.

[28] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: algorithm and applications," *Future Generation Computer Systems*, vol. 97, pp. 849–872, 2019.

[29] K. Zervoudakis and S. Tsafarakis, "A mayfly optimization algorithm," *Computers & Industrial Engineering*, vol. 145, Article ID 106559, 2020.

[30] D. Yu, T. Zhang, G. He, S. Nojavan, K. Jermsittiparsert, and N. Ghadimi, "Energy management of wind-PV-storage-grid based large electricity consumer using robust optimization technique," *Journal of Energy Storage*, vol. 27, Article ID 101054, 2020.

[31] P. Akbary, M. Ghiasi, M. R. R. Pourkheranjani, H. Alipour, and N. Ghadimi, "Extracting appropriate nodal marginal prices for all types of committed reserve," *Computational Economics*, vol. 53, no. 1, pp. 1–26, 2019.

[32] H. Khodaei, M. Hajiali, A. Darvishan, M. Sepehr, and N. Ghadimi, "Fuzzy-based heat and power hub models for cost-emission operation of an industrial consumer using compromise programming," *Applied Thermal Engineering*, vol. 137, pp. 395–405, 2018.

[33] R. Kumar, R. Arora, V. Bansal et al., *Accurate Prediction of COVID-19 Using Chest X-Ray Images through Deep Feature Learning Model with Smote and Machine Learning Classifiers*, MedRxiv, 2020.

[34] M. Aminu, N. A. Ahmad, and M. H. M. Noor, "Covid-19 detection via deep neural network and occlusion sensitivity maps," *Alexandria Engineering Journal*, vol. 60, no. 5, pp. 4829–4855, 2021.

[35] K. K. Singh, M. Siddhartha, and A. Singh, "Diagnosis of coronavirus disease (COVID-19) from chest X-ray images using modified XceptionNet," *Romanian Journal of Information Science and Technology*, vol. 23, no. 657, pp. 91–115, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 3694723, 10 pages
https://doi.org/10.1155/2021/3694723

Research Article

A Method for Optimal Detection of Lung Cancer Based on Deep Learning Optimized by Marine Predators Algorithm

Xinrong Lu,[1] **Y. A. Nanehkaran**,[2] **and Maryam Karimi Fard**[3]

[1]*Gannan University of Science & Technology, Ganzhou 341000, Jiangxi, China*
[2]*Informatics School, Xiamen University, Xiamen 361005, Fujian, China*
[3]*Non-Communicable Diseases Research Center, Rafsanjan University of Medical Sciences, Rafsanjan, Iran*

Correspondence should be addressed to Maryam Karimi Fard; karimifard.m@rums.ac.ir

Received 4 July 2021; Revised 24 July 2021; Accepted 4 August 2021; Published 17 August 2021

Academic Editor: Anastasios D. Doulamis

Lung cancer is the uncontrolled growth of cells in the lung that are made up of two spongy organs located in the chest. These cells may penetrate outside the lungs in a process called metastasis and spread to tissues and organs in the body. In this paper, using image processing, deep learning, and metaheuristic, an optimal methodology is proposed for early detection of this cancer. Here, we design a new convolutional neural network for this purpose. Marine predators algorithm is also used for optimal arrangement and better network accuracy. The method finally applied to RIDER dataset, and the results are compared with some pretrained deep networks, including CNN ResNet-18, GoogLeNet, AlexNet, and VGG-19. Final results showed higher results of the proposed method toward the compared techniques. The results showed that the proposed MPA-based method with 93.4% accuracy, 98.4% sensitivity, and 97.1% specificity provides the highest efficiency with the least error (1.6) toward the other state of the art methods.

1. Introduction

Lung cancer is the uncontrolled and abnormal growth of cells that starts in one or two lungs. Abnormal cells do not grow in healthy tissues; they divide rapidly and form tumors. Primary lung cancer has its roots in the lungs, while secondary lung cancer begins elsewhere in the body, spreading from one part of the body to another and reaching the lungs [1]. The onset of lung cancer in the patient's body is indicated in most cases by early signs. The growing number of lung diseases in today's industrialized communities doubles the need for modern methods for accurate and early diagnosis.

Among lung diseases, lung cancer is still recognized as one of the most dangerous cancers. One-third of cancer deaths are due to lung cancer. About 80% of patients remain in the best condition for five years after being diagnosed with this type of cancer. Air pollution is one of the main causes of this disease. Early diagnosis of lung disease will have a major impact on the likelihood of definitive treatment of the disease. Major diagnostic methods for lung cancer include imaging, radiography and CT scan, biopsy, bronchoscopy, and examination of the breast mucosa. Pulmonary nodule is a small, round, opaque mass that forms inside the lung tissue. In other words, nodules are spherical radiographic opacities less than 30 mm in diameter. So far, various researches have been done in order to identify and describe lung diseases. Due to the removal and high number of radiographic images of the lungs, as well as its complex and uneven structure, it is difficult for a specialist physician to diagnose nodules from vessels, wounds, etc. A computer diagnostic aid system is a system that assists a physician in diagnosing a disease. Such systems are used as an intelligent tool that expresses a second option to the radiologist, which shows suspicious situations in the images to the radiologist and thus helps the radiologist to make the most accurate diagnosis. The basic idea is not to leave the diagnosis to a machine, but to use it because it increases the sensitivity of the work and reduces the rate of positive error. Several works were introduced in this subject.

Hussain et al. [2] presented a method based on multiscale sample entropy (MSE) with a mean and KD-tree algorithmic approach, multiscale fuzzy entropy (MFE),

refined composite multiscale fuzzy entropy (RCMFE), and multiscale permutation entropy (MPE) for the diagnosis of the lung cancer. The results showed that the MFE-based texture features with standard deviation provide the highest accuracy with $1.95E-50$ P value. Simulation results indicated that the developed measures using RCMFE outdone the others in dynamics analysis of the lung cancer.

Lakshmanaprabu et al. [3] proposed an automatic detection system for the classification of lung cancer in computed tomography (CT) images of lungs. They analyzed the CT scan of lung images based on linear discriminate analysis (LDA) and optimal deep neural network (ODNN). Features that were extracted from the CT lung images were then reduced by an LDR to reduce the features dimension. The system was a binary classification to show if the data is benign or malignant. The ODNN was then performed to CT images and optimized by a modified gravitational search algorithm (MGSA) to provide a method with higher accuracy.

Wang et al. [4] introduced a weakly supervised technique for efficient and fast diagnosis of the lung cancer from the images. They first used a patch-based fully convolutional network (FCN) to save discriminative blocks and offer illustrative deep features with high efficiency. Afterward, diverse feature aggregation and context-aware block selection policies were used to make globally holistic WSI descriptor. Then, it fed into a random forest classifier. Final results showed the system efficiency.

Shakeel et al. [5] proposed another technique for the diagnosis of the lung cancer. They provided a developed diagnosis system based on misclassification reduction. Weighted mean histogram equalization was used for noise elimination of the input images, and improved profuse clustering technique (IPCT) was used for images quality enhancement. Spectral features were extracted from the region of the interest, and then, deep learning was used for lung cancer prediction. Simulation results have some advantages and disadvantages that showed the proper efficiency of the proposed method. As can be observed from the literature, several works are proposed for efficient diagnosis of the lung cancer. However, using each of them has its advantages and disadvantages. In this study, we attempt to use deep learning technique and metaheuristic technique for providing a method with higher efficiency for the diagnosis of the lung cancer.

The main contributions of the present study are highlighted as follows:

(1) Presenting a new method for lung cancer diagnosis from the lung CT scan images

(2) Proposing a new structure for convolutional neural network as an efficient tool for the cancer diagnosis

(3) Optimizing the convolutional neural network based on a newly introduced metaheuristic, called marine predators algorithm

2. Methodology

Recently, convolutional neural networks (CNNs) have been turned into one of the most popular technologies in the medical imaging technology. Most of the application of deep learning in the cancer screening is based on CNNs. Convolutional neural network is a deep learning algorithm that receives the input image and assigns importance (learnable weights and biases) to each of the objects/aspects in the image and is able to distinguish them from each other. The CNN algorithm requires less "preprocessing" than other classification algorithms. While the primary method filters are manually engineered, the CNN acquires the ability to learn these filters/specifications with sufficient training [6]. The CNN architecture is similar to the connection pattern of "neurons" in the human brain and is inspired by the organization of the "visual cortex" in the brain. Each neuron responds to stimuli only in a limited area of the visual field known as the "Receptive Field." A set of these fields overlap to cover the entire visual area. In this research, the application of convolutional neural network on the lung cancer diagnosis has been investigated. To determine the cancerous or healthiness of this disease, we used deep neural networks based on MATLAB software. The graphical abstract of the proposed method is given in Figure 1.

2.1. Preprocessing Step. The first stage in the proposed method is to preprocess the input data to remove all cases that can imply bad effects on the CNN. The preprocessed images are then injected to a CNN with specific architecture and training, and test on the image data begins. The lung images dataset even with the best imaging contain some lights and noises which should be eliminated. These noises may even affect the precision of the final classifier. Pixels with high frequencies have high destructive effects on the images which can be removed with adding a low-pass filter.

2.1.1. Noise Removal. As mentioned before, in medical imaging, noise removal from the input images is a noteworthy task. Using noise removal stage should be done by leaving the images edges unchanged and keeping the image sharpness as much as possible. Median filtering is a popular low-pass filter such that each output pixel is established from the average brightness values of the neighboring pixels based on the input pixels [7]. In median filtering, the amount of a pixel has been achieved by the middle amount of the neighboring pixels. Since the center filter has low sensitivity to throw values, it can eliminate these points with no image resolution decreasing. This filter also decreases the light intensity variance, while keeping the shape of the edges and their position [8]. This filter by $m \times n$ neighborhood sorts in ascending order, with central element selection of the sorted values and replacing by the center pixels. The median filter also can remove the salt and pepper noises easily [9]. Therefore, in this study, we utilized this filter as a preprocessing stage of the input images. In median filtering, pixels are replaced by the median value of their neighbors, i.e.,

$$y_{(m,n)} = \text{median}\left(x_{i,j}:\ (i,j) \in \tau\right), \tag{1}$$

where τ defines the nearby neighbors in (m, n). Figure 2 shows a simple example of performing the noise removal on

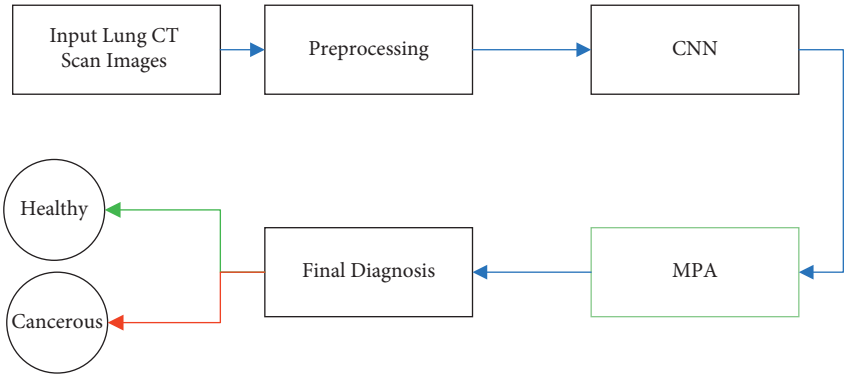
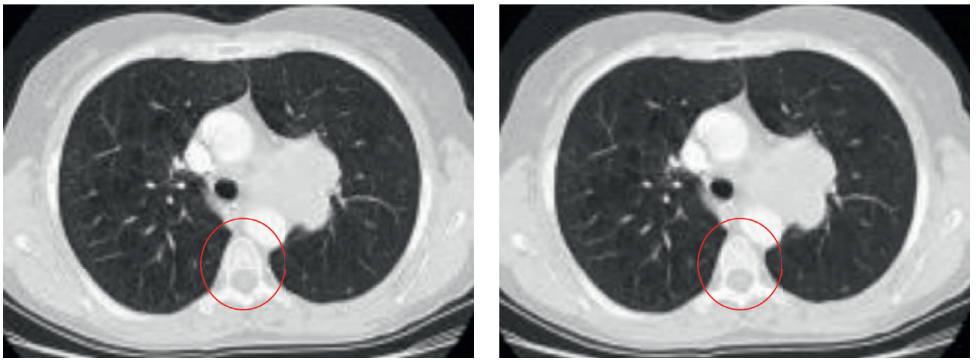


Figure 1: The graphical abstract of the proposed method.



(a) (b)

Figure 2: A simple example of performing the noise removal on the lung images based on median filtering.

the lung images based on median filtering. It should be noted that the size of filter in this example is 5×5.

As can be observed from Figure 1, the method provides a small filtering on the image, while it keeps the other details.

2.1.2. Image Level Balancing. After noise removing, normalization is required to scale the acquired images between 0 and 1 for simplifying the complexity of the dataset. Here, we used min-max normalization method. The present study employs 250×250 scale normalization for this problem. With assuming a grayscale image with n dimension and the following limitation, i.e., $I_n = [X \subseteq R^n] \longrightarrow [a, \ldots, b]$, the normalized image, I^*, is formulated as follows:

$$I^* = a_{\text{new}} + \frac{b_{\text{new}} - a_{\text{new}}}{b - a} \times (I - a), \tag{2}$$

where a and b represent the intensity values of the image, $I^* = [X \subseteq R^n] \longrightarrow [a_{\text{new}}, \ldots, b_{\text{new}}]$, and a_{new} and b_{new} signify the intensity values for the normalized image [10].

2.2. Convolutional Neural Network. In the architecture of this research, some images with the same scale are needed for processing. Therefore, all of the images have been rescaled to 227×227 before training by the CNN. In Figure 3, a simple example of this method on the considered database is shown.

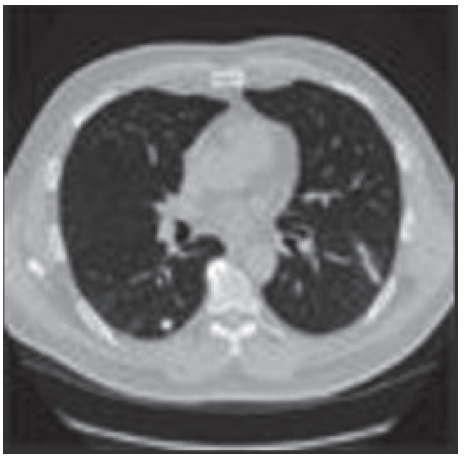


Figure 3: An example of preprocessed image.

The network training is indeed minimizing the error function between the network real outputs and the network desired outputs. This is done by improving the free parameters of the network, i.e., weights and biases [11]. The method of training in this study is supervised. In this method, a supervisor handles the training behavior and trains the network the correct way of learning. In other words, some examples of inputs and outputs are presented to the network [12]. Afterward, the output of the network is compared with the desired output to achieve the amount of

the error value. Then, the weights and biases are selected to minimize this value. The training of the network parameters can be established based on two ways: training after each training sample which is called "sequential," or the training updated after applying all of the training samples which is called "batch."

The first method needs less memory, but, it has less stability because each training sample can address the network parameters into a new direction. In the second method, although it needs more memory for storing the parameters, it will have more stability. Therefore, in the following, we used the batch mode. In this study, we use batches with scale 32 to train the database images. In the deep neural network's architecture, 3 numbers of "convolutional" layers and three numbers of "max pooling" layers have been used. Figure 4 shows the total architecture.

The convolutional layer contains the main kernel of the CNN, and output mass can be as a 3D mass of neurons. Convolution is a mathematical operation for the process of the signal that is utilized for applying the convolutional operations to the inputs by the neurons. The most important parameter in the convolutional layers is their filter size where, in the proposed model, the three layers of 32×32, 64×64, and 128×128 have been used. After the convolutional layer, to reduce the special size (depth), pooling layer has been utilized. This helps to decrease the parameters numbers and to increase the network speed. The pooling layers reduce the number of the output layers of the filter. In this study, we used 2×2 filter for this purpose. Indeed, the main idea is to subsample the input image to reduce the complexity cost, memory, and the number of parameters in the network.

The size reduction of the input image also decreases the sensitivity of the neural network. In the pooling layer, like convolutional layer, each neuron is connected to the output of some neurons. To reduce the size and increase the computation speed, pooling layer for sampling is used. In this research, the pooling layer with a 2×2 window steps on the image. From the four existing pixels in this window, the maximum pixel is selected and transferred to the next layer. The method of max pooling is shown in Figure 5.

After each convolutional layer, an activation layer is used where its target is to introduce nonlinear operations to a system that completely computed the linear operations in the computational layers [13]. This research uses RELU layers for this purpose because the network due to the computational efficiency can train faster without changing in the accuracy. The performance profile of the RELU activation function is shown in Figure 6. The RELU layer applies the function $f(x)$ to all of the input images and converts all of the negative activations into the zero. With using this layer, the nonlinear features of the model and the network have been increased with no effect on the convolutional layers.

By considering the "dropout" layer during the training, based on a definite probability, the output of some neurons has become zero where, by this operation, a different network is accessible, and this network when facing with independence to other neurons discovers and uses strong features. Therefore, the dropout technique is used to prevent the overfitting problem [14]. Because the parameters of a fully-connected deep network are a lot, the training will be slow, and during the training step, the probability for premature convergence has been increased. Therefore, among the fully-connected layers, the dropout layer has been used for the parameter's reduction. Figure 7 shows the method of performing the dropout layer.

After convolutional and pooling layers, lots of feature data with low sizes will have been produced. With connecting these layers to a Softmax classifier, due to the fact that the input images marks in training stage are determined, the network can be trained for classification of all images by entering all of the trained images and their marks. During the training process, the system is looking for the best unknown parameters, especially filter weights and the layers coefficients such that the minimum error value has been achieved for the classification. With considering the output "Flatten" layer, the convolutional layers that are a multidimensional tensor converted to a 1D tensor, and finally using the "Optimizer RMSprop" that is used for evaluating several activation functions in tensor flow, the weights optimization has been done. In the test stage, the remained images from the data that are not used for training utilized for test the network. In this condition, the output of the layers is used as the demanded image feature vector. At last, the image feature vector is compared with the feature matrix, and the deep neural network uses some layers for understanding some parts of data; however, for the classification of data, we should have a collect of probabilities for the final decision [15]. Softmax is a popular function that is used for normalizing the probability values in a standard range (0 to 1).

As aforementioned, "Optimizer RMSprop" method has been done for optimal determining of the weights. The optimization operation in the system is based on minimizing the cross-entropy [16] which can be mathematically formulated as follows:

$$L = \sum_{j=1}^{N} \sum_{i=1}^{M} -d_j^i \log z_j^i, \tag{3}$$

where N defines the samples number, $d_j = (0, \ldots, 0, \underbrace{1, \ldots, 1}_{k}, 0, \ldots, 0)$ signifies the desired output vector, and z_j represents the obtained output vector of the m^{th} class which is formulated below:

$$z_j^{(i)} = \frac{e^{f_j}}{\sum_{i=1}^{M} e^{f_i}}. \tag{4}$$

Equation (3) can be extended by considering a weight penalty as η as follows:

$$L = \sum_{j=1}^{N} \sum_{i=1}^{M} -d_j^i \log z_j^i + \frac{1}{2} \eta \sum_{K} \sum_{O} \omega_{k,l}^2, \tag{5}$$

where O signifies the total number of layers, K describes the connections in layer l, and ω_k defines the connection weight.

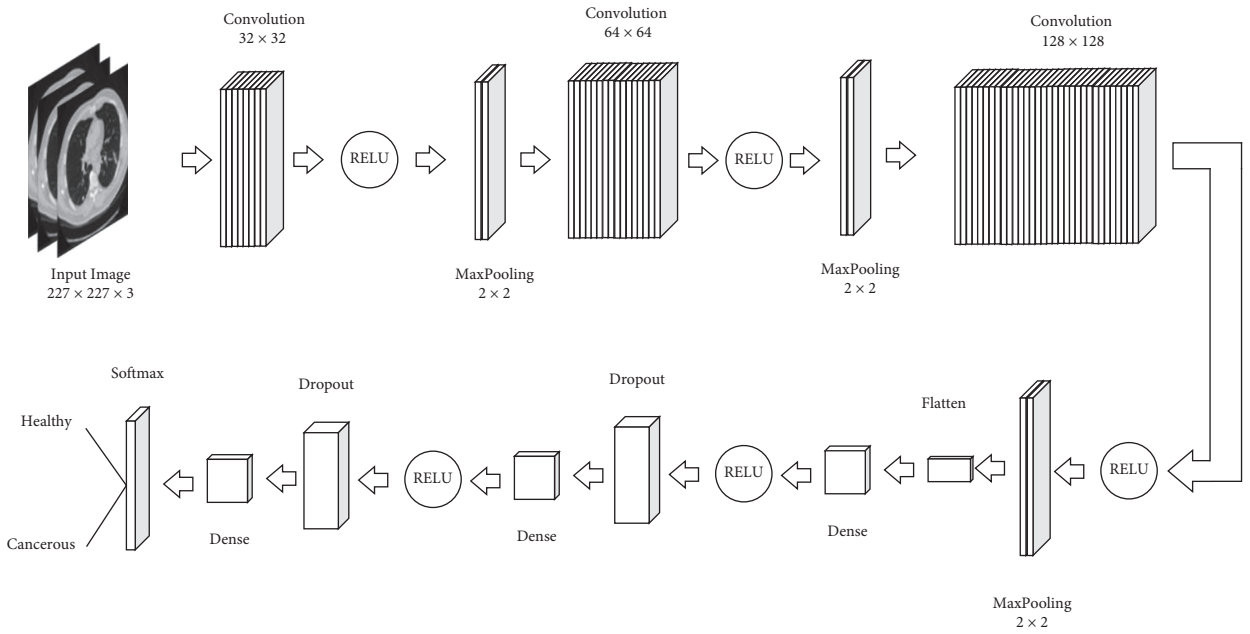


Figure 4: The proposed CNN model for the studied case.

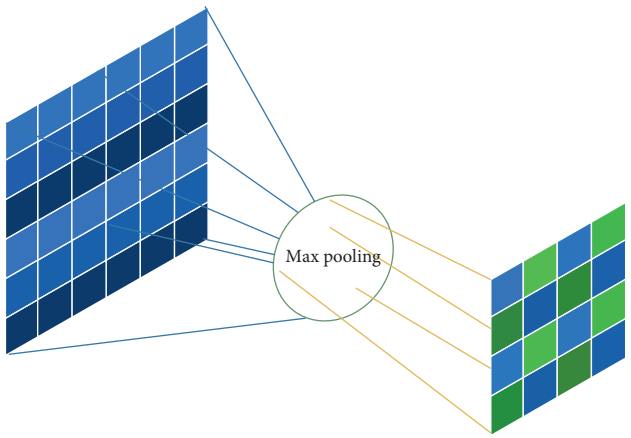


Figure 5: The method of max pooling.

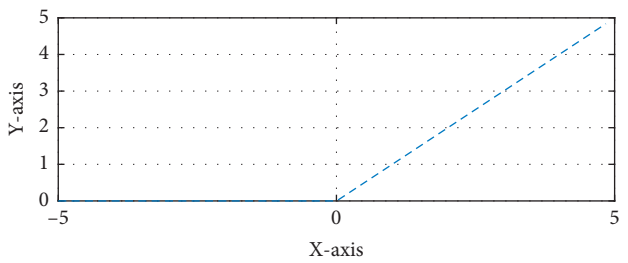


Figure 6: The performance profile of the RELU activation function.

Numerous research works are presented for optimal arrangement of the CNNs [17]. In the present research, we also work on a new optimal technique based on meta-heuristic for this purpose.

In this study, we first determine the minimum (min) and the maximum (max) limitations of the algorithm to avoid the system errors, such that min is set 2 as the minimum value that is acceptable for max pooling, and max defines the size of the

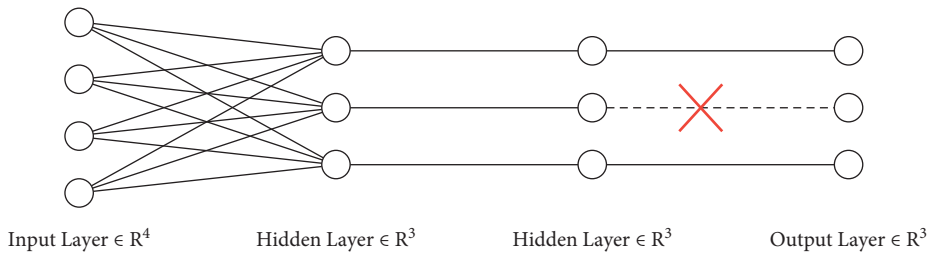


Figure 7: The method of performing the dropout layer.

sliding window. It should be noted that the input data value should be larger than the sliding window. The initial population is set 100, and the hyperparameters settings of the CNN are described 10 integer values. Then, the solutions are assessed. In this study, we used the half-value precision to the optimized CNN as the cost function. It should be noted that the overall methodology needs high computational cost because all of the population members need to be trained on the lung cancer dataset with performing backpropagation algorithm. In this study, two features are selected for optimizing as follows:

$$\begin{aligned} W &= \{w_1, w_2, \ldots, w_p\}, \\ A &= \{a_1, a_2, \ldots, a_A\}, \\ w_n &= \{w_{1n}, w_{2n}, \ldots, w_{Ln}\}, \\ b_n &= \{b_{1n}, b_{2n}, \ldots, b_{Ln}\}, \quad l = 1, 2, \ldots, Ln = 1, 2, \ldots, A, \end{aligned} \tag{6}$$

where l defines the layer index, n signifies the number of agents, w_{in} describes the value of the weight in layer i, and L and A represent the total number of layers and the total number of agents, respectively. The error cost function for the CNN is provided as follows:

$$E = \frac{1}{T} \sum_{i=1}^{T} \sum_{j=1}^{k} \left(d_{ji} - o_{ji}\right)^2, \tag{7}$$

where k describes the number of output layers, T signifies the number of training samples, and d_{ji} and o_{ji} represent the desired and the output value of the CNN. This method has been easily trapped into the local minimum that can be resolved by using the metaheuristics. In addition, metaheuristics do not need backward phase as a high computational cost process. In this study, marine predators algorithm has been used for this purpose which is explained in the following section.

2.3. Marine Predators Algorithm. In 2020, Faramarzi et al. proposed a new metaheuristic algorithm, called marine predators algorithm (MPA) for solving the optimization problems [18]. During the MPA, to model the relation between prey and marine predators, the predators utilize foraging methodology and Lévy random movement and Brownian behavior. If the attraction sense for the prey in the hunting region is high, predators employ Brownian method; if not so, if attraction sense of prey is low, Lévy method has been used. Ecological problems such as fish aggregating devices (FADs) effects or eddy formation alternates the marine predator's behavior [19]. In MPA, at first, initial n numbers of population for the prey are defined with considering a lower and an upper bound, i.e.,

$$\vec{X}_i = \vec{X}_{\min} + \overrightarrow{\text{rand}}(0, 1) \otimes \left(\vec{X}_{\max} - \vec{X}_{\min}\right), \quad i = 1, \ldots, n, \tag{8}$$

where rand defines d-dimensional random values that are uniformly distributed in the range between 0 and 1, $\vec{X}_i$ signifies the i^{th} candidate in the initial population for prey, and $\vec{X}_{\min}$ and $\vec{X}_{\max}$ describe the d-dimensional vectors of the lower and the upper bounds for the search space.

Afterward, the cost value for the candidates has been measured based on cost function. Based on the survival of the fittest mechanism, the candidate which has the best cost value has been selected as the top candidate for the solution. Lastly, the top member vector has been replicated n times to achieve the elite matrix as follows:

$$\text{Elite} = \begin{bmatrix} X^B_{1,1} & X^B_{1,2} & \cdots & X^B_{1,d} \\ X^B_{2,1} & X^B_{2,2} & & X^B_{2,d} \\ \vdots & & \ddots & \vdots \\ X^B_{n,1} & X^B_{n,2} & \cdots & X^B_{n,d} \end{bmatrix}, \tag{9}$$

where n describes the population size, d describes the problem dimension, and $\vec{X}^B$ describes the vector for the top member of the predator population. The prey candidates vector is defined as follows:

$$\text{prey} = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,d} \\ X_{2,1} & X_{2,2} & & X_{2,d} \\ \vdots & & \ddots & \vdots \\ X_{n,1} & X_{n,2} & \cdots & X_{n,d} \end{bmatrix}, \tag{10}$$

where $X_{i,j}$ describes the j^{th} dimension for the i^{th} prey.

It should be noted that elite and prey matrices contain two significant characteristics of the MPA. The foraging process in the MPA has been divided into three main portions that are described in the following. The first part is high-velocity ratio. Here, this mechanism simulates the exploration term of the algorithm and resembles the time when the prey moves quicker than the predator. The mathematical model of the exploration phase for MPA is as follows.

If Iteration < (1/3) ∗ Max_Iteration,

$$\begin{aligned}\overrightarrow{\text{Step}_i} &= \overrightarrow{R}_B \otimes \left(\overrightarrow{\text{Elite}_i} - \overrightarrow{R}_B \otimes \overrightarrow{\text{Prey}_i}\right),\\ \overrightarrow{\text{Prey}_i} &= \overrightarrow{\text{Prey}_i} + P * \overrightarrow{R} \otimes \overrightarrow{\text{Step}_i},\end{aligned} \tag{11}$$

where $\overrightarrow{R}_B$ describes d-dimensional vector including some random values based on Brownian motion, the operator $\otimes$ signifies the entry-wise multiplication, R describes a uniformly distributed value in the range [0, 1], and P is a constant value equal to 0.5.

The second part contains unit velocity ratio. In this part, both prey and predator move equally. This is the intermediate part of the optimization and contains both exploitation and exploration stages where the exploration phase regularly becomes the exploitation phase. Here, the population has been divided into two portions. Half of it has been utilized for exploration, and the other half has been utilized for exploitation. If $(1/3) \times \text{Max\_Iteration} < \text{Iteration} < (2/3) \times \text{Max\_Iteration}$, for the first half, we have

$$\begin{aligned}\overrightarrow{\text{Step}_i} &= \overrightarrow{R}_L \otimes \left(\overrightarrow{\text{Elite}_i} - \overrightarrow{R}_L \otimes \overrightarrow{\text{Prey}_i}\right),\\ \overrightarrow{\text{Prey}_i} &= \overrightarrow{\text{Prey}_i} + P * \overrightarrow{R} \otimes \overrightarrow{\text{Step}_i}.\end{aligned} \tag{12}$$

To the other population half candidates,

$$\begin{aligned}\overrightarrow{\text{Step}_i} &= \overrightarrow{R}_B \otimes \left(\overrightarrow{R}_B \otimes \overrightarrow{\text{Elite}_i} - \overrightarrow{\text{Prey}_i}\right),\\ \overrightarrow{\text{Prey}_i} &= \overrightarrow{\text{Elite}_i} + P * CF \otimes \overrightarrow{\text{Step}_i},\end{aligned} \tag{13}$$

where R_L signifies a random vector based on Levy-Flight (Lévy motion) and CF represents an adaptive parameter to control the predator motion step as follows:

$$CF = \left(1 - \frac{\text{Iteration}}{\text{Max\_Iteration}}\right)^{(2*(\text{Iteration}/\text{Max\_Iteration}))}. \tag{14}$$

The third part includes low-velocity ratio which is accomplished when the prey moves with lower speed than the predators and the exploitation stage is finished. If $\text{Iteration} > (2/3) \times \text{Max\_Iteration}$,

$$\overrightarrow{\text{Step}_i} = \overrightarrow{R}_L \otimes \left(\overrightarrow{R}_L \otimes \overrightarrow{\text{Elite}_i} - \overrightarrow{\text{Prey}_i}\right), \tag{15}$$

and in MPA, marine predators spend 80% of their time looking for prey and the remaining 20% looking for more extensively for a diverse environment with various prey distribution. This is modeled by the following equation:

$$\overrightarrow{\text{Prey}_i} = \begin{cases} \overrightarrow{\text{Prey}_i} + CF * \left[\overrightarrow{X}_{\min} + \overrightarrow{R} \otimes \left(\overrightarrow{X}_{\max} - \overrightarrow{X}_{\min}\right)\right] \otimes \overrightarrow{U} if r \le \text{FADs},\\ \overrightarrow{\text{Prey}_i} + [\text{FADs} * (1 - r) + r] * \left(\overrightarrow{\text{Prey}_{r1}} - \overrightarrow{\text{Prey}_{r2}}\right) if r > \text{FADs}, \end{cases} \tag{16}$$

where $\overrightarrow{U}$ describes a binary vector with initial value equal to zero and FADs are considered 0.2, r describes a random value between 0 and 1, and r_1 and r_2 define two members of the prey population which are selected randomly.

3. Dataset Description

Since the diagnosis capability analysis is so important before using it as a tool, the method efficiency should be evaluated by applying it to a dataset. In this study, the method is tested based on the Reference Image Database to Evaluate Therapy Response (RIDER) dataset [20]. This project was first originated in 2004 as a collaboration among the NCI's Center for Bioinformatics, the Cancer Research and Prevention Foundation (CRPF), the NCI's Cancer Imaging Program, and the National Institute of Biomedical Imaging and Bioengineering (NIBIB) with information technology support from the Radiological Society of North America (RSNA). The RIDER dataset has been downloaded from the Cancer Imaging Archive [20]. The RIDER Lung CT Collection includes images from repeated lung CT scans ("coffee break experiment") for 32 subjects detected with non-small-cell lung cancer. The collection also comprises a spreadsheet of recognized lesions where each of them experienced 2 CT studies on the same scanner and the same imaging protocol. In Figure 8, some samples of RIDER Lung CT are shown.

4. Evaluation Criteria

The accuracy measurement indicator is a criterion that assesses the capability of the model in the relation of the results with the existing information features. Figure 9 shows the matrix of cluttered elements.

Equation (17) shows the accuracy criteria formulation:

$$\text{accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{FN} + \text{TP} + \text{FP}}. \tag{17}$$

In statistics, there are two accuracy and correctness for analyzing the result of the classification method; in other words, when we can divide the data into two groups of positive and negative, the accuracy of the results of a test which divide the information into these two classes can be measurable and definable by the sensitivity analysis [21]. The sensitivity is a ratio of the positive cases that experiment marks them correctly as positive cases. The sensitivity of a test depends only on the test nature and the tested sample. Nevertheless, by only using the sensitivity, we cannot interpret a test. The equation of the sensitivity is given as follows:

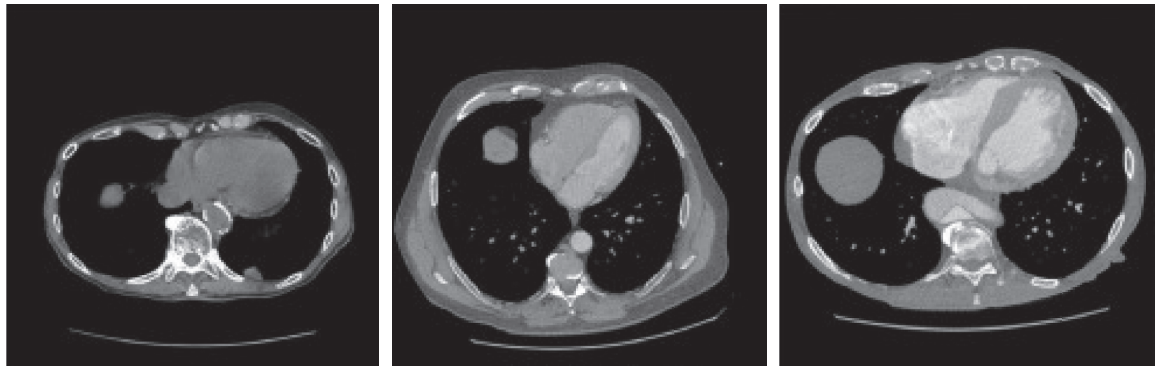


Figure 8: Some samples of RIDER Lung CT.

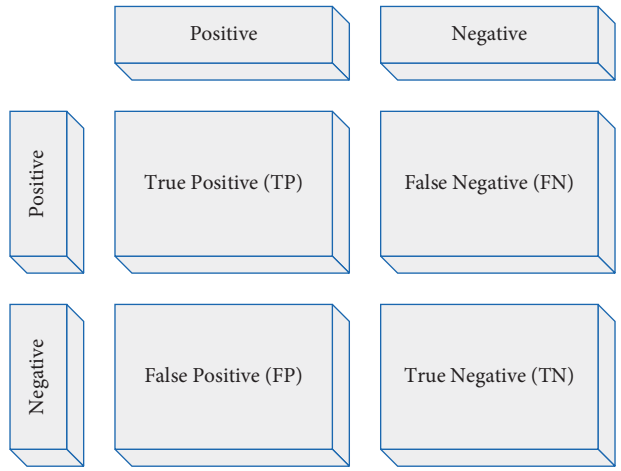


Figure 9: The matrix of cluttered elements.

$$\text{sensitivity (recall)} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{18}$$

The classification error criteria are completely opposite of its classification accuracy criteria which are obtained by the following formula:

$$\text{error} = \frac{\text{FN} + \text{FP}}{\text{TN} + \text{FN} + \text{TP} + \text{FP}}, \tag{19}$$

where its minimum value is zero, i.e., when the maximum and the best performance have been achieved.

5. Results

The deep training tools for the identification of the key features from the complicated dataset indicate the importance of them in this modern medical century. In the proposed method in this study, 80% of the images from the dataset are randomly selected for training and 20% of remaining images are used for test, while there is no interconnection between the training and test data.

The training data with batch size 32 are sent to the deep neural network, where this network is trained during 200 iterations. The proposed model by using the sequential training technique and extracting the high-level features provides better performance than the other methods. Table 1 indicates the simulation results of the proposed method and its comparison with others.

The results from Table 1 show that the proposed method has higher accuracy than the other compared methods. Figure 10 shows the convergence profile of the accuracy for the diagnosis of the cancer.

As can be observed, the proposed method is converged in iteration 166.

6. Discussions

Recently, deep learning is widely used in different fields of machine vision, such as image classification, object detection, and segmentation. The images classification can be also accurately performed by using deep neural networks on the databases. There are several works about automatic detection of cancer in machine vision and deep learning. These research attempts provide worthy results in this area. In this study, we utilized a designed convolutional neural network for the lung cancer detection. The results of the proposed method are compared with some other pretrained CNN ResNet-18, GoogLeNet, AlexNet, and VGG-19. The proposed method has been trained based on two methods: classic (Optimizer RMSprop) method and a new

Table 1: The simulation results of the proposed method and its comparison with others.

Method	Accuracy	Sensitivity	Specificity	Error
Proposed method/MPA	93.4	98.4	97.1	1.6
Optimizer RMSprop/optimizer RMSprop	88.0	92.9	88.3	12.0
ResNet-18	89.3	93.4	68.5	6.6
GoogLeNet	68.2	71.5	48.2	31.8
AlexNet	88.6	92.7	72.7	11.4
VGG-19	82.7	91.6	73.6	17.3

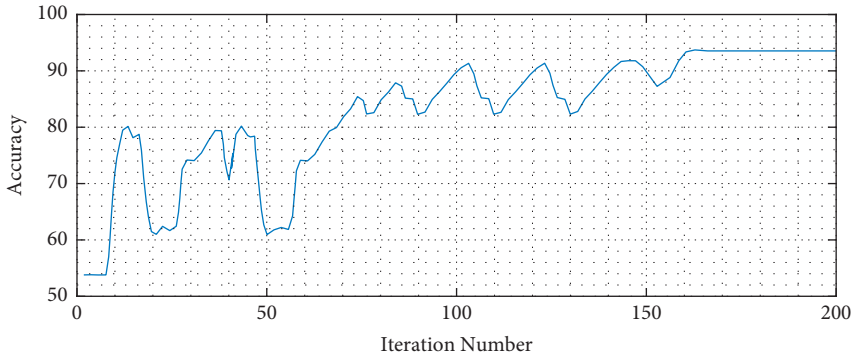


Figure 10: The convergence profile of the accuracy for the diagnosis of the cancer.

metaheuristic-based (MPA) method. The results showed that the proposed method based on MPA with 93.4% accuracy has the highest accuracy value.

7. Conclusions

Due to the reduction of the living standards such as sedentary lifestyles, poor eating habits, and increased smoking, cancer rates have risen dramatically in the last century, prompting researchers and scientists to take steps to treat this dangerous disease. Research by scientists has shown that the early detection of this disease makes it easier to treat, and the risk of death from this disease is reduced. In this article, we will discuss a new method for early detection of lung cancer. Here, an automatic system based on deep neural networks was proposed for optimal diagnosis of the CT-based lung images, in medical images. The proposed method used a metaheuristic technique, called marine predators algorithm for optimal arrangement and better network accuracy. Due to the extraction of the high-level features based on deep networks in the proposed method, the classification and diagnosis accuracy were high. Moreover, the feature vector size was reduced, and in the initial steps, its accuracy was increased that accordingly reduced the saving space and increased the speed and accuracy. In the future work, we study modifying the proposed method to provide the system with higher speed to be feasible in the real-time application.

Data Availability

The data that support the findings of this study are available at https://wiki.cancerimagingarchive.net/display/Public/RIDER+Lung+CT.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] R. Navid, A. Mohsen, K. Maryam et al., "Computer-aided diagnosis of skin cancer: a review," *Current Medical Imaging*, vol. 16, no. 7, pp. 781–793, 2020.

[2] L. Hussain, W. Aziz, A. A. Alshdadi, M. S. Ahmed Nadeem, I. R. Khan, and Q.-U.-A. Chaudhry, "Analyzing the dynamics of lung cancer imaging data using refined fuzzy entropy methods by extracting different features," *IEEE Access*, vol. 7, pp. 64704–64721, 2019.

[3] S. Lakshmanaprabu, S. N. Mohanty, K. Shankar, N. Arunkumar, and G. Ramirez, "Optimal deep learning model for classification of lung cancer on CT images," *Future Generation Computer Systems*, vol. 92, pp. 374–382, 2019.

[4] X. Wang, H. Chen, C. Gan et al., "Weakly supervised deep learning for whole slide lung cancer image analysis," *IEEE transactions on cybernetics*, vol. 50, no. 9, pp. 3950–3962, 2019.

[5] P. M. Shakeel, M. A. Burhanuddin, and M. I. Desa, "Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks," *Measurement*, vol. 145, pp. 702–712, 2019.

[6] Y. Zhou, Y. Lu, and Z. Pei, "Accurate diagnosis of early lung cancer based on the convolutional neural network model of the embedded medical system," *Microprocessors and Microsystems*, vol. 81, Article ID 103754, 2021.

[7] Z. Zhang and W.-C. Hong, "Electric load forecasting by complete ensemble empirical mode decomposition adaptive noise and support vector regression with quantum-based dragonfly algorithm," *Nonlinear Dynamics*, vol. 98, no. 2, pp. 1107–1136, 2019.

[8] D. Sharifrazi, R. Alizadehsani, M. Roshanzamir et al., "Fusion of convolution neural network, support vector machine and

Sobel filter for accurate detection of COVID-19 patients using X-ray images," *Biomedical Signal Processing and Control*, vol. 68, Article ID 102622, 2021.

[9] S. Song, H. Jia, and J. Ma, "A chaotic electromagnetic field optimization algorithm based on fuzzy entropy for multilevel thresholding color image segmentation," *Entropy*, vol. 21, no. 4, p. 398, 2019.

[10] G. Guo and N. Razmjooy, "A new interval differential equation for edge detection and determining breast cancer regions in mammography images," *Systems Science & Control Engineering*, vol. 7, no. 1, pp. 346–356, 2019.

[11] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Analysis and Applications*, vol. 24, pp. 1–14, 2021.

[12] P. M. Shakeel, M. Burhanuddin, and M. I. Desa, "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier," *Neural Computing & Applications*, pp. 1–14, 2020.

[13] R. Ranjbarzadeh, A. B. Kasgari, S. J. Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Scientific Reports*, vol. 11, no. 1, pp. 1–17, 2021.

[14] Y. Song, S. Zheng, L. Li et al., "Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021.

[15] M. B. Tahir, M. A. Khan, K. Javed et al., "Recognition of apple leaf diseases using deep learning and variances-controlled features reduction," *Microprocessors and Microsystems*, Article ID 104027, 2021.

[16] B. Van Merriënboer, D. Bahdanau, V. Dumoulin et al., "Blocks and fuel: frameworks for deep learning," 2015, https://arXiv.org/abs/1506.00619.

[17] L. Xie and A. Yuille, "Genetic CNN," in *Proceedings of the 2017 IEEE International Conference on Computer Vision*, pp. 1379–1388, Venice, Italy, October 2017.

[18] A. Faramarzi, M. Heidarinejad, S. Mirjalili, and A. H. Gandomi, "Marine predators algorithm: a nature-inspired metaheuristic," *Expert Systems with Applications*, vol. 152, Article ID 113377, 2020.

[19] Z. Liu, R. Liao, H. Ma et al., "Classification of marine microalgae using low-resolution Mueller matrix images and convolutional neural network," *Applied Optics*, vol. 59, no. 31, pp. 9698–9709, 2020.

[20] N.C.C. (NCI), "RIDER lung CT," 2014, https://wiki.cancerimagingarchive.net/display/Public/RIDER+Lung+CT.

[21] U. Jamil, A. Sajid, M. Hussain, O. Aldabbas, A. Alam, and M. U. Shafiq, "Melanoma segmentation using bio-medical image analysis for smarter mobile healthcare," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 4099–4120, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 8430565, 14 pages
https://doi.org/10.1155/2021/8430565

Research Article

Time-Frequency Analysis of EEG Signals and GLCM Features for Depth of Anesthesia Monitoring

Seyed Mortaza Mousavi,[1] **Akbar Asgharzadeh-Bonab**,[2] **and Ramin Ranjbarzadeh**[3]

[1]*Department of Biomedical Engineering, Urmia Medical Sciences University, Urmia, Iran*
[2]*Department of Electrical and Computer Engineering, Urmia University, Urmia, Iran*
[3]*Department of Telecommunications Engineering, Faculty of Engineering, University of Guilan, Rasht, Iran*

Correspondence should be addressed to Seyed Mortaza Mousavi; mousavi.m@umsu.ac.ir

Received 15 June 2021; Revised 26 July 2021; Accepted 4 August 2021; Published 12 August 2021

Academic Editor: Navid Razmjooy

One of the important tasks in the operating room is monitoring the depth of anesthesia (DoA) during surgery, and noninvasive techniques are very popular. Hence, we propose a new scheme for DoA monitoring considering the time-frequency analysis of electroencephalography (EEG) signals and GLCM features extracted from them. To this end, at first, the time-frequency map (TFM) of each channel of each EEG is computed by smoothed pseudo-Wigner–Ville distribution (SPWVD), where the EEG signal used in this paper is recorded in 15 channels. After that, we consider the gray-level co-occurrence matrix (GLCM) to obtain the content of TFM, and after that, four features such as homogeneity, correlation, energy, and contrast are obtained for each GLCM. Finally, after the selection of efficient features using the minimum redundancy maximum relevance (MRMR) method, the K-nearest neighbor (KNN) classifier is utilized to determine the DoA. Here, we consider the three states, namely, deep hypnotic, surgical anesthesia, and sedation and awake states according to bispectral index (BIS), and each EEG epoch is classified to these states. We also employ data augmentation to enhance the training phase and increase accuracy. We obtain the accuracy and confusion matrix of the proposed method. We also analyze the effects of a number of gray levels of GLCM, distance measure in KNN classifier, and parameters of data augmentation on the performance of the proposed method. Results indicate the efficiency of the proposed method to determine the DoA during surgery.

1. Introduction

General anesthesia (GA) is a necessary state for many surgical procedures. There are several essential features of anesthesia which are displayed by patients. Some of these features are the lack of movement, awareness, and recall of the surgical intervention as well as unresponsiveness to painful stimuli [1, 2]. As overly light anesthesia is the most common cause of awareness, anesthesiologists use several indicators to measure the depth of anesthesia (DoA). A continuum of progressive central nervous system (CNS) depression and decreased responsiveness to stimulation is referred to as DoA or depth of hypnosis [3, 4].

One of the most forlorn and fearsome senses is awareness during anesthesia. It is a complication with potential long-term psychological consequences like post-traumatic irritability, stress, and anxiety [5]. Monitoring DoA is a solution to this issue. The most important task in the operational room is preventing excessive DoA or awareness and improving patients' outcomes which can be performed by precise drug delivery to the patients. An accurate evaluation of DoA can help us to this end [6].

Numerous approaches and devices were designed and produced to assess DoA. These devices work based on clinical/conventional monitoring and/or brain electrical activity monitoring, and each of them has its special drawbacks. Previous studies demonstrate that since CNS is the final target for GA drugs, the electroencephalogram (EEG) signal can be more informative for the determination of DoA compared to those works just based on simple vital

signs. BIS [7], auditory evoked potential (AEP) monitors [8], entropy [9], and narcotrend [10] are some of EEG-based commercially DoA monitors. It turned out that these devices are not exactly accurate and suffer from several drawbacks.

Recently, several works aimed at introducing new methods to measure the DoA. Bayesian techniques were employed in [11] to the assessment of DoA, where the limiting large-sample normal distribution was considered, and it was shown that the maximum a posteriori (MAP) values increase gradually as the anesthesia states change from awake to light, moderate, and deep anesthesia. Distinguishing awake states from GA using EEG signals was the aim of authors in [12]. They extracted 11 features from EEG signals such as entropy, fractal, and spectral, and then, efficient ones were selected. It was found that entropies including permutation and sample, Beta-index, and detrended fluctuation analysis yield the highest accuracy with a classifier based on the adaptive neuro-fuzzy inference with linguistic hedges. Quasi-periodicities in EEG were used for analyzing the variations of DoA in [13]. To this end, phase-rectified signal averaging was employed. The results indicated that this method achieves better results than the sample and permutation entropies.

The six features including beta ratio, spectral edge frequency, and four bands of spectral energy were extracted from the EEG signal, and then, the decision tree classifier was used to determine the DoA in [14]. The authors considered the four classes for DoA as deep, moderate, and light versus awake state. In [15], at first, the noise removal from EEG signals was performed by Hurst's method, and the maximum of Hurst's ranges was considered as EEG response. Then, it was shown that maximum PSD can be used to distinct transitions of DoA states. Atomic decomposition was considered in [16] to decompose the EEG signals. Then, several features were extracted from decomposed subbands, and the SVM classifier discriminates between awake and sedated states.

The near-infrared spectroscopy (NIRS) signals were considered for recording the cerebral hemodynamic variables in [17] to monitor the DoA. The authors proposed to measure the sample entropy to describe the complexity information of cerebral hemodynamic variables. The multimodal system, which simultaneously uses the EEG and NIRS signals, was considered in [18] to monitor the DoA. It was shown that, with the EEG + NIRS signals, the clinically important transition from the awake to deep state can be detected, while transition in a clinical trial cannot be detected by BIS.

Ordinal power spectral density (O-PSD) was introduced in [19] for measuring the DoA. A deep neural network (DNN) named AnesNet was introduced in [20] to quantify the DoA. The raw EEG signals were given to a convolutional neural network (CNN) with convolution, pooling, and fully connected layers to determine the DoA. In [21], short-time Fourier transform (STFT) was employed for obtaining PSD, and CNN determines the DoA according to the given PSD. The wavelet transform was employed in [22] for analyzing the DoA from EEG signals. For this purpose, extracted features were clustered using a classifier based on the wavelet. Also, the specifications of eigenvectors were considered for extracting the specs of the midlatency auditory evoked EEG under anesthesia.

Since EEG signals present nonlinear characteristics in anesthesia conditions [23], we present a method for DoA monitoring considering the time-frequency analysis of EEG signals. We employ the smoothed pseudo-Wigner–Ville distribution (SPWVD) [24] to obtain a time-frequency map (TFM) of EEG epochs. In order to obtain the characteristics of TFM, we employ the gray-level co-occurrence matrix (GLCM) and, then, extract four features from it. Since all extracted features are not informative, we employ minimum redundancy maximum relevance (MRMR) to select the efficient ones from the feature vector. Finally, the K-nearest neighbor classifier (KNN) determines the depth of anesthesia. We also utilize the data augmentation by adding zero-mean white Gaussian noise to training samples in order to enhance the generality of the trained classifier. Hence, the contributions of this research can be summarized as follows:

(i) Employing SPWVD for obtaining the TFM of EEG signals
(ii) Employing GLCM features in order to describe the time-frequency content
(iii) Feature selection by MRMR algorithm to reduce the complexity of classification
(iv) Employing data augmentation to increase the generality of KNN classifier
(v) Obtaining the accuracy and confusion matrix of the proposed scheme
(vi) Analyzing the accuracy for different distance measures as well as different number of gray levels and augmentation parameters

Following this introduction, Section 2 describes the recorded EEG signals which are used in this research. Section 3 explains the proposed for monitoring DoA in detail. Section 4 contains the results and eventually. Section 5 presents the concluding remarks and the directions of future works.

2. Data

Six female participants aging in the range 26–72 years old, with a mean of 45.5 years old, were contributed in order to record the required data. These participants were scheduled for elective gynecological surgeries. It should be mentioned that this research was approved by the Institutional Research Ethics Committee. According to the American Association of Anesthesiology physical status classification, all patients were in ASA I and ASA II. All patients confirmed informed consent before initiation of data recording [6].

In order to avoid unnecessary delays in the surgical program, the preoperation (Pre-Op) period started about an hour before surgery, and then, spontaneous EEGs were recorded for 5 minutes. After that, the patients were transferred to the operation room and electrodes of the BIS device were attached before starting the operational (OP)

period. Spontaneous EEG signals were recorded during maintenance and emergence periods of anesthesia. A long recording was made ten minutes before the anesthesia was stopped and waking up. Table 1 contains the duration of spontaneous EEG and BIS data recording during surgery for different patients [6].

The EEG electrodes were montaged according to the 10/20 standard to include 15 channels, namely, Fp1, Fp2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P3, Pz, and P4 [6]. Simultaneously, BIS was recorded in a parallel manner using the available anesthesia monitor (Aspect Medical Systems) and was considered as a reference. BIS index is in the range [0, 100], where 0 denotes full cortical silence and 100 is the fully awake state. The appropriate state for adequate surgical anesthesia is the BIS level between 40 and 60 [2].

The EEG signals were segmented into the epochs of 30 seconds, which have 50% overlap with each other. The average of BIS during each epoch was computed, and each was labeled as deep hypnotic (D), surgical anesthesia (A), and sedation and awake (S) considering the average BIS of an epoch. The epochs with the average BIS smaller than the 40 are labeled as D. The average BIS between 40 and 60 labels the epoch as A. The epochs with average BIS greater than 60 are called S. In Table 2, the numbers of epochs from each label for different subjects are reported.

Table 1: EEG recording duration for different patients.

Patient number	1	2	3	4	5	6
Duration (minutes)	90	140	140	52	175	105

Table 2: The number of epochs from each label for different patients.

Label	Patient					
	1	2	3	4	5	6
D	65	125	101	31	242	303
A	336	304	333	158	275	169
S	36	207	227	73	165	74
Total	437	636	661	262	682	546

3. Proposed DoA Monitoring

Figure 1 presents the procedure of the proposed scheme for DoA classification. As shown, the proposed scheme generally determines the DoA in three steps including preprocessing, feature extraction and selection, and classification, which are explained in detail in the following.

3.1. Preprocessing. At first, artifacts and corrupted BIS data were identified and removed manually from raw recorded signals. The frequency content of the cleaned data is in the range [0, 300] Hz and has a maximum amplitude of about 100 μV. At first, a high-pass filter with a cutoff frequency of 0.5 Hz was employed to remove the disturbances at very low frequencies. Also, to remove the high-frequency noise, a low-pass filter with a cutoff frequency of 70 Hz was used. Furthermore, a notch filter with a null frequency of 50 Hz was employed to remove power supply noise. Since the maximum informative frequencies of EEG signals are less than 60 Hz, we consider the decimation to reduce the sampling frequency from 1000 Hz to 100 Hz, which reduces the computational complexity [21]. One epoch from states of overdeep, surgical anesthesia, and sedation and awake in different channels is shown in Figures 2–4 , respectively.

3.2. Feature Extraction. The proposed feature extraction in this paper employs the three steps as follows:

(i) Time-frequency analysis

(ii) GLCM features

(iii) Feature selection

In the following, we present each step in more detail.

3.2.1. Time-Frequency Analysis. It was mentioned that EEG signals are not stationary and should be analyzed as pseudostationary signals. Therefore, traditional frequency-domain analysis tools such as Fourier transform cannot help to characterize it in the frequency domain. Hence, we should consider time-frequency transforms in order to obtain the frequency-domain content.

The first step for computing the WVD of a signal $x(t)$ is obtaining its analytical signal as

$$y(t) = x(t) + j\widehat{x}(t), \tag{1}$$

where the Hilbert transform of $x(t)$, i.e., $\widehat{x}(t)$, is defined as

$$\widehat{x}(t) = H[x(t)] = \frac{1}{\pi}\int_{-\infty}^{\infty} x(t)\frac{1}{t-\tau}\,\mathrm{d}\tau. \tag{2}$$

Accordingly, the Winger distribution is defined as

$$W_x(t,f) = \int_{-\infty}^{\infty} y^*\left(t-\frac{\tau}{2}\right)y\left(t+\frac{\tau}{2}\right)e^{-j2\pi f\tau}\mathrm{d}\tau, \tag{3}$$

where $(\cdot)^*$ denotes the conjugate operation and $j=\sqrt{-1}$.

The Wigner–Ville distribution provides the energy density in the time-frequency domain. This distribution is one of the Cohen class distributions. Distribution smoothing should be applied to diminish the cross-term, which is the zero-density energy components of Winger distribution. Reducing the cross-term enhances the accuracy of distribution. In order to improve WVD, a smoothing function, either frequency or time smoothing function, can be added [25, 26].

Considering the frequency smoothing function $h(\tau)$, the pseudo-WVD (PWVD) is defined as

$$\mathrm{PWVD}(t,f) = \int_{-\infty}^{\infty} h(\tau)y\left(t+\frac{\tau}{2}\right)y^*\left(t-\frac{\tau}{2}\right)e^{-j2\pi f\tau}\mathrm{d}\tau. \tag{4}$$

As the smoothing function $g(t)$ is added in the time domain, the smoothed PWVD (SPWVD) is obtained as

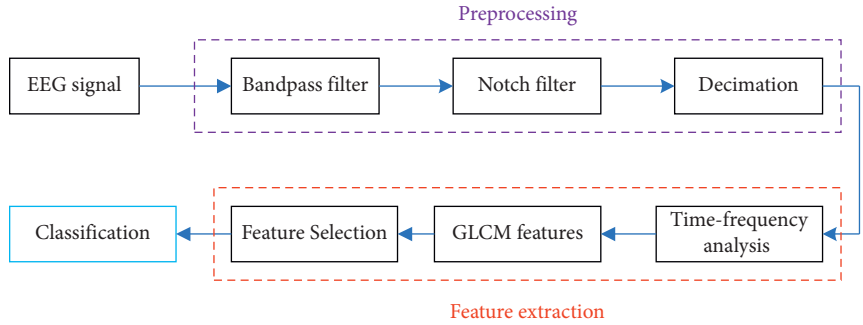


Figure 1: The procedure of the presented method in this paper for DoA classification.

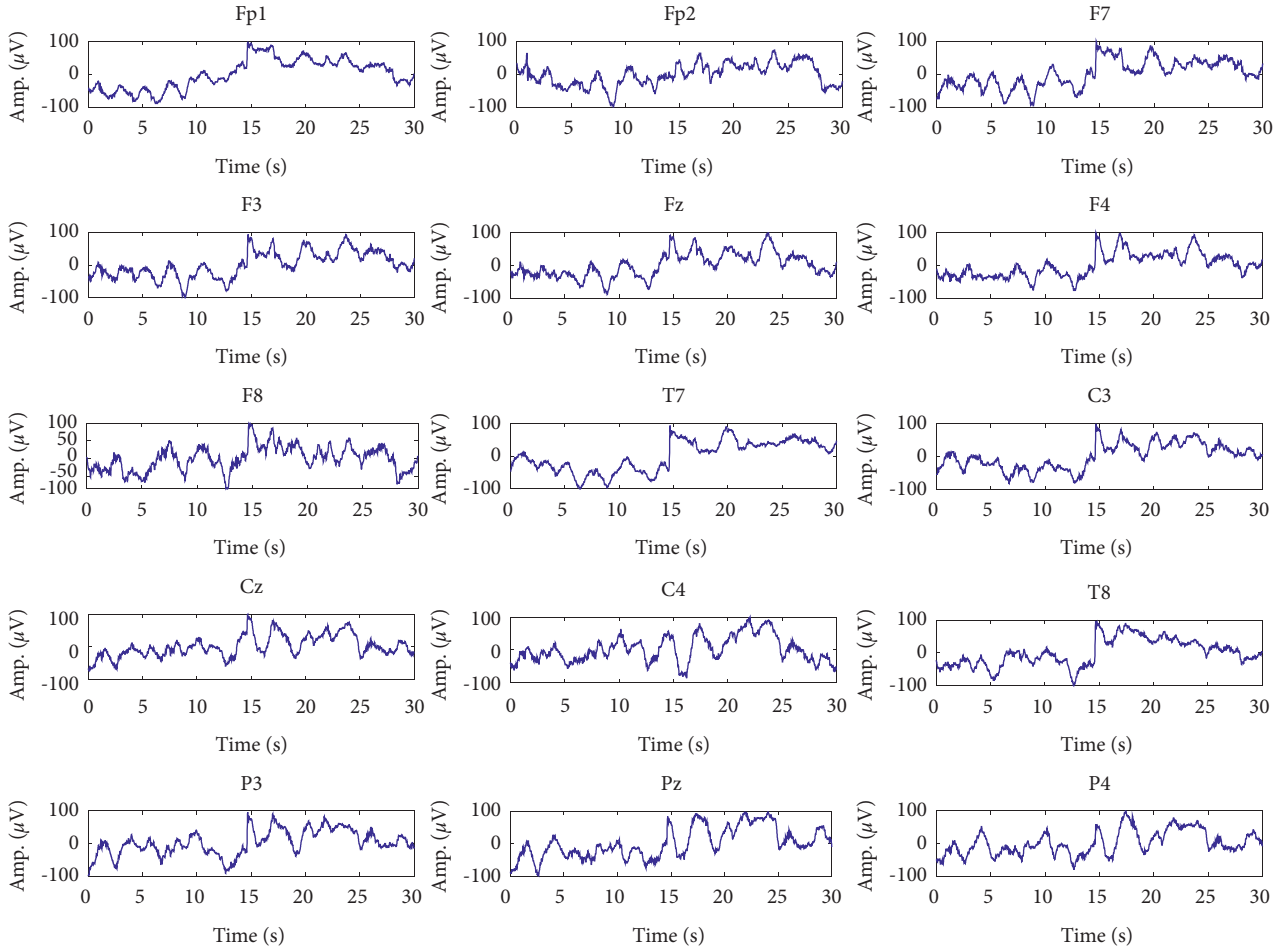


Figure 2: Signals of different channels in overdeep (D) class.

$$\mathrm{SPWVD}(t,f)=\int_{-\infty}^{\infty}h(\tau)\left(\int_{-\infty}^{\infty}g(t)y\left(t+\frac{\tau}{2}\right)y^{*}\left(t+\frac{\tau}{2}\right)\mathrm{d}t\right)e^{-j2\pi f\tau}\mathrm{d}\tau. \tag{5}$$

It should be mentioned that both $h(\tau)$ and $g(t)$ are rectangular windows.

EEG signals are recorded in 15 channels; hence, SPWVD should be calculated for each channel separately. The SPWVDs for different channels of one epoch from each class of subject 1 are shown in Figures 5–7 for deep hypnotic (D), surgical anesthesia (A), and sedation and awake (S) classes, respectively. The results show that the values of SPWVD are negligible for frequencies higher than 8 Hz. In other words, EEG signals have meaningful content in the delta and theta bands. Therefore, the SPWVD is given for the frequency range [0 8] Hz for better representation of EEG variations in

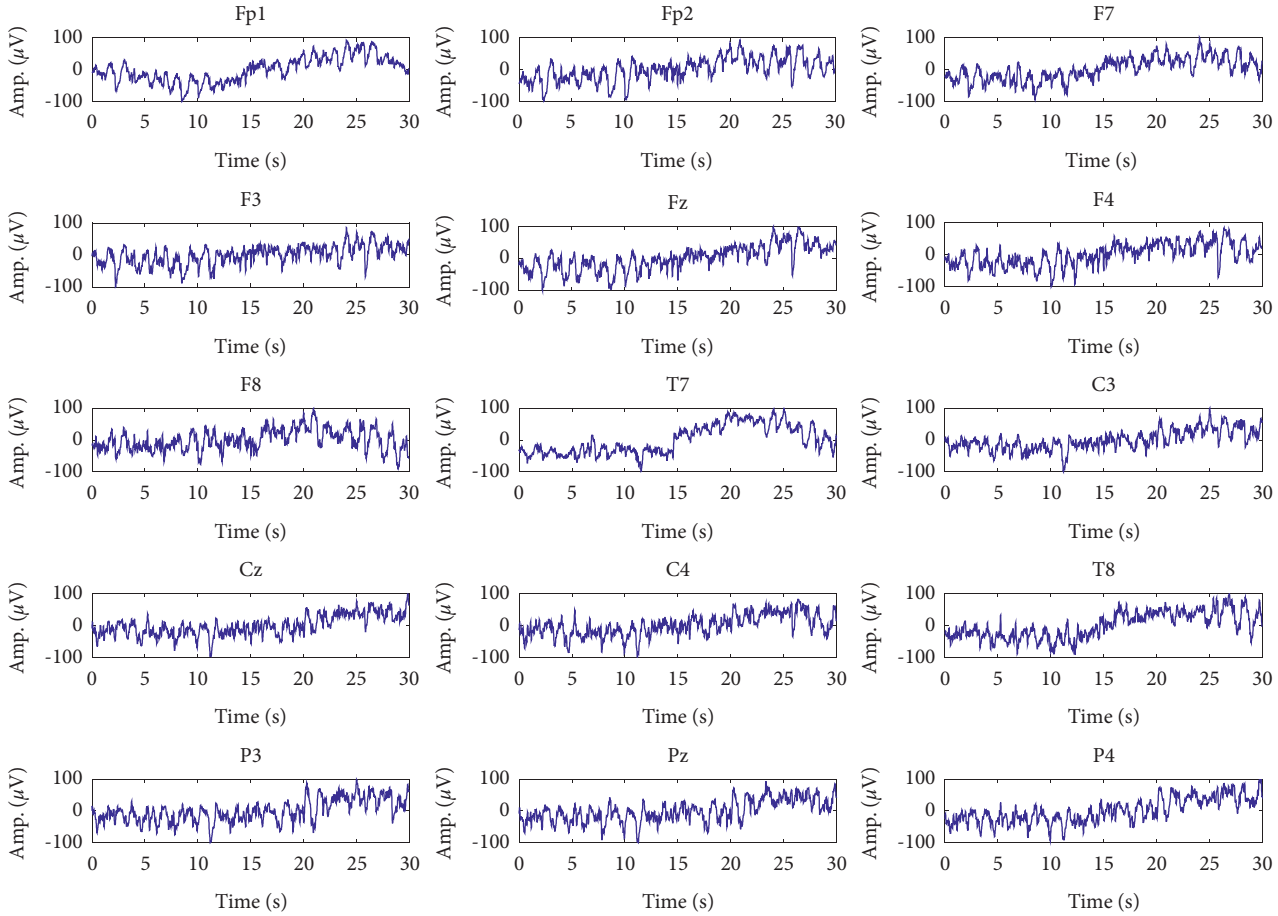


Figure 3: Signals of different channels in surgical anesthesia (A) class.

the time-frequency plane. It is observed that the deep hypnotic epoch demonstrates the minimum activity compared to the surgical anesthesia and sedation and awake epochs. Because in this case person has the lowest activity and as observed EEG rhythms with frequencies lower than two Hz are dominant. Comparing Figure 6 with Figures 5 and 7 depicts that the EEG epoch from the surgical anesthesia class has higher activity in the frequency range in the band [0 8] Hz compared to surgical anesthesia and sedation and awake states. The activity in higher frequencies reduces in sedation and awake compared to the surgical anesthesia state, but it is higher than the deep hypnotic state.

$$E_k = \sum_{i=1}^{L}\sum_{i=1}^{L} \mathbf{G}_k^2(i,j),$$
$$h_k = \sum_{i=1}^{L}\sum_{i=1}^{L} \frac{\mathbf{G}_k(i,j)}{1+|i-j|}, \tag{7}$$

3.2.2. GLCM Features. As shown, the EEG epochs have different time-frequency contents during different levels of GA. Hence, we can use texture analysis methods to describe the time-frequency content. Several texture-based methods were introduced to obtain the content of images such as local binary pattern (LBP) [27–30], autocorrelation function (ACF) [31], binary Gabor pattern (BGP) [32], gray-level co-occurrence matrix (GLCM) [28, 33], and local spiking pattern (LSP) [34].

Different combinations of gray levels within the image can be described by GLCM which can be useful in the identification of the different regions of interest in the images. GLCM extracts the texture features considering the second-order relationship between reference and neighboring pixels [35, 36]. GLCM develops the co-occurrence matrix by comparing the pixel values of neighboring pixels, where the number of rows and columns of the matrix is equal to a number of gray levels. After computing GLCM for kth channel of EEG signal, i.e., $\mathbf{G}_k$, with L levels, four features are extracted from it including contrast (c_k), correlation (r_k), energy (e_k), and homogeneity (h_k) which are computed as follows [37]:

$$c_k = \sum_{i=1}^{L}\sum_{i=1}^{L} (i-j)^2 \mathbf{G}_k(i-j),$$
$$r_k = \sum_{i=1}^{L}\sum_{i=1}^{L} \mathbf{G}_k(i-j)\left(\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{\sigma_i^2\sigma_j^2}}\right). \tag{6}$$

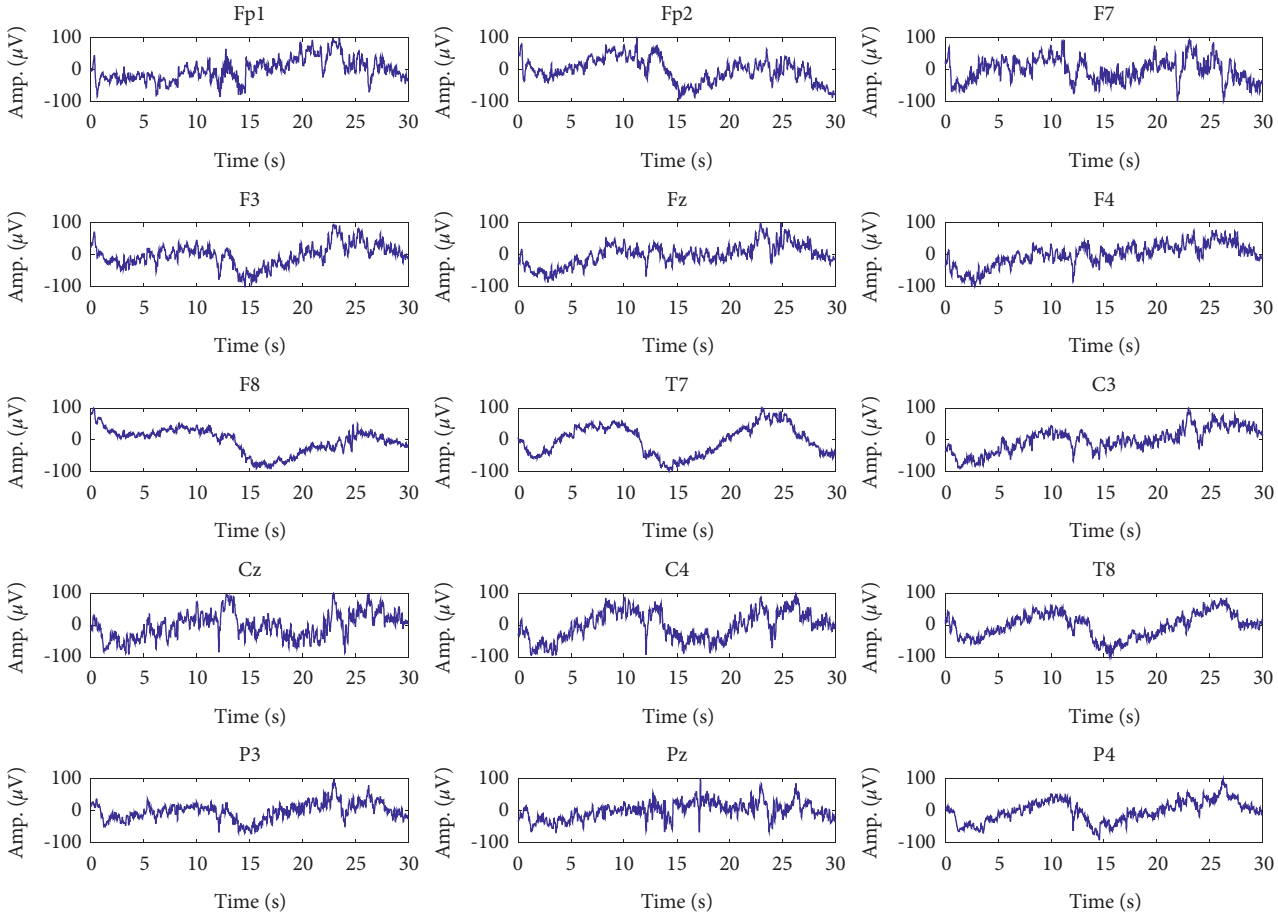


Figure 4: Signals of different channels in sedation and awake (S) class.

where

$$\mu_i = \sum_{j=1}^{L}\sum_{i=1}^{L} i\mathbf{G}_k(i,j),$$
$$\mu_j = \sum_{j=1}^{L}\sum_{i=1}^{L} j\mathbf{G}_k(i,j),$$
$$\sigma_i^2 = \sum_{j=1}^{L}\sum_{i=1}^{L} (i-\mu_i)^2\mathbf{G}_k(i,j), \quad (8)$$
$$\sigma_j^2 = \sum_{j=1}^{L}\sum_{i=1}^{L} \left(j-\mu_j\right)^2\mathbf{G}_k(i,j).$$

Let the 4×1 vector $v_k = [c_k, r_k, e_k, h_k]^T$ denote the feature vector for each channel; hence, the feature vector of each epoch considering all channels with 60 features is obtained as follows:

$$\mathbf{f}_{60\times 1} = \left[\mathbf{v}_1^T, \ldots, \mathbf{v}_{15}^T\right]^T. \quad (9)$$

3.2.3. Feature Selection. Selecting the best features among all obtained features plays a key role in the signal and image processing tasks [30, 38, 39]. In the previous part, we presented each EEG epoch with 60 features, but all computed features are not meaningful. Hence, significant features should be selected from the vector $\mathbf{f}$. In this paper, we adopt the MRMR method [40]. Its goal is to find the set of features such that selected features are mutually and maximally dissimilar which is achieved by minimizing the redundancy and maximization of the relevance of the selected features to the actual classes. MRMR considers the pairwise mutual information of features and mutual information of a feature and actual classes to quantify the redundancy and relevance.

Let r_h and e_h denote the relevance of $\mathbf{h}$ with respect to a response $\mathbf{y}$ and the redundancy of $\mathbf{h}$, respectively, which are defined as follows:

$$r_h = \frac{1}{|\mathbf{h}|}\sum_{x\in\mathbf{h}} I(x,\mathbf{y}), e_h = \frac{1}{|\mathbf{h}|^2}\sum_{x,z\in\mathbf{h}} I(x,z), \quad (10)$$

where $|\mathbf{h}|$ is the cardinality of set $\mathbf{h}$ and $I(A,B)$ is the mutual information of two sets A and B which is computed as follows:

$$I(A,B) = \sum_{i,j} P\left(A=a_i, B=b_j\right)\log\frac{P\left(A=a_i, B=b_j\right)}{P\left(A=a_i\right)P\left(B=b_j\right)}. \quad (11)$$

The selected feature set $\mathbf{h}$ should maximize the r_h and minimize the e_h. There are $2^{|\mathbf{f}|}$combinations and finding the

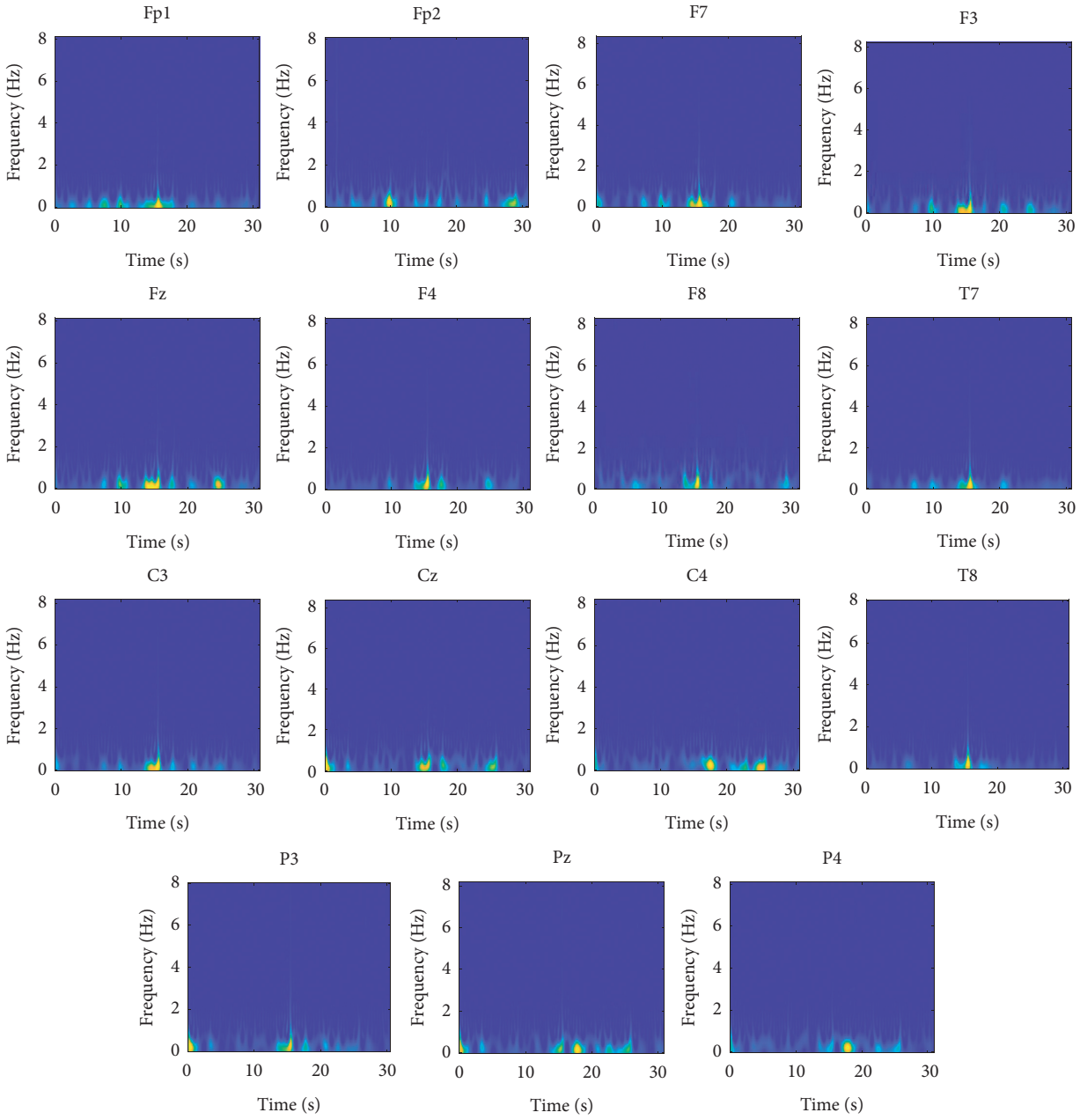


Figure 5: The SPWVD of different channels in overdeep (D) class.

optimal set $\mathbf{h}$ requires them all. But, MRMR employs the forward addition scheme by using the mutual information quotient (MIQ) value to rank the features. MIQ value is defined as

$$q(f) = \frac{r_f}{e_f}, \tag{12}$$

where r_f and e_f denote the relevance and redundancy of feature f, respectively, and defined as

$$r_f = I(f, \mathbf{y}),$$
$$e_f = \frac{1}{|\mathbf{h}|} \sum_{z \in \mathbf{h}} I(f, z). \tag{13}$$

The steps by which MRMR selects the features in the steps are as follows.

(1) Initialize the set of selected features as $\mathbf{h}$ =.

(2) Select the feature with the largest relevance, $h_1 = \text{argmax}_{|\mathbf{f}|} r_f$ and update $\mathbf{h}$ as $\mathbf{h} = \{h_1\}$.

(3) Find the features with nonzero relevance and zero redundancy in the complement of $\mathbf{h}$, i.e., $\mathbf{h}^c$.

If $\mathbf{h}^c$ does not include a feature with nonzero relevance and zero redundancy, go to step 5.
Otherwise, select the feature with the largest relevance, $h = \text{argmax}_{|\mathbf{h}^c|, r_h = 0} r_f$ and update $\mathbf{h}$ as $\mathbf{h} = \{\mathbf{h}, h\}$

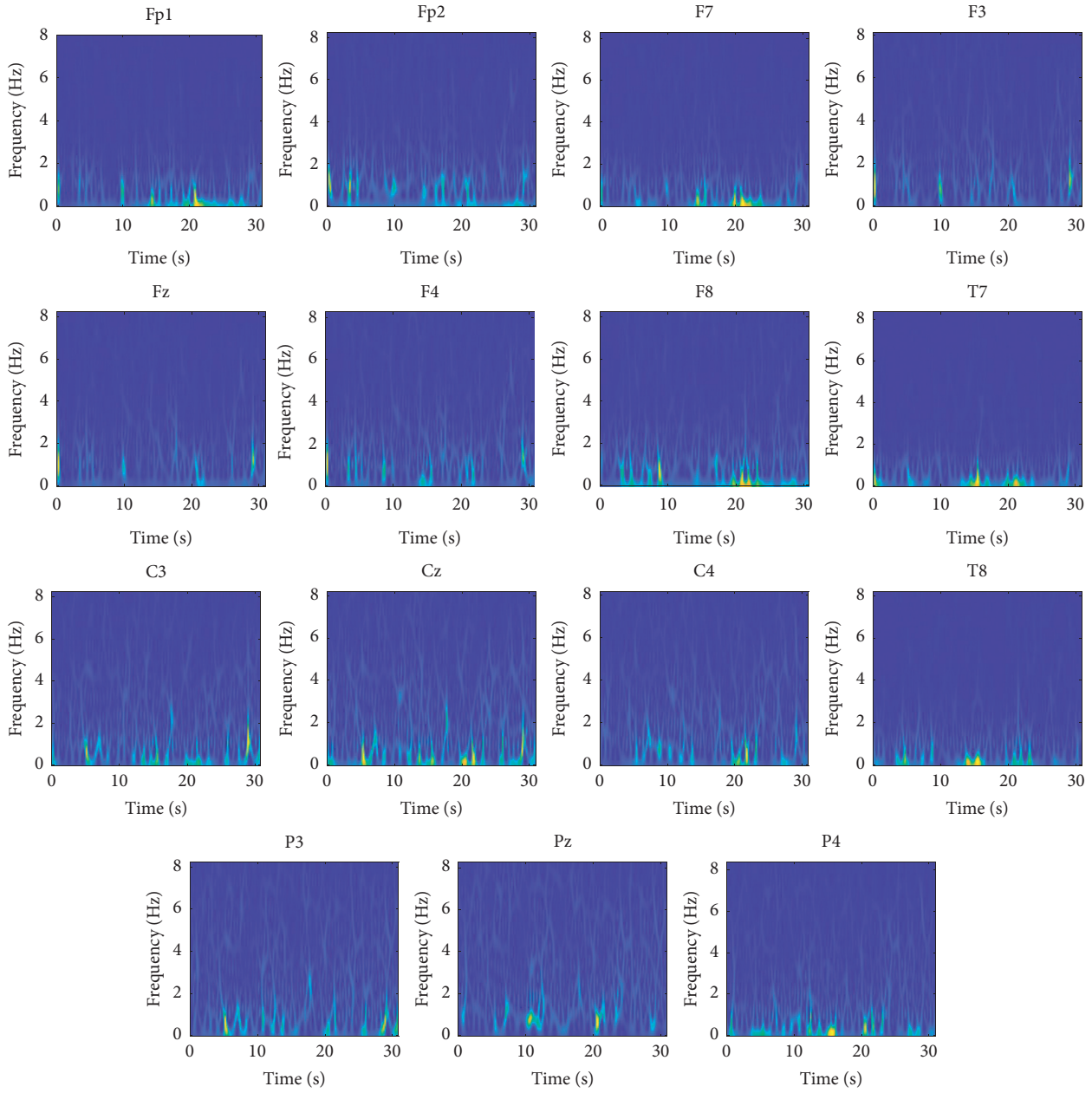


Figure 6: The SPWVD of different channels in surgical anesthesia (A) class.

(4) Step 3 is repeated until the redundancy is greater than zero for all features in $\mathbf{h}^c$.

(5) Select the feature that has the largest MIQ value with nonzero relevance and nonzero redundancy in $\mathbf{h}^c$ and update $\mathbf{h}$ as $\mathbf{h} = \{\mathbf{h}, h\}$

$$h = \arg\max_{\mathbf{h}^c} q_h = \arg\max_{\mathbf{h}^c} \frac{I(h, \mathbf{y})}{(1/|\mathbf{h}|)\sum_{z\in\mathbf{h}} I(x, z)}. \quad (14)$$

(6) Repeat step 5 until the relevance is zero for all features in $\mathbf{h}^c$.

(7) Add the features with zero relevance to $\mathbf{h}$ in random order.

3.3. Classification. The nonlinear and commonly used KNN classifier is considered in this paper to classify the features obtained from the feature selection step. KNN classifies the samples based on the distance between the unknown features and the features of training samples. It considers the labels of K as the most similar neighbors to predict the class of the training samples and considers the label of the class with the greatest number of samples among them [41–44]. In this study, we consider several distance measures to compute the similarity between the test and training samples.

Consider two feature vectors with N features as $\mathbf{v} = \{v_1, \ldots, v_N\}^T$ and $\mathbf{w} = \{w_1, \ldots, w_N\}^T$. The different distance metrics are calculated as follows:

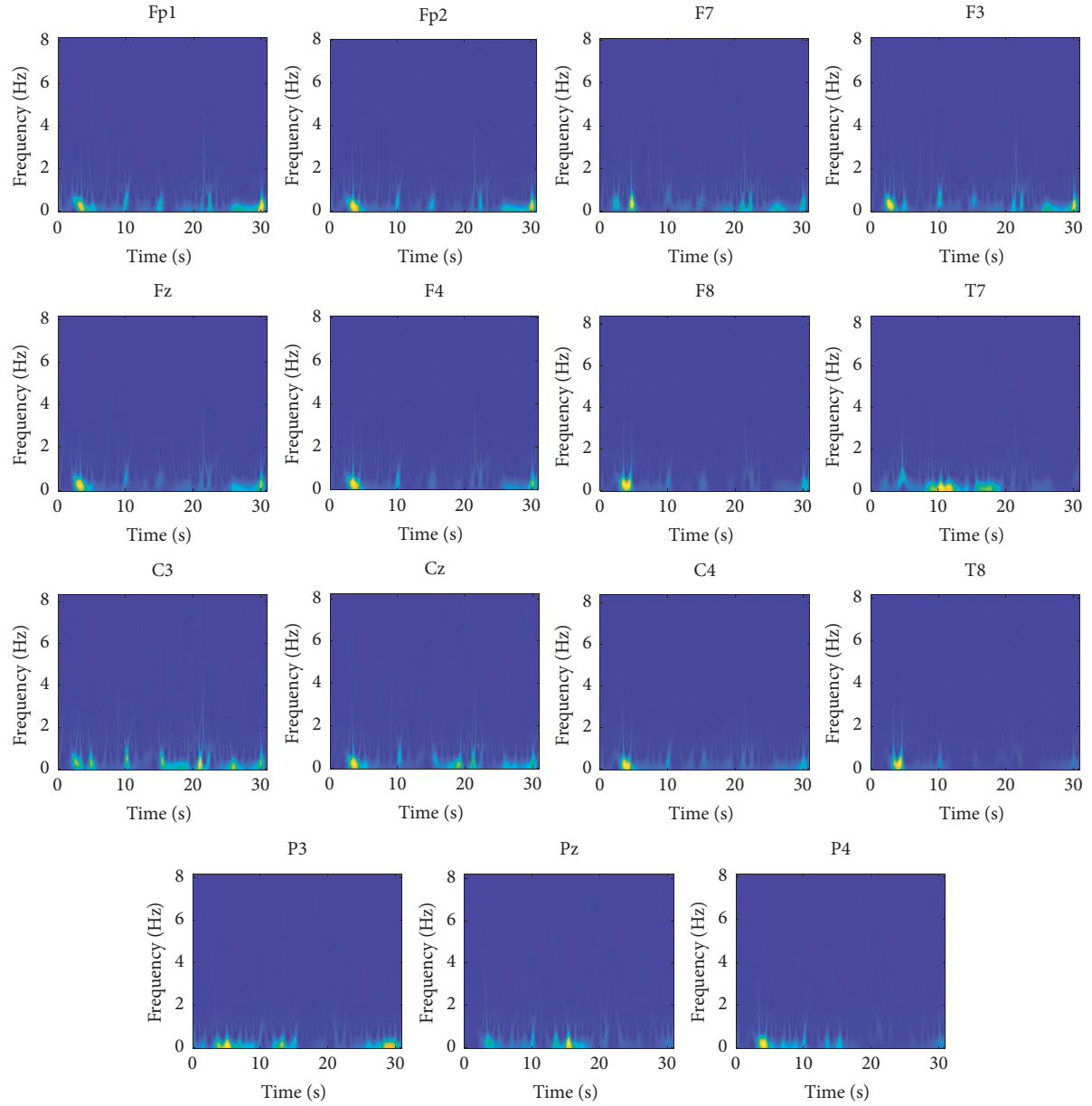


Figure 7: The SPWVDs of different channels in sedation and awake (S) class.

(i) Euclidean distance:

$$d_E = \sqrt{(\mathbf{v}-\mathbf{w})^T(\mathbf{v}-\mathbf{w})}. \tag{15}$$

(ii) Standardized Euclidean distance:

$$d_{\mathrm{SE}} = \sqrt{(\mathbf{v}-\mathbf{w})^T\mathbf{V}^{-1}(\mathbf{v}-\mathbf{w})}, \tag{16}$$

where V denotes the diagonal $N \times N$ matrix. The ith diagonal element of V is $S^2(i)$, where S is a vector of scaling factors for each dimension.

(iii) Minkowski distance:

$$d_{mi} = \sqrt[p]{\sum_{i=1}^{N} |v_i - w_i|^p}. \tag{17}$$

(iv) Chebyshev distance:

$$d_{\mathrm{ch}} = \max_i \left\{ |v_i - w_j| \right\}. \tag{18}$$

(v) City block distance:

$$d_{\mathrm{cb}} = \sum_{i=1}^{N} |v_i - w_i|. \tag{19}$$

(vi) Cosine distance:

$$d_{\mathrm{cs}} = 1 - \frac{\mathbf{v}^T\mathbf{w}}{\sqrt{\left(\mathbf{v}^T\mathbf{v}\right)\left(\left(\mathbf{w}^T\mathbf{w}\right)\right)}}. \tag{20}$$

(vii) Correlation distance:

$$d_{\mathrm{cr}} = 1 - \frac{(\mathbf{v}-\overline{v})^T(\mathbf{w}-\overline{w})}{\sqrt{(\mathbf{v}-\overline{v})^T(\mathbf{v}-\overline{v})}\sqrt{(\mathbf{w}-\overline{w})^T(\mathbf{w}-\overline{w})}}, \tag{21}$$

where $\overline{v}$ and $\overline{w}$ are the mean of vectors $\mathbf{v}$ and $\mathbf{w}$, which are computed as follows:

$$\overline{v} = \frac{1}{N}\sum_{i=1}^{N} v_i, \qquad \overline{w} = \frac{1}{N}\sum_{i=1}^{N} w_i. \tag{22}$$

(viii) Mahalanobis distance:

$$d_{\mathrm{ma}} = \sqrt{(\mathbf{v}-\mathbf{w})^T\mathbf{C}^{-1}(\mathbf{v}-\mathbf{w})}, \tag{23}$$

where $\mathbf{C}$ denotes covariance matrix.

It should be noted that, for $p = 1$, $p = 2$, and $p = \infty$, Minkowski and the city block distances are the same, Euclidean distance, and Chebyshev distances, respectively.

4. Results

Here, we provide the obtained results to demonstrate the robustness of the proposed DoA classification, where a 10-fold cross-validation method is used to obtain the accuracy. In this way, the available data is randomly partitioned into 10 nonoverlapping parts with an equal number of samples, and training and testing procedure is repeated 10 times. At each time, nine parts are considered as training samples and one part is used for testing. Also, training and test are performed for each subject separately, and the results are reported for each subject.

In order to increase the number of training samples and enhancing the generality of the trained classifier, the data augmentation scheme is presented in [45]. This method increases the number of training samples by adding zero-mean Gaussian noise with standard deviation σ to them. It should be noted that data augmentation is only performed on training samples, and test samples are those selected from original samples. The classification accuracy is evaluated for several augmented multiples and noise variances.

4.1. Classification Accuracy. The classification accuracy of the proposed DoA classification is presented in Table 3 considering the features extracted from each channel of recorded EEG signals as well as considering whole features extracted from all channels. It is observed that the selection of the channel has considerable effect accuracy. These results are obtained considering augmented multiple and noise variance equal to 35 and 0.1, respectively. It should be noted that the performance of the proposed method is presented for different values of the augmented multiple and noise variances in the following subsections. Also, the number of gray levels is set to 16, and Mahalanobis distance is considered for measuring the distance between training and test samples.

It is observed that channel T7 provides the highest average accuracy which is equal to 75.54%. Also, subjects 4 and 6 have the highest accuracy as 67.93% and 82.41%, respectively. On the other side, channel Fz yields the lowest average accuracy at 65.63%, where subjects 4 and 2 have the lowest and highest accuracy equal to 52.29% and 77.98%, respectively.

According to reported results, considering whole extracted features from all channels enhances the classification accuracy considerably. In this case, the lowest and highest accuracy results belong to subjects 3 and 5 with 93.34% and 96.92%, respectively, and an average result of 95.32% is obtained. These results indicate the efficiency of the proposed method for the classification of EEG signals in order to determine DoA.

Table 4, in the revised version, compares the performances of several TFMs and classifiers. We considered the PWVD, WVD, and STFT as well as KNN, SVM, random forest, and decision tree classifiers. It is observed that the pair (SPWVD, KNN) outperforms the other pairs of TFMs and classifiers. For all TFMs, KNN provides the best accuracy, and then SVM yields good performance. Among TFMs, SPWVD and WVD have the highest and lowest accuracy, respectively.

4.2. Confusion Matrix. Table 5 presents the confusion matrix of the proposed method for the classification of DoA for different subjects. It is observed that the proposed method provides efficient accuracy for all subjects in the presence of biased data. It is noticeable that the proposed method does not classify the deep hypnotic as sedation and awake epoch and vice versa, which indicates the efficiency of the proposed method. The misclassifications of epochs from classes deep hypnotic and sedation and awake are classified as class surgical anesthesia. Also, most misclassifications of epochs from surgical anesthesia are classified as sedation and awake. The sensitivity of different classes is also provided in Table 4 which indicates the classification accuracy of each class in different subjects.

4.3. The Effect of the Number of Gray Levels. The effect of the number of gray levels in computing GLCM on the performance of the proposed DoA classification is presented in Table 6 in which the classification accuracy is obtained considering all channels. It is observed that, for all subjects, accuracy increases considerably as the number of gray levels increases from four to 16 while increasing the number of gray levels from 16 to 64 reduces the accuracy. Hence, 16 gray levels are considered in order to extract texture-based features from time-frequency images obtained from SPWVD.

Table 3: Classification accuracy of the proposed DoA classification.

Channel	Subject						Average	Min	Max
	1	2	3	4	5	6			
1 (Fp1)	77.80	65.72	68.68	55.72	74.48	83.51	70.98	55.72	83.51
2 (Fp2)	66.36	73.58	64.44	70.99	67.44	75.64	69.74	64.44	75.64
3 (F7)	65.21	74.37	64.75	61.83	77.41	74.35	69.65	61.83	77.41
4 (F3)	67.27	79.08	63.08	57.25	66.42	69.41	67.08	57.25	79.08
5 (Fz)	64.53	77.98	62.02	52.29	70.67	66.30	65.63	52.29	77.98
6 (F4)	69.56	63.36	73.97	50.76	74.34	82.41	69.07	50.76	82.41
7 (F8)	79.40	66.03	63.84	60.67	59.97	73.62	67.25	59.97	79.40
8 (T7)	73.22	79.24	69.74	67.93	79.03	84.06	75.54	67.93	84.06
9 (C3)	64.98	60.84	76.85	70.61	68.47	84.79	71.09	60.84	84.79
10 (Cz)	69.10	79.24	77.91	67.93	57.33	78.20	71.62	57.33	79.24
11 (C4)	67.73	71.06	68.38	53.81	76.39	81.86	69.87	53.81	81.86
12 (T8)	67.96	72.32	56.88	67.93	62.17	69.04	66.05	56.88	69.04
13 (P3)	73.68	64.62	72.31	54.19	62.31	78.75	67.64	54.19	78.75
14 (Pz)	69.56	63.99	66.26	68.32	73.61	74.35	69.35	63.99	74.35
15 (P3)	77.34	76.25	75.34	51.52	67.88	81.68	71.67	51.25	81.68
All	94.5	94.65	93.34	95.8	96.92	96.7	95.32	93.34	96.92

Table 4: Performance comparison between different TFMs and classifiers.

TFM	Classifier			
	KNN	SVM	Random forest	Decision tree
SPWVD	95.32	94.54	93.49	93.86
PWVD	94.01	93.46	92.34	92.80
WVD	93.24	92.52	91.35	92.00
STFT	93.70	93.11	91.59	92.18

Table 5: Confusion matrix of the proposed DoA classification.

		Predicted class			Sensitivity (%)
		D	A	S	
Actual class for subject 1	D	60	5	0	92.31
	A	4	320	12	95.24
	S	0	3	33	91.67
Actual class for subject 2	D	119	6	0	95.2
	A	7	285	12	93.75
	S	0	9	198	95.65
Actual class for subject 3	D	94	7	0	93.07
	A	12	311	10	93.39
	S	0	15	212	93.39
Actual class for subject 4	D	30	1	0	96.77
	A	2	151	5	95.57
	S	0	3	70	95.89
Actual class for subject 5	D	235	7	0	97.11
	A	4	264	7	96
	S	0	3	162	98.18
Actual class for subject 6	D	293	10	0	96.69
	A	2	163	4	96.45
	S	0	2	72	97.29

4.4. Accuracy of Distance Measures. The accuracy of the proposed method considering different distance measures considered in the KNN classifier is given in Table 7. As mentioned earlier, the Mahalanobis distance provides the highest accuracy of 95.32%, and after that, the Chebyshev distance reaches the accuracy of 93.26%. Also, the standardized Euclidean distance has the lowest accuracy of 90.73%.

4.5. The Effect of Augmentation. The effect of the augmentation multiple and noise variances in data augmentation on

Table 6: The effect of the number of gray levels of GLCM.

Number of gray levels	Subject						Average	Min	Max
	1	2	3	4	5	6			
4	87.87	86.16	87.74	85.49	90.89	89.56	87.95	85.49	90.89
8	90.85	92.45	91.22	91.22	93.11	91.75	91.77	90.85	92.45
16	94.5	94.65	93.34	95.8	96.92	96.7	95.32	93.34	96.92
32	93.13	94.03	92.58	94.27	95.61	95.78	94.23	92.58	95.78
64	92.22	93.55	92.13	92.36	93.41	94.87	93.09	92.13	94.87

Table 7: Accuracy of different distance measures.

Distance measure	Subject						Average	Min	Max
	1	2	3	4	5	6			
Euclidean	93.59	91.98	93.34	89.69	94.87	92.67	92.69	89.69	94.87
Standardized Euclidean	89.02	91.04	91.38	92.37	89.69	90.84	90.73	89.02	92.37
Mahalanobis	94.50	94.65	93.34	95.80	96.92	96.70	95.32	93.34	96.92
City block	88.79	93.55	91.98	92.37	95.31	93.96	92.66	88.79	95.31
Minkowski	90.85	92.77	93.65	90.84	90.91	93.41	92.07	90.85	93.65
Chebyshev	94.28	92.92	90.47	90.46	94.72	96.70	93.26	90.46	96.7
Cosine	90.62	93.40	89.11	93.51	90.32	93.04	91.67	89.11	93.51
Correlation	92.45	93.87	89.86	91.60	96.92	93.96	93.11	89.86	96.92
Average	91.76	93.02	91.64	92.08	93.71	93.91			
Min	89.02	91.04	89.11	89.69	89.69	90.84			
Max	94.50	94.65	93.34	95.80	96.92	96.7			

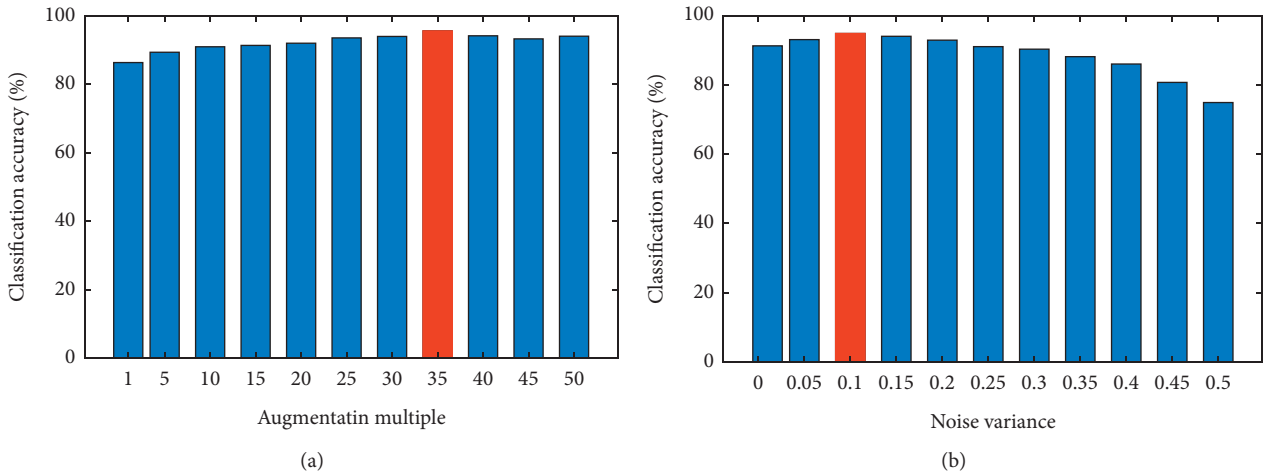


Figure 8: Accuracy of the proposed method for different values of (a) augmentation multiple and (b) noise variance.

the performance of the proposed DoA classification is shown in Figure 8, where the average accuracy of all subjects is reported. As observed from Figure 8(a), increasing the augmentation multiple from 1 to 35 increases the accuracy, but after that, accuracy reduces. Hence, the augmentation multiple is set to 35 in all results reported previously in this section. When the greater number of augmented data increases, it is possible that there is outlier data which reduces the accuracy. Also, it is seen from Figure 8(b) that increasing the noise variance from 0.1 to 0.1 increases the accuracy, but after that, accuracy reduces; therefore, we choose the variance of 0.1.

5. Conclusion

In this paper, we presented a noninvasive method for monitor the DoA based on time-frequency analysis of EEG signals recorded in 15 channels considering the electrode placement in 10/20 standard. EEG signals were recorded from six subjects and were partitioned into epoch with the length of 30 seconds, where consequent epochs have 50% overlap with each other. The TFM of each channel was calculated using SPWVD. Obtained TFM showed that frequencies higher than 8 Hz have near-zero amplitude, and we can remove them from TFM. Then, GLCM was employed to

describe the gray content of each TFM. Contrast, correlation, energy, and homogeneity were calculated for each GLCM; hence, the feature vector of each epoch was constructed by 60 features. The redundant features were removed by the MRMR algorithm and KNN classified the remaining feature to determine the DoA. The results showed that the proposed method achieves the average accuracy of 95.32% with 16 gray levels and Mahalanobis distance and the minimum and maximum accuracy among subjects are 93.34% and 96.92%, respectively. The accuracy with high mean and low variation among subjects indicates the efficiency of the proposed method. We also analyzed the effect of the parameters of data augmentation, which indicates that augmentation multiple and noise variance equal to 35 and 0.1, respectively, achieve the highest accuracy.

Employing the methods based on deep learning and transfer learning can be considered as future works. Convolutional neural networks (CNNs) to classify the TFM are obtained from time-frequency analysis. From the viewpoint of time-series analysis, we can use long short-term memory (LSTM) networks to determine the DoA.

Data Availability

Data used to support the study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] B. Musizza and S. Ribaric, "Monitoring the depth of anaesthesia," *Sensors*, vol. 10, no. 12, pp. 10896–10935, 2010.

[2] J. Bruhn, P. S. Myles, R. Sneyd, and M. M. R. F. Struys, "Depth of anaesthesia monitoring: what's available, what's validated and what's next?," *British Journal of Anaesthesia*, vol. 97, no. 1, pp. 85–94, 2006.

[3] S. Pilge and G. Schneider, "BIS and state entropy of the EEG - comparing apples and oranges," *British Journal of Anaesthesia*, vol. 115, no. 2, pp. 164–166, 2015.

[4] M. Cascella, "Impact of anesthetics on brain electrical activity and principles of pEEG-based monitoring during general anesthesia," *General Anesthesia Research*, Humana Press Inc., vol. 150, pp. 35–56, , Totowa, NJ, USA, 2020.

[5] P. Bischoff and I. Rundshagen, "Unerwünschte wachheit während der narkose," *Deutsches Arzteblatt*, vol. 108, no. 1–2, pp. 1–7, 2011.

[6] S. M. Mousavi, A. Adamoğlu, T. Demiralp, and M. G. Shayesteh, "A wavelet transform based method to determine depth of anesthesia to prevent awareness during general anesthesia," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 354739, 13 pages, 2014.

[7] C. Rosow and P. J. Manberg, "Bispectral index monitoring," *Anesthesiology Clinics of North America*, vol. 19, no. 4, pp. 947–966, 2001.

[8] D. Rani and S. Harsoor, "Depth of general anaesthesia monitors," *Indian Journal of Anaesthesia*, vol. 56, no. 5, pp. 437–441, 2012.

[9] A. Barthelme, R. Wiesmayr, and W. Utschick, "Model order selection in DoA scenarios via cross-entropy based machine learning techniques," in *Proceedings of the ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4622–4626, Barcelona, Spain, May 2020.

[10] S. Kreuer and W. Wilhelm, "The narcotrend monitor," *Best Practice & Research Clinical Anaesthesiology*, vol. 20, no. 1, pp. 111–119, 2006.

[11] T. Nguyen-Ky, P. Wen, and Y. Li, "Consciousness and depth of anesthesia assessment based on bayesian analysis of EEG signals," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 6, pp. 1488–1498, 2013.

[12] A. Shalbaf, M. Saffar, J. W. Sleigh, and R. Shalbaf, "Monitoring the depth of anesthesia using a new adaptive neurofuzzy system," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 3, pp. 671–677, 2018.

[13] Q. Liu, Y.-F. Chen, S.-Z. Fan, M. F. Abbod, and J.-S. Shieh, "Quasi-periodicities detection using phase-rectified signal averaging in EEG signals as a depth of anesthesia monitor," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1773–1784, 2017.

[14] W. Saadeh, F. H. Khan, and M. A. B. Altaf, "Design and implementation of a machine learning based EEG processor for accurate estimation of depth of anesthesia," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 4, pp. 658–669, 2019.

[15] T. Nguyen-Ky, H. D. Tuan, A. Savkin, M. N. Do, and N. T. T. Van, "Real-time EEG signal classification for monitoring and predicting the transition between different anaesthetic states," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 5, pp. 1450–1458, 2021.

[16] S. B. Nagaraj, L. M. McClain, E. J. Boyle et al., "Electroencephalogram based detection of deep sedation in ICU patients using atomic decomposition," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 12, pp. 2684–2691, 2018.

[17] G. Wang, Z. Liu, Y. Feng et al., "Monitoring the depth of anesthesia through the use of cerebral hemodynamic measurements based on sample entropy algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 3, pp. 807–816, 2020.

[18] U. Ha, J. Lee, M. Kim, T. Roh, S. Choi, and H.-J. Yoo, "An EEG-NIRS multimodal SoC for accurate anesthesia depth monitoring," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 6, pp. 1830–1843, 2018.

[19] H.-I. Choi, G.-J. Noh, and H.-C. Shin, "Measuring the depth of anesthesia using ordinal power spectral density of electroencephalogram," *IEEE Access*, vol. 8, pp. 50431–50438, 2020.

[20] Y. Park, S.-H. Han, W. Byun, J.-H. Kim, H.-C. Lee, and S.-J. Kim, "A real-time depth of anesthesia monitoring system based on deep neural network with large EDO tolerant EEG analog front-end," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 4, pp. 825–837, 2020.

[21] Q. Liu, J. Cai, S.-Z. Fan et al., "Spectrum analysis of EEG signals using CNN to model patient's consciousness level based on anesthesiologists' experience," *IEEE Access*, vol. 7, pp. 53731–53742, 2019.

[22] C. Guo, J. Yu, L. Wu, Y. Liu, C. Jia, and Y. Xie, "Analysis and feature extraction of EEG signals induced by anesthesia monitoring based on wavelet transform," *IEEE Access*, vol. 7, pp. 41565–41575, 2019.

[23] A. Khosla, P. Khandnor, and T. Chand, "A comparative analysis of signal processing and classification methods for

different applications based on EEG signals," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 2, pp. 649–690, 2020.

[24] I. Matias, N. Garcia, S. Pirbhulal et al., "Prediction of atrial fibrillation using artificial intelligence on electrocardiograms: a systematic review," *Computer Science Review*, vol. 39, Article ID 100334, 2021.

[25] E. Ebrahimzadeh, S. M. Alavi, A. Bijar, and A. Pakkhesal, "A novel approach for detection of deception using Smoothed Pseudo Wigner-Ville Distribution (SPWVD)," *Journal of Biomedical Science and Engineering*, vol. 6, no. 1, pp. 8–18, 2013.

[26] S. Pikula and P. Beneš, "A new method for interference reduction in the smoothed pseudo wigner-ville distribution," *International Journal on Smart Sensing and Intelligent Systems*, vol. 7, no. 5, 2020.

[27] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognition*, vol. 42, no. 3, pp. 425–436, 2009.

[28] R. Ranjbarzadeh, S. B. Saadi, and A. Amirabadi, "LNPSS: SAR image despeckling based on local and non-local features using patch shape selection and edges linking," *Measurement*, vol. 164, Article ID 107989, 2020.

[29] N. Karimi, R. Ranjbarzadeh Kondrood, and T. Alizadeh, "An intelligent system for quality measurement of golden bleached raisins using two comparative machine learning algorithms," *Measurement*, vol. 107, pp. 68–76, 2017.

[30] R. Ranjbarzadeh, S. Jafarzadeh Ghoushchi, M. Bendechache et al., "Lung infection segmentation for COVID-19 pneumonia based on a cascade convolutional network from CT images," *BioMed Research International*, vol. 2021, Article ID 5544742, 16 pages, 2021.

[31] G. N. Srinivasan and G. Shobha, "Statistical texture analysis," 2021, https://publications.waset.org/10009941/statistical-texture-analysis.

[32] H. Hadizadeh, "Multi-resolution local Gabor wavelets binary patterns for gray-scale texture description," *Pattern Recognition Letters*, vol. 65, pp. 163–169, 2015.

[33] S. Mamli and H. Kalbkhani, "Gray-level co-occurrence matrix of Fourier synchro-squeezed transform for epileptic seizure detection," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 1, pp. 87–99, 2019.

[34] S. Du, Y. Yan, and Y. Ma, "Local spiking pattern and its application to rotation- and illumination-invariant texture classification," *Optik*, vol. 127, no. 16, pp. 6583–6589, 2016.

[35] J. Virmani, G. P. Singh, Y. Singh, and Kriti, "PNN-based classification of retinal diseases using fundus images," in *Sensors for Health Monitoring*, pp. 215–242, Elsevier, Amsterdam, Netherlands, 2019.

[36] N. Razmjooy, V. V. Estrela, and H. J. Loschi, "A study on metaheuristic-based neural networks for image segmentation purposes," in *Data Science*, pp. 25–49, CRC Press, Boca Raton, FL, USA, 2019.

[37] Z. M. Arthurs, P. D. Bishop, L. E. Feiten, M. J. Eagleton, D. G. Clair, and V. S. Kashyap, "Evaluation of peripheral atherosclerosis: a comparative analysis of angiography and intravascular ultrasound imaging," *Journal of Vascular Surgery*, vol. 51, no. 4, pp. 933–939, 2010.

[38] Q. Tian, Y. Wu, X. Ren, and N. Razmjooy, "A new optimized sequential method for lung tumor diagnosis based on deep learning and converged search and rescue algorithm," *Biomedical Signal Processing and Control*, vol. 68, Article ID 102761, 2021.

[39] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.

[40] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," *Journal of Bioinformatics and Computational Biology*, vol. 3, no. 2, pp. 185–205, 2005.

[41] M. Bendechache, N.-A. Le-Khac, and M.-T. Kechadi, "Hierarchical aggregation approach for distributed clustering of spatial datasets," in *Proceedings of the2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pp. 1098–1103, ICDMW, Barcelona, Spain, December 2016.

[42] M. Bendechache, "Study of distributed dynamic clustering framework for spatial data mining," 2019, https://oatd.org/oatd/record?record=handle%5C%3A10197%5C%2F10614.

[43] R. Ranjbarzadeh and S. B. Saadi, "Automated liver and tumor segmentation based on concave and convex points using fuzzy c-means and mean shift clustering," *Measurement*, vol. 150, Article ID 107086, 2020.

[44] S. Jafarzadeh Ghoushchi, M. Khazaeili, A. Amini, and E. Osgooei, "Multi-criteria sustainable supplier selection using piecewise linear value function and fuzzy best-worst method," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 2, pp. 2309–2325, 2019.

[45] F. Wang, S.-H. Zhong, J. Peng, J. Jiang, and Y. Liu, "Data augmentation for eeg-based emotion recognition with deep convolutional neural networks,"vol. 10705, pp. 82–93, in *Proceedings of the International Conference on Multimedia Modeling*, vol. 10705, , LNCS, Bangkok, Thailand, February 2018, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 2096208, 9 pages
https://doi.org/10.1155/2021/2096208

Review Article

Pivotal Role of Quantum Dots in the Advancement of Healthcare Research

Pawan K. Tiwari,[1] Mugdha Sahu,[1] Gagan Kumar,[1] and Mohsen Ashourian [2]

[1]*Department of Physics, Birla Institute of Technology, Mesra, Ranchi 835215, Jharkhand, India*
[2]*Department of Electrical Engineering, Majlesi Branch, Islamic Azad University, Isfahan 8631656451, Iran*

Correspondence should be addressed to Mohsen Ashourian; ashourian@gmail.com

Received 24 May 2021; Accepted 31 July 2021; Published 9 August 2021

Academic Editor: Suresh Manic

The quantum dot is a kind of nanoparticle whose dimension is smaller than the size of a typical nanoparticle ranging from tens of nanometers to a few hundredths of nanometers. The quantum mechanical behavior associated with the quantum dot displays different optical and electronic properties, enabling the quantum dot to find potential applications in a multitude of areas such as solar cells, light-emitting diodes, lasers, and biomedical applications. The objective of this investigation is to explore its fundamentals, synthesis, and applications, especially in the healthcare domain. We have discussed the quantum dot synthesis techniques using chemical methods, namely, wet-chemical methods and vapor-phase methods and plasma processing methods, namely, an ion sputtering method and plasma-enhanced chemical vapor deposition method. We have thoroughly investigated the application of quantum dots in imaging, diagnostics, and gene therapy areas. A significant outcome of this review is to propose quantum dots as a new modality in the treatment of cancer and gene therapeutics in the healthcare domain and the potentials of artificial intelligence to improve their performance via the applications of neural networks.

1. Introduction

From solar panels to medical applications, quantum dots are receiving discernible attention in today's world due to their unparalleled and cutting-edge scope [1, 2]. Quantum dots are colloidal semiconductor nanoparticles that exhibit a distinctive set of optical and transport properties due to their spatial confinement regime, also known as the quantum confinement effect. In bulk semiconductors, the presence of multiple atoms causes splitting of electronic energy levels which when grouped forms an energy band. The most filled band, known as the valence band, is at lower energy, and the mostly empty band known as the conduction band is at relatively higher energy. The valence band and conduction band are forbidden by an energy gap, called a bandgap. In order to excite a valence electron to the conduction band, the applied radiation of energy radiation must be equivalent to the forbidden energy of the bandgap. After absorption of suitable energy, an electron (e−) can jump into conduction from the valence band which causes the formation of vacant space in the valence band, known as a hole (h+). This pair of electrons (e−) and hole (h+) can be perceived as a hydrogen-like species and is known as an exciton. These excitons for a specific semiconductor bear a separation between the electron (e−) and hole (h+) which is called the exciton Bohr radius. In quantum dots, the excitons are confined to a much smaller volume of the semiconductor material which is in order of its exciton Bohr radius. This results in less splitting of the energy bands and leads to a quantum confinement region [3]. Such a region of electron-hole pairs in various dimensions within a material and the electronic energy bands associated with it are discrete and quantized. The size and the composition of the quantum dots can be altered to allow the energy levels and the bandgap to be fine-tuned to specific desired energies. Quantum dots are bandgap tunable by their dimension (diameter ranging typically from 2 to 10 nm) which means that their optical and electrical properties can be engineered to meet specific biomedical applications [4].

Nowadays, these quantum dots have been the subject of numerous exceptional reviews. One important attribute to note is that quantum dots are semiconductor nanocrystals. One of the most important characteristics of quantum dots is the conductive region of semiconductor heterostructure that comprises the precisely regulated quantity of excess electrons which are constrained to a small region by the application of external electromagnetic radiations. These parameters can be altered to achieve the desired configuration of the quantum dot to meet the requirements of different applications. The energy levels of quantum dots are mainly determined by two factors, first, the size of the quantum dots, and second, the intensity of external electromagnetic fields under which the corresponding quantum dots are exposed. Moreover, quantum dots can also be used in the field of quantum computation and artificial intelligence (AI). The quantum dots can be fabricated on the layer of semiconductor heterostructure to construct a three-dimensional matrix of quantum registers. The fact that quantum dots can easily be drafted and controlled according to application requirements makes it advantageous, particularly for quantum neural processing networks [5].

The structure of these nanocrystals comprises a fluorescent semiconductor 'core' which has excellent optical properties (i.e., emission and absorption spectrum) due to coating with another semiconductor material which acts as passivation for "core" and is hence known as "shell". The shell has a generally greater bandgap than the core shell that passivates the surface of the core and inhibits nonradiative recombination processes, which can act as a trap state of electrons in the conduction band, and as a consequence, it can lead to a reduction in the fluorescence efficiency, also referred to as photobleaching. Thus, the passivation of the core from the shell is important to improve its photostability and increase quantum yield (brightness) and also to reduce chemical attacks. As a result, the optical properties of quantum dots get less sensitive to changes in their surroundings. This characteristic of core-shell quantum dots is widely used in applications for biological imaging, chemical sensing, and optoelectronics [3].

2. Fabrication of Quantum Dots

Fabrication of quantum dots can be led by various techniques according to the core requisites of the application. These methods of synthesis of quantum dots can be broadly categorized into two approaches, bottom-up approach and top-down approach [6]. The bottom-up synthesis technique approach includes self-organized processes such as colloidal synthesis, ion sputtering, plasma-assisted epitaxy, and others. Alternatively, the top-down approach techniques include ion implantation, molecular beam epitaxy (MBE), X-ray lithography, and electron-beam lithography [7].

2.1. Fabrication Using Various Mechanisms. Figure 1 shows methods of fabrication of quantum dot. A bulk semiconductor is thinned to make quantum dots in the top-down method. Electron-beam lithography, reactive-ion etching,

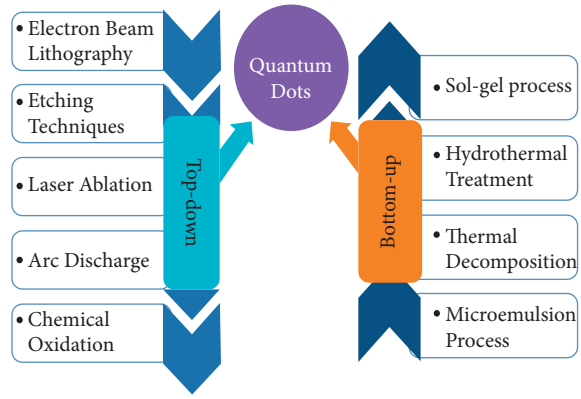


Figure 1: Methods of fabrication of quantum dots.

and/or wet-chemical etching are widely used to create quantum dots with a diameter of about 30 nm. For efficient quantum confinement effect tests, optimized dimensions with defined packing geometries are possible. Alternatively, arrays of zero-dimension quantum dots have been created using focused ion or laser beams. These fabrication processes, however, have significant disadvantages. These processes include the integration of impurities into quantum dots and the patterning of structural flaws [8–14].

Another well-known technique, etching, is important in nanoscale fabrication. Dry etching involves embedding a reactive gas species into an etching chamber and applying a radio frequency voltage to create a plasma that breaks down the gas particles into more volatile parts [7]. These high-kinetic-energy species collide with the surface and form an unstable reaction species, which is used to etch a patterned sample [8]. The etching method for ions is known as reactive-ion etching (RIE). The precise etching of the substrate is done using a masking pattern. The fabrication of GaAs/AlGaAs quantum structures as small as 40 nm has been detailed using RIE and a blend of boron trichloride and argon [7, 8].

Various techniques are used to synthesize quantum dots in the bottom-up approach, and they can be narrowly divided into wet-chemical and vapor-phase approaches. Wet-chemical methods such as microemulsion, sol-gel, competitive reaction chemistry, hot-solution decomposition, and electrochemistry are all classified as (1) wet-chemical methods. Vapor-phase methods include self-assembly of nanostructures in material grown by molecular beam epitaxy (MBE), ion sputtering, liquid metal ion sources, and aggregation of gaseous monomers [5].

2.2. Wet-Chemical Methods. Wet-chemical methods are generally used to prepare colloidal quantum dots. These methods are based on the conventionally performed precipitation methods with more precisely controlled parameters for single or compounded solutions. Invariably, both homogeneous nucleation (condensation and crystallization) and controlled growth of quantum dots are involved in the precipitation methods. Quantum dots of the desired shape, composition, and dimension can be accomplished by altering,

layer thickness, temperature, electrostatic potential, micelle or stabilizers forming, forerunner concentrations, concentrations of anionic and cationic agents, and solvent [5].

However, there are some inherent drawbacks, such as relative instability, associated with the colloidal synthesis of quantum dots. When chemically synthesized quantum dots are coupled with devices, stabilization is mandatory. The stabilization of quantum dots usually involves the construction of a core-shell structure and passivation of quantum dot nanocrystals. Lately, surface chemistry methods of immersing quantum dots in polymer/glass matrices have aroused interest. By this modification, the dangling bonds of quantum dots are blocked. In both quantum dots/polymer and quantum dots/glass structures, distinct optical properties have been shown. As a result, they are used in configurable lasers to support fluorescent materials in order to achieve maximum color emissions [5].

2.3. Vapor-Phase Methods. The vapor-phase methods of synthesis of nanocrystals (quantum dots) are broadly based on the phenomenon of homogeneous nucleation and deposition of vapor-phase composition and formation of the solid particle which is to be restricted to nanoparticulated size and shape. The fundamental cause for the nucleation and deposition of vapor-phase composition is the thermal unitability of vapor-phase composition which is due to self-generated surrounding conditions. The whole process of synthesis of quantum dots occurs in a special type of chamber, also called process chamber, to alter the surrounding conditions such as temperature and pressure. The topic name 'vapor-phase syntheses' indicates the involvement of the concept of supersaturated vapor. In the state of supersaturation, the vapor-phase composition is on the verge of transition to condense phase. The process of phase transition may also involve a chemical reaction. Since vapor-phase methods are performed in the process chamber, the state of the supersaturated vapor phase can be generated thermodynamically or by altering pressure [6]. When the level of saturation is high enough for the nucleation vapor phase, particles will nucleate homogeneously through the kinetics of chemical reaction or condensation. Once the molecules or atoms start to nucleate, the remaining vapor-phase molecules get to react chemically or condense on the nucleated particle to relieve the latent supersaturation, and rather than further nucleation, growth of the quantum dots nanocrystal occurs. Therefore, to produce small quantum dots, the degree of supersaturation of the vapor phase must be high. Further, when the growth of quantum dots reaches the desired size, further growth can be obstructed by immediately quenching the system by eliminating the source of supersaturation. In spite of the fact that quantum dots can be self-assembled utilizing vapor-phase techniques to produce quantum dots arrays without a layout, the change within the estimate of quantum dots frequently results in inhomogeneous optoelectronic properties [14, 15].

2.4. Plasma Processing of Quantum Dots. Figure 2 shows the characterization of plasma. Low-pressure plasmas are an intriguing source for the meticulous preparation of group IV and III–V nanocrystals. Plasma is a quasi-neutral gas made up of a

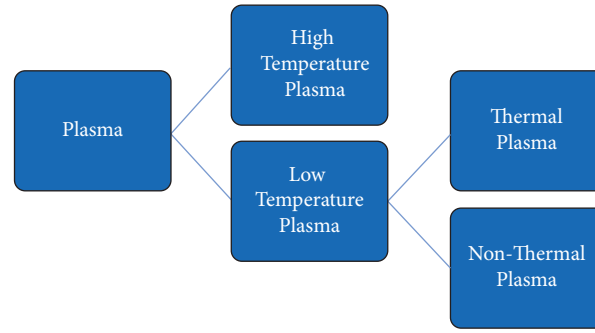


Figure 2: Characterization of plasma.

number of dissimilar entities including electrons, positive and negative ions, free radicals, gas atoms, and molecules in their ground or higher states [6]. This ionized gas can occur at a wide range of temperatures and is formed at low pressure or atmospheric pressure and can occur in a wide temperature range. It is achieved by binding energy to a gaseous environment using different methods such as mechanical, thermal, chemical, radiant, nuclear, applying voltage, injecting electromagnetic waves, and a combination of these methods to divide the gaseous constituent particles into a populace of ions, electrons, charge-neutral gas molecules, and other particles. Plasmas are classified into two types: (1) high temperature or fusion plasmas, in which all electrons, ions, and neutral species are in thermal equilibrium, and (2) low-temperature plasmas. Low-temperature plasmas branch out into thermal plasma and are referred to as quasi-equilibrium plasmas. They are in a state of local thermal equilibrium (LTE). Nonthermal plasma (NTP), also known as a nonequilibrium plasma or cold plasma, is another kind of plasma [16].

The wet-chemical synthesis routes in colloidal solutions are well established, according to mainstream investigations in the field of quantum dots conducted on II-VI and IV-VI chalcogenide semiconductors. Group IV semiconductor nanocrystals (NCs), such as Si or Ge, and several III–V semiconductor nanocrystals (NCs), such as GaAs, GaN, and InAs, are difficult to make in solution due to their stronger covalent bonding and therefore higher melting points. As a second route, electron-beam lithography enables the fabrication of structures with nanometer dimensions but has a severe drawback of the low amount of serial lithographic methods. Therefore, the fabrication of well-defined quantum dots must be conducted by various plasma processes. A further benefit of plasma synthesis over liquid phase methods is that the resulting material is devoid of ligands or surfactants. Depending on the application, these ligands may be added later. This makes it easier to integrate plasma-produced NCs into devices because it eliminates the need for the intermediate ligand exchanges that are normally needed for II-VI and IV-VI NCs. Ion sputtering and Plasma-Enhanced Chemical Vapor Deposition (PECVD) are two plasma-assisted techniques for producing quantum dots (QDs) [16, 17].

2.5. Ion Sputtering Method. One of the most cost-effective and efficient methods of fabrication of highly dense and homogeneous nanoparticles is the surface sputtering

technique, which is broadly a Physical Vapor Deposition (PVD) of bulk semiconductors (cathode). It is a phenomenon in which the semiconductor surface (target material) is bombarded by a beam of energetic particles of plasma, partially ionized gas (commonly Ar). This technique is also known as ion sputtering; this approach involves the utilization of phenomena occurring on the surfaces of bulk semiconductors exposed in the vacuumed chamber to the beam of energetic particles, especially ions. These ions are generated as a spatially clustered beam from gas discharged plasma, and for directed flow, they are accelerated to predetermined energies by the electric field, to have sufficient energy to get projected onto the substrate (Anode). The process of sputtering is commonly performed inside a special device known as an evacuated chamber. For ion beam milling and ion assistance, energy ranging from ~100 eV to 5 keV is used [16]. The sputtering mechanism involves besieging the solid semiconductor surface with energetic gaseous particles (usually Ar+) which are accelerated under high voltage. These energetic gas particles (plasma) encountered with the solid semiconductor's surface particles or sometimes whole atoms of the semiconductor material are expelled out and impelled toward the target substrate and develop a really strong bond. Mechanical forces firmly bind the ensuing film deposition formed by the projected ions to the substrate (Anode), but in certain situations, an alloy or chemical bond may also be formed.

2.6. Plasma-Enhanced Chemical Vapor Deposition (PECVD) Method. Another reliable and renowned synthesis method that has been gaining popularity recently due to its distinctive applications unlike the different stages of microelectronic circuit fabrication is the Plasma-Enhanced Chemical Vapor Deposition (PECVD) method. PECVD is a method for the fabrication of large-scale ICs (integrated circuits) and thin film transistors. In Plasma-Enhanced Chemical Vapor Deposition, the reactive gases are decomposed via electrical discharge. This enables film deposition at lower temperatures than CVD [16]. In CVD, the substrate is heated to produce thermal energy that is required to control chemical reactions. Thermal energy imparts the energy necessary to break bonds. In PECVD, a plasma is used to decompose the gas molecules for film deposition. Thus, plasma is here essentially an energy source replacing thermal energy and can perform the same operations at lower temperatures as conventional processes achieve at elevated temperatures. Therefore, it enables low-temperature processing to be conducted for several synthesis steps. Besides lowering the deposition temperature, PECVD provides superior control over film properties through plasma effects. Thus, when opposed to the traditional CVD process, PECVD offers twofold advantages: low processing temperature and the flexibility of film properties. A low processing temperature meets the small thermal expenditure criterion for many manufacturing applications, while the flexibility of film properties allows for customizing properties to achieve the desired device characteristics. Now, Plasma-Enhanced Chemical Vapor Deposition is being used in upcoming applications such as the synthesis of silicon nanocrystals due to its versatility.

3. Applications of Quantum Dots in Versatile Medical Domains

Due to their various beneficial properties, such as fluorescence, solvation in water, biocompatibility, low toxicity, compact size and ease of modification, rational scale-up processing, and versatile conjugation with other nanoparticles, quantum dots have risen as a potential material in a variety of biomedical applications. As a result, they have become a popular option for a variety of biomedical applications, including drug nanocarriers, therapeutic genes, photosensitizers, and antibacterial particles. Their capabilities must be proven in multifunctional diagnostic levels, cellular and bacterial bioimaging, and the advancement of theragnostic nanomedicine. They will be used shortly for identifying various types of cancer cells, atomic components of disease, and current drug action mechanisms, applying them to intracellular/extracellular considerations, and developing previously untested biochemical assay strategies.

3.1. Diagnostics and Bioimaging. Currently, significant research is aimed to use the distinct optical properties of quantum dots in diagnostics and bioimaging. In the past decades, the most common method of labeling biological entities has been with organic fluorophores. But organic fluorophores have some limitations such as the instability of organic fluorophores when exposed to photoirradiation results in degradation and photobleaching over time. This implies that it does not continuously fluoresce for an extended period of time [18]. Typically, conventional organic fluorophores have a very limited excitation range which allows simultaneous excitation of several organic fluorophores very challenging, resulting in incompatibility in multicolor applications. Dyes are more receptive to their immediate surrounding conditions, such as changes in pH level, causing nonradiative recombination processes [3, 19].

On the other hand, quantum dots have unique optical properties that make them ideal for biological imaging and tagging applications for several reasons [20]. First, quantum dots have a higher extinction coefficient and typically have a broad excitation range and narrow emission peak. This allows multiplex bioimaging with a single source [20]. Second, quantum dots have high quantum yields (QYs), which helps in producing high contrast imaging [18, 21]. Third, they show less photobleaching compared to conventional dyes. Fourth, their size-tunable property of absorption and emission spectrum allows imaging and tagging multiple targets at the same time. Fifth, due to the narrow emission spectrum and minimal interference of the spectrum, quantum dots can be used for the same assay with minimal spectral-overlapping with each other. Sixth, the inorganic properties of quantum dots make them less toxic than conventional organic dyes [22]. Seventh, quantum dots

can be functionalized with different bioactive agents to affect the dynamics and functioning of the target biomolecules.

Quantum dots that have a near-infrared (650–900 nm) emission spectrum can be used to prevent any interference from the autofluorescence emitted by cells, hemoglobin, and water. The absorption coefficients and scattering effects of cells, hemoglobin, and water are lower in the near-infrared region. Intravital microscopy is usually performed with light, but imaging deeper tissues necessitate the use of near-infrared light. Moreover, under ultraviolet excitation, quantum dots are more photostable than organic molecules and have more saturated fluorescence [6].

Quantum dots have promising applications in cellular labeling, deep-tissue imaging, and assay labeling due to their unique optical and surface properties, which make them an excellent alternative to conventionally used organic dyes. Besides, their unique optical properties make them excellent fluorescence resonance energy transfer (FRET) donors [21]. FRET is the distance-dependent phenomenon based on the nonradiative transmission of photoexcitation energy from donor to acceptor through dipole-dipole interaction. FRET from quantum dots to organic acceptor molecules allows real-time sensing and recognition of the target [21].

The bioassays and biomedical investigation can be broadly categorized into two parts in vitro and in vivo. In in vitro, bioassays are performed in a controlled and artificial environment whereas, in in vivo, biological systems or biomolecular bioassays are investigated in the presence of living organisms [23]. The selective quantum dots and their purpose of use are categorized into in vivo and in vitro in Tables 1 and 2.

3.2. Drug Research. In the last two decades, and even currently, there has been a lot of research focused on targeted and rate-controlled drug delivery. One of the main concerns in medical treatment is the target and rate-controlled drug delivery to lower the side effects of conventional chemotherapy and other treatment methods. The targeting could be by utilizing pH-level maintenance paramagnetic nanoparticle controlling, ultrasonic waves interaction, and temperature control to localize and target the delivery of the drug to the diseased cells [21]. Unlike conventional organic fluorophores, the multivalent nature of quantum dots enables different ligands to attach to the same quantum dot. In addition, the surface chemistry of the quantum dot may confer additional characteristics to the quantum dot. For example, a quantum dot surface can be coated with a paramagnetic substance and can be used in MRI [6, 21].

The surface properties and high surface area of quantum dots make it suitable for conjugation with different kinds of biomolecules in order to localize the drug delivery to the target cell and can also be traced because of the unique optical and electrochemical properties of quantum dots. To reduce the toxicity of therapeutic agents, it is crucial to penetrate through the cell membrane for the targeted delivery of therapeutic agents or drugs. Also, the surface properties of quantum dots allow bioconjugation with some proteins and peptides, which are found to be capable of penetrating cell membranes. Cell-penetrating peptides (CPPs) are the most widely used biomolecules for penetration through the cell membranes and are compatible with most of the therapeutic agents, nucleic acids, proteins, and liposomes [24]. CPPs are used for internalization to cells of quantum dots which comprises therapeutic agents or drugs. Moreover, some studies also show that quantum dots can be internalized to cells without any biofictionalization by optimizing the size, charge, and surface properties of the respective quantum dots [21, 22]. Studies show that the positively charged particle can easily penetrate through the cell membrane, and the larger ones are maintained to the cytoplasm region while smaller ones can penetrate through the nucleus [19]. Once they localize to their respective target, the drug release can be automatically initiated in a pH-controlled manner. Tumor cells have less pH than the corresponding cells [21]. Moreover, drug release can also be triggered externally from electromagnetic radiation which can change the thermal state quantum dot [18]. Quantum dots should be selectively conjugated with biomolecules according to the respective target. Quantum dots can be used as specific biomolecule sensors, like sugar maltose; they can be surface conjugated with protein which has a binding affinity to maltose. Likewise, to target special antigens, quantum dots should be conjugated with an appropriate antibody [21, 24]. The most studied bioconjugation of quantum dots is a surface fictionalization of foliate to target foliate receptors overexpressed in many types of tumor cells [24]. Recent researches have shown the capability of quantum dot for targeted drug delivery and tracing the delivery process of the drugs simultaneously. Encapsulation of quantum dots is crucial to reduce toxicity and to prevent drug release before attaching to the target cell membrane [3, 19]. Thus, quantum dots are encapsulated in biodegradable polymer like the most appropriate polymer to encapsulate anticancer drug which is chitosan [24]. The size-tunable emission spectrum of quantum dots makes it beneficial for simultaneous multicolor imaging of different targets in the body [6, 18, 23].

3.3. Gene Therapy. Numerous studies are being conducted by scientists across the globe to modify the genes, carry out gene replacement of the defective ones and novel ways to cure or prevent diseases, and improve health quality. In the past few decades, with the rising need for development in healthcare research in gene therapy, tremendous capabilities have been shown in a variety of applications. Its prominence has been evidently portrayed in several manifestations of developing concepts such as CRISPR-based genome editing and epigenetic research. The growing prominence of researches based on the science of metabolomics, proteomics, and transcriptomics will further expedite the imminent breakthrough of gene therapy [25, 26].

Gene therapy has been quite instrumental in the treatment of numerous health issues, such as cancers or tumors, neurological malfunctions, dermatological disorders, cardiovascular disorders, ophthalmologic disorders, inner ear disorders, and transmittable diseases [27]. However,

Table 1: Quantum dots in *in vitro* bioimaging [6].

Quantum dots	Purpose	Emission/size of QDs
$CdSe/CdS/SiO_2$	Mouse fibroblast cell imaging	550 nm & 630 nm
CdSe/ZnS	Biological detection/sensing	1–4 nm
$CdSe/ZnS/SiO_2$	Phagokinetic track imaging	554 nm & 626 nm
CdSe/ZnS	Tumor vasculature and lung endothelium imaging	<10 nm

Table 2: Quantum dots in *in vivo* bioimaging [6].

Quantum dots	Purpose	Emission/size of QDs
CdSe/ZnS	Tumor vasculature and lung endothelium imaging	<10 nm
CdSe/CdSe	Cancer cell lymph nodes imaging	Near infrared
CdSe/ZnS	Maltose binding protein	560

nowadays, gene therapy is being utilized in a broader and unlimited aspect. Small interfering RNA, antisense oligonucleotides, and microRNA are various other therapeutic nucleic acid materials that have been included within the conventions of quality therapy [25, 28]. Several viral vectors have been proven to be effective in carrying out these tasks. However, due to some poisonous immunological impacts shown by these approaches, quantum dots have been developed as a safe and compelling nonviral gene vector within the domain of gene therapy [25, 29].

Gene therapy necessitates the use of multifunctionalized and biocompatible delivery platforms penetrating through the cell membranes. The optical and surface properties and high biocompatible surface area of quantum dots make it suitable for conjugation with different kinds of fictionalizers, so quantum dots can function as multifuctionalized delivery platforms for gene therapy [25]. Owing to the impressive size-tunable properties of absorption and emission spectrum and high quantum yield, quantum dots act as a competent FRET donor. Quantum dots utilized as delivery platforms are optically traceable and competent in tying different bioligands, such as DNA, RNA, cell-penetrating peptides (CPPs), and nuclear localization groupings (NLS), without aggravating the natural function [30]. Quantum dot-conjugated oligonucleotide arrangements that are joined through surface carboxylic acid groups may be focused on to tie with DNA or mRNA [31, 32]. Various experiments held by scientists revealed that when siRNA is internalized into a cell in conjugation with quantum dots, the synthesis of test protein in the cell gets reduced by 98 percent. Additionally, the significant role of the quantum dots in RNA technology is to detect mRNA molecules by utilizing in situ hybridization (ISH) and also for the combination of siRNA with RNA treatment applications. In the ISH technique, quantum dots are successfully used to examine the expression of unambiguous mRNA transcripts in mouse midbrain regions [33].

4. Artificial Intelligence in Quantum Dot Biomedical Applications

Artificial intelligence provides new potentials for biomedical applications such as the prediction of patients' responses to therapies, the early diagnosis of a variety of disorders, and giving specific treatment options to the patients [32–46]. The support vector machine (SVM) learning method provides us with the lesser root mean square errors (RMSE) which is a result of optimization of the weights, W, calculated through the expression

$$W = \sum_{j=1}^{N} \left(\alpha_j - \alpha_j^*\right) X_j, \tag{1}$$

where α and X are variables [47].

Artificial neural network (ANN) is applied for modeling the relationships between the in vitro data and in vivo effects [48] in asthmatic patients receiving the monodisperse aerosols of salbutamol sulfate.

ANN has been used widely to find cause-and-effect relationships and prediction of in vitro and in vivo data (IVIVC) correlations [47–53]. The ANN helps to find the interpolation of the pharmacokinetic parameters and constructing complex relationships among the constituent parameters [53].

In [54], a multilayer perceptron (MLP) feedforward network has been used for modeling the effects of causal factors on in vitro release profile of theophylline. The authors examined several training algorithms besides the experimental datasets. They found the gradient descent back-propagation algorithms as the best training algorithms for modeling and prediction of drug release profiles [54]. Using ANN results in improvement of coating levels and reducing the amount of chitosan-pectin complex results in the retarded drug release. As compared to another causal factor, coating weight gain, chitosan-pectin in the coating solution played a significantly influential role in determining dissolution profiles [54].

ANN and genetic algorithm are used to predict and optimize in vitro proliferation mineral medium for G × N15 rootstock [55]. The authors were able to determine a suitable culture medium formulation to achieve the best in vitro productivity [55].

ANN is used in [56] to predict formulations of sustained-release tablets, solubility, and ratios of hydroxypropyl methylcellulose/dextrin for several tablet formulations and their accumulation and in vitro release at various sampling times. A close agreement was found between ANN-

predicted and experimental findings [56] indicating the capability of ANN models for the development of formulations with suitable physicochemical properties. The trained ANN model has been able to predict optimal compositions of tablet formulations based on the proper dissolution-time profiles in vitro and release profiles in vivo [56].

An ANN and pharmacokinetic simulations were used in the design of controlled release formulations with predictable in vitro and in vivo behavior [57].

The optimized ANN model was used for the prediction of formulations based on two desired target in vitro dissolution-time profiles and two desired bioavailability profiles [57].

In [58], the authors investigate the effects of formulation factors on the in vitro release profile of diclofenac sodium from matrix tablets using the design of experiment (DOE). Formulations of diclofenac sodium tablets, with Carbopol 71G as matrix substance, were optimized by an artificial neural network [58].

In [59], the authors use the Generative Adversarial Network (GAN) framework to synthesize calcium imaging data, as a method to scale up or augment the amount of neuronal population activity data. They validated the method on artificial data with known ground-truth and synthesize data mimicking real two-photon calcium (Ca^{2+}) imaging data as recorded from the primary visual cortex of a behaving mouse [59, 60]. Their results show that the GAN is capable of synthesizing realistic fluorescent calcium indicator signals similar to those imaged in the somata of neuronal populations of behaving animals. WaveGAN architecture with the Wasserstein distance training objective is used in their system development [61]. They found that Calcium GAN is able to learn the underlying distribution of the data. The authors then fitted a developed model to imaging data from the primary visual cortex of behaving mice. Importantly, it is possible to show that the statistics of the synthetic spike trains match the statistics of the recorded data.

In summary, AI-based models along with the conventional methods speed up the development and optimization of controlled release drug delivery systems [62]. They also help in the evaluation of the effects of process and formulation variables on the delivery system. The prevalent ANN in the biomedical applications could be ARTMAP, Bayesian Feedforward, Backpropagation, Hopfield, Neurofuzzy, and Resilient Propagation [63].

5. Conclusion and Future Work

The unique features of quantum dots which include a higher excitation coefficient broader absorption range, narrow and symmetric emission peak, and considerably good resistance to photobleaching, and many more have cemented the role of these versatile nanoparticles in the burgeoning fields of nanoscience and nanotechnology. Although quantum dots are not designed to assist everywhere, the application of quantum dots transcends beyond these artificial domains. In the future, quantum dots are going to be widely used as tools for diagnosing cancer cells. These dots can be intended for localized delivery and rate-controlled release of anticancer drugs or therapeutic agents to very specific parts of the body. Due to the property to target single organs, unlike other conventional drugs which come with unpleasant side effects, this makes them precise and more advantageous than other traditional processes like chemotherapy. These nanoparticles will also assist to categorize a wide spectrum of molecular mechanisms of various diseases as well as novel drug action mechanisms. They will be also used for observing and studying all the intracellular and extracellular activities.

In comparison to the conventional use of organic fluorophore dyes, quantum dots such as CdSe have a higher quantum yield (approximately twenty times) and better resistance to photobleaching over time, making them more suitable for long-term imaging applications. Moreover, quantum dots are excellent for multiplex bioimaging which cannot be accomplished by organic fluorophore dyes; since quantum dots have a wide absorption spectral range, their emission spectrum wavelength can be readily altered in the range of 400 nm to 4000 nm by tuning their size. Quantum dots are more cost-effective for drug delivery than gold nanoparticles, and they can also serve as carrier modules of externally guided nanobots.

However, quantum dots have also some drawbacks. In some in vitro research, degradation of quantum dots has been seen over time, which could lead to toxicities in the relevant living organisms. Quantum dots are extremely toxic and therefore must be necessarily encapsulated with a stable polymer shell. Additionally, in the core-shell quantum dot, the optical properties of the core can be obstructed by the shell and also make the regulation of the size of the quantum dots difficult.

When conjugated with targeting agents and photosensitizers, a new category of mutable multifunctional nanoparticles can be developed that might prove as an aid to diagnostics or therapy. Also, the preparation of high-quality quantum dots from compounds that are nontoxic such as carbon and silicon might reveal a superior nanocrystal which might be of paramount clinical relevance. In imaging applications and bioanalytical chemistry assays, quantum dots have recently been shown to be a promising alternative to molecular fluorophores. But these innovative trends are evident that quantum dots have a futuristic scope and will become leading fluorescent reporters in biology and medicine over the next decade. A thoughtful, vigilant, and ground-breaking approach to deal with these nifty nanoparticles might bring unparalleled miracles in the field of healthcare research.

Data Availability

The data used to support the findings of this study were obtained from the results of a literature review, and the authors do not have any new data.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] V. Dunjko and H. J. Briegel, "Machine learning & artificial intelligence in the quantum domain: a review of recent progress," *Reports on Progress in Physics*, vol. 81, no. 7, Article ID 074001, 2018.

[2] B. J. Shastri, A. N. Tait, T. Ferreira de Lima et al., "Photonics for artificial intelligence and neuromorphic computing," *Nature Photonics*, vol. 15, no. 2, pp. 102–114, 2021.

[3] A. Fontes, R. B. De Lira, M. Seabra, T. G. Da Silva, A. G. De Castro Neto, and B. S. Santos, "Quantum dots in biomedical research," *Biomedical Engineering—Technical Applications in Medicine*, pp. 269–290, 2012.

[4] J. S. Lee, Y. H. Youn, I. K. Kwon, and N. R. Ko, "Recent advances in quantum dots for biomedical applications," *Journal of Pharmaceutical Investigation*, vol. 48, no. 2, pp. 209–214, 2018.

[5] M. V. Altaisky, N. N. Zolnikova, N. E. Kaputkina, V. A. Krylov, Y. E. Lozovik, and N. S. Dattani, "Towards a feasible implementation of quantum neural networks using quantum dots," *Applied Physics Letters*, vol. 108, no. 10, Article ID 103108, 2016.

[6] D. Bera, L. Qian, T.-K. Tseng, and P. H. Holloway, "Quantum dots and their multimodal applications: a review," *Materials*, vol. 3, no. 4, pp. 2260–2345, 2010.

[7] A. Valizadeh, H. Mikaeili, M. Samiei, S. M. Farkhani, N. Zarghami, M. Akbarzadeh et al., "Quantum dots: synthesis, bioapplications, and toxicity," *Nanoscale Research Letters*, vol. 7, no. 1, pp. 480–514, 2012.

[8] J. Yang, T. Ling, W. Wu et al., "A top–down strategy towards monodisperse colloidal lead sulphide quantum dots," *Nature Communications*, vol. 4, p. 1695, 2013.

[9] J. Zhang, H. Wang, Y. Xiao et al., "A simple approach for synthesizing of fluorescent carbon quantum dots from tofu wastewater," *Nanoscale Research Letters*, vol. 12, p. 611, 2017.

[10] X. Xu, R. Ray, Y. Gu et al., "Electrophoretic analysis and purification of fluorescent single-walled carbon nanotube fragments," *Journal of the American Chemical Society*, vol. 126, no. 40, pp. 12736-12737, 2004.

[11] H. X. Zhao, L. Q. Liu, Z. D. Liu, Y. Wang, X. J. Zhao, and C. Z. Huang, "Highly selective detection of phosphate in very complicated matrixes with an off-on fluorescent probe of europium-adjusted carbon dots," *Chemical Communications*, vol. 47, no. 9, pp. 2604–2606, 2011.

[12] S. C. Ray, A. Saha, N. R. Jana, and R. Sarkar, "Fluorescent carbon nanoparticles: synthesis, characterization, and bioimaging application," *Journal of Physical Chemistry C*, vol. 113, no. 43, pp. 18546–18551, 2009.

[13] X. Li, H. Wang, Y. Shimizu, A. Pyatenko, K. Kawaguchi, and N. Koshizaki, "Preparation of carbon quantum dots with tunable photoluminescence by rapid laser passivation in ordinary organic solvents," *Chemical Communications*, vol. 47, pp. 932–934, 2010.

[14] A. A. Guzelian, U. Banin, A. V. Kadavanich, X. Peng, and A. P. Alivisatos, "Colloidal chemical synthesis and characterization of InAs nanocrystal quantum dots," *Applied Physics Letters*, vol. 69, no. 10, pp. 1432–1434, 1996.

[15] B. O. Dabbousi, J. Rodriguez-Viejo, F. V. Mikulec et al., "(CdSe)ZnS core–shell quantum dots: synthesis and characterization of a size series of highly luminescent nanocrystallites," *The Journal of Physical Chemistry B*, vol. 101, no. 46, pp. 9463–9475, 1997.

[16] V. K. Khanna, "Plasma synthesis of quantum dots," *Nanoscience and Nanotechnology Letters*, vol. 4, pp. 275–290, 2013.

[17] K. I. Hunter, J. T. Held, K. A. Mkhoyan, and U. R. Kortshagen, "Nonthermal plasma synthesis of core/shell quantum dots: strained Ge/Si nanocrystals," *ACS Applied Materials & Interfaces*, vol. 9, no. 9, pp. 8263–8270, 2017.

[18] M. H. Ko, S. Kim, W. J. Kang et al., "In vitro derby imaging of cancer biomarkers using quantum dots," *Small (Weinheim an der Bergstrasse, Germany)*, vol. 5, no. 10, pp. 1207–1212, 2009.

[19] J. Wang, "Cd-containing quantum dots for biomedical imaging," in *The World Scientific Encyclopedia of Nanomedicine and Bioengineering I*, vol. 4, pp. 111–158, World Scientific, Nano Imaging: From Fundamental Principles to Translational Medical Applications, 2016.

[20] S. Muralidhara, K. Malu, P. Gaines, and B. M. Budhlall, "Quantum dot encapsulated nanocolloidal bioconjugates function as bioprobes for in vitro intracellular imaging," *Colloids and Surfaces B: Biointerfaces*, vol. 182, Article ID 110348, 2019.

[21] A. Nemati, "Diagnostic and drug delivery applications "a brief review"," *Iranian Journal of Materials Science and Engineering*, vol. 17, pp. 1–12, 2020.

[22] X. T. Zheng, A. Ananthanarayanan, K. Q. Luo, and P. Chen, "Glowing graphene quantum dots and carbon dots: properties, syntheses, and biological applications," *Small*, vol. 11, no. 14, pp. 1620–1636, 2015.

[23] E. Asadi, A. Chimeh, S. Hosseini et al., "A review of clinical applications of graphene quantum dot-based composites," *Journal of Composites and Compounds*, vol. 1, no. 1, pp. 31–40, 2019.

[24] I. D. Tomlinson, M. R. Warnement, and S. J. Rosenthal, "Quantum dots in medical chemistry," in *Quantum Dots: Research, Technology and Applications*, R. W. Knoss, Ed., pp. 243–259, 2009.

[25] L. F. Barnes, *Applications of Quantum Dots in Gene Therapy*, Florida State University, Tallahassee, FL, USA, 2010.

[26] H. Xu, Z. Li, and J. Si, "Nanocarriers in gene therapy: a review," *Journal of Biomedical Nanotechnology*, vol. 10, no. 12, pp. 3483–3507, 2014.

[27] E. Abbasi, T. Kafshdooz, M. Bakhtiary et al., "Biomedical and biological applications of quantum dots," *Artificial Cells, Nanomedicine, and Biotechnology*, vol. 44, no. 3, pp. 885–891, 2016.

[28] V. G. Reshma and P. V. Mohanan, "Quantum dots: applications and safety consequences," *Journal of Luminescence*, vol. 205, pp. 287–298, 2019.

[29] P. Devi, S. Saini, and K.-H. Kim, "The advanced role of carbon quantum dots in nanomedical applications," *Biosensors and Bioelectronics*, vol. 141, Article ID 111158, 2019.

[30] D. Shao, Q. Zeng, Z. Fan et al., "Monitoring HSV-TK/ganciclovir cancer suicide gene therapy using CdTe/CdS core/shell quantum dots," *Biomaterials*, vol. 33, no. 17, pp. 4336–4344, 2012.

[31] D. Gerion, W. J. Parak, S. C. Williams, D. Zanchet, C. M. Micheel, and A. P. Alivisatos, "Sorting fluorescent nanocrystals with DNA," *Journal of the American Chemical Society*, vol. 124, no. 24, pp. 7070–7074, 2002.

[32] P. Chan, T. Yuen, F. Ruf, J. Gonzalez-Maeso, and S. C. Sealfon, "Method for multiplex cellular detection of mRNAs using quantum dot fluorescent in situ hybridization," *Nucleic Acids Research*, vol. 33, no. 18, p. e161, 2005.

[33] H. Soo Choi, W. Liu, P. Misra et al., "Renal clearance of quantum dots," *Nature Biotechnology*, vol. 25, no. 10, pp. 1165–1170, 2007.

[34] B. Bostan, R. Greiner, D. Szafron, and P. Lu, "Predicting homologous signaling pathways using machine learning," *Bioinformatics*, vol. 25, no. 22, pp. 2913–2920, 2009.

[35] H. Chen, O. Engkvist, Y. Wang, M. Olivecrona, and T. Blaschke, "The rise of deep learning in drug discovery," *Drug Discovery Today*, vol. 23, no. 6, pp. 1241–1250, 2018.

[36] A. W. Przybyszewski, M. Kon, S. Szlufik, A. Szymanski, P. Habela, and D. M. Koziorowski, "Multimodal learning and intelligent prediction of symptom development in individual Parkinson's patients," *Sensors*, vol. 16, no. 9, p. 1498, 2016.

[37] D. Romeo-Guitart, J. Forés, M. Herrando-Grabulosa et al., "Neuroprotective drug for nerve trauma revealed using artificial intelligence," *Scientific Reports*, vol. 8, p. 1879, 2018.

[38] M. H. Bahari, M. Mahmoudi, A. Azemi, M. M. Mirsalehi, and M. Khademi, "Early diagnosis of systemic lupus erythmatosus using ANN models of dsDNA binding antibody sequence data," *Bioinformation*, vol. 5, no. 2, pp. 58–61, 2010.

[39] S. M. Pedersen, J. S. Jørgensen, and J. B. Pedersen, "Use of neural networks to diagnose acute myocardial infarction. II. A clinical application," *Clinical Chemistry*, vol. 42, no. 4, pp. 613–617, 1996.

[40] X. Deng, K. Li, and S. Liu, "Preliminary study on application of artificial neural network to the diagnosis of Alzheimer's disease with magnetic resonance imaging," *Chinese Medical Journal*, vol. 112, no. 3, pp. 232–237, 1999.

[41] K. Rudzki, M. Hartleb, T. Sadowski, and J. Rudzka, "Focal liver disease: neural network-aided diagnosis based on clinical and laboratory data," *Gastroenterologie Clinique et Biologique*, vol. 21, no. 2, pp. 98–102, 1997.

[42] M. G. Penedo, M. J. Carreira, A. Mosquera, and D. Cabello, "Computer-aided diagnosis: a neural-network-based approach to lung nodule detection," *IEEE Transactions on Medical Imaging*, vol. 17, no. 6, pp. 872–880, 1998.

[43] G. Mello, E. Parretti, A. Ognibene et al., "Prediction of the development of pregnancy-induced hypertensive disorders in high-risk pregnant women by artificial neural networks," *Clinical Chemistry and Laboratory Medicine*, vol. 39, no. 9, pp. 801–805, 2001.

[44] H. Hirose, T. Takayama, S. Hozawa, T. Hibi, and I. Saito, "Prediction of metabolic syndrome using artificial neural network system based on clinical data including insulin resistance index and serum adiponectin," *Computers in Biology and Medicine*, vol. 41, no. 11, pp. 1051–1056, 2011.

[45] J. E. Arle, K. Perrine, O. Devinsky, and W. K. Doyle, "Neural network analysis of preoperative variables and outcome in epilepsy surgery," *Journal of Neurosurgery*, vol. 90, no. 6, pp. 998–1004, 1999.

[46] P. Schmid, M. B. Wischnewsky, O. Sezer, R. Böhm, and K. Possinger, "Prediction of response to hormonal treatment in metastatic breast cancer," *Oncology*, vol. 63, no. 4, pp. 309–316, 2002.

[47] H. T. Huynh, J. J. Kim, and Y. G. Won, "Non-linear estimation methods for hematocrit density based on changing pattern of transduced anodic current curve," *WSEAS Transactions on Information Science and Applications*, vol. 5, pp. 1541–1550, 2008.

[48] M. De Matas, Q. Shao, M. F. Biddiscombe, S. Meah, H. Chrystyn, and O. S. Usmani, "Predicting the clinical effect of a short acting bronchodilator in individual patients using artificial neural networks," *European Journal of Pharmaceutical Sciences*, vol. 41, no. 5, pp. 707–715, 2010.

[49] M. De Matas, Q. Shao, C. H. Richardson, and H. Chrystyn, "Evaluation of in vitro in vivo correlations for dry powder inhaler delivery using artificial neural networks," *European Journal of Pharmaceutical Sciences*, vol. 33, no. 1, pp. 80–90, 2008.

[50] M. E. Brier, J. M. Zurada, and G. R. Aronoff, "Neural network predicted peak and trough gentamicin concentrations," *Pharmaceutical Research*, vol. 12, no. 3, pp. 406–412, 1995.

[51] P. Veng-Pedersen and N. B. Modi, "Application of neural networks to pharmacodynamics," *Journal of Pharmaceutical Sciences*, vol. 82, no. 9, pp. 918–926, 1993.

[52] J. V. S. Gobburu and W. H. Shelver, "Quantitative structure-pharmacokinetic relationships (QSPR) of beta blockers derived using neural networks," *Journal of Pharmaceutical Sciences*, vol. 84, no. 7, pp. 862–865, 1995.

[53] J. A. Dowell, A. Hussain, J. Devane, and D. Young, "Artificial neural networks applied to the in vitro-in vivo correlation of an extended-release formulation: initial trials and experience," *Journal of Pharmaceutical Sciences*, vol. 88, no. 1, pp. 154–160, 1999.

[54] A. Ghaffari, H. Abdollahi, M. R. Khoshayand, I. S. Bozchalooi, A. Dadgar, and M. Rafiee-Tehrani, "Performance comparison of neural network training algorithms in modeling of bimodal drug delivery," *International Journal of Pharmaceutics*, vol. 327, no. 1-2, pp. 126–138, 2006.

[55] U. Mandal, V. Gowda, A. Ghosh et al., "Optimization of metformin HCl 500 mg sustained release matrix tablets using artificial neural network (ANN) based on multilayer perceptrons (MLP) model," *Chemical and Pharmaceutical Bulletin*, vol. 56, no. 2, pp. 150–155, 2008.

[56] X. H. Wei, J. J. Wu, and W. Q. Liang, "Application of an artificial neural network in the design of sustained-release dosage forms," *Yao Xue Xue Bao*, vol. 36, no. 9, pp. 690–694, 2001, in Chinese.

[57] Y. Chen, T. W. McCall, A. R. Baichwal, and M. C. Meyer, "The application of an artificial neural network and pharmacokinetic simulations in the design of controlled-release dosage forms," *Journal of Controlled Release*, vol. 59, no. 1, pp. 33–41, 1999.

[58] I. Branka, I. Svetlana, C. Nebojša, P. Aleksandra, T. Svetlana, and D. Zorica, "Application of design of experiments and multilayer perceptrons neural network in the optimization of diclofenac sodium extended release tablets with Carbopol® 71G," *Chemical and Pharmaceutical Bulletin*, vol. 58, no. 7, pp. 947–949, 2010.

[59] A. Y. Zhang and T. Y. Fan, "Optimization of calcium alginate floating microspheres loading aspirin by artificial neural networks and response surface methodology," *Beijing Da Xue Xue Bao Yi Xue Ban*, vol. 42, no. 2, pp. 197–201, 2010.

[60] B. M. Li, T. Amvrosiadis, N. Rochefort, and A. Onken, "CalciumGAN: a generative adversarial network model for synthesizing realistic calcium imaging data of neuronal populations," 2020, arXiv e-prints arXiv-2009.

[61] P. Bojanowski, A. Joulin, D. Lopez-Paz, and A. Szlam, "Optimizing the latent space of generative networks," 2017, arXiv preprint arXiv:1707.05776.

[62] P. Hassanzadeh, F. Atyabi, and R. Dinarvand, "The significance of artificial intelligence in drug delivery system design," *Advanced Drug Delivery Reviews*, vol. 151-152, pp. 169–190, 2019.

[63] R. Nayak, L. C. Jain, and B. K. H. Ting, "Artificial neural networks in biomedical engineering: a review," *Computational Mechanics-New Frontiers for the New Millennium*, pp. 887–892, 2001.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 7714351, 14 pages
https://doi.org/10.1155/2021/7714351

Research Article

Presentation of a Segmentation Method for a Diabetic Retinopathy Patient's Fundus Region Detection Using a Convolutional Neural Network

Amin Valizadeh,[1] **Saeid Jafarzadeh Ghoushchi**,[2] **Ramin Ranjbarzadeh**,[3] **and Yaghoub Pourasad**[4]

[1]*Department of Mechanical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran*
[2]*Department of Industrial Engineering, Urmia University of Technology (UUT), P.O. Box 57166-419, Urmia, Iran*
[3]*Department of Telecommunications Engineering, Faculty of Engineering, University of Guilan, Rasht, Iran*
[4]*Department of Electrical Engineering, Urmia University of Technology (UUT), P.O. Box 57166-419, Urmia, Iran*

Correspondence should be addressed to Saeid Jafarzadeh Ghoushchi; s.jafarzadeh@uut.ac.ir

Received 6 June 2021; Revised 30 June 2021; Accepted 18 July 2021; Published 27 July 2021

Academic Editor: Navid Razmjooy

Diabetic retinopathy is characteristic of a local distribution that involves early-stage risk factors and can forecast the evolution of the illness or morphological lesions related to the abnormality of retinal blood flows. Regional variations in retinal blood flow and modulation of retinal capillary width in the macular area and the retinal environment are also linked to the course of diabetic retinopathy. Despite the fact that diabetic retinopathy is frequent nowadays, it is hard to avoid. An ophthalmologist generally determines the seriousness of the retinopathy of the eye by directly examining color photos and evaluating them by visually inspecting the fundus. It is an expensive process because of the vast number of diabetic patients around the globe. We used the IDRiD data set that contains both typical diabetic retinopathic lesions and normal retinal structures. We provided a CNN architecture for the detection of the target region of 80 patients' fundus imagery. Results demonstrate that the approach described here can nearly detect 83.84% of target locations. This result can potentially be utilized to monitor and regulate patients.

1. Introduction

The most prevalent cause of vision loss in people with diabetes is diabetic retinopathy; however, additional causes of harm include various eye disorders in the retinal and nonretinal visions including glaucoma, macular degeneration associated with age, and neuropathy in the vascular and cataract vision. The intensity and severity of diabetic retinopathy are often diagnosed by visual examination of the fundus by direct inspection and analyzing color pictures by an ophthalmologist. This technique is costly and time-demanding, given the enormous number of diabetes patients globally [1]. The severity of diabetic retinopathy and the diagnosis of primary illnesses are both relatively subjective, with agreement amongst qualified professionals documented variably in prior research [2, 3]. Furthermore, 75% of diabetic retinopathy patients reside in poor areas with a lack of experts and diagnostic infrastructure [4]. To prevent the development of avoidable eye illnesses, global area-based prediction systems have been established; however, large-scale diabetic retinopathy exists for such systems to detect and cure retinopathy successfully on an individual basis. As a consequence, millions of people throughout the world continue to experience vision loss due to a lack of effective diagnosis and eye care.

Automated techniques for identifying retinal disease utilizing stained fundus pictures have been developed to solve the inadequacies of conventional diagnostic approaches [5]. Decentralized technicians may check numerous patients objectively and without relying on physicians using such a gadget, reducing the burden of qualified specialists. Previous techniques to the automated

detection of diabetic retinopathy, on the other hand, have substantial disadvantages that impede widespread use. Since the majority of these methods are based on short data sets of around 500 pictures in specific clinical situations, they try to diagnose large-scale diabetic retinopathy in real-world heterogeneous fundus data sets [6]. Techniques generated from certain data sets may not be applicable to fundus pictures (gained from other clinical researches and employing various types of fundus cameras, eye-opening methods, or both). Furthermore, many of these algorithms for diagnosing diabetic retinopathy depend on the manual extraction of features, which seeks to explain the prediction of anatomical structures in the fundus, such as the optic disc or blood vessels, using accurate manual characteristics. Though these manually tuned features may be applied to individual fundus data sets, they are still used to detect diabetic retinopathy using fundus photos of people who meet the prototype's demographics [7]. Although general objective features, such as acceleration of feature detection and orientation gradient histogram, have been examined as a nonspecific technique for characterizing diabetic retinopathy, these methods do not tackle weaker and disproportionate features that can be found in diabetic retinopathy. It does not go into detail on small changes in retinopathy severity [8].

Researchers have paid a lot of consideration to artificial-intelligence-based approaches in recent times, and they have been able to produce good outcomes in a variety of areas, particularly in the field of computer vision, and in some situations, they have even been able to compete with humans [9]. The preparation of specialized characteristics is another significant aspect in this domain. What has been typical in prior computer vision systems is that researchers first do extensive study to uncover certain characteristics in the input and then utilize these characteristics to execute their intended processing operations. The procedure of developing these characteristics is time-consuming and does not always provide satisfactory outcomes. The features are automatically found through artificial intelligence and data mining technologies, and then the required action is done utilizing these features. This approach offers a high level of accuracy, which is why so much study is being done in this area. As a result, in order to take advantage of the strong capabilities that these approaches give researchers in addressing artificial intelligence challenges, it is vital to get more familiar with them and address them in order to tackle their difficulties in departments [10]. In diabetic patients with diabetic retinopathy, the regular growth-based prognosis is a crucial and critical element of patient therapy. The cost of therapy is heavily influenced by the accuracy and timeliness of this care [11]. If discovered early, compensatory therapy for diabetic retinopathy is available, and it is a vital step. The weighting and location of several characteristics are included in the diabetic retinopathy (DR) classification. This takes a lot of time for doctors. Computers may gain classifications considerably faster after being trained, allowing them to aid clinicians in quick classification. Diabetic retinopathy is defined by morphological lesions in the retina that are related to abnormal retinal blood flow [12]. These lesions have a geographical distribution that can predict disease development and contain risk variables in the early stages of the disease.

This paper's contribution is completely exposed to the early identification of diabetic retinopathy (DR) by preprocessing the fundus retinal picture with the blood vessels separated. As deep learning models are able to extract crucial features from the input image automatically, we used a deep learning structure to segment the image to corroborate our guesses. To show the outcome, the preprocessed picture is fed into the trained CNN. The advantage of utilizing a trained CNN is that it can provide a faster diagnostic and report than an expert can. In diabetics, a yearly retinal examination and early identification of DR can significantly reduce the chance of visual loss.

2. Literature Review

The existence of lesions in the retina, which are produced by the illness, can help ophthalmologists forecast DR. Because of a lack of experience and equipment in some locations where diabetes is common and DR diagnosis and treatment is virtually always required, the suggested strategy is effective. The number of individuals suffering from diabetes is rising, and ophthalmologists are racing to avert blindness, yet DR infrastructure and professionals are in short supply. Gaussian smoothing, morphological top-hat filtering, and contrast enhancement are examples of image enhancement techniques. Enhancement is employed initially to boost contrast and minimize noise, and then adaptive local thresholding is utilized to perform the segmentation job [13]. Sundaram et al. suggested a hybrid segmentation strategy that employs methods such as morphology, multiscale vessel enhancement, and image fusion to emphasize blood vessels, that is, area-based morphology and thresholding [14]. Zhao et al. developed an infinite active contour model for autonomously segmenting retinal blood vessels, based on hybrid region information from the image for microscopic vascular structures [15]. Jiang et al. suggested a global thresholding-based morphological technique, in which capillaries are discovered through centerline detection, to find vessels quickly and correctly. Rodrigues et al. employed the wavelet transform and mathematical morphology to accomplish vascular segmentation, where tubular features of blood vessels were exploited to locate retinal veins and arteries [16]. To improve segmentation accuracy, Sazak et al. presented a retinal blood vessel image-enhancing approach. They employed the mathematical morphology patterns multiscale bowler-hat transform, in which vessel-like formations are recognized by thresholding after merging distinct structural elements [17].

Savelli et al. presented a new approach for segmenting vessels that compensated for lighting. To eliminate haze and shadow noise, a dehazing approach was utilized as a preprocessing approach, and classification was accomplished using a CNN trained on 800,000 patches with a dimension of 27×27 (the decision pixel was deemed the center pixel) [18]. Girard et al. created a fast machine learning strategy for segmenting vessels using a U-Net-inspired CNN for

semantic segmentation, where the encoder and decoder provide down- and up-sampling of the image, accordingly. [19]. Hu et al. suggested a technique based on a CNN and conditional random fields (CRFs) for segmenting retinal vessels. This approach is divided into two phases: first, a multiscale CNN architecture with better cross-entropy loss function was used to the picture, and then CRFs were used to enhance the final output [20]. DeepVessel, a software that combines deep learning and CRFs, was presented by Fu et al. To learn rich hierarchical representations from pictures, a multiscale and multilevel CNN is employed [21]. Soomro et al. suggested a semantic segmentation network centered on deep learning and influenced by the well-known SegNet architecture. Grayscale data were created in the first phase using principal component analysis (PCA). The vessels were extracted. Semantic segmentation convolutional neural networks are used in the second step. After that, post-processing was used to fine-tune the segmentation [22].

Jin et al. suggested a deep neural network based on a deformable U-Net. The network includes deformable convolutions, and an up-sampling operator is employed to improve the picture resolution and extract more accurate feature information [23]. Pixel CNN with batch normalization (PixelBNN), developed by Leopold et al., is centered on U-Net and pixel CNN and uses preprocessing to resize, decrease the dimension, and improve the picture [24]. Dense U-Net was employed as a semantic segmentation network for vascular segmentation by Wang et al. [25], with random transformations employed for data augmentation to increase the dense network's patch-based training effectiveness. For retinal vascular segmentation, Feng et al. suggested a cross-connected CNN (CcNet). Only the green channel of the fundus picture is used to train the CcNet; cross-connections and fusion of multiscale characteristics boost the network's performance [26]. Nevertheless, deep networks were employed in prior studies, which incorporated a large number of trainable parameters that boosted the network's complexity. To address these issues, this study offers a dual residual stream-based Vess-Net, which is not as deep as traditional semantic segmentation networks but provides good segmentation with a small number of trainable parameters and layers. The technique uses machine learning in the segmentation process to aid in the detection of retinopathy. Das et al. employed a CNN to train a diabetic retinopathy classifier and conduct classification. The CNN's classification structure is made up of a combination of squeeze and stimulation and bottleneck layers, one for each class, as well as a convolution and pooling layer architecture for classification between the two classes. When compared to standard systems, experimental findings suggest that the suggested method produces better outcomes. When tested on the DIARETDB1 data set [27], the model had an accuracy of 98.7% and a precision of 97.2%.

Shanthini et al. proposed a DR detection approach based on threshold segmentation. This approach used pixel-based segmentation to classify the foreground and background of the input retinal picture. A two-layer CNN is used to supplement the layer assessment process, which reduces false positives during classification. This procedure is followed in order to determine the exact detection of the retina's target region. Furthermore, the segment-based CNN (S-CNN) corrects the diagnostic fault using two hidden layers to distinguish between threshold and normalized circumstances based on categorization. The suggested technique is effective in improving detection accuracy, sensitivity, and true positives [28]. For the segmentation and localization of OD and fovea centers, Hasan et al. suggested the DRNet, an end-to-end encoder-decoder network. They suggested a skip link, called the residual skip connection, in the DRNet to compensate for the spatial information lost owing to encoder pooling. The suggested skip connection does not immediately concatenate low-level feature maps from the encoder's starting layers with the matching same scale decoder, unlike the U-Net's previous skip connection [29]. Other methods include genetic algorithm [30], two-path CNN [31], cascade CNN [32], patch shape selection method for feature extraction [33], FCM with mean-shift clustering method [34, 35], GARCH feature selection method [36], particle swarm algorithm for reducing surface roughness [37], and fatigue detection [38].

To identify and categorize DR from color fundus pictures, Jayanthi et al. utilized a novel particle swarm optimization (PSO) algorithm-based CNN model dubbed the PSO-CNN model. Preprocessing, feature extraction, and classification are the three steps of the proposed PSO-CNN algorithm. The PSO-CNN model is simulated using a benchmark DR database, and the experimental results show that the PSO-CNN model outperforms all other approaches by a substantial margin [39] (Table 1).

3. Methods and Materials

3.1. Fundus Photography. According to Singh et al. [5], fundus photography is a technique for photographing the back of the eye, which includes the retina, optic disc, and macula. Trained medical specialists can use fundus photography to observe and analyze the severity of the condition. For example, fundography pictures from DR-disease-selected data are shown in Figure 1.

3.2. Convolutional Neural Network. In this part, we will look into a convolutional neural network (CNN) model. One of the learning networks inspired by the perceptron neural net can be considered a neural network. This dynamic network has three layers: an input layer, an output layer, and a highly concealed layer. First, the problem's image or data are detected and put into the algorithm. The output layer's hidden weights would subsequently manifest themselves in a variety of ways [55–59]. A classification or recognition technique is used when the output has many numerical components, such as a binary number or index (e.g., image classification, normal = 1, and abnormal = 2). The findings are weighted after several photos have been trained. When a fresh picture, different than the training pictures, is introduced to the algorithm, the shape of the images is recognized. For example, if we train the computer with a matrix of diverse images, such as pictures of benign

Table 1: Literature review of DR images segmentation.

Methods	Strength	Weakness
Vessel segmentation using thresholding [40–42]	Approximation of vessel pixels using a simple approach	When the vessel pixel values are closer to the background, false points are recognized
Fuzzy-based segmentation [43]	With consistent pixel values, it works great	To increase the responsiveness of blood vessels, extensive preprocessing is necessary
Active contours [44]	For detection of real boundaries, a better approximation is used	Processes that are iterative and time-consuming are necessary
Vessel tubular properties-based method [16]	Approximation of vessel-like structures is excellent	Pixel discontinuities put restrictions on how far you can go
Line detection-based method [45]	The removal of the backdrop aids in the reduction of fake skin-like pixels	
Random forest classifier-based method [46]	To identify pixels, use a lighter technique	Before categorization, many transformations are required to produce features
Patch-based CNN [47]	Better categorization	Training and testing take a long time to complete
SVM-based method [48]	Training time is reduced	To create a feature vector, preprocessing approaches involving many photos are used
Extreme machine learning [49]	Machine learning has a lot of differentiating factors	To create distinguishable features, morphology and other traditional procedures are required
Mahalanobis distance classifier [49]	Training is a simple operation	To compute relevant features, preprocessing overhead is still necessary
U-Net-based CNN for semantic segmentation [19]	The boundaries are nicely preserved by the U-Net construction	Preprocessing in gray scale is necessary
Multiscale CNN [20]	Multireceptive fields improve learning	In other circumstances, tiny vessels are not recognized
CNN with CRFs [20]	Faster segmentation is achieved using a CNN with a few layers	CRFs are difficult to compute
SegNet-inspired method [22]	The design of encoders and decoders creates a unified network layer topology	PCA is used to prepare data for training purposes
CNN with visual codebook [50]	For correlation with ground truth representation, a 10-layer CNN is used	There is no end-to-end training and testing mechanism
CNN with quantization and pruning [51]	Convolutions that have been pruned improve the network's efficiency	The number of trainable parameters increase as layers are fully coupled
Three-stage CNN-based deep learning method [52]	A compelling representation is provided by the fusion of multifeature images	The use of three CNNs necessitates a higher level of computing power and expense
Modified U-Net with dice loss [53]	Dice loss provides good results with unbalanced classes	PCA is used to prepare data for training purposes
Deformable U-Net-based method [23]	When classes are unequal, dice loss delivers acceptable results	Patch-based training and testing take a lot of time
PixelBNN [24]	Pixel is a term that refers to CNN is well-known for its ability to forecast pixels with spatial dimensions	CLAHE is used for preprocessing
Dense U-Net-based method [25]	The use of a dense block can help solve the problem of disappearing gradients	Patch-based training and testing take a lot of time
Cross-connected CNN (CcNet) [26]	Layer cross-connections provide features more power	Preprocessing is used to create a complex architecture
Vess-Net [54]	With fewer layers, robust segmentation is possible	To properly train the network, augmented data is required

or malignant cancers, Alzheimer's, sarcoma, or brain tumors. Depending on the weights acquired, the approach indicates the type of disease. The CNN's base is the convolutional sublayer, and its output matrix is a three-dimensional neuron matrix. For a better understanding, typical neural networks are considered. In typical neural networks, each layer was nothing more than a list (one-dimensional as a rectangle) of neurons, each of which generated its own output and gradually amassed a sequence of outputs corresponding to each neuron. Instead of a single number, we are presented with a three-dimensional list (one cube) in which the neurons in the CNN are organized in three dimensions. As a consequence, this cube's output is a three-dimensional matrix as well [60, 61].

In traditional architecture, placing a pooling layer between multiple consecutive levels is a common approach. This layer seeks to minimize the number of variables and measurements in the grid, resulting in a reduced matrix (input) size by overfitting the display (width and height). The pooling layer works on each depth cut of the input matrix individually. The max-pooling option expands or contracts the size of the position. The activation process of artificial neural networks determines the node's output node or "neuron" based on the input or group of inputs. In the following node, this output is known as the input [62]. This occurs before a solution to the problem is identified. The results are transformed into a goal range, such as 0 to 1 or −1 to 1 (depending on the activation mechanism used). The

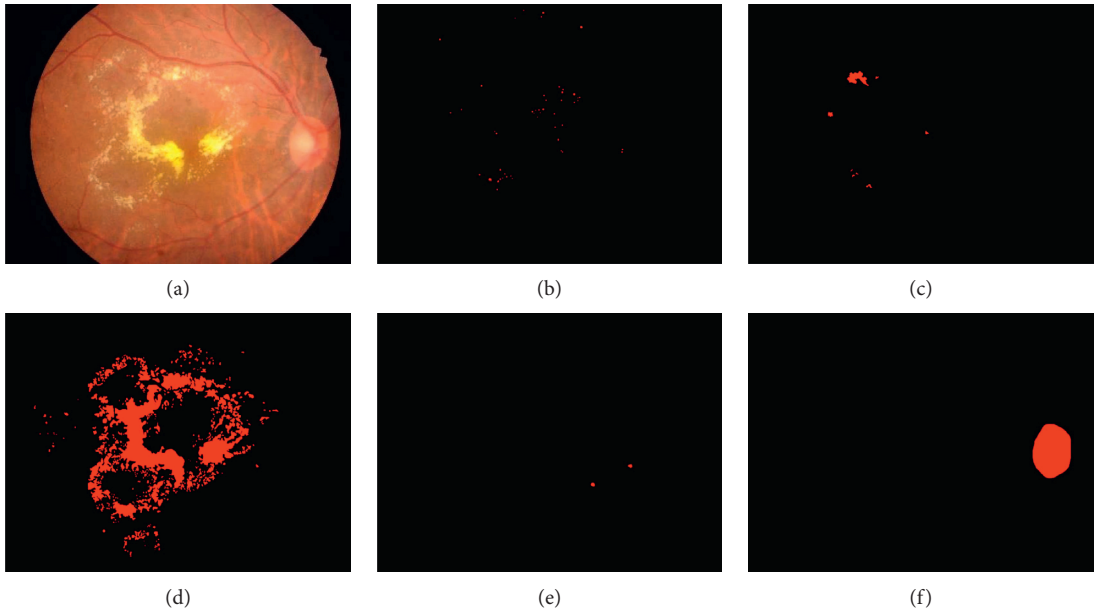


Figure 1: Different classes of disorders in DR disease [24]: (a) original image, (b) microaneurysms, (c) haemorrhages, (d) hard exudates, (e) soft exudates, and (f) optic disc.

logistic activation function, for example, transforms all inputs to true absolute ranges of 0 to 1. Optimizing your weight is another strategy to reduce weight. The rectified linear unit (ReLU) was used in this paper for the following functions [63]:

$$f(x) = \begin{cases} 0 & x < 0, \\ x & x \geq 0. \end{cases} \tag{1}$$

Some activation functions, such as Softmax [49], are not limited to a single variable and may be applied to a vector or several variables employed in this paper:

$$f_i(\vec{x}) = \frac{e^{x_i}}{\sum_{j=1}^{J} e^{x_j}}. \tag{2}$$

A batch normalization layer is introduced to the network to normalize input data, speed up the training phase, and minimize network sensitivity between convolutional layers and nonlinearities. To create an aberrant image augmentation, a dropout layer on the fully connected layers is frequently utilized. We employed fully related layers at the end of the concealed layer, demonstrating to differentiate pictures. The output of the deep learning layer is a fully connected layer that guides the classification choice.

3.3. Receiver Operating Characteristic (ROC) Curve. The ROC curve is constructed by comparing the true-positive rate (TPR) to the false-positive rate (FPR) under various threshold conditions. The TPR is sometimes referred to as adaptability, recall, or likelihood of detection in machine learning. Starting on the left side of the ROC, we can observe that the FPR and TPR are both 0. (This suggests that the threshold line, which indicates the most test results, is long.) It is better to begin with the most test results and work your way up from there. The measurable and descriptive consistency of the outcomes of a measure separates knowledge into these two groups. The data may be split into positive and negative classes using sensitivity and attribute tests. The number of positive samples that can be adequately tested as positive is referred to as sensitivity. The number of negative examples that are correctly categorized as negative (positive = particular illness and negative = other cases) are referred to as specificity. The sensitivity of splitting the percentages of TP items into the number of true-positive and false-negative examples in mathematical terms [55, 64] is given as follows:

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{3}$$

The confusion matrix is the function of the approaches outlined in the field of artificial intelligence. This type of presentation is common in supervised learning algorithms, but it is also common in unsupervised learning. The expected value is shown in each column of the matrix. Assume that each row includes a valid (true) example. The name of this matrix is also given, allowing for mistakes and tampering with the result. This matrix is also characterized as a contingency matrix or an error matrix outside of artificial intelligence.

A machine learning classification model may be employed to predict the data point's actual class or its chance of belonging to multiple classes. The latter allows us to have more influence over the outcome. To understand the classifier's outcome, we may choose our own threshold. Depending on whether we want to reduce the number of false negatives or false positives, one of these criteria will almost certainly produce a better outcome than the others. The measurements alter when the threshold values vary. We may create a variety of confusion matrices and compare the

results. It, however, would not be a wise decision. Alternatively, we can create a plot between some of these indicators so that we can quickly see which threshold is producing the best results.

The receiver operator characteristic (ROC) curve is a binary classification issue assessment measure. It is a probability curve that compares the TPR with the FPR at various threshold values, successfully distinguishing the signal from the noise. The area under the curve (AUC) is a summary of the ROC curve that measures a classifier's capability to discriminate between classes. The AUC indicates how well the model distinguishes between positive and negative classes. The higher the AUC number, the better. In a ROC curve, a higher X-axis value indicates that there are more false positives than true negatives. A higher Y-axis value indicates that there are more true positives than false negatives, whereas a lower Y-axis value indicates that there are fewer true positives. As a result, the threshold selection is based on the capacity to strike a balance between false positives and false negatives (Figure 2).

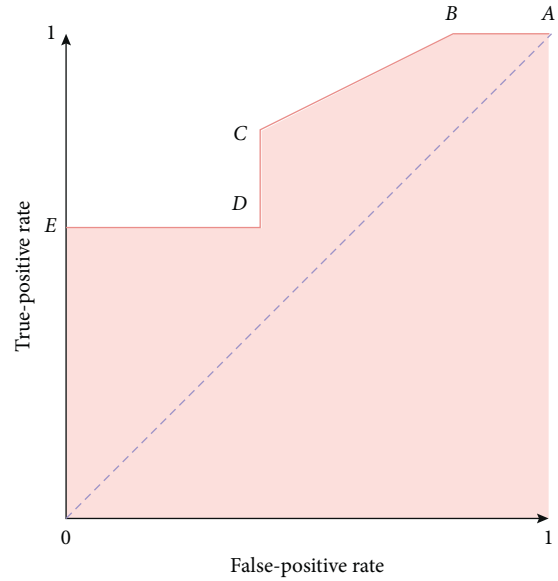


Figure 2: The ROC curve sample.

4. Results and Discussion

4.1. Data Collection. IDRiD (Indian Diabetic Retinopathy Image Dataset), the database used in this research, is the first of its kind to be representative of an Indian population. Furthermore, at the pixel level, this is the only data set that comprises both typical diabetic retinopathic lesions and normal retinal structures [61]. For each picture, this data set contains information on the severity of diabetic retinopathy and diabetic macular degeneration. This is appropriate for the creation and testing of image analysis algorithms for diabetic retinopathy early detection.

4.2. Blood Vessel Extraction Process. In this part, the suggested CNN for automated detection of DR was discussed. Figure 1 depicts the workflow of our suggested approach. First, the green channel preprocesses the input fundus retinal picture, and then CLAHE is utilized to improve the output from the green channel. The blood vessels are then extracted from the retinal fundus picture using multiple morphological techniques and binarization thresholding. Finally, the suggested CNN, which was constructed in MATLAB, is used to evaluate the segmented blood vessel in order to determine the severity of the DR automatically. Preprocessing methods are used to enhance the original image in order to increase the success rate of the planned task. Preprocessing procedures are used before segmenting blood vessels to ensure that the properties of the original picture are not altered. A preprocessing consists of five processes to eliminate these defects and provide more acceptable pictures for extracting the features: (1) green channel separation, (2) augmentation, (3) morphological procedures, (4) background subtraction, and (5) thresholding. Also, the histogram equalization approach generally improves the overall contrast of many pictures, primarily when a small number of intensity values represent the image. The intensities can also be better dispersed on the histogram with this modification, employing the whole range of intensities equitably. It enables locations with poor local contrast to get a boost in contrast. It is accomplished by histogram equalization, which effectively spreads out the densely packed intensity values that decrease picture contrast. Figure 3 depicts a flow diagram for the segmentation of blood vessels procedure. The approaches are utilized to function well for the retinal pictures and are based on a trial-and-error method process. The following is a quick description of each step:

Step 1: Green channel separation

Because of red and blue channels of an RGB image suffer from noise and poor visual quality, the green channel is used to segment blood vessels. To increase the contrast of retinal images, the red and blue elements of the image are eliminated before processing. In the preprocessing step, the green channel is used to offer the best vasculature and contrast between the optic disc and the retinal tissue. As a consequence, the input image's green pixel values are extracted and saved as a vector.

Step 2: Contrast enhancement

Contrast-limited histogram equalization is a contrast enhancement technique that separates a big region into a number of little parts of similar size and works on each one separately to improve contrast. On the green channel, CLAHE is employed to help collect crucial blood vessel data. This method improves the quality of the image dynamically. It means that each part of the image is enhanced separately that causes no saturation in colors being observed. In Figure 4, the CLAHE is used to a fundus image.

Step 3: Morphological processing

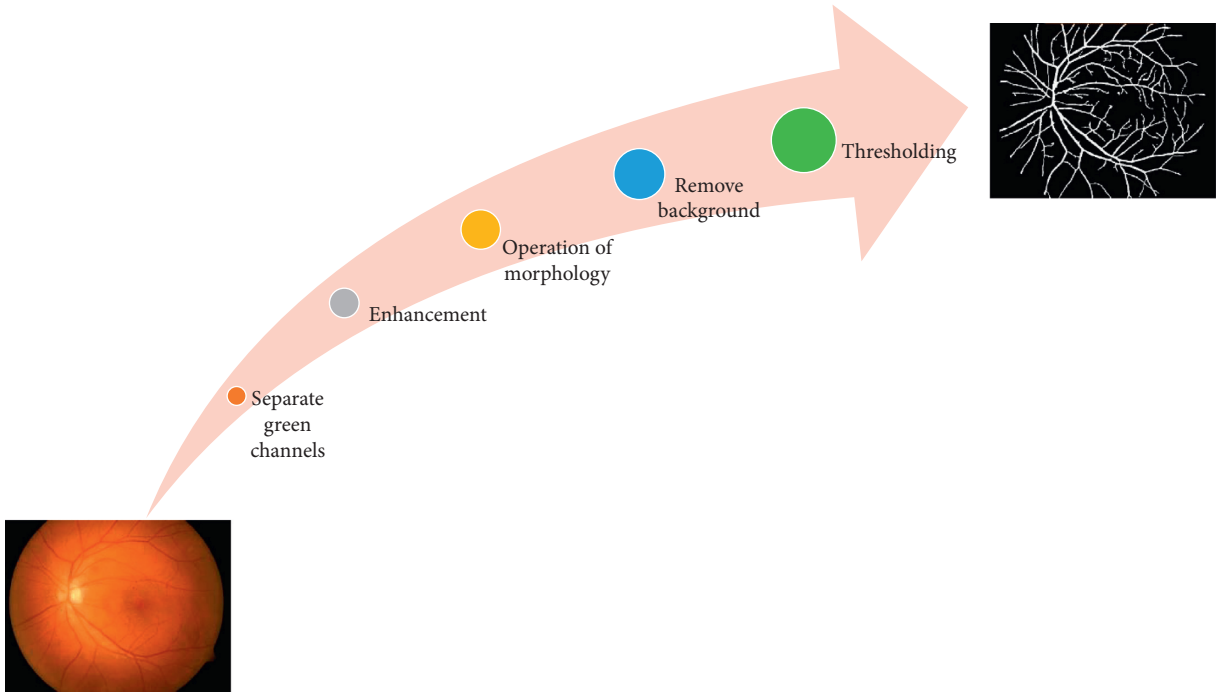


Figure 3: The process of blood vessel extraction from fundus images.

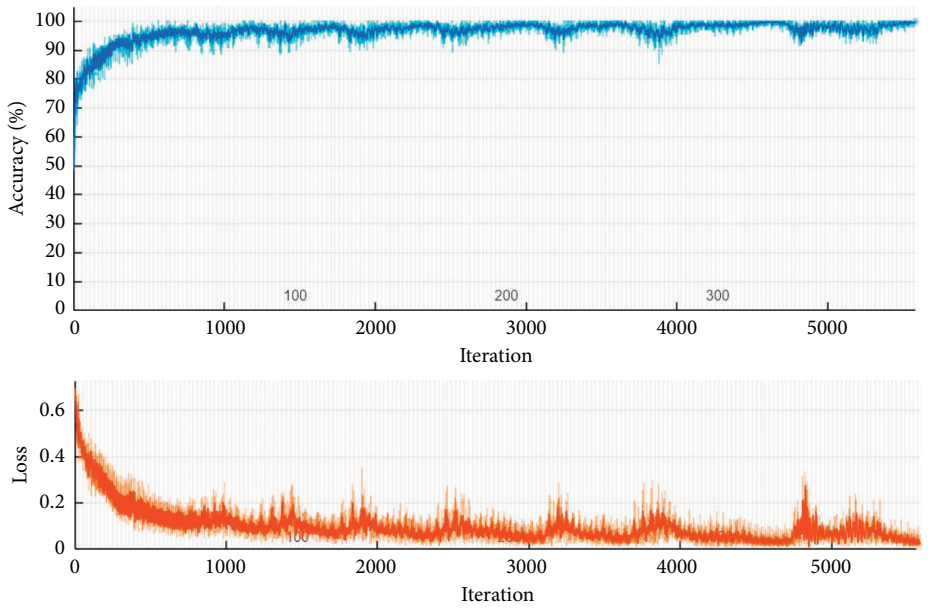


Figure 4: The training process of the CNN approach.

The basic procedures of erosion and dilation are employed in mathematical morphology to investigate the relationship between a picture and a certain structuring element (SE). Dilation, erosion, opening, and closure are the most common procedures utilized here. This can be implemented by convolving (dot product) two sets of pixels, one of which contains the SE or kernel and the other of which contains the picture to be processed. Opening and closing are the two fundamental operators. Furthermore, dilation and erosion enlarge and reduce the size of an object inside the image, respectively. The SE, which is a matrix of just 0s and 1s with arbitrary size and form employed to examine the input picture, is a key element of these two procedures. The morphological opening and closing procedures are used to smooth the vessel's edge and the surrounding area, remove tiny holes, and fill in contour gaps. Algorithms combine the foregoing methods in the proposed work to recognize edges, eliminate noise and background, and discover specific forms in photos.

There is not any problem when the input image has a different size, since the kernel size can be defined based on the size of aiming objects.

Step 4: Background subtraction

This strategy can be used to reduce anomalies in the backdrop of an image so that the foreground elements may be inspected more easily. The recommended method removes the backdrop by deleting some defined objects from the contrast-enhanced image.

Step 5: Thresholding

The practice of eliminating extraneous information from an image is known as thresholding. By eliminating all gray-level information from the fundus photographs, the blood vessels are converted to binary pixels. It is critical to distinguish between foreground blood vessels and background data. Thresholding is utilized to bring out hidden details as a result, selecting the appropriate threshold value is crucial, as a low number may reduce the size or amount of these items, whilst a high number may include extra background data. It is used to make a binary image with 1 (blood vessel) or 0 (background) pixel values so that numerical data may be read and fed into ML techniques.

4.3. Results of Segmentation Using the CNN Method. The findings of the CNN segmentation method are presented in this section. The detection method's architecture is depicted in Figure 5. It has 11 layers, including three convolutional layers. The input pictures are 256×256 grayscale fundus pictures taken from the eyes of DR patients, and the output layer includes ground truth from the IDRiD data set. The target tissue of the patient's eye is identified with 255 in these photos, whereas other spots are represented with zero values.

Colored pictures in the input picture in Figure 5 are converted to a dark-gray zone. It is difficult to recognize this area with the aid of a brain-computer. Given the possibility of deep learning approaches for segmenting regions with a broader range of hues than photographs with a similar color palette, more diverse colors have a better potential for segmentation. The target location is first discovered and labeled by medication or an automated method in this research. As a result, the target region's ground truth pictures are stored in the output layer of the proposed architecture. The presented architecture consists of 11 sublayers as shown in Table 2.

The segmentation results are depicted in Figure 6. The input picture of the DR eyes is shown in the first column of Figure 6. The second row, on the other hand, displays the ground truth image of the output layer. Seventy percent of pictures are utilized to train the network, and 30% are utilized to test the findings to begin the process. In the third column of Figure 6, the results of target region detection are displayed. We can see that the observations and the result are nearly identical. To improve the outcomes, all small regions from the findings are removed and replaced with a probability contour (fourth column). According to the findings, the suggested architecture is capable of detecting the target region with near-perfect accuracy. Figure 4 illustrates the training method.

Theoretically, the outcomes of segmentation algorithms are supplied with performance criteria. The ROC curve depicts the true-positive rate versus the false-positive rate. In the segmentation process, each image has its own set of criteria. Whether plots are shown with a higher true-positive rate and a smaller false-positive rate, data show that most photographs of ROC curves have an almost high performance (Figure 7(a)). To better understand the ROC curve using realistic figures, we provide the area under the curve (AUC) value. This criterion assesses whether values in the vicinity of one indicate high performance or not. In Figure 7(b), 92% of the images are in the high-performance range. This is something that can also be deduced from Figure 7(b). More than 90% of photographs are segmented accurately, with a 70% accuracy rate. At the completion of the process, the average amount of accuracy for all photographs is 83.84% (Figure 7(c)). The last requirement is the Jaccard value (Figure 7(d)), which depicts the overlapping regions of the generated image and ground facts. This number has ranged around 0.4 for many of the photos that have been produced. This lower score is due to the presence of black spots in the final photos when compared to ground truth data.

Table 3 compares the results of several techniques presented in the literature to standard incoherence evaluations. This approach achieves the best AUC and accuracy in terms of numbers but the minor sensitivity. This suggested approach obtains a mean sensitivity of 0.9% and accuracy of 83.8% in its overall performance.

5. Discussion

The most frequent way for ophthalmologists to diagnose DR is through a dilated eye examination. Fluorescein angiography, optical coherence tomography (OCT), and fundus photography are more ways to diagnose illness. The blood circulation and arterial anomalies are imaged during fluorescein angiography after an intravenous infusion of contrast dye. OCT is used to assess retinal anatomy, size, and edema (i.e., retinal swelling). Generally, DR diagnosis is arbitrary and requires a retina expert who has completed advanced training in diagnosis and grading. Visual evaluation and manual measurements of changes in retinal vasculature and layers are considered complicated duties. Regrettably, due to the lack of qualified eye-care experts and tertiary eye-care facilities, many diabetes patients seek to see a retina specialist only after they have symptomatic vision loss when their disease has progressed and is mainly permanent. As a result, an essential clinical motive is developing an objective and noninvasive diagnostic method to identify and evaluate DR properly at a preliminary phase.

CNNs have recently been used to diagnose diabetic retinopathy (DR) by evaluating fundus pictures, and their efficiency in segmentation and localization tasks has been demonstrated. DR is a severe consequence of diabetes that

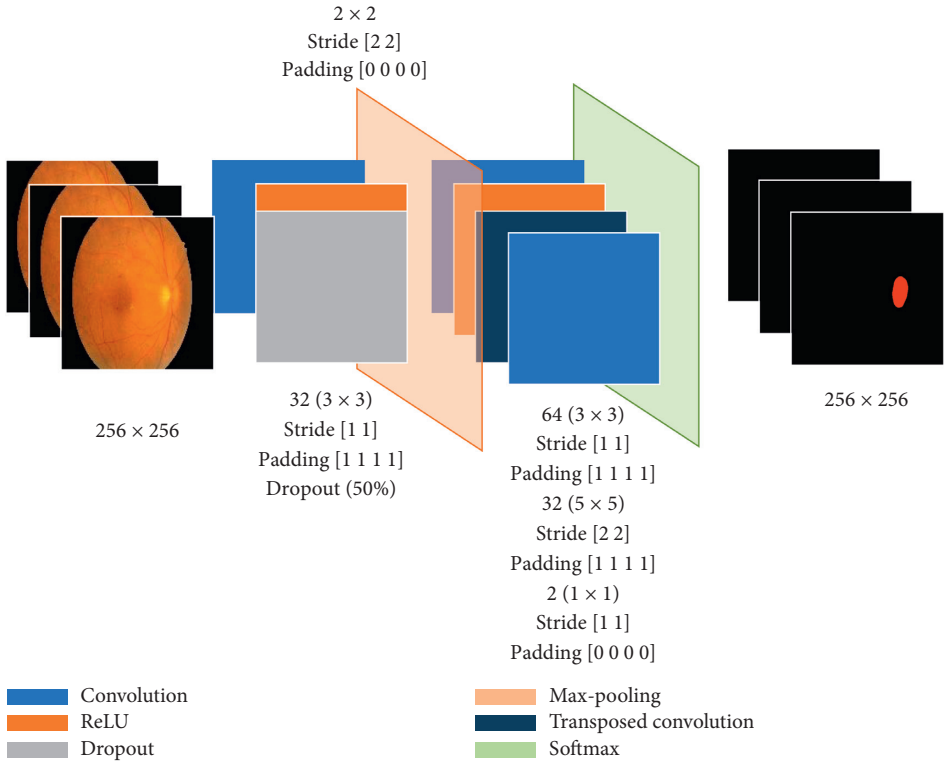


Figure 5: The architecture of the CNN method for segmentation of fundus images.

Table 2: The architecture of presented CNN layers.

No.	Layer name	Properties
1	Image input	256 × 256 × 1 images with "zerocenter" normalization
2	Convolution	32 (3 × 3 × 1) convolutions with stride [1 1] and padding [1 1 1 1]
3	ReLU	
4	Dropout	50% dropout
5	Max pooling	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
6	Convolution	64 (3 × 3 × 32) convolutions with stride [1 1] and padding [1 1 1 1]
7	ReLU	
8	Transposed convolution	32 (4 × 4 × 64) transposed convolutions with stride [2 2] and cropping [1 1 1 1]
9	Convolution	2 (1 × 1 × 32) convolutions with stride [1 1] and padding [0 0 0 0]
10	Softmax	
11	Pixel classification	Dice loss

can lead to vision loss and possibly blindness. Such models provide good classification performance for the items in the training data set, but their use is limited to specialized areas such as DR detection. A broad range of complicated characteristics and their localizations within the picture is used to diagnose abnormal indications in fundoscopy. Each layer of CNN creates a new input image size by gradually extracting the most distinguishing features. Several approaches for improving accuracies, such as dimension reduction and feature augmentation, were presented in the state-of-the-art methods. Nonetheless, deep-learning-based DR detection research regularly indicates good performance in severe cases, whereas mild case detection remains difficult. Because of the probable absence of the early phase of DR, this restriction impedes the broader use of utterly automated mass-screening, perhaps leading to more complex condition future development.

A number of the study's flaws have been discovered. First, only small-to-moderate data set sizes were employed in the investigation because of the scarcity of DR pictures. To increase the data set, the necessary data augmentation techniques were used, such as rotation, horizontal/vertical flipping, and so on. Furthermore, when training the classifiers, the default hyperparameters were used. It was thought to be the most acceptable practice in the area. Nonetheless, tests with different optimizers were carried out. Lastly, the "black-box" aspect of deep-learning-based solutions is widely criticized, resulting in operator opposition to broader use.

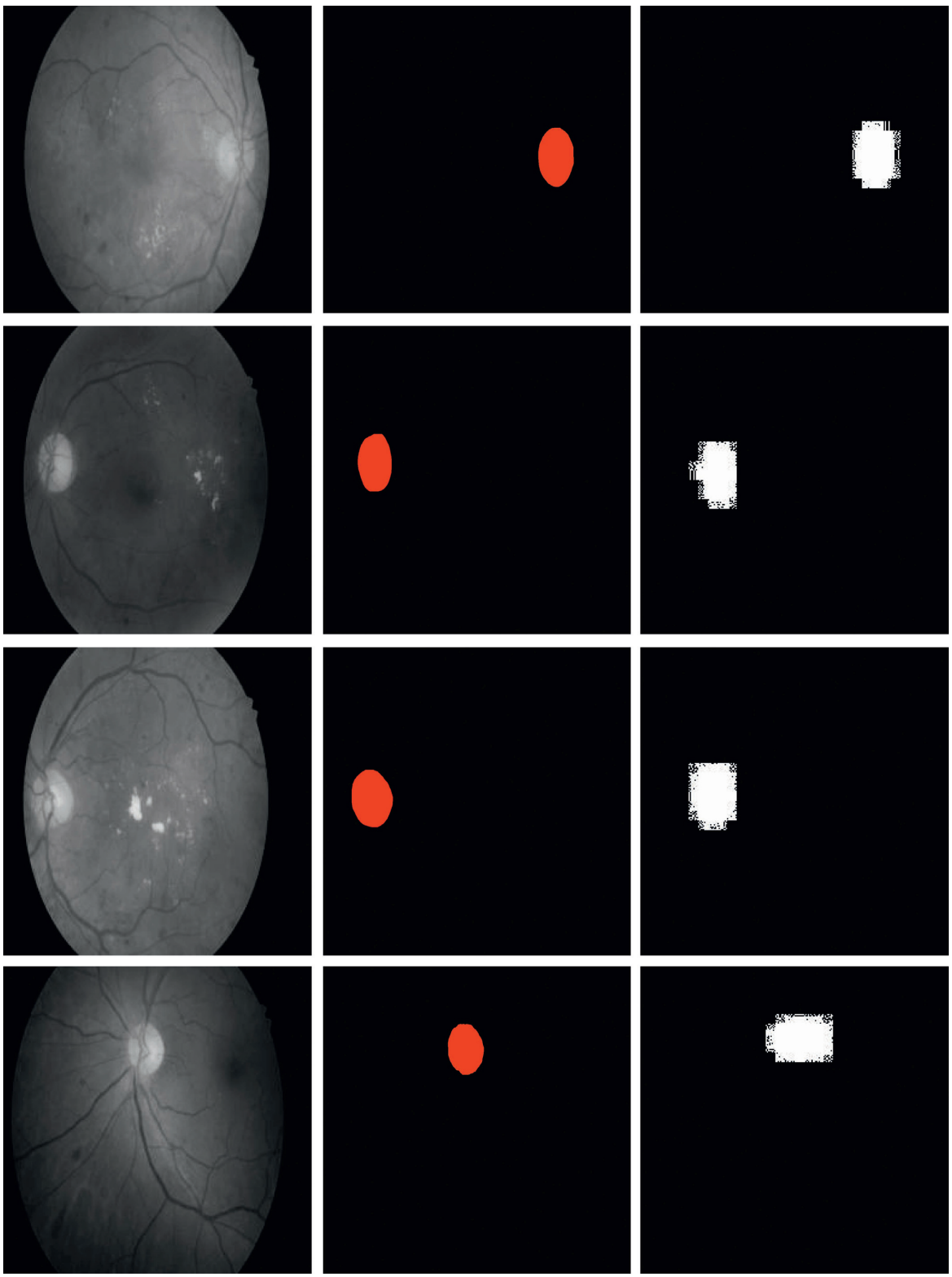


Figure 6: Examples of segmentation of fundus images of DR patients.

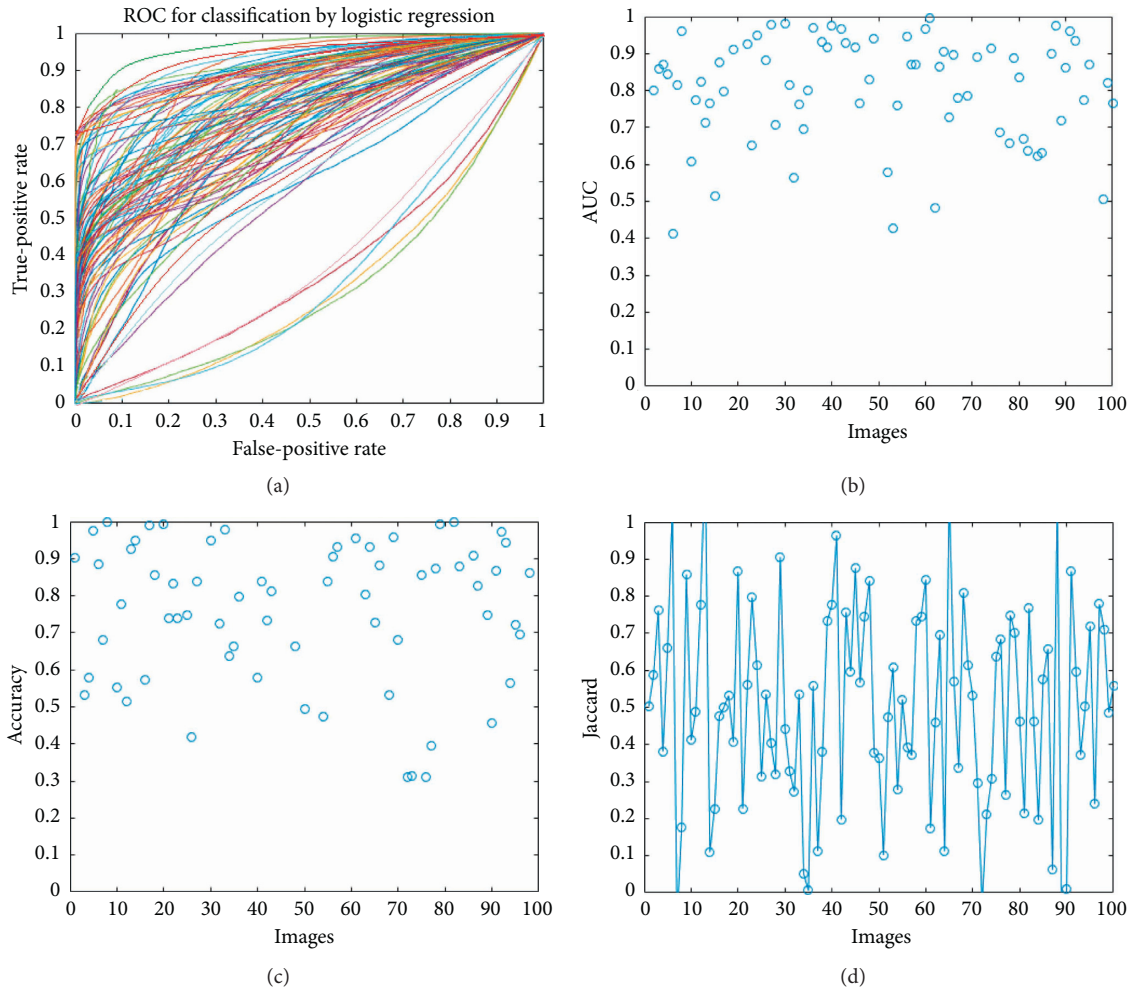


Figure 7: Results of segmentation: (a) ROC cure, (b) AUC criteria, (c) accuracy of segmentation, and (d) Jaccard criteria.

Table 3: Comparison between presented methods and the state-of-the-art methods.

Authors	Method	Sensitivity	Accuracy
Li et al. [65]	Cross-modality method	0.757	0.953
Christodoulidis et al. [66]	Multiscale tensor voting	0.851	0.948
Aslani and Sarnel [67]	Multiscale Gabor wavelet	0.755	0.951
Vega et al. [68]	Lattice neural networks	0.744	0.941
Jebaseeli et al. [69]	Average value	0.806	0.995
Presented method	CNN	0.891	0.838

6. Conclusion

The most prevalent cause of vision loss in diabetic patients is diabetic retinopathy. Other retinal and nonretinal visual issues, such as glaucoma, age-related macular degeneration, and neuropathy, can also cause visual deterioration or loss. Nonarterial optic ischemic neuropathy (NAION) and cataracts are examples of vascular vessels. Refractive errors, contrast sensitivity, direct light, and compliance amplitude should all be evaluated when a diabetic complains of visual problems. Physicians who treat diabetic patients utilize these vision issues to guarantee quick referral and treatment to prevent vision impairment, which can have a substantial impact on everyday living, especially for people who drive. Developing approaches that do not require the assistance of a doctor, even with smart technology such as phones, is critical. These procedures can also aid doctors in making an accurate and timely diagnosis of the condition. The investigation continues in the next section of the paper with the presentation of a segmentation approach for detecting target regions in human eyes. We devised a CNN architecture for pixel categorization of patients' eye fundus pictures in order to achieve our aim. For the output layer, we utilized preprocessed ground truth images. Seventy percent of the photos were utilized for training data, with the remaining photos being utilized to validate the provided architecture. The provided approach can practically recognize the target region of eye pictures with high accuracy of 83.84%,

according to the results. This network may also be used to treat patients better by separating the target region from their eye pictures. It is past time for artificial intelligence to be used in medicine to assist doctors in making better and faster diagnoses.

Data Availability

The data set IDRiD (Indian Diabetic Retinopathy Image Dataset) is available online at https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] M. N. Ozieh, K. G. Bishu, C. E. Dismuke, and L. E. Egede, "Trends in health care expenditure in U.S. Adults with diabetes: 2002-2011," *Diabetes Care*, vol. 38, no. 10, pp. 1844–1851, 2015.

[2] L. Sellahewa, C. Simpson, P. Maharajan, J. Duffy, and I. Idris, "Grader agreement, and sensitivity and specificity of digital photography in a community optometry-based diabetic eye screening program," *Clinical Ophthalmology (Auckland, N.Z.)*, vol. 8, pp. 1345–1349, 2014.

[3] P. Ruamviboonsuk, N. Wongcumchang, P. Surawongsin, E Panyawatananukul, and M Tiensuwan, "Screening for diabetic retinopathy in rural area using single-field, digital fundus images," *Journal of the Medical Association of Thailand*, vol. 88, pp. 176–180, 2005.

[4] L. Guariguata, D. R. Whiting, I. Hambleton, J. Beagley, U. Linnenkamp, and J. E. Shaw, "Global estimates of diabetes prevalence for 2013 and projections for 2035," *Diabetes Research and Clinical Practice*, vol. 103, no. 2, pp. 137–149, 2014.

[5] M. R. K. Mookiah, U. R. Acharya, C. K. Chua, C. M. Lim, E. Y. K. Ng, and A. Laude, "Computer-aided diagnosis of diabetic retinopathy: a review," *Computers in Biology and Medicine*, vol. 43, no. 12, pp. 2136–2155, 2013.

[6] Z. Li, J. Jiang, H. Zhou et al., "Development of a deep learning-based image eligibility verification system for detecting and filtering out ineligible fundus images: a multicentre study," *International Journal of Medical Informatics*, vol. 147, Article ID 104363, 2021.

[7] R. J. Winder, P. J. Morrow, I. N. McRitchie, J. R. Bailie, and P. M. Hart, "Algorithms for digital image processing in diabetic retinopathy," *Computerized Medical Imaging and Graphics*, vol. 33, no. 8, pp. 608–622, 2009.

[8] D. Sidibé, I. Sadek, and F. Mériaudeau, "Discrimination of retinal images containing bright lesions using sparse coded features and SVM," *Computers in Biology and Medicine*, vol. 62, pp. 175–184, 2015.

[9] S. Kulkarni, N. Seneviratne, M. S. Baig, and A. H. A. Khan, "Artificial intelligence in medicine: where are we now?" *Academic Radiology*, vol. 27, no. 1, pp. 62–70, 2020.

[10] Y. Mori, H. Neumann, M. Misawa, S. E. Kudo, and M. Bretthauer, "Artificial intelligence in colonoscopy: now on the market. What's next?" *Journal of Gastroenterology and Hepatology*, vol. 36, 2021.

[11] H. Endo, S. Kase, M. Saito et al., "Choroidal thickness in diabetic patients without diabetic retinopathy: a meta-analysis," *American Journal of Ophthalmology*, vol. 218, pp. 68–77, 2020.

[12] S. Roy and D. Kim, "Retinal capillary basement membrane thickening: role in the pathogenesis of diabetic retinopathy," *Progress in Retinal and Eye Research*, vol. 82, Article ID 100903, 2020.

[13] L. Câmara Neto, G. L. B. Ramalho, J. F. S. Rocha Neto, R. M. S. Veras, and F. N. S. Medeiros, "An unsupervised coarse-to-fine algorithm for blood vessel segmentation in fundus images," *Expert Systems with Applications*, vol. 78, pp. 182–192, 2017.

[14] R. Sundaram, R. KS, P. Jayaraman, and V. B, "Extraction of blood vessels in fundus images of retina through hybrid segmentation approach," *Mathematics*, vol. 7, no. 2, p. 169, 2019.

[15] Y. Zhao, L. Rada, K. Chen, S. P. Harding, and Y. Zheng, "Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images," *IEEE Transactions on Medical Imaging*, vol. 34, no. 9, pp. 1797–1807, 2015.

[16] L. C. Rodrigues and M. Marengoni, "Segmentation of optic disc and blood vessels in retinal images using wavelets, mathematical morphology and Hessian-based multi-scale filtering," *Biomedical Signal Processing and Control*, vol. 36, pp. 39–49, 2017.

[17] Ç. Sazak, C. J. Nelson, and B. Obara, "The multiscale bowler-hat transform for blood vessel enhancement in retinal images," *Pattern Recognition*, vol. 88, pp. 739–750, 2019.

[18] B. Savelli, A. Bria, A. Galdran et al., "Illumination correction by dehazing for retinal vessel segmentation," in *Proceedings of the 30th International Symposium on Computer-Based Medical Systems*, pp. 219–224, Thessaloniki, Greece, June 2017.

[19] F. Girard, C. Kavalec, and F. Cheriet, "Joint segmentation and classification of retinal arteries/veins from fundus images," *Artificial Intelligence in Medicine*, vol. 94, pp. 96–109, 2019.

[20] K. Hu, Z. Zhang, X. Niu et al., "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function," *Neurocomputing*, vol. 309, pp. 179–191, 2018.

[21] H. Fu, Y. Xu, S. Lin, D. W. Kee Wong, and J. Liu, "DeepVessel: retinal vessel segmentation via deep learning and conditional random field," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 132–139, Athens, Greece, October 2016.

[22] T. A. Soomro, A. J. Afifi, J. Gao et al., "Boosting sensitivity of a retinal vessel segmentation algorithm with convolutional neural network," in *Proceedings of the International Conference on Digital Image Computing: Techniques and Applications*, pp. 1–8, Sydney, Australia, November 2017.

[23] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, "DUNet: a deformable network for retinal vessel segmentation," *Knowledge-Based Systems*, vol. 178, pp. 149–162, 2019.

[24] H. Leopold, J. Orchard, J. Zelek, and V. Lakshminarayanan, "PixelBNN: augmenting the PixelCNN with batch normalization and the presentation of a fast architecture for retinal vessel segmentation," *Journal of Imaging*, vol. 5, no. 2, p. 26, 2019.

[25] C. Wang, Z. Zhao, Q. Ren, Y. Xu, and Y. Yu, "Dense U-net based on patch-based learning for retinal vessel segmentation," *Entropy*, vol. 21, no. 2, p. 168, 2019.

[26] S. Feng, Z. Zhuo, D. Pan, and Q. Tian, "CcNet: a cross-connected convolutional network for segmenting retinal vessels using multi-scale features," *Neurocomputing*, vol. 392, 2019.

[27] S. Das, K. Kharbanda, S. M, and R. Raman, "Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy," *Biomedical Signal Processing and Control*, vol. 68, Article ID 102600, 2021.

[28] A. Shanthini, G. Manogaran, G. Vadivu, K. Kottilingam, P. Nithyakani, and C. Fancy, "Threshold segmentation based multi-layer analysis for detecting diabetic retinopathy using convolution neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 136, pp. 1–15, 2021.

[29] M. K. Hasan, M. A. Alam, M. T. E. Elahi, S. Roy, and R. Martí, "DRNet: segmentation and localization of optic disc and Fovea from diabetic retinopathy image," *Artificial Intelligence in Medicine*, vol. 111, Article ID 102001, 2021.

[30] S. J. Ghoushchi, R. Ranjbarzadeh, A. H. Dadkhah, Y. Pourasad, and M. Bendechache, "An extended approach to predict retinopathy in diabetic patients using the genetic algorithm and fuzzy C-means," *BioMed Research International*, vol. 2021, Article ID 5597222, 13 pages, 2021.

[31] A. Aghamohammadi, R. Ranjbarzadeh, F. Naiemi, M. Mogharrebi, S. Dorosti, and M. Bendechache, "TPCNN: two-path convolutional neural network for tumor and liver segmentation in CT images using a novel encoding approach," *Expert Systems with Applications*, vol. 183, Article ID 115406, 2021.

[32] R. Ranjbarzadeh, S. Jafarzadeh Ghoushchi, M. Bendechache et al., "Lung infection segmentation for COVID-19 Pneumonia based on a cascade convolutional network from CT images," *BioMed Research International*, vol. 2021, Article ID 5544742, 16 pages, 2021.

[33] R. Ranjbarzadeh, S. B. Saadi, and A. Amirabadi, "LNPSS: SAR image despeckling based on local and non-local features using patch shape selection and edges linking," *Measurement*, vol. 164, Article ID 107989, 2020.

[34] R. Ranjbarzadeh and S. Baseri Saadi, "Corrigendum to "Automated liver and tumor segmentation based on concave and convex points using fuzzy c-means and mean shift clustering," *Measurement*, vol. 151, Article ID 107230, 2020.

[35] R. Ranjbarzadeh and S. B. Saadi, "Automated liver and tumor segmentation based on concave and convex points using fuzzy c-means and mean shift clustering," *Measurement*, vol. 150, Article ID 107086, 2020.

[36] A. Hamzenejad, S. Jafarzadeh Ghoushchi, V. Baradaran, and A. Mardani, "A robust algorithm for classification and diagnosis of brain disease using local linear approximation and generalized autoregressive conditional heteroscedasticity model," *Mathematics*, vol. 8, no. 8, p. 1268, 2020.

[37] M. Shirmohammadi, S. J. Goushchi, and P. M. Keshtiban, "Optimization of 3D printing process parameters to minimize surface roughness with hybrid artificial neural network model and particle swarm algorithm," *Progress in Additive Manufacturing*, vol. 6, no. 2, pp. 199–215, 2021.

[38] A. Sharifi, M. Ahmadi, M. A. Mehni, S. Jafarzadeh Ghoushchi, and Y. Pourasad, "Experimental and numerical diagnosis of fatigue foot using convolutional neural network," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 14, pp. 1–13, 2021.

[39] J. Jayanthi, T. Jayasankar, N. Krishnaraj, N. B. Prakash, A. Sagai Francis Britto, and K. Vinoth Kumar, "An intelligent particle swarm optimization with convolutional neural network for diabetic retinopathy classification model," *Journal of Medical Imaging and Health Informatics*, vol. 11, no. 3, pp. 803–809, 2021.

[40] R. J. Chalakkal and W. H. Abdulla, "Improved vessel segmentation using curvelet transform and line operators," in *Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, pp. 2041–2046, Honolulu, HI, USA, November 2018.

[41] F. F. Wahid and G. Raju, "A dual step strategy for retinal thin vessel enhancement/extraction," in *Proceedings of the Amity International Conference on Artificial Intelligence*, pp. 666–671, Dubai, UAE, February 2019.

[42] A. T. U. Akyas, J. A. Abhin, J. G. Geo, and K. B. V. Santhosh, "Automated system for retinal vessel segmentation," in *Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies*, pp. 717–722, Coimbatore, India, April 2018.

[43] S. S. Kar and S. P. Maity, "Blood vessel extraction and optic disc removal using curvelet transform and kernel fuzzy c-means," *Computers in Biology and Medicine*, vol. 70, pp. 174–189, 2016.

[44] Y. Zhao, J. Zhao, J. Yang et al., "Saliency driven vasculature segmentation with infinite perimeter active contour model," *Neurocomputing*, vol. 259, pp. 201–209, 2017.

[45] D. Pandey, X. Yin, H. Wang, and Y. Zhang, "Accurate vessel segmentation using maximum entropy incorporating line detection and phase-preserving denoising," *Computer Vision and Image Understanding*, vol. 155, pp. 162–172, 2017.

[46] J. Zhang, Y. Chen, E. Bekkers et al., "Retinal vessel delineation using a brain-inspired wavelet transform and random forest," *Pattern Recognition*, vol. 69, pp. 107–123, 2017.

[47] J. H. Tan, U. R. Acharya, S. V. Bhandary, K. C. Chua, and S. Sivaprasad, "Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network," *Journal of Computational Science*, vol. 20, pp. 70–79, 2017.

[48] E. Tuba, L. Mrkela, and M. Tuba, "Retinal blood vessel segmentation by support vector machine classification," in *Proceedings of the 27th International Conference Radioelektronika*, pp. 1–6, Brno, Czech Republic, April 2017.

[49] X. Wang, X. Jiang, and J. Ren, "Blood vessel segmentation from fundus image by a cascade classification framework," *Pattern Recognition*, vol. 88, pp. 331–341, 2019.

[50] P. Chudzik, B. Al-Diri, F. Calivá, and A. Hunter, "DISCERN: generative framework for vessel segmentation using convolutional neural network and visual codebook," in *Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5934–5937, Honolulu, HI, USA, July 2018.

[51] M. Hajabdollahi, R. Esfandiarpoor, K. Najarian, N. Karimi, S. Samavi, and S. M. R. Soroushmeh, "Low complexity convolutional neural network for vessel segmentation in portable retinal diagnostic devices," in *Proceedings of the 25th IEEE International Conference on Image Processing*, pp. 2785–2789, Athens, Greece, October 2018.

[52] Z. Yan, X. Yang, and K.-T. Cheng, "A three-stage deep learning model for accurate retinal vessel segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1427–1436, 2019.

[53] T. A. Soomro, O. Hellwich, A. J. Afifi, M. Paul, J. Gao, and L. Zheng, "Strided U-Net model: retinal vessels segmentation using dice loss," in *Proceedings of the Digital Image Computing: Techniques and Applications*, pp. 1–8, Canberra, Australia, December 2018.

[54] M. Arsalan, M. Owais, T. Mahmood, S. W. Cho, and K. R. Park, "Aiding the diagnosis of diabetic and hypertensive retinopathy using artificial intelligence-based semantic segmentation," *Journal of Clinical Medicine*, vol. 8, no. 9, p. 1446, 2019.

[55] M. Ahmadi, A. Sharifi, M. Jafarian Fard, and N. Soleimani, "Detection of brain lesion location in MRI images using convolutional neural network and robust PCA," *International Journal of Neuroscience*, vol. 10, pp. 1–12, 2021.

[56] Z. Xu, F. R. Sheykhahmad, N. Ghadimi, and N. Razmjooy, "Computer-aided diagnosis of skin cancer based on soft computing techniques," *Open Medicine*, vol. 15, no. 1, pp. 860–871, 2020.

[57] A. Hu and N. Razmjooy, "Brain tumor diagnosis based on metaheuristics and deep learning," *International Journal of Imaging Systems and Technology*, vol. 31, no. 2, pp. 657–669, 2021.

[58] Q. Tian, Y. Wu, X. Ren, and N. Razmjooy, "A New optimized sequential method for lung tumor diagnosis based on deep learning and converged search and rescue algorithm," *Biomedical Signal Processing and Control*, vol. 68, Article ID 102761, 2021.

[59] N. Razmjooy and S. Razmjooy, "Skin melanoma segmentation using neural networks optimized by quantum invasive weed optimization algorithm," in *Proceedings of the Metaheuristics and Optimization in Computer and Electrical Engineering*, pp. 233–250, Springer, Cham, Switzerland, July 2021.

[60] M. Ahmadi, A. Sharifi, S. Hassantabar, and S. Enayati, "QAIS-DSNN: tumor area segmentation of MRI image with optimized quantum matched-filter technique and deep spiking neural network," *BioMed Research International*, vol. 2021, Article ID 6653879, 16 pages, 2021.

[61] S. Dorosti, S. Jafarzadeh Ghoushchi, E. Sobhrakhshankhah, M. Ahmadi, and A. Sharifi, "Application of gene expression programming and sensitivity analyses in analyzing effective parameters in gastric cancer tumor size and location," *Soft Computing*, vol. 24, no. 13, pp. 9943–9964, 2020.

[62] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.

[63] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Scientific Reports*, vol. 11, no. 1, pp. 10930–11017, 2021.

[64] M. Q. H. Abadi, S. Rahmati, A. Sharifi, and M. Ahmadi, "HSSAGA: designation and scheduling of nurses for taking care of COVID-19 patients using novel method of Hybrid Salp Swarm Algorithm and Genetic Algorithm," *Applied Soft Computing*, vol. 108, Article ID 107449, 2021.

[65] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 109–118, 2015.

[66] A. Christodoulidis, T. Hurtut, H. B. Tahar, and F. Cheriet, "A multi-scale tensor voting approach for small retinal vessel segmentation in high resolution fundus images," *Computerized Medical Imaging and Graphics*, vol. 52, pp. 28–43, 2016.

[67] S. Aslani and H. Sarnel, "A new supervised retinal vessel segmentation method based on robust hybrid features," *Biomedical Signal Processing and Control*, vol. 30, pp. 1–12, 2016.

[68] R. Vega, G. Sanchez-Ante, L. E. Falcon-Morales, H. Sossa, and E. Guevara, "Retinal vessel extraction using lattice neural networks with dendritic processing," *Computers in Biology and Medicine*, vol. 58, pp. 20–30, 2015.

[69] T. J. Jebaseeli, C. A. Deva Durai, and J. D. Peter, "Retinal blood vessel segmentation from diabetic retinopathy images using tandem PCNN model and deep learning based SVM," *Optik*, vol. 199, Article ID 163328, 2019.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 9651957, 12 pages
https://doi.org/10.1155/2021/9651957

Research Article

Skin Cancer Detection Using Kernel Fuzzy C-Means and Improved Neural Network Optimization Algorithm

Jia Huaping,[1] **Zhao Junlong**,[2] **and A. M. Norouzzadeh Gil Molk**[3]

[1]*College of Computer, Weinan Normal University, Weinan, Shaanxi, China*
[2]*Rehabilitation Medicine Department, Weinan Central Hospital, Weinan, Shaanxi, China*
[3]*Department of Computer Engineering, University of Guilan, Rasht, Iran*

Correspondence should be addressed to A. M. Norouzzadeh Gil Molk; norouzzadeh.ali@guilan.ac.ir

Received 20 May 2021; Revised 20 June 2021; Accepted 6 July 2021; Published 19 July 2021

Academic Editor: V. Rajinikanth

Early diagnosis of malignant skin cancer from images is a significant part of the cancer treatment process. One of the principal purposes of this research is to propose a pipeline methodology for an optimum computer-aided diagnosis of skin cancers. The method contains four main stages. The first stage is to perform a preprocessing based on noise reduction and contrast enhancement. The second stage is to segment the region of interest (ROI). This study uses kernel fuzzy C-means for ROI segmentation. Then, some features from the ROI are extracted, and then, a feature selection is used for selecting the best ones. The selected features are then injected into a support vector machine (SVM) for final identification. One important part of the contribution in this study is to propose a developed version of a new metaheuristic, named neural network optimization algorithm, to optimize both parts of feature selection and SVM classifier. Comparison results of the method with 5 state-of-the-art methods showed the approach's higher superiority toward the others.

1. Introduction

Cancer is fundamentally an uncontrolled cell division disease. Its growth is usually associated with a group of variations in the action of the cell cycle regularized. The inhibitors of the cell cycle stop cell division while situations are not proper; therefore, so little action of these inhibitors can cause cancer. Cancer cells disregard the signals that must cause cells to prevent dividing. For example, when common cells growing in a culture medium are surrounded on all sides by adjacent cells, they no longer divide (contact inhibition). On the contrary, cancer cells divide into layers in a mass and stack on top of each other, and contact inhibition does not prevent them from growing. One of the most dangerous cancers around the world is skin cancer. The skin, as the first cell layer of the body that is in connection with the outside environment, can suffer from many injuries and diseases. The most common cancer in the United States is skin cancer, which happens in the tissues of the largest section of the skin body. Skin cancer generally occurs on the outer layer of the skin, which may first appear as a swelling, bulge, or different section of the skin [1]. The skin blocks heat, sunlight, sores, and infections; it also warms the body and stores fat water in the body and produces vitamin D. The skin owns 2 principal layers: the superficial layer of the skin (epidermis) and the inner skin (dermis) [2]. The skin's surface layer (epidermis) is mainly composed of flat, scaly cells named Squamous cells. Round cells, called basal cells, are located beneath the superficial cells of the squamous melanocyte in the surface layer [3]. Prolonged exposure to sunlight enhances melanoma skin cancer's risk over a lifetime. One of the most dangerous kinds of cancer is Melanoma, which is caused by damage caused by overexposure to the sun and some other factors. Melanoma can be diagnosed with skin biopsies. Figure 1 shows the statistical information of the lead cancers in quantity and death value in 2019 [4].

If melanoma is diagnosed and treated early, it can be treated, and late diagnosis can lead to the patient's death. Only experienced physicians can diagnose melanoma on

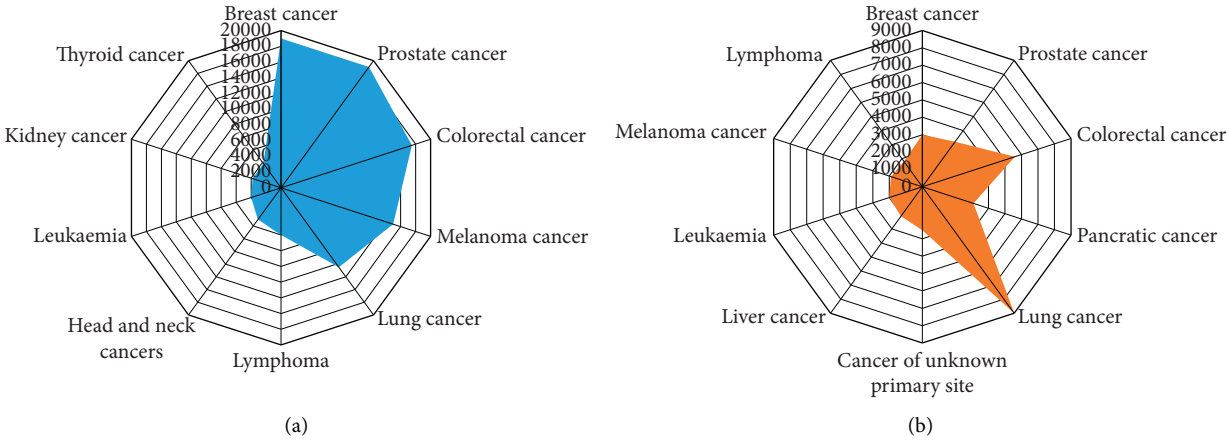


Figure 1: The statistical information of the lead cancers in (a) quantity and (b) death value in 2019 [4].

time using appropriate tools and histological reports. One of the devices used to diagnose melanoma is dermatoscopy. With this tool, changes in the pigmentation of skin lesions in diseases can be evaluated. With the development of science in recent years, digital dermatoscopy has been replaced by conventional dermatoscopes with the ability to capture and store skin images. With the development of digital dermatoscopes, the development of an algorithm for the diagnosis of melanoma was considered by researchers, and then, other methods were proposed to enhance the accuracy of skin lesion diagnosis. Some of the selected researches are explained in the following.

Astorino et al. [5] suggested an approach for melanoma diagnosis. The method uses multiple instance learning (MIL) approach on some dermoscopy images to categorize between melanomas and common nevi. The method was then analyzed by comparing it with some state-of-the-art classification methods, like support vector machine. Simulation outcomes demonstrated that the suggested approach accomplishes better precision. A leave-one-out authentication is carried out on a dataset and showed the method's higher efficiency.

Mohamed et al. [6] suggested an image processing approach for melanoma diagnosis. The method used a simple pipeline image preprocessing technique including methods like image conversion, noise reduction, and hair removal. Also, the melanoma area is threshold to segment the area, and then, feature extraction was used for taking out the main features of the image using the ABCD rule.

Barros et al. [7] proposed a real-time cancer diagnosis system based on hardware designing by the field-programmable gate arrays (FPGA). A multilayer perceptron artificial neural networks (ANN) is utilized for this purpose. The proposed image processing is used to extract the characteristics of the skin from the images and classify them based on the features and using the ANN classifier. Simulation results of the proposed approach are confirmed by an open-access database. Ultimate results of the suggested approach were then compared with the results of another hardware-based technique using ARM A9 microprocessor to demonstrate the method's proper performance.

Santos and Espitia [8] presented an approach for the diagnosis of uveal melanoma (UM), which is a sort of intraocular cancer. The suggested technique combined iris segmentation methods and designed a method for UM diagnosis based on neural networks and fuzzy logic. Simulation results of the study indicated 96.04% accuracy for the artificial neural networks and 76% correct classification for the fuzzy logic system.

Wang [9] proposed an all-inclusive method to provide a proper segmentation technique. The method is applied based on deep convolution networks on the hyperspectral pathology images. The study employed a 3D fully convolutional network, called Hyper-net, to obtain the best performances of segmentation from the hyperspectral pathology images. The loss function was then modified for improving the model sensitivity. Ultimate outcomes demonstrated that the suggested approach has higher performance than the 2D models in the aim of segmentation.

The present research suggests a novel optimized pipeline methodology for automatic detection of skin cancer detection. The method includes four main phases. The first phase is preprocessing of the input image to prepare it for the main processing. The first phase here includes two parts: noise reduction and contrast enhancement. The next phase is to segment the region of interest. Afterwards, the features are extracted, and the best ones selected for injecting to a support vector machine (SVM) as the final phase. This investigation employs an improved version of a new metaheuristic, called neural network optimization algorithm, to optimize both parts of feature selection and SVM classifier. This can improve the accuracy of the system as can be seen in the next sections. The main contribution of the paper has been highlighted as follows:

(i) An early diagnosis system for skin cancer from dermoscopy images is proposed

(ii) Kernel fuzzy C-means is used to segment the region of interest (ROI)

(iii) Some features from the ROI are extracted for the final diagnosis

(iv) An improved neural network algorithm is used for optimal feature selection

(v) Support vector machine (SVM) is used for final classification

(vi) Comparison results are compared with 5 state-of-the-art methods

Figure 2 illustrates the flowchart of the suggested approach.

2. Image Preprocessing

2.1. Noise Reduction. Ordinary nonpolarized light is used in medical images. Since, according to "Fresnel relations," the reflection of light from the surface of matter and its underlying layers and scattering from uneven and rough surfaces is affected, as well as a function of the state of polarization of light, it is possible to eliminate some annoying and destructive reflected or diffused lights by using a suitable filter to achieve a high signal to noise ratio. Therefore, the first step for starting the image analysis of the cancer images is to eliminate, or, at least, reduce, this noise to have a high quality of images for processing. Different methods have been introduced for this purpose. One simple and popular method for noise reduction is Median filtering. However, this filter is nonlinear; it has some significant advantages such as

keeping the main edges of the image after filtering. In the median filter, the pixels have been substituted by the median value of their neighbors, i.e.,

$$Z_{m,n} = \text{median}\left(y_{i,j}:\ (i,j) \in \beta\right), \tag{1}$$

where β signifies the neighborhood of the considered pixel in position (m,n).

The present study uses a 3×3 mask for the median filter. This selection is based on trial and error and may be different for other databases. Thus, it should be the point that using a higher mask can improve the probability of edges losing. Figure 3 shows an example of using the median filtering for noise reduction of the input images with considering 0.1 noise density salt and pepper noises.

2.2. Contrast Enhancement. In some situations, due to different reasons, like the sufficient user experience's lack in imaging and the bad quality of measurement devices and sensors utilized in them, the contrast of the images has lessened, such that its intensity gets darkened, or overexposed. To correct this deficiency from the images, we need to implement the image contrast enhancement. Therefore, in this study, we can use an image contrast enhancement step on the low-quality images to improve the image contrast to simplify the segmentation step. The present study considers global contrast enhancement based on Lookup Table to reach this purpose [10]. The method of global contrast enhancement is formulated as follows:

$$\text{Out}_H = \frac{\text{In}_H - \text{Min}_H}{\text{Max}_H - \text{Min}_H}, \tag{2}$$

where Out_H signifies the improved output image, and Min_H and Max_H demonstrate the minimum and the maximum levels of gray values of the original image histogram, orderly. We used an 8-bit lookup table for the method. A simple contrast enhancement on a medical image is shown in Figure 4.

3. Medical Image Processing

3.1. Segmentation Based on Kernel-Based Fuzzy C-Means. One of the significant issues in medical image processing is the image segmentation of the medical image into its components. Image segmentation describes the success or the final failure of image analysis methods. However, there is no general method for successful segmentation of all medical images, and it still has good research areas due to its wide application. The accuracy of this study is crucial in areas such as medicine, which helps preserve and protect human life. Due to the wide range of applications, it provides image segmentation and the use of methods in various fields.

The importance of this part especially in mammographic images is too high, because medical images have a naturally low quality, which makes them a complicated problem for mass segmentation. This research uses an enhanced multistep fuzzy C-means method for performing on the skin cancer images for their mass segmentation. Based on this method, the first step is to generate superpixels using a simple linear iterative clustering (SLIC) algorithm [11]. The SLIC algorithm generates superpixels based on CIE LAB color space in 5D space. Afterwards, texture and color properties are employed to separate the superpixels and finally to find the image elements.

Then, the color feature from the superpixels has been extracted and has been arranged, such that, for all images in HSV space, three separate histograms for S, H, and V channels have been considered. To reduce time complexity, the histograms have been quantized into 8, 4, and 2 subdistances. Afterward, to find the texture features, the NSCT algorithm has been used. Then, the generated data is clustered by a fuzzy kernel, and all superpixels with cluster tags have been placed in one cluster. This study uses kernel fuzzy C-means (KFCM) algorithm employing segmentation. This technique offers a kernel-based version of fuzzy C-means to evaluate the data point's distance from the center of clusters. The kernel function in this study is achieved as follows:

$$K = \exp\left(\frac{-x - y^2}{\rho}\right). \tag{3}$$

First, the membership function has been evaluated. Then, based on a similarity measure of kernel fuzzy, the belonging of each data sample to the clusters has been

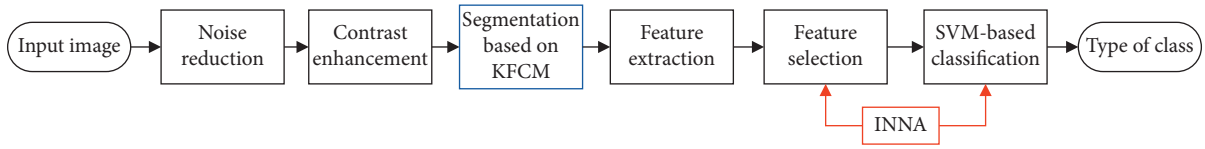
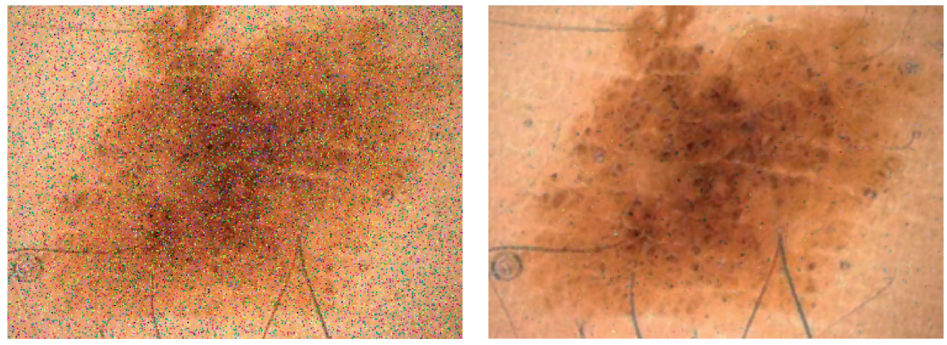
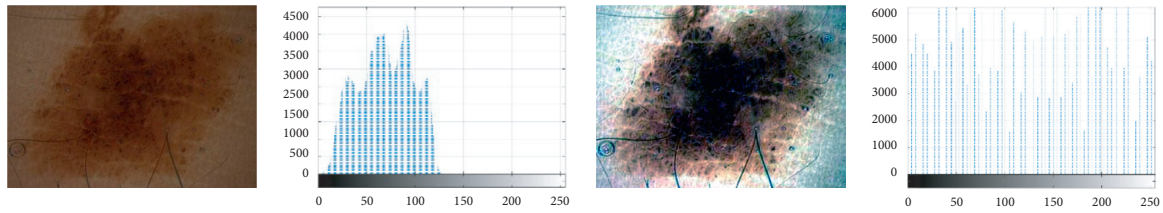


Figure 2: The flowchart of the suggested approach.



(a) (b)

Figure 3: Image noise reduction: (a) image plus 0.1 salt and pepper noises and (b) image after noise reduction with median filtering.



(a) (b) (c) (d)

Figure 4: (a) Original Image with low contrast. (b) Histogram of (A). (c) Contrast improvement of (A). (d) Histogram of (C).

achieved. The pseudocode of the algorithm is given as follows:

(1) Apply KFCM to cluster the set of objects and generating a U membership matrix [12]

(2) In all elements, x_i, x_j, the number of t nearest neighbor is found.

(3) If x_i, x_j are not t nearest neighborhood, $W_{ij} = 0$.

(4) Else if they belong to a cluster, $W_{ij} = 1$.

(5) Else, $W_{ij} = \exp(ln2 \times (u_i \oplus u_j))$, where $\oplus$ defines the exclusive OR that totally indicates the overlap between two fuzzy sets.

(6) Here, diagonal matrix, D, is formulated as follows:

$$D_{ij} = \sum_{j=1}^{n} W_{ij}. \tag{4}$$

And, it is normalized as follows:

$$L = D^{-1/2} \mathrm{W} \mathrm{D}^{1/2}. \tag{5}$$

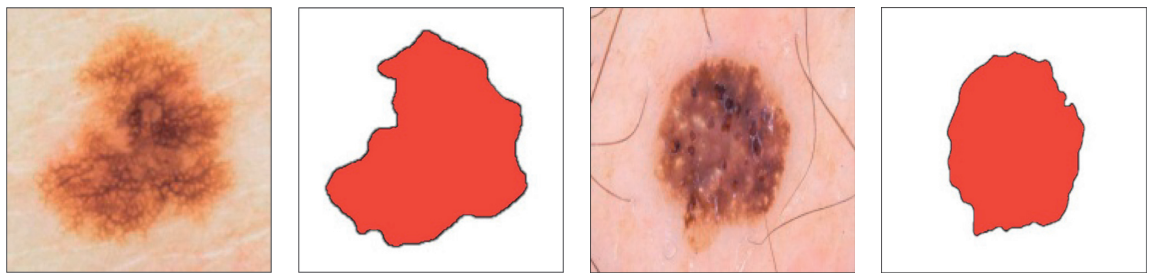
The number of K bigger eigenvector than L (the first vector) is found (the first vector is selected), and the matrix $P = [p^1, p^2, \ldots, p^k]$ has been formed, and then, the algorithm normalizes the rows in the P matrix to form matrix Y.

(7) Each line of Y has considered a point in space R^k and at then, the final clustering has been established by the K-means algorithm.

Figure 5 shows two sample examples of skin cancer segmentation based on the proposed KFCM methodology.

4. Improved Neural Network Algorithm

4.1. Optimization. Generally speaking, in most applications of engineering, optimization is a vital subject. Optimization is to make the best decision to get the optimal (minimum or maximum) result for the considered problem. Several methods have been introduced for optimization. However, classic methods as exact methods can solve these problems, and they fail in some cases that the problem is nonlinear or complicated. To overcome this problem, another technique, called metaheuristic, has been introduced [13]. Metaheuristics include a set of optimization algorithms that are inspired by nature, human behaviors, animals' competitions, etc. The main advantage of using metaheuristic algorithms is



(a) (b) (c) (d)

Figure 5: Two sample examples of skin cancer segmentation based on the proposed KFCM methodology.

that they use random structure instead of using gradient methods. This simplifies the optimization process [14]. Furthermore, the most complicated problems are that their number is increasing day by day and they cannot be solved by the classic methods, but these methods, based on their stochastic nature, can find a near-optimal result in a logical time. There are different types of these algorithms like ant lion optimizer (ALO) algorithm [15], chimp optimization (CO) algorithm [16], Harris Hawks optimization [17], and world cup optimization (WCO) algorithm [18]. Recently, a new metaheuristic algorithm, called neural network algorithm (NNA), is introduced, which is inspired by the concepts of biological nervous systems and artificial neural networks (ANNs). Based on [19], each ANN includes some artificial neurons, which are inspired by the biological nervous systems. The relationships among units principally specify the network function. The NNA approach is illustrated in the following.

4.2. Neural Network Algorithm. Like any other metaheuristic algorithm, the NNA starts with an initial population that is called the pattern. In NNA, just like ANN, the main idea is to update the pattern population to minimize error among the forecasted data and the desired output data. Here, the best solution is the desired output that can be updated in each iteration for achieving the minimum error amount by moving the pattern population in the direction of the desired solution. In the following, the algorithm methodology is briefly explained.

4.2.1. Initialization. The first operation in this algorithm, like any other metaheuristic, is to generate some random population (decision variables) for initial evaluation. By considering a D dimensional problem, the considered pattern solution vector will be an array of $1 \times D$, representing input data in the NNA. Moreover, by considering the D dimension and N_{pop} several random candidates, the initial pattern population can be considered as follows:

$$X = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_D^1 \\ x_1^2 & x_2^2 & x_3^2 & x_D^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_D^{N_{\text{pop}}} & x_D^{N_{\text{pop}}} & \cdots & x_D^{N_{\text{pop}}} \end{bmatrix}, \tag{6}$$

where the matrix X is made randomly between the minimum and maximum limitations of a problem.

And, the cost value is achieved as follows:

$$C_i = f\left(x_1^i, x_2^i, \ldots, x_D^i\right), \tag{7}$$

where f describes the objective value.

After evaluating the objective value of each pattern solution, the best one is selected as the best pattern solution. The NNA has N_{pop} input data and D dimension with only one target data. Figure 6 shows this structure.

After defining the target solution (X^T) among the other pattern solutions, its weight (W^T) has to be chosen from the population of weight (weight matrix).

In an ANN, the neurons connect with dendrite using a simple summation. The output is connected to the input layers based on weighted (w) interconnection. Initial weights are random values in this algorithm, and then, they have been updated based on some equations to provide the minimum network error. The initial weights in the algorithm are as follows:

$$W(t) = \begin{bmatrix} w_{11} & \cdots & w_{i1} & \cdots & w_{N_{\text{pop}}1} \\ w_{12} & \cdots & w_{i2} & \cdots & w_{N_{\text{pop}}2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{1N_{\text{pop}}} & \cdots & w_{iN_{\text{pop}}} & \cdots & w_{N_{\text{pop}}N_{\text{pop}}} \end{bmatrix}_{N_{\text{pop}} \times N_{\text{pop}}}, \tag{8}$$

where W includes uniformly distributer random values in the range [0, 1]. The first and the second subscripts of weight relate to the pattern solution and participated with the other solutions of the pattern. All pattern solutions have their weight value to generate a new candidate solution. These weights have also a constraint that should not exceed 1 and mathematically formulated as follows:

$$\sum_{j=1}^{N_{\text{pop}}} w_{ij}(t) = 1, \qquad w_{ij} \in U(0,1), \tag{9}$$

where $i, j = 1, 2, \ldots, N_{\text{pop}}$.

Weight is randomly distributed values in the range between 0 and 1, where their aggregate in a solution of the pattern must not be more than one, which is because of the bias control of the movement and to generate the new

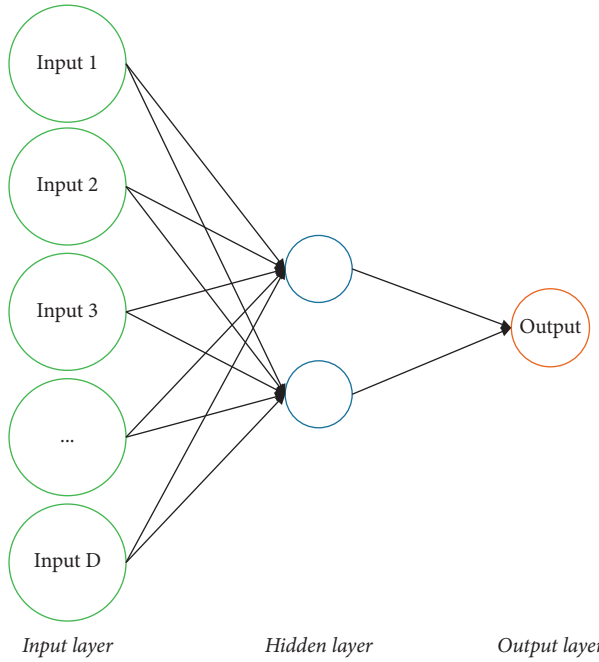


Figure 6: The structure of NNA.

pattern solutions. If the constraint is not considered, the probability of sticking in the local optimum by the values of the weight will be increased. After weight matrix formation, new pattern solutions (X^N) are evaluated as follows:

$$X_j^{\text{New}}(t+1) = \sum_{i=1}^{N_{\text{pop}}} w_{ij}(t) \times X_i(t),$$
$$X_i(t+1) = X_i(t) + X_i^{\text{New}}(t+1). \tag{10}$$

After achieving the new pattern solutions using the best weight amount that is named target weight, this updating is established by the following equation:

$$W_i^{\text{updated}}(t+1) = W_i(t) + 2 \times \text{rand} \times \left(W^{\text{Target}}(t) - W_I(t)\right). \tag{11}$$

Another operator for updating in NNA is the bias operator. This operator is employed for modifying the exploration part of the algorithm by using a specific percentage of the candidates in the new population ($X_i^{\text{New}}(t+1)$) and updated weight matrix ($W_i^{\text{updated}}(t+1)$). The pseudocode of this operator is as follows:

For I = 1 to N_{pop}
If rand $\leq \gamma$
N_{b} = round $(D \times \gamma)$
For j = 1: N_{b}
X_{Input}(I, integer rand [0, D]) = LB + (UB−LB) × rand.
End for
N_{wb} = round $(N_{\text{pop}} \times \gamma)$
For j = 1: N_{wb}
$W^{\text{Updated}}(j$, integer rand $[0, N_{\text{pop}}])$ = U (0,1).
End for
End if
End for

Here, LB and UB represent the problem minimum and maximum bounds, orderly. γ describes the modification factor that defines the candidates' percentage. $\gamma = 1$ at first, and then, it decreases gradually during the process based on any decreasing formulation as follows:

$$\gamma(t+1) = 0.99 \times \gamma(t). \tag{12}$$

In the NNA, there is also a transfer function operator to transfer the new candidates to the new updated positions to provide better value against the target solution. The solution reimproved by giving the novel pattern solutions to the best solution direction (target solution). The transfer function operator (TF) can be formulated as follows:

$$X_i^*(t+1) = \text{TF}\left(X_i(t+1)\right) = X_i(t+1) + 2 \times \text{rand} \times \left(X^{\text{Target}}(t) - W_i(t+1)\right). \tag{13}$$

The pseudocode of the collaboration between bias and TF operators is tabulated below:

For i = 1 to N_{pop}
If rand $\leq \beta$
Bias operator
Else (rand $> \beta$)
Apply the transfer function operator
End if
End for

4.3. Improved Neural Network Algorithm. However, the neural network algorithm provides good results to solve different applications of optimization problems [20–22], and it is sometimes stuck in the premature convergence that gives a high impact on the solution accuracy. In this investigation, to modify this drawback, 2 mechanisms are considered. The first mechanism is to employ the Chaos theory. This mechanism explores novel situations, which are dynamic discrete-time, i.e.,

$$\beta_{m+1} = f(\beta_m), \quad m = 0, 1, 2, \ldots. \tag{14}$$

In this research, the logistic map function has been used as the modification mechanism. This mechanism could be mathematically modeled as follows:

$$\beta_{j,i+1,q} = \rho \times \beta_{j,i,q}\left(1 - \beta_{j,i,q}\right), \tag{15}$$

where i signifies the populations' number, $\beta_{j,i,q}$ defines the value for the i^{th} chaotic iteration, ρ describes a constant, which is set 4, q demonstrates the iteration number, j illustrates the generators' number in the system, and β_0 describes the initial value of β_i with a random value between 0 and 1 [23, 24]. With assuming the above explanation, the set of initial variables w_{ij} is rewritten as follows:

$$w_{ij} = L_j + \left(U_j - L_j\right) \times \beta_{j,i,q}. \tag{16}$$

This study also used Lévy flight (LF) as a second modification. The LF is a popular technique for improving optimization algorithms [25]. The LF uses random motion to control the local searching. This is obtained by the following equations:

$$\begin{aligned} Le(w) &\approx \frac{1}{w^{1+\tau}}, \\ w &= \frac{A}{\sqrt[\tau]{|B|}}, \\ \sigma^2 &= \left\{ \frac{\Gamma(1+\tau)}{\tau\Gamma\left((1+\tau)/2\right)} \times \frac{\sin(\pi\tau/2)}{2^{(1+\tau)/2}} \right\}^{2/\tau}, \end{aligned} \tag{17}$$

where τ represents the Lévy flight index in the range [0, 2] (here, $\tau = 1.5$ [26]), w describes the step size, $A, B \sim N(0, \sigma^2)$, $\Gamma(.)$ describes Gamma function, and the samples generate a Gaussian distribution with σ^2 variance and zero mean value. By assuming the LF mechanism, the updated equation for the new pattern solutions based on the best weight value is achieved as follows:

$$\begin{aligned} W_i^{\text{updated}}(t+1) = W_i(t) + 2 \times \text{rand} \times \text{Le}(\delta) \\ \times \left(W^{\text{Target}}(t) - W_I(t)\right). \end{aligned} \tag{18}$$

4.4. Algorithm Validation. To assess the effectiveness of the proposed improved neural network algorithm, six standard benchmark functions have been utilized, and the results are compared with some new state-of-the-art algorithms, containing multiverse optimizer (MVO) [27], moth-flame optimization (MFO) algorithm [28], world cup optimizer (WCO) [19], and the original neural network algorithm (NNA). The parameter settings of the algorithms are as follows: for MVO: wormhole existence probability (WEP); $\text{WEP}_{\min} = 0.2$; $\text{WEP}_{\max} = 1$; Coefficient $(P) = 6$. For MFO: logarithmic spiral shaped $(b) = 1$; for WCO: $\alpha = 0.5$, playoff = 4%. The population size for whole algorithms is set to 100, and their maximum iteration is set to 100. Table 1 tabulates the studied benchmark functions for the analysis. The benchmark functions are considered with 30 dimensions.

To validate the algorithms based on the studied functions, four statistical indicators including minimum amount (Min), maximum amount (Max), mean amount (Mean), and standard deviation amount (Std) are employed, and the results are indicated in Table 2.

According to Table 2, the outcomes of the suggested INNA have the proper outcomes in terms of lower, higher, and amount. Because it provides the minimum value for these values that is the main purpose of these benchmark functions, therefore, the proposed INNA has the highest accuracy. Furthermore, the proposed method with a minimum standard deviation value provides the highest reliable solution for the studied benchmark functions.

5. Feature Extraction and Selection Based on INNA

Feature selection is one of the most important methods and techniques for data preprocessing and data mining. Due to the introduction of new programs for large data mining, media information retrieval, and medical data processing that requires the processing of large volumes of data, it is important to limit the number of features. The purpose of feature extraction here is to take out the important characteristics from the region of interest in skin cancer images to simplify the diagnosis process. In this paper, 19 different features have been used for extraction from the image. Table 3 indicates the used features for the extraction.

Therefore, here, we utilized a feature selection technique to reduce and select some more important features and to eliminate the low-cost features. Different methods have been introduced for this purpose. Metaheuristics are a kind of new technique for feature selection.

Feature selection using metaheuristics is one of the subsets of feature extraction and is discussed in various fields of machine learning and data mining. In general, this problem does not have a definite solution, and so far, no precise method has been proposed to solve it. Various classical approaches have been proposed to these problems, but usually, the quality of their solutions is generally not very good. In contrast, intelligent optimization methods can provide far better solutions to these problems. Therefore, one of the effective and constructive methods in solving feature selection problems and related issues is the use of metaheuristic optimization methods and evolutionary algorithms.

Here, we employed a cost function for optimum selection of the features using four indicators, including false positive (FP), false negative (FN), true positive (TP), and true negative (TN). The cost function is formulated as follows:

$$\text{fitness} = \frac{(\text{TP} \times \text{TN}) - (\text{FP} \times \text{FN})}{\sqrt{((\text{TN} + \text{FP}) \times (\text{TP} + \text{FP}) \times (\text{TP} + \text{FN}) \times (\text{TN} + \text{FN}))}}. \tag{19}$$

The idea is to minimize the abovementioned function with selecting the high cost and important features. Based on the designed INNA, the above function has been minimized, and less-cost features are eliminated from the processing. The next step is to employ an efficient method for classifying these features.

After performing the binary optimization technique on the features, the following features with higher effectiveness are selected for training: elongation, area, mean, correlation, entropy, and elongation. This selection is performed based on a metaheuristic technique based on a stochastic nature.

6. Skin Cancer Image Classification

Classification is the process of separating an input image into predetermined classes. Image classification is considered the last part of the diagnosis in medical image analysis. This step

Table 1: The applied test functions for the analysis.

Function	Formula	Minimum	Limitation
Rosenbrock	$f_1 = \sum_{i=1}^{D-1}\left\{100\left(x_{i+1} - x_i\right)^2 + \left(x_i - 1\right)^2\right\}$	0	$[-30, 30]^D$
Sum squares	$f_2 = \sum_{i=1}^{D} i x_i^2$	0	$[-10, 10]^D$
Step 2	$f_3 = \sum_{i=1}^{D} \left(x_i + 0.5\right)^2$	0	$[-100, 100]^D$
Schwefel 2.22	$f_4 = \sum_{i=1}^{D} \lvert x_i\rvert - \prod_{i=1}^{D} \lvert x_i\rvert$	0	$[-10, 10]^D$
Schwefel 1.2	$f_5 = \sum_{i=1}^{D} \left(\sum_{j=1}^{i} x_j\right)^2$	0	$[-100, 100]^D$
Chung Reynolds	$f_6 = \left(\sum_{i=1}^{D} x_i^2\right)^2$	0	$[-100, 100]^D$

Table 2: The simulation results of the comparative algorithms on the studied benchmark functions, moth-flame optimization (MFO) algorithm [28], world cup optimizer (WCO) [19], and the original neural network algorithm (NNA).

Algorithm		f_1	f_2	f_3	f_4	f_5	f_6
MVO [27]	Min	17.58	0.0024	0.0031	19.38	0	$6.37e-8$
	Max	$7.29e+3$	436.15	$3.86e+5$	$2.37e+3$	$5.39e-7$	$4.19e-8$
	Mean	$2.96e+3$	280.16	$4.19e+4$	67.25	$6.15e-9$	$1.27e-7$
	Std	$5.79e+2$	95.37	$6.37e+4$	11.62	$3.19e-8$	$3.07e-7$
MFO [28]	Min	28.14	2.57	0.051	6.29	$4.38e-6$	8.39
	Max	129.08	4.18	9.27	53.49	0.015	110.35
	Mean	80.46	23.09	4.52	30.17	0.02	35.08
	Std	22.46	1.64	1.73	3.28	0.01	24.63
WCO [19]	Min	7.39	$5.13e-5$	$5.92e-5$	2.57	$1.94e-8$	$6.38e-9$
	Max	$1.38e+2$	0.264	0.017	4.12	$7.51e-8$	$7.29e-8$
	Mean	105.37	0.0573	0.024	3.28	$6.48e-9$	$1.18e-9$
	Std	14.83	0.0873	$3.28e-5$	1.23	$4.29e-9$	$3.19e-8$
NNA [29]	Min	4.91	$6.29e-21$	$9.75e-9$	0.017	$9.37e-16$	$5.9e-17$
	Max	45.22	$34.39e-18$	$1.19e-8$	4.62	$2.58e-12$	$5.19\ e-16$
	Mean	13.83	$6.30e-18$	$5.94e-8$	0.532	$3.42e-13$	$6.67e-16$
	Std	5.29	$3.62e-21$	$6.53e-9$	0.42	$1.97e-13$	$7.50e-17$
INNA	Min	5.16	$15.26e-22$	$6.37e-11$	$6.38e-15$	$7.26e-7$	$6.71e-39$
	Max	111.57	$6.76e-20$	$4.29e-10$	$3.95e-14$	0.0234	$9.43e-35$
	Mean	24.13	$2.48e-20$	$1.17e-10$	$1.09e-14$	$3.11e-5$	$4.57e-6$
	Std	12.28	$6.19e-21$	$5.64e-11$	$9.64e-15$	$11.97e-7$	$9.55e-36$

Table 3: Utilized features for the extraction.

Parameter	Equation	Parameter	Equation
Elongation	$\sqrt{\text{number of tumor pixels}/2 \times (\text{Max radius})^2}$	Compactness	$(2 \times \sqrt{\text{number of tumor pixels} \times \pi})/\text{Area}$
Area	$\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j)$	Mean	$1/\text{MN}\left(\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j)\right)$
Perimeter	$\sum_{I=1}^{M}\sum_{j=1}^{N} B_p(i,j)$	Correlation	$\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j) - \mu_r\mu_c/\sigma_r\sigma_c$
Variance	$1/\text{MN}\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\left(p(i,j) - \mu\right)\right)$	Solidity	Area/convex area
Invariant moments	$\varphi_1 = \eta_{20} + \eta_{02}$ $\varphi_2 = \left(\eta_{20} - \eta_{02}\right)^2 + 4\eta_{11}^2$ $\varphi_3 = \left(\eta_{30} - 3\eta_{12}\right)^2 + \left(3\eta_{21} - \mu_{03}\right)^2$	Entropy	$-\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j)\log p(i,j)$
Elongation	$2\sqrt{\text{Area}}/a\sqrt{\pi}$	Eccentricity	$2a^{-1}\left(a^2 - b^2\right)^{0.5}$
Rectangularity	$\text{Area}/b \times a$	Energy	$\sum_{i=1}^{M}\sum_{j=1}^{N} p^2(i,j)$
Irregularity index	$4\pi \times \text{area}/\text{perimeter}^2$	Standard deviation	$\text{Variance}^{1/2}$
Form factor	Area/a^2		

MN describes the image size, B_p defines the external side length for the boundary pixel, $p(i, j)$ defines the pixels intensity amount at position (i, j), μ and σ describe the mean value and the standard deviation, orderly, and a and b present the major axis and the minor axis, respectively. However, some of the above features have a high impact, and some others have a low impact on feature extraction.

can decrease the time for the cancer diagnosis process by decreasing the search space in it. A popular and useful method in classification is to use a support vector machine (SVM). The SVM seeks the best hyperplane that operates as a multipart (here two-part) data separator in the input space. For reaching the constraints of the optimization problem, the support vector machine has been employed to evaluate the normal vector, w, in the hyperplane, the bias b, and the slack variable, η for incorrectly assigned training patterns that support the generalization, which is defined as follows:

$$y = \min \frac{w^2}{2} + C \times \sum_{i=1}^{n} \eta_i. \tag{20}$$

With separation function:

$$y_i (q \times x_i + b) \geq 1 - \eta_i. \tag{21}$$

Feature space is highly affected by the separating function in equation (21), so that a function selection needs to be performed correctly to generate optimal output. Here, for minimizing equation (20), the suggested INNA has been utilized. Accordingly, the support vector machine uses a kernel that is employed to alter data to the feature space. In this research, three different kernels including linear, polynomial, radial basis function (RBF), and sigmoid have been used. Figure 7 shows a typical SVM.

As can be explained, the SVM technique is a proper tool for classifying the images. Medical images are also a part of this event. Therefore, after performing feature extraction on the images for decreasing the operation complexity, the skin cancer features for each image can be injected into the SVM, and the SVM will present the final classification, which will be healthy or cancer.

7. Evaluation and Results

7.1. Implementation Details. As mentioned before, the proposed skin cancer detection system contains different modules that are generated in MATLAB R2019b environment. The modules include image segmentation, image acquisition, feature extraction and choice, and the last classification. In the image acquisition stage for dermoscopic images, American Cancer Society (ACS) [30] and PH^2 [31] databases are used:

(i) The ACS database: the ACS database includes 68 pairs of XLM and TLM images that are integrated by the Nervoscope system. Experimental analysis is established by a 16 GB RAM Intel Core i7 processor. This database is downloaded one by one from https://www.cancer.org/cancer/skin-cancer.html.

(ii) The PH^2 database: dermoscopic images of this database are collected from the Dermatology Service of Hospital Pedro Hispano (Matosinhos, Portugal) under the same conditions. These images are 8-bit RGB color images with a resolution of 768×560 pixels. The image database includes a total of 200 dermoscopic images of melanocytic lesions, including 80 common nevi, 80 atypical nevi, and 40 melanomas. This database can be downloaded from: https://www.fc.up.pt/addi/ph2%20database.html.

The training and the test data for the benchmarks are set to 80% and 20%, respectively.

7.2. Results. As previously mentioned, the recommended simulations have been carried out to the ACS dataset to investigate the system efficiency. For numerical analysis of the suggested methodology, three popular measurement indicators for classification problems have been used. These metrics are accuracy (ACC), precision (PR), and sensitivity (SN). By comparing the ground-truth with the output mask, the classification indicators are evaluated as follows:

$$\begin{aligned} \mathrm{ACC} &= \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \times 100, \\ \mathrm{PR} &= \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \times 100, \\ \mathrm{SN} &= \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100, \end{aligned} \tag{22}$$

where TP, TN, FP, and FN denote the true positive, true negative, false positive, and false negative.

So, by analyzing the efficiency of the INNA-SVM method by considering different kernels based on the abovementioned indicators, we have the following.

It can be observed from Table 4 that the linear kernel of the proposed INNA-SVM, in addition to simplicity, results in the best values for all indicators for both databases: sensitivity, accuracy, and precision. Therefore, this kernel is utilized here for the classification of the features. For a comprehensive investigation about the proposed method's efficiency, five evaluation measurements including specificity (SP), positive predictive value (PPV), negative predictive value (NPV), F1 score, and Matthews correlation coefficient (MCC) are considered, which are mathematically as follows:

$$\begin{aligned} \mathrm{SP} &= \frac{\mathrm{TN}}{\mathrm{FP} + \mathrm{TN}}, \\ \mathrm{PPV} &= \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \\ \mathrm{NPV} &= \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FN}}, \\ F1\ \mathrm{score} &= \frac{2 \times \mathrm{TP}}{2 \times \mathrm{TP} + \mathrm{FP} + \mathrm{FN}}, \\ MCC &= \frac{\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}}{\sqrt{(\mathrm{TP} + \mathrm{FP})(\mathrm{TP} + \mathrm{FN})(\mathrm{TN} + \mathrm{FP})(\mathrm{TN} + \mathrm{FN})}}. \end{aligned} \tag{23}$$

The general proposed method has been compared with five new methods, and the comparison results have been indicated in Table 5. The compared methods include Astorino's method [5] based on multiple instance learning (MIL), Hassan's method [6] based on a simple pipeline image preprocessing technique, Barros's method [7] based on hardware designing by the multilayer perceptron, Santos's method [8] based on neural networks and fuzzy logic, and Wang's method [9] based on deep convolution networks. The results are achieved from the mean value of both databases.

According to Table 5, the F1-score of the suggested method with 84.39% that defines its accuracy based on the precision and recall of the test data is the highest among all of

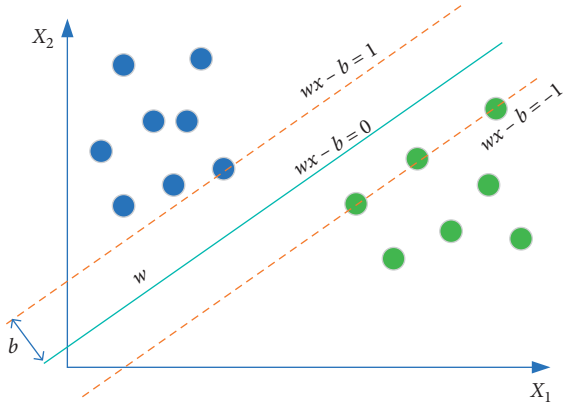


Figure 7: A typical model of an SVM.

Table 4: The classification results of INNA-SVM method based on different kernels in two databases.

Database	Kernel	ACC (%)	Pr (%)	SN (%)
ACS [30]	Linear	64.19	65.18	64.19
	RBF	37.57	38.09	37.57
	Polynomial	46.29	38.61	46.29
	Sigmoid	40.16	41.53	40.16
PH^2 [31]	Linear	70.00	69.00	71.00
	RBF	42.00	43.00	42.00
	Polynomial	51.00	45.00	51.00
	Sigmoid	42.50	46.00	53.00

Table 5: The performance comparison between the suggested method and the analyzed methods.

Method	MCC (%)	SP (%)	PPV (%)	NPV (%)	*F*1-score (%)
The proposed method	92.35	88.67	81.43	92.66	81.69
Astorino's method [5]	72.29	87.49	81.37	86.72	69.29
Hassan's method [6]	75.46	80.76	69.35	85.87	63.49
Barros's method [7]	83.65	60.34	68.37	83.98	70.81
Santos's method [8]	85.74	60.41	71.95	86.69	76.76
Wang's method [9]	88.67	80.77	74.49	83.71	82.92

the compared methods. Also, the value of MCC in the proposed method with 94.21% is the highest among the others, and because it uses all of the analysis terms, it shows the method's higher efficiency than the others. Also, the higher value of the NPV and PPV with 95.57% and 85.09% against the other methods shows higher condition occurrence of the method to control the likelihood of a test identifying cancer toward the others. Finally, the higher SP value based on the proposed method reports higher occurrence-independent results of the algorithm.

8. Conclusions

Skin cancer is one of the most common cancers, and malignant melanoma is the most invasive and deadliest kind of this cancer. Early detection of this cancer by physicians has a high effect on the treatment of this cancer. Recently, several machine vision-based techniques are used for helping physicians with accurate early detection. This paper proposed a new approach for optimal skin cancer detection. The first section in the proposed method was to establish a preprocessing technique including noise reduction and contrast enhancement. Then, the region of interest (ROI) was segmented by a kernel fuzzy C-means segmentation method. The features of the ROI were then extracted, and the most important features were selected optimally. Afterward, the selected features were injected into an optimal classifier using support vector machine (SVM) for classification. The main contribution of this study was to present a developed version of a new metaheuristic, called neural network optimization algorithm, to optimize both parts of feature selection and SVM classifier in the system. The superiority of

the proposed method was proved by performing a comparison among the suggested approach and five state-of-the-art methods.

Data Availability

The data are available at https://www.cancer.org/.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] R. Navid, "Computer-aided diagnosis of skin cancer: a review," *Current Medical Imaging*, vol. 16, no. 7, pp. 781–793, 2020.

[2] S. Ahuja, B. K. Panigrahi, N. Dey, V. Rajinikanth, and T. K. Gandhi, "Deep transfer learning-based automated detection of COVID-19 from lung CT scan slices," *Applied Intelligence*, vol. 51, no. 1, pp. 571–585, 2021.

[3] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.

[4] A. Costa, Y. Kieffer, A. Scholer-Dahirel et al., "Fibroblast heterogeneity and immunosuppressive environment in human breast cancer," *Cancer Cell*, vol. 33, no. 3, pp. 463–479, 2018.

[5] A. Astorino, A. Fuduli, P. Veltri, and E. Vocaturo, "Melanoma detection by means of multiple instance learning," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 12, no. 1, pp. 24–31, 2020.

[6] S. H. Mohamed, S. D. Noha, A. M. Tarek, and H. E. May, "Detection of melanoma using image processing techniques driven tds index," *Journal of Environmental Science*, vol. 46, no. 2, pp. 1–46, 2019.

[7] W. K. P. Barros, D. S. Morais, F. F. Lopes, M. F. Torquato, R. D. M. Barbosa, and M. A. C. Fernandes, "Proposal of the CAD system for melanoma detection using reconfigurable computing," *Sensors*, vol. 20, no. 11, p. 3168, 2020.

[8] D. F. Santos and H. E. Espitia, "Detection of uveal melanoma using fuzzy and neural networks classifiers," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 4, pp. 2213–2223, 2020.

[9] Q. Wang, "Identification of melanoma from hyperspectral pathology image using 3D convolutional networks," *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 218–227, 2020.

[10] R. Navid, B. S. Mousavi, and F. Soleymani, "A real-time mathematical computer method for potato inspection using machine vision," *Computers & Mathematics with Applications*, vol. 63, no. 1, pp. 268–279, 2012.

[11] K.-S. Kim, D. Zhang, M.-C. Kang, and S.-J. Ko, "Improved simple linear iterative clustering superpixels," in *Proceedings of the 2013 IEEE International Symposium on Consumer Electronics (ISCE)*, pp. 259–260, IEEE, Hsinchu, Taiwan, June 2013.

[12] M. R. P. Ferreira and F. A. T. de Carvalho, "Kernel fuzzy c-means with automatic variable weighting," *Fuzzy Sets and Systems*, vol. 237, pp. 1–46, 2014.

[13] R. Navid and R. Saeid, "Skin melanoma segmentation using neural networks optimized by quantum invasive weed optimization algorithm," in *Metaheuristics and Optimization in Computer and Electrical Engineering*, pp. 233–250, Springer, Berlin, Germany, 2021.

[14] D. B. Mehdi, "A new improved model of marine predator algorithm for optimization problems," *Arabian Journal for Science and Engineering*, Springer, Berlin, Germany, 2021.

[15] M. Mani, O. Bozorg-Haddad, and X. Chu, "Ant lion optimizer (ALO) algorithm," in *Advanced Optimization by Nature-Inspired Algorithms*, pp. 105–116, Springer, Berlin, Germany, 2018.

[16] M. Khishe and M. R. Mosavi, "Chimp optimization algorithm," *Expert Systems with Applications*, vol. 149, Article ID 113338, 2020.

[17] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: algorithm and applications," *Future Generation Computer Systems*, vol. 97, pp. 849–872, 2019.

[18] N. Razmjooy, M. Khalilpour, and M. Ramezani, "A new meta-heuristic optimization algorithm inspired by FIFA world cup competitions: theory and its application in PID designing for AVR system," *Journal of Control, Automation and Electrical Systems*, vol. 27, no. 4, pp. 419–440, 2016.

[19] Z. Zhang, "Artificial neural network," in *Multivariate Time Series Analysis in Climate and Environmental Research*, pp. 1–35, Springer, Berlin, Germany, 2018.

[20] Y. Zhang, Z. Jin, and Y. Chen, "Hybrid teaching-learning-based optimization and neural network algorithm for engineering design optimization problems," *Knowledge-Based Systems*, vol. 187, Article ID 104836, 2020.

[21] M. Elsisi, K. Mahmoud, M. Lehtonen, and M. M. F. Darwish, "An improved neural network algorithm to efficiently track various trajectories of robot manipulator arms," *IEEE Access*, vol. 9, pp. 11911–11920, 2021.

[22] Y. Zhang, Z. Jin, and Y. Chen, "Hybridizing grey wolf optimization with neural network algorithm for global numerical optimization problems," *Neural Computing and Applications*, vol. 32, no. 14, pp. 10451–10470, 2020.

[23] D. Yang, G. Li, and G. Cheng, "On the efficiency of chaos optimization algorithms for global optimization," *Chaos, Solitons & Fractals*, vol. 34, no. 4, pp. 1366–1375, 2007.

[24] C. Rim, S. Piao, G. Li, and U. Pak, "A niching chaos optimization algorithm for multimodal optimization," *Soft Computing*, vol. 22, no. 2, pp. 621–633, 2018.

[25] Y. Cao, Y. Wu, L. Fu, K. Jermsittiparsert, and N. Razmjooy, "Multi-objective optimization of a PEMFC based CCHP system by meta-heuristics," *Energy Reports*, vol. 5, pp. 1551–1559, 2019.

[26] X. Li, P. Niu, and J. Liu, "Combustion optimization of a boiler based on the chaos and Lévy flight vortex search algorithm," *Applied Mathematical Modelling*, vol. 58, pp. 3–18, 2018.

[27] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-verse optimizer: a nature-inspired algorithm for global optimization," *Neural Computing and Applications*, vol. 27, no. 2, pp. 495–513, 2016.

[28] S. Mirjalili, "Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm," *Knowledge-based Systems*, vol. 89, pp. 228–249, 2015.

[29] A. Sadollah, H. Sayyaadi, and A. Yadav, "A dynamic metaheuristic optimization model inspired by biological nervous systems: neural network algorithm," *Applied Soft Computing*, vol. 71, pp. 747–782, 2018.

[30] ACS, "Skin cancer database," 2019, https://www.cancer.org/.

[31] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. Marcal, and J. Rozeira, "PH 2-A dermoscopic image database for research and benchmarking," in *Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5437–5440, IEEE, Osaka, Japan, July 2013.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 5396327, 11 pages
https://doi.org/10.1155/2021/5396327

Research Article

Automatic Breast Tumor Diagnosis in MRI Based on a Hybrid CNN and Feature-Based Method Using Improved Deer Hunting Optimization Algorithm

Weitao Ha[1] **and Zahra Vahedi**[2]

[1]*College of Computer, Weinan Normal University, Weinan, Shaanxi, China*
[2]*Department of Electrical and Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran*

Correspondence should be addressed to Zahra Vahedi; zahra.vahedi@srbiau.ac.ir

Received 15 May 2021; Revised 12 June 2021; Accepted 6 July 2021; Published 16 July 2021

Academic Editor: Navid Razmjooy

Breast cancer is an unusual mass of the breast texture. It begins with an abnormal change in cell structure. This disease may increase uncontrollably and affects neighboring textures. Early diagnosis of this cancer (abnormal cell changes) can help definitively treat it. Also, prevention of this cancer can help to decrease the high cost of medical caring for breast cancer patients. In recent years, the computer-aided technique is an important active field for automatic cancer detection. In this study, an automatic breast tumor diagnosis system is introduced. An improved Deer Hunting Optimization Algorithm (DHOA) is used as the optimization algorithm. The presented method utilized a hybrid feature-based technique and a new optimized convolutional neural network (CNN). Simulations are applied to the DCE-MRI dataset based on some performance indexes. The novel contribution of this paper is to apply the preprocessing stage to simplifying the classification. Besides, we used a new metaheuristic algorithm. Also, the feature extraction by Haralick texture and local binary pattern (LBP) is recommended. Due to the obtained results, the accuracy of this method is 98.89%, which represents the high potential and efficiency of this method.

1. Introduction

The term "cancer" refers to the abnormal growth of some cells. These cells can spread and invade other sections of the human body. The occurrence of this in the breast texture as a mass is called breast cancer. Breast cancer occurs mainly in women, but it may be observed in men, too. After lung cancer, the second cause of death in women is breast cancer. Generally, there are two types of breast cancer based on primary origin: primary tumors and secondary tumors. The primary tumor originates from breast texture cells. In metastatic tumors, the cells become cancerous in another section of the body and spread to the breast through the lymphatic system or bloodstream.

Based on the American Cancer Society (ACS) statistics, the incidence rate for new cases of this disease is 125.3 for females each year (per 100,000 men and women). The death rate value is 20.3 for females per 100,000 men and women each year. The rates are given age-arranged and according to 2012–2016 cases and 2013–2017 death cases [1].

Invasive Ductal Carcinoma (IDC) is the most communal type of this cancer. It begins from cells that put the milk duct into the breast. Then destroys the duct wall and spreads to the nearby breast textures. At this point, it spreads (metastasize) through the lymph system and stream of the blood. Many women with breast cancer diseases are treated by (1) hormone therapy and (2) chemotherapy. Similarly, targeted therapy or radiation is the local treatment. Sometimes, a combination of these treats is used. Early diagnosis of this disease can be very important to cure.

The breast imaging methods commonly used at this time are mammograms, ultrasound, breast MRI, breast

tomosynthesis (3D mammography), positron emission tomography (PET), computed tomography (CT) scan, optical imaging tests, electrical impedance imaging (EIT) scans, contrast-enhanced mammography (CEM), and, at last, the chest X-rays. These techniques are used for careful observation of significant things like the shape, size, location, the exact kind of cancer, more details about the stage of cancer or how fast it is growing, and metabolism of breast tumors. Sometimes, a combination of these methods is used for a more accurate diagnosis.

According to recent researches, MRI may locate certain small breast lesions that are occasionally missed by the mammography method. Therefore, it can be a useful diagnostic tool. Nowadays, computer-aided diagnosis (CAD) based on MRI images is used to detect tumors. Hence, this efficient method is more important. Indeed, combining the CAD systems with MRI images is caused to decrease the useless data and aided in fast detection of the tumor. Recently, artificial intelligence based on CAD has been used to improve detection.

Formerly, mammography and MRI image processing for breast tumor diagnosis were based on machine learning techniques and extraction of geometric features. In deep-learning algorithms, the convolutional neural networks (CNNs) are speedily becoming a prevalent technique to process medical images. Deep learning is hierarchical learning and one of the subbranches of machine learning which is for learning high-level data summaries. This emerging method has been noticed more in the artificial intelligence field.

In previous articles, several techniques were presented for breast tumor diagnosis. For instance, Hu et al. [2] proposed a method for feature extractions of (1) (DCE)-MRI and (2) (T2w) sequences to improve breast cancer detection. (DCE)-MRI sequence is the dynamic contrast-enhanced and T2 is the weighted (T2w) MRI sequence for each MR study. Based on the mentioned features, this method was used as a pretrained convolutional neural network (CNN) for classification and final detection between benign and malignant tumors. In conclusion, feature fusion using DCE (P value < 0.001) (95% confidence intervals) had statistically better performance.

Ibrahim et al. [3] introduced a segmentation approach for breast tumors in thermal images. They used the chaotic salp swarm algorithm (CSSA) to this. This segmentation algorithm uses the quick-shift technique which clusters the breast thermal image pixels to reach the optimal superpixels. The final results showed that removing the extra parts of the image and keep the breast area. This leads to improving the detection accuracy (92%).

Ibraheem et al. [4] presented a median 2D filter which used to preprocess breast cancer images. Feature extraction was performed by the DWT (discrete wavelet transform) method and then reduced to 13 features. Eventually, a support vector machine (SVM) has been utilized to detect the cancerous mass. Simulation and test results have shown 98.03% accuracy.

Navid et al. [5] recommended a method that uses a threshold-based WCO optimization algorithm. WCO is a metaheuristic algorithm inspired by the FIFA World Cup challenge. Then, the Kapur approach was used to define an objective function. Finally, the candidate solutions were selected from random samples of the search space in the image histogram.

Toğaçar et al. [6] introduced novel deep learning which was developed based on the convolutional neural network. They proposed a model called BreastNet to improve the quality of the classification. The BreastNet was built based on attention modules. The data has been processed by augmentation techniques. The image features are exchanged by various augmentation techniques like shift, rotation, flip change, and brightness. Also, they used the hypercolumn method for the accurate classification of the data. Other sections of the BreastNet pattern model include the pooling, convolutional, dense blocks, and residual. The method obtained 98.80% accuracy.

Alanazi et al. [7] proposed a CNN method that analyzes the hostile ductal carcinoma tissue regions in whole-slide images (WSIs) to automatic detection of the breast cancer. The suggested system using CNN Model 3 obtains 87% accuracy. The five-layer CNN in Model 3 is best suited for this detection. The paper studies the presented technique that applies various convolutional neural network (CNN) architectures to automatically detect breast cancer, comparing the results with those from machine learning (ML) algorithms.

Ma et al. [8] presented that a 1D-CNN model was developed and trained for classification. The Fisher discrimination analysis (FDA) and support vector machine (SVM) classifiers were trained and tested with the same spectral data for comparison. The best classification performance, namely, the overall diagnostic accuracy of 92%, the sensitivity of 98%, and the specificity of 86%, has been achieved by using the 1D-CNN model. Table 1 indicates some recent studies in breast tumor identification.

In this study, our main purpose is to provide a twofold system for better diagnosis. The proposed approach includes a CNN model optimized with an improved algorithm, which is performed with a texture feature-based technique as a sequential method. Then, the results are combined to achieve the best result. This computer system reduces the complexity and leads to improving computational performance. Additionally, it solves the problems of the previous literature to achieve the best results [9]. Figure 1 is an overview of the proposed method.

The main contributions of this study are highlighted as follows:

(i) Optimal comprehensive approach for the automatic detection of breast cancer by the CAD

(ii) Hybrid method to improve the classification performance and efficiency

(iii) A noise reduction and normalization of the data

(iv) CNN classifier amended by Balanced Deer Hunting Optimization Algorithm (BDHOA) for classification

(v) Haralick texture and local binary pattern for feature extraction

Table 1: Recent studies in breast tumor identification.

Authors	Year	Approach	Dataset	Average accuracy (%)
Ibraheem et al. [4]	2019	Feature extracted with DWT and then reduced to 13 features Classification: CAD-SVM	Reference image database to evaluate response (RIDER)	98.03
Navid et al. [5]	2020	Segmentation WCO based on image thresholding method	MIAS	87.5
Toğaçar et al. [6]	2020	Classification: hypercolumn technique	BreakHis	98.80
Ibrahim et al. [3]	2020	Segmentation based on a chaotic Salp swarm algorithm (CSSA)	Mastology research with Infrared image (DMR-IR)	92
Hu et al. [2]	2020	Feature: DCE and T2w sequences Classification: CNN- SVM	mpMRI	95
Alanazi et al. [7]	2021	Classification: CNN- SVM	Kaggle 162 H&E	87
Ma et al. [8]	2021	Classification: 1D-CNN model	Spectral data	98

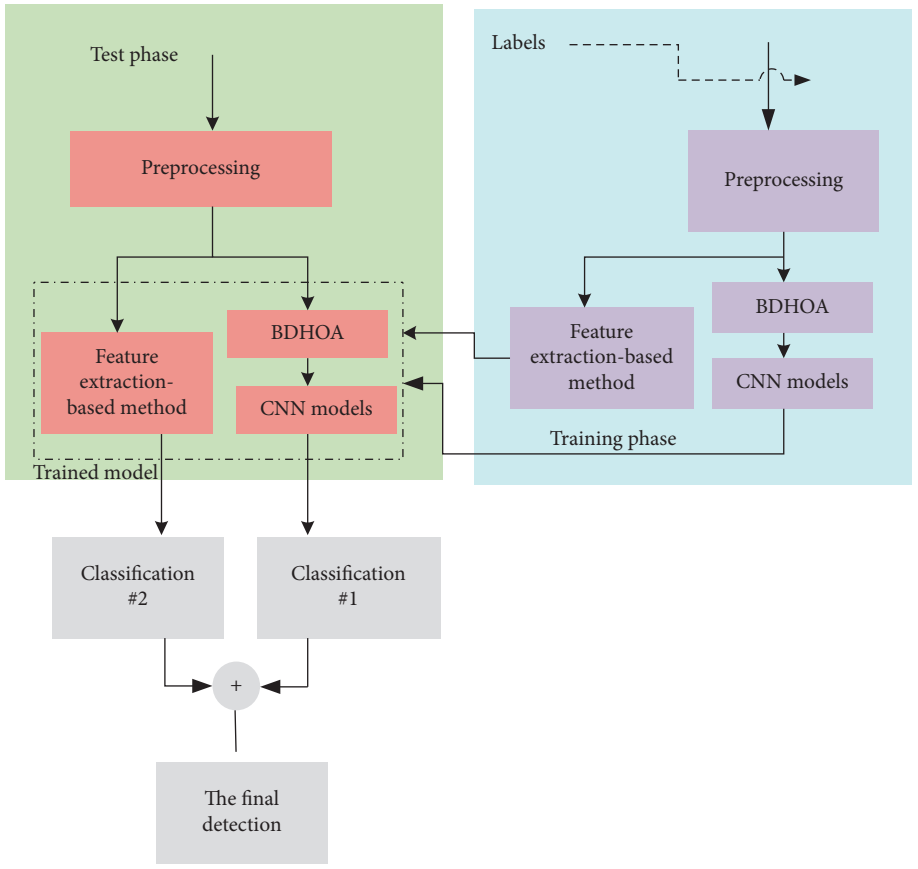


Figure 1: The graphical abstract of the proposed method.

(vi) Independent component analysis for dimension reduction of the features

2. Image Preprocessing

First, input breast MRI image data should be simplified and prepared for the next steps. Thus, in the first step, normalization is applied. Hence, the data intensity values are normalized by the min-max method in the scale range 0-1. Here, the size of the image is 250×250. Then, the noise reduction method is used to eliminate undesired distortions. Noise reduction is the most important phase of preprocessing. In recent studies, partial equations have been used to reconstruct images. Also, MR images have problems such as electromagnetic (EM) noise emitted from circuits. The main cause of noise in MRI imaging can be of two types: (1) hardware and (2) subject (physiological noise, body motions, cardiac pulsation, respiratory motions, etc.).

To overcome the noise problems of the breast MR images, they must be filtered. Acoustic noise is the main noise in MRI. So, noise removal is important in medical image processing. In this regard, an Intelligent Hybrid Filter is

used. This fuzzy-based filter is utilized to eliminate the noise of images. This filter is used in particular for the preprocessing of medical images [8]. The procedure of this filter is summarized as follows [7]:

(1) The noisy image is passed in parallel from four noise removal filters

(2) X is the input image, and X_0, X_1, X_2, and X_3 are the output of the filters

(3) The output of the filters enters the fuzzy-neural system as input

(4) Finally, Y is the output of the fuzzy-neural system and is the final improved image

3. Convolutional Neural Networks

CNN is an abbreviation form of convolutional neural networks. It is one of the branches of deep neural networks. Also, it is highly accurate in image processing, classification, and segmentation. The CNNs are mainly used in machine learning for visual or speech analysis and diagnosis.

Convolutional networks were inspired by biological processes of the (connections between neurons) cat's visual cortex. It is a significant approach in deep learning, wherein several layers are trained purposefully and powerfully [10]. It is more efficient because of its accuracy and fast operation. In computer vision, CNN is one of the most important methods.

Generally, all models of CNN contain three key parts: (1) convolution layer, (2) pooling layer, and (3) fully connected layer, where each layer has a definite task. Also, CNN has two training steps. At first, the input image has been injected into the CNN with a simple dot multiplication between the input and the neuron parameters, which is followed by a convolutional multiplication during each layer. The network error is computed from the output to network training. For this purpose, it compares the network output by using a loss error function and correct solution and computing the error rate. Then, the phase of backpropagation starts based on the amount of calculated error rate, where the derivative of each parameter is obtained using the chain rule.

Also, these components change based on their effect on the network error [9]. After updating the parameters, the next stage is feedforward. These steps are repeated several times in sufficient numbers until the network training is completed. The learning is used to get a certain number of kernel matrixes. In this case, gradient descent was utilized for the selection of the optimal network weights. In network, a ReLU (rectified linear unit) function with $f(z) = \max(z, 0)$ is utilized to activate neurons. The output scale is intensely reduced by max pooling [10].

The training error is evaluated to adapt the weight of the neuron and obtain the desired output. The backpropagation step minimizes the cross-entropy loss [11].

The loss is with the following formula:

$$\mathrm{CE} = \sum_{j=1}^{N}\sum_{i=1}^{M} -D_j^{(i)} \log z_j^{(i)}, \tag{1}$$

where D_j defines the achieved output vector for the m^{th} class such that $d_j = (0, \ldots, 0, \underbrace{1, \ldots, 1}_{k}, 0, \ldots, 0)$ explains the desired output vector and $z_j^{(i)}$ demonstrate the Softmax function which is shown by the following formula:

$$z_j^{(i)} = \frac{e^{f_j}}{\sum_{i=1}^{M} e^{f_i}}, \tag{2}$$

where M describes the sample number, and a weight penalty (ρ) has been utilized for extending the function by storing the values of the weight large; that is,

$$L = \sum_{j=1}^{N}\sum_{i=1}^{M} -D_j^{(i)} \log z_j^{(i)} + \frac{1}{2}\rho \sum_{K}\sum_{L} V_{k,l}^2, \tag{3}$$

where connection weight is indicated by V_k, k in layer l and L shows the layer's totality number, and K illustrates the layer for l connections. Given that the designed layouts of the CNN are based on tests and errors, it also has problems. In recent years, various types of optimal automatic approaches have been presented for extending the network using bioinspired optimization algorithms [12].

4. Deer Hunting Optimization Algorithm (DHOA)

One of the steps is optimization, which is the process of obtaining the "best available" values of a problem. Sometimes, conventional classical optimization algorithms are not able to solve problems correctly and quickly [13]. To overcome this issue, there is a new technique called metaheuristics for fast solutions of the problems such as NP-hard (nondeterministic polynomial-time hard). Metaheuristics can be imitated based on different phenomena from the animals hunting behavior to humankind's social behaviors [14]. In some cases, algorithms are also improved to find the best optimum response. For example, the harmony search algorithm, dolphin swarm algorithm, genetic algorithm, symbiotic organism search, and the world-cup optimization algorithm are used to solve the various types of complex problems [15]. Besides, Yin and Navid suggested a modern bioinspired algorithm that is inspired by hunting deer [16].

The deer's features make their hunting process more difficult. An important feature of deer is their vision. Its visual feature is five times stronger than man's vision. The other remarkable feature of the deer is its sense of smell. This sense in deer is sixty times stronger than human's smell sense.

The deer snores loudly and walks heavily when it realizes the danger of this. This reaction can let another deer know. The deer can also detect the extreme frequency of the sounds well. In the following text, the deer hunting system has been described in detail.

4.1. Initialization. The metaheuristic deer hunting algorithm starts with the set known as the hunter, which is a group of random populations, which is defined as follows:

$$Z = \{y_1, y_2, \ldots, y_n\}, \quad 1 < j \leq n, \tag{4}$$

where the number of hunters or types of solutions is indicated by n. Also, Z refers to the total hunter population.

4.2. Initializing the Parameters. The second stage involves quantifying the main components, the angle of the position of the deer and the angle of the wind. Space is considered a circle. Thus, the wind angle is written in the formula of a circle:

$$\theta_j = 2\pi\alpha, \tag{5}$$

where α is a random integer within the limitation $\{0, 1\}$ and j describes the present iteration. Also, the angle of the deer location is defined as follows:

$$\varphi_j = \pi + \beta. \tag{6}$$

Here, β shows the angle of the wind.

4.3. Position Propagation. During the first iteration, it is usually not possible to find the best solution for the algorithm [17]. However, after generating a random integer and evaluating the cost function from it, the best integer is considered as the optimal solution value [18]. Here, we assumed two parameters, including leader position (z_l) as the initial best location of the hunter and the successor position (z_s) as the succeeding hunter position.

4.3.1. Propagation Based on the Leader's Position. To get the best position using the initial repetition, the entire population tries to reach the best position by updating their location. Hence, the "encircling behavior" is mathematically formulated by the following equation:

$$Z_{j+1} = Z_L - k \times S^w \times \left| L \times Z_L - Z_j \right|. \tag{7}$$

In this formula, Z_j and Z_{j+1} indicate the present and later locations, S^w shows the random integer based on wind velocity in the scope [0, 2], and the coefficient vectors are denoted by L and k, where the formula is written as follows:

$$k = 0.25 \times \log\left(I + \frac{1}{I_{\max}}\right)\gamma, \tag{8}$$

$$L = 2 \times \delta, \tag{9}$$

where $I_{\max}$ illustrates the peak of repetition, and in the range $[-1, 1]$, γ is the random component.

δ is the random integer in the range from 0 to 1. Figure 2 presents the updating position of Z^*, where (Z, Y) shows the primary location of the hunter that can be updated depending on the prey location.

The updated status will be ongoing to achieve the best situation (Z^*, Y^*) based on the L and K. The hunters go to the place where the leader is located. If the leader's move was not successful, the hunter stays in his previous position.

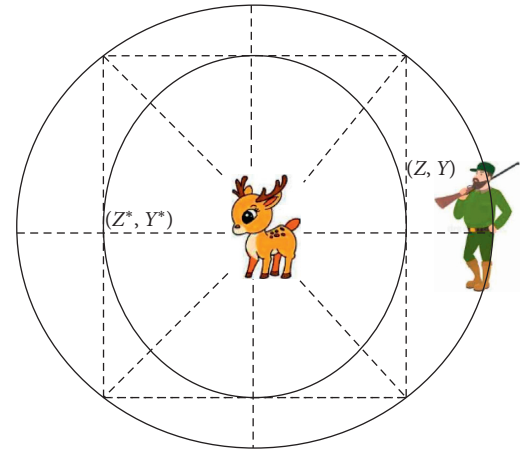


Figure 2: Updating the best position Z^*.

The updating of position is according to (9), when $S^w < 1$. Indeed, hunters can move in all directions regardless of the position angle. Therefore, according to (9) and (10), the hunters can update their locations in each random position.

4.3.2. Propagation Based on the Position Angle. Also, we can expand the space of solving way by considering the location angle. Angle assessment is very important to determine the position of the hunter. So, the successful attack should not be visible to the prey. The visualization of the deer angle (prey) formula is as follows:

$$u_j = \frac{1}{8} \times \pi \times \alpha. \tag{10}$$

Due to the difference between the angle of the wind and the angle at which the prey is seen, u is the parameter that is considered for updating the angle of position:

$$C_j = \beta_j - u_j, \tag{11}$$

where β illustrates the angle of wind blowing.

Then, to update the position angle parameter,

$$\varphi_{j+1} = \varphi_j + C_j. \tag{12}$$

After obtaining the angle of location, the new location can be calculated using the following formula:

$$Z_{j+1} = Z_j - S^w \times \left| \cos\left(\varphi_{j+1}\right) \times Z_l - Z_i \right|. \tag{13}$$

The prey does not see the hunter because of the view angle.

4.4. Propagation Based on the Position of the Successor. To use the exploration, it is possible to adjust L in the behavior of encircling. According to the random first search, the integer of vector L cannot be taken into account as more than 1. Thus, the successor location is used for providing a new update of the best solution. An exploration updating formula is given as follows:

$$Z_{j+1} = Z_s - k \times S^w \times \left| L \times Z_L - Z_j \right|, \tag{14}$$

where Z_s explains the successor location of hunters at any moment. In this algorithm, the location of the hunters is updated by the best solution in each repetition. The best solution is obtained while $|L| \geq 1$. If $|L| < 1$, one of the hunters is randomly selected. This method creates an L switch, which can modify the mode of the algorithm between the exploitation and exploration stages.

Stuck into the local optimum is a shortage of the original DHO algorithm [19]. In the following, a new modification has been proposed to recover this problem.

4.5. The Balanced DHO Algorithm. Here, Lévy flight (LF) is used to evolve the DHO algorithm. Lévy flight is a method that solves the problem of component convergence defect. Lévy flight creates a random walking system that helps to control the local search correctly. The formula is given in the following:

$$\begin{aligned} Le(D) &\approx D^{-1-\mu}, \\ D &= \frac{R}{|T|^{1/\mu}}, \\ \sigma^2 &= \left\{ \frac{\Gamma(1+\mu)}{\mu \Gamma((1+\mu)/2)} \frac{\sin(\pi\mu/2)}{2^{(1+\mu)/2}} \right\}^{2/\mu}, \end{aligned} \tag{15}$$

where $0 < \delta \leq 2$, $R \sim N(0, \sigma^2)$ and $T \sim N(0, \sigma^2)$, $\Gamma(.)$ presented the gamma, D describes the size of the step, the index for Lévy is denoted by μ, and $R/T \sim N(0, \sigma\ 2)$ describes that Gaussian distribution is used for generating the samples, where the mean value and the variance value are zero and σ2, respectively. Here, $\mu = 3/2$.

According to the Lévy flight system, the following equation is the new enhanced hunter location:

$$Z_{j+1}^l = Z_{j+1} + (Z^* - RG) \times Le(\rho), \tag{16}$$

where Z_{j+1}^l shows the new location for the agent of search Z_{j+1} and

$$\begin{aligned} R &= u(2 \times r - 1), \\ G &= EZ'(t) - Z(t), \end{aligned} \tag{17}$$

where u is limited within [0, 2] and r indicates a random integer at the range from 0 to 1. $Z'(t)$ represents a random location vector selected for the current population.

For providing guaranteed best solution candidates, fitter agents are kept:

$$\overrightarrow{G}_{el} = \begin{cases} Z_{j+1}^l & F\left(Z_{j+1}^l\right) > F\left(Z_{j+1}\right) \\ Z_{j+1} & \text{if not} \end{cases}, \tag{18}$$

The following diagram shows the balanced DHOA (BDHOA), which illustrates the steps of this process is given in Figure 3.

5. Validation of the BDHO Algorithm

Here, four benchmarks are proposed to analyze the (BDHO) algorithm. Also, several metaheuristic algorithms have been compared with BDHOA. To do this, the benchmarks have been validated on the balanced DHOA, ant colony optimization algorithm (ACO) [20], gray wolf optimization algorithm (GWO) [21], and grasshopper optimization algorithm (GOA) [22, 23], and particle swarm optimization (PSO) [24]. The original DHOA is also given in the table to show the capabilities of BDHOA.

We have simulated by Matlab R2016b with a laptop configuration of 2.20 GHz CPU and 6.00 GB RAM. In this section, the first benchmark function is Rastrigin. Its constraint is [−512, 512] and has the dimension of (30–50) that can be mathematically formulated as follows:

$$F_1(z) = 10 \times G + \sum_{j=1}^{G} \left(z_j^2 - 10 \times \cos\left(2\pi z_j\right) \right). \tag{19}$$

The second benchmark function is Rosenbrock, which is within [−2.045, 2.045] and has a dimension of 30 to 50. This benchmark can also be formulated as follows:

$$F_2(z) = \sum_{j=1}^{G-1} \left(100\left(z_j^2 - z_{j+1}\right) + \left(z_j - 1\right)^2 \right). \tag{20}$$

The third benchmark function with a dimension of 30–50 is within [−10, 10], which is called Ackley. The following formula is related to this benchmark:

$$\begin{aligned} F_3(z) = 20 + e - 20 \times \exp &\left(-0.2 \sqrt{\frac{1}{G} \sum_{j=1}^{G} z_j^2} \right) \\ &- \exp\left(\frac{1}{G} \sum_{j=1}^{G} \cos\left(2\pi z_j\right) \right). \end{aligned} \tag{21}$$

Sphere benchmark function is the fourth which has [−512, 512] constraint and 30–50 dimensions.

The formula is shown in the following:

$$F_4(z) = \sum_{j=1}^{G} z_j^2. \tag{22}$$

The comparison result, according to (1) mean deviation (MD) and (2) standard deviation (SD), has been demonstrated in Table 2.

The table above shows that the mean deviation and standard deviation in the BDHOA method are less and this result is convenient. Also, it can be observed that BDHOA gives the best results compared with the original DHOA. Due to that, it can be useful for obtaining an optimum solution.

5.1. Breast Tumor Classification Based on the Proposed Method. The CNN training is mostly performed according to the backpropagation. To overcome the issue of stuck in

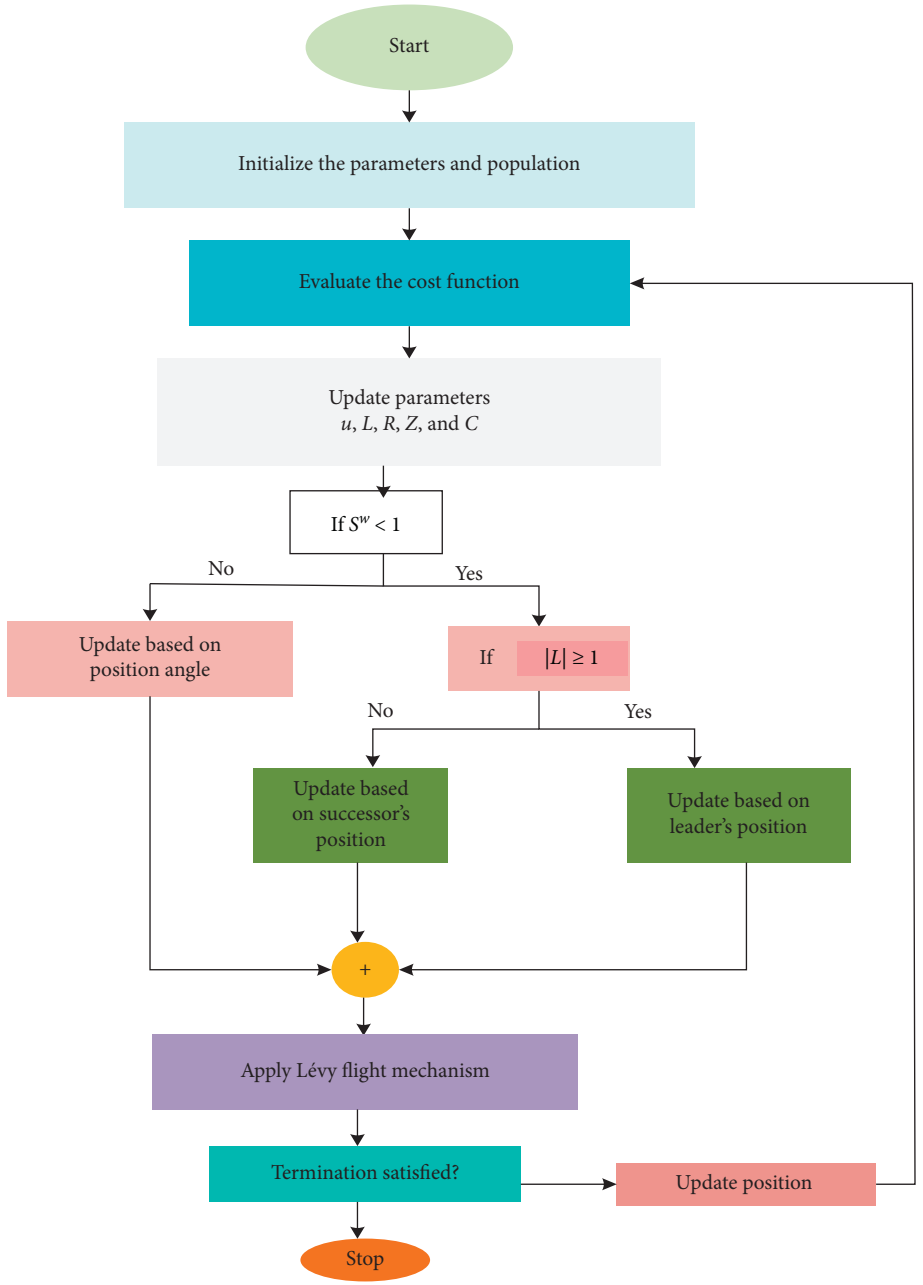


Figure 3: The flowchart diagram of the improved BDHO algorithm.

the local optimum, several methods are presented. In this section, to reduce the network error, the proposed BDHO algorithm is employed instead of the backpropagation approach. The purpose of using this metaheuristic algorithm for CNN is to minimize the value of the mean square error (MSE) function. The formula of the MSE function is as follows:

$$\mathrm{MSE} = \frac{1}{T}\sum_{j=1}^{M}\sum_{i=1}^{N}\left(a_j^i - b_j^i\right)^2, \tag{23}$$

where a_j^i represents the j^{th} for network output and b_j^i illustrates the j^{th} of the desired integer of the CNN during the t period. N signifies the value for layers of the output and M indicates the data number in the formula. CNN technique can be very useful in the rapid detection of breast tumors in MR images. In this study, CNN classification is used with two models of CNN and different classifiers [5]. This classification includes (1) extraction of the features and (2) features dimension decrease. This is briefly exhibited in the following.

6. Extracting Features

In image processing, feature extraction means converting image data into usable information for the next stages [25]. This is performed by extracting some general or particular features of the input image [26]. Among some types of

Table 2: Analysis of compared methods by dimensions of 30.

Benchmark		BDHOA	DHOA	ACO [14]	GWO [15]	Goa [16]	PSO [24]
f_1	MD	0.00	3.20	67.19	69.80	1.58	4.58
	SD	0.00	3.6	1.50	9.12	3.14	1.13
f_2	MD	6.53	9.58	40.12	89.3	15.22	12.43
	SD	1.88	2.26	24.10	46.15	5.47	2.71
f_3	MD	0.00	$5.23e-16$	$4.11e-3$	9.47	$2.92e-4$	$7.12e-9$
	SD	0.00	$2.8e-6$	$1.58e-2$	3.10	$1.5e-3$	0.00
f_4	MD	0.00	$1.32e-9$	$1.25e-4$	$5.29e-5$	$7.26e-9$	$8.97e-12$
	SD	0.00	$15.30e-18$	$3.65e-5$	$4.89e-5$	$2.73e-9$	$10.14e-15$

feature extraction techniques, the texture technique gives more information with details on the spatial arrangement and intensities of colors. Also, this method has some fans in medical imaging. To do feature extraction in this research, two features are used: (1) Haralick features and (2) local binary pattern (LBP). In the following, these two methods are briefly explained.

6.1. Local Binary Pattern Features. The operator of LBP selects binary integers of the pixels, then compares these values with their neighbor pixels and decimal numbers, and finally, encodes the surrounding local structure of each pixel.

In the binary labeling step, the resultant strictly negative have been encoded with 0 and the other values encoded with 1. The achieved binary numbers (codes) from the LBP feature are in the clockwise direction. The final extracted binary values are assumed as local binary patterns codes. The final values extracted are binary and called local binary patterns or LBP codes.

6.2. Haralick Texture Features. The Haralick feature is a statistical feature that is evaluated from the gray level cooccurrence matrices (GLCM). The purpose is to evaluate the matrix and computes the neighboring gray level cooccurrence in the input image. The GLCM explains the information about a square matrix in the region of interest (ROI) that illustrates a correlation among the reference pixel with a presented intensity integer and the pixels around it that are located in various directions. In this study, four directions, in 0°, 45°, 90°, and 135°, have been employed for the pixels, and the average integers have been used as last the Haralick features.

6.3. Dimension Reduction of the Features Based on ICA. At this stage, given that the data volume in the feature is high, to achieve the desired volume, the data reduction method is used which also leads to simplification. To do this, feature dimensions are reduced by the independent component analysis (ICA) method. The ICA is a computational methodology for tight-fitting of the concealed factors, which underlie a series of signals. ICA introduces a reproductive model for the large database of MRI. By considering the ICA property, which is a method to separate the blind source, and assuming that the subfactors are non-Gaussian signals and also these subcomponents are free (independent) from each other, the ICA is a very powerful algorithm for analyzing and evaluating principal parameters. The difference is that ICA's ability is to find the underlying sources, even if the classic methods lead to failing. In this algorithm, measurements are given as an array of time series. The phrase Blind Source Separation (BSS) is employed for characterizing the breast waves recorded by several sensors. Finally, the input of the classifier is the data image, which is divided into two sections: training and testing images. After injecting into the classifier, the classifier trains them and predicts the label for test images. Figure 4 indicates a diagram of the feature extraction-based method.

6.4. Final Simulations. The final step is to classify the obtained results from the proposed feature-based techniques and the CNN model. The main goal is to propose an accurate and efficient method to detect breast tumors from MR images. The approach of classification is briefly described in the following.

At the first step, the diagnosis results of hybrid technique (proposed CNN and feature extraction-based) have been collected. After that, the results of the suggested CNN are checked out. If the output is labeled as a tumor, the output will be exhibited as cancer. Otherwise, if the output of the presented CNN is labeled as healthy, the features of the MR image are checked out again according to the feature extraction-based method. In this condition, if the output image has been diagnosed as a tumor, the output will be labeled as cancer, or else, it will be diagnosed as healthy.

7. Results and Discussion

7.1. Database Description. This method aims to quickly detect the breast tumors in MRI by using MATLAB R2016b software with a system configuration of 2.20 GHz CPU and 6.00 GB RAM. The main idea is to design an optimized CNN (convolutional neural network) to achieve promising results. To validate this technique, it is implemented on the DCE-MRI dataset, which is usually used for analyses of classification efficiency. The DCE-MRI dataset includes a set of 219 breast MR images that is collected from 105 different patients with breast cancer (angiosarcoma, inflammatory, DCIS, ILC, and LCIS) (55 tumor-like and 50 non-tumor-like malignant lesions), and 114 DCE-MRI were detected to be normal. In MATLAB, image size is 512×512 pixels. The presented scheme (optimized CNN) and feature extraction-based method are designed for analyzing the MR images.

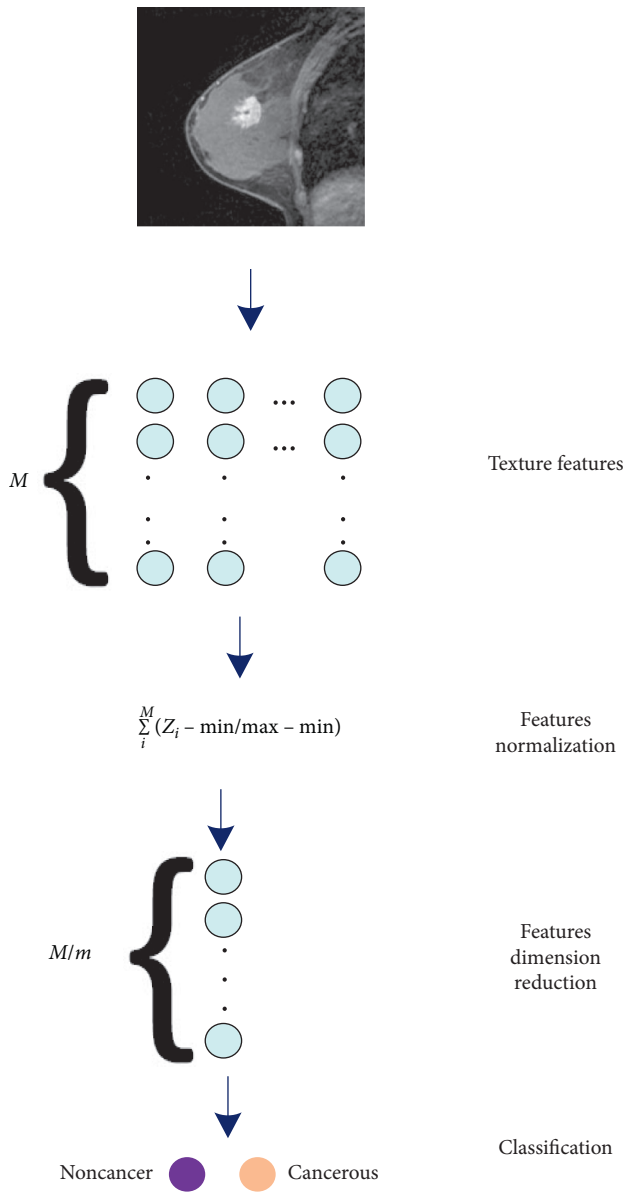


Figure 4: A schematic of the feature extraction-based method.

7.2. Simulation Results. There are several types of performance analysis to evaluate classification. One of these analyzes is accuracy. The accuracy determines the proportion of the correctly classified image number to the total image number. The results of the analysis of accuracy for the studied methods, including the feature-based method, optimized CNN, and the hybrid method (feature-based and optimized CNN), are indicated in Table 3.

As can be observed from Table 3, the data presents that the highest efficiency has been achieved when hybrid feature-based/optimized CNN has been utilized for classification. The accuracy of classification can be considered as an efficient indicator to determine the performance of the method when the test dataset contains equal numbers of samples from the classes. The results indicate the efficiency of the proposed system in the rapid diagnosis and timely treatment of the patient.

Table 3: Results of the accuracy according to the methods in breast tumor diagnosis on the DCE-MRI database.

Method	Accuracy	Variance (%)
Feature-based	93.15	0.6
Optimized CNN	98.65	0.5
Hybrid feature-based/optimized CNN	98.89	0.4

Table 4: Confusion matrix of the hybrid feature-based and optimized CNN structure.

		Predicted		
		Angiosarcoma	Inflammatory	DCIS
Actual	Angiosarcoma	721	12	10
	Inflammatory	27	1393	3
	DCIS	3	0	896

To get more evaluations, confusion matrices have been used for performance analysis of the breast tumor classification. A confusion matrix is a table with two dimensions which is usually employed for determining the classification efficiency and performance, on a test set to define the true values. Table 4 illustrates a sample confusion matrix for hybrid feature-based/optimized CNN. This table is based on an investigation on breast cancer: angiosarcoma, inflammatory, and ductal carcinoma in situ (DCIS).

Several indicators have been used for determining the efficiency of the classifier, particularly for each cancer tumor class [27]. The critical indicators in the classification report are specificity, precision, and sensitivity which are obtained from the following formulas:

$$\begin{aligned} \text{precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{sensitivity} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{specificity} &= \frac{\text{TN}}{\text{TP} + \text{FP}}, \end{aligned} \tag{24}$$

where the following parameters show some classified cases: FP, false positives; TP, true positives; FN false negatives; and TN, true negatives.

Table 5 shows the final results of using the proposed technique once the optimized and the feature-based method classifier are together for the detection goal.

In Table 5, the integer of the specificity for all the datasets is high, which illustrates the correct identifying samples without the specific disease. The proposed method is also compared with two types of well-known methods. Table 6 provides a comprehensive comparison, according to some different state of the art for classification techniques [28]. In Table 6, it is clear that the precision parameter in the proposed method is better and higher than other methods.

Briefly, it is observed that once using the suggested method, the value of system efficiency indicators (precision, sensitivity, and specificity) is increased.

Table 5: Results of the proposed technique for detection.

Tumor type	Precision	Sensitivity	Specificity
Angiosarcoma	95.16	95.79	97.42
Inflammatory	97.00	97.07	98.55
DCIS	98.25	98.30	99.15

Table 6: The comparison of precision for comparative approaches on the DCE-MRI dataset.

	Angiosarcoma	Inflammatory	DCIS
Mahmuda Rahman	93.06	90.98	96.50
Amira Mofreh Ibraheem	92.70	97.09	92.84
Proposed method	93.99	98.00	96.95

8. Conclusions

A new comprehensive approach was proposed for the automatic detection of breast tumors. The method is a hybrid model, including an optimized design of a convolutional neural network and feature extraction-based technique to improve the classification efficiency. In this study, preprocessing steps are applied, which eliminate noise and simplifies classification. Additionally, it leads to an increase in the quality of the dataset. Thus, the value is also normalized. Also, the feature extraction-based method was based on Haralick texture features. This method was used with independent component analysis (ICA) to reduce the dimension of the features. Simulations were performed according to the DCE-MRI dataset. The results were compared by various states of the method. Also, other methods were compared to indicate the system's efficiency. It is also possible to increase the accuracy of the study by using a variety of other metaheuristic algorithms. Furthermore, deep convolutional neural network model can be used for further research, to classify the breast cancer images. The final satisfactory results stated the advantage of the suggested approach toward the other methods. In the future, we will examine the proposed technique on a different dataset. The proposed method can be generalized to the design of high-performance computer-aided diagnosis systems for other medical imaging tasks in the future.

Data Availability

The dataset images can be found at https://wiki.cancerimagingarchive.net/display/Public/QIN+Breast+DCE-MRI.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Sung, J. Ferlay, R. L. Seigal et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal For Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[2] Q. Hu, H. M. Whitney, and M. L. Giger, "A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI," *Scientific Reports*, vol. 10, no. 1, pp. 10536–10611, 2020.

[3] A. Ibrahim, S. Mohammed, H. A. Ali, and S. E. Hussein, "Breast cancer segmentation from thermal images based on chaotic Salp swarm algorithm," *IEEE Access*, vol. 8, pp. 122121–122134, 2020.

[4] A. M. Ibraheem, K. H. Rahouma, and H. F. Hamed, "Automatic MRI breast tumor detection using discrete wavelet transform and support vector machines," in *Proceedings of 2019 Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, November 2019.

[5] R. Navid, V. V. Estrela, and H. J. Loschi, "Entropy-based breast cancer detection in digital mammograms using world cup optimization algorithm," *International Journal of Swarm Intelligence Research (IJSIR)*, vol. 11, no. 3, pp. 1–18, 2020.

[6] M. Toğaçar, K. B. Ozkurt, B. Ergen, and Z. Comert, "BreastNet: a novel convolutional neural network model through histopathological images for the diagnosis of breast cancer," *Physica A: Statistical Mechanics and Its Applications*, vol. 545, Article ID 123592, 2020.

[7] S. A. Alanazi, M. M. Kamruzzaman, M. N. I. Sarkar et al., "Boosting breast cancer detection using convolutional neural network," *Journal of Healthcare Engineering*, vol. 2021, Article ID 5528622, 11 pages, 2021.

[8] D. Ma, L. Shang, J. Tang, Y. Bao, J. Fu, and J. Yin, "Classifying breast cancer tissue by Raman spectroscopy with one-dimensional convolutional neural network," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 256, Article ID 119732, 2021.

[9] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.

[10] M. Heidari, "An optimal machine learning model for breast lesion classification based on random projection algorithm for feature optimization," in *Proceedings of Medical Imaging 2021: Imaging Informatics for Healthcare, Research, and Applications*, February 2021.

[11] V. Vijayasarveswari, "Reliable early breast cancer detection using artificial neural network for small data set," in *Journal of Physics: Conference Series*vol. 1755, IOP Publishing, 2021.

[12] S. Parvathavarthini and D. Deepa, "A hybrid artificial neural network classifier based on feature selection using binary dragonfly optimization for breast cancer detection," in *Proceedings of IOP Conference Series: Materials Science and Engineering*, April 2021.

[13] Y. Meraihi, A. B. Gabis, S. Mirjalili, and A. Ramdane-Cherif, "Grasshopper optimization algorithm: theory, variants, and applications," *IEEE Access*, vol. 9, pp. 50001–50024, 2021.

[14] Z. Yuan, W. Wang, H. Wang, and M. Ashourian, "Parameter identification of PEMFC based on Convolutional neural network optimized by balanced deer hunting optimization algorithm," *Energy Reports*, vol. 6, pp. 1572–1580, 2020.

[15] D. Izci, "A novel improved atom search optimization algorithm for designing power system stabilizer," *Evolutionary Intelligence*, pp. 1–15, 2021, inpress.

[16] Z. Yin and R. Navid, "PEMFC identification using deep learning developed by improved deer hunting optimization algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 40, no. 2, 2020.

[17] N. Razmjooy, Z. Vahedi, V. V. Estrela, R. Padilha, and A. C. B. Monteiro, "Speed control of a DC motor using PID

controller based on improved whale optimization algorithm," in *Metaheuristics and Optimization in Computer and Electrical Engineering*, pp. 153–167, Springer, Berlin, Germany, 2021.

[18] R. Navid, "A new design for robust control of power system stabilizer based on moth search algorithm," in *Metaheuristics and Optimization in Computer and Electrical Engineering*, pp. 187–202, Springer, Berlin, Germany, 2021.

[19] H. Ketabi, A. Ekhlasi, and H. Ahmadi, "A computer-aided approach for automatic detection of breast masses in digital mammogram via spectral clustering and support vector machine," *Physical and Engineering Sciences in Medicine*, vol. 44, no. 1, pp. 1–14, 2021.

[20] M. Hamim, I. L. Moudden, M. D. Pant et al., "A hybrid gene selection strategy based on fisher and ant colony optimization algorithm for breast cancer classification," *International Journal of Online and Biomedical Engineering*, vol. 17, no. 2, 2021.

[21] S. Kumar and M. Singh, "Breast cancer detection based on feature selection using enhanced grey Wolf optimizer and support vector machine algorithms," *Vietnam Journal of Computer Science*, vol. 8, no. 2, pp. 177–197, 2021.

[22] N. Safdarian and S. Yousefian Dezfoulinejad, "Mammographic image processing for classification of breast cancer masses by using support vector machine method and grasshopper optimization algorithm," *Journal of Arak University of Medical Sciences*, vol. 23, no. 2, pp. 246–263, 2020.

[23] R. Navid, "Skin color segmentation based on artificial neural network improved by a modified grasshopper optimization algorithm," in *Metaheuristics and Optimization in Computer and Electrical Engineering*, pp. 169–185, Springer, Berlin, Germany, 2021.

[24] R. Habibi, "Svm performance optimization using PSO for breast cancer classification," *Budapest International Research in Exact Sciences (BirEx) Journal*, vol. 3, no. 1, pp. 28–41, 2021.

[25] M. M. Mafarja and S. Mirjalili, "Hybrid whale optimization algorithm with simulated annealing for feature selection," *Neurocomputing*, vol. 260, pp. 302–312, 2017.

[26] G. Guo and N. Razmjooy, "A new interval differential equation for edge detection and determining breast cancer regions in mammography images," *Systems Science & Control Engineering*, vol. 7, no. 1, pp. 346–356, 2019.

[27] Z. Vahedi and R. Seyfi, "An image watermarking approach to combat geometric attacks using hybrid DWT, DCT and SVD Method," *World Essays Journal*, vol. 7, no. 1, pp. 36–45, 2019.

[28] A. Z. Khuzani, "Applying a new feature fusion method to classify breast lesions," in *Medical Imaging 2021: Computer-Aided Diagnosis*vol. 11597, International Society for Optics and Photonics, 2021.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 6078524, 11 pages
https://doi.org/10.1155/2021/6078524

Research Article

Lung Cancer Diagnosis Based on an ANN Optimized by Improved TEO Algorithm

Rong Shan[1] **and Tahereh Rezaei**[2]

[1]*College of Computer, Weinan Normal University, Weinan, Shaanxi, China*
[2]*Neuroscience Research Center, Shiraz, Iran*

Correspondence should be addressed to Tahereh Rezaei; dr.tahereh.rz@gmail.com

Received 24 May 2021; Revised 22 June 2021; Accepted 9 July 2021; Published 16 July 2021

Academic Editor: Navid Razmjooy

A quarter of all cancer deaths are due to lung cancer. Studies show that early diagnosis and treatment of this disease are the most effective way to increase patient life expectancy. In this paper, automatic and optimized computer-aided detection is proposed for lung cancer. The method first applies a preprocessing step for normalizing and denoising the input images. Afterward, Kapur entropy maximization is performed along with mathematical morphology to lung area segmentation. Afterward, 19 GLCM features are extracted from the segmented images for the final evaluations. The higher priority images are then selected for decreasing the system complexity. The feature selection is based on a new optimization design, called Improved Thermal Exchange Optimization (ITEO), which is designed to improve the accuracy and convergence abilities. The images are finally classified into healthy or cancerous cases based on an optimized artificial neural network by ITEO. Simulation is compared with some well-known approaches and the results showed the superiority of the suggested method. The results showed that the proposed method with 92.27% accuracy provides the highest value among the compared methods.

1. Introduction

The proliferation of lung diseases in today's industrialized societies doubles the need for modern methods of accurate and early diagnosis. Among lung diseases, lung cancer is still recognized as one of the most dangerous cancers. Cancer means the abnormal growth, and sometimes proliferation of cells in the body. All cancers have an uncontrolled growth pattern and a tendency to detach from the source and metastasize [1]. A normal lung cell may become a lung cancer cell for no apparent reason, but in most cases, the transformation is the result of repeated exposure to carcinogens such as alcohol and tobacco. The appearance and function of cancer cells are different from normal cells. A mutation or change in the DNA or genetic material of a cell occurs [2]. DNA is responsible for controlling the appearance and function of cells. When a cell's DNA changes, that cell differentiates from the healthy cells next to it and no longer does the body's normal cells. This altered cell separates from its neighboring cells and does not know when it should stop growing and die [3]. In other words, the altered cell does not follow the internal commands and signals that other cells are in control of and acts arbitrarily instead of coordinating with other cells [4]. One-third of all cancer deaths are due to lung cancer. About 80% of patients have five years left in their best condition after being diagnosed with this type of cancer [5]. Based on a survey by the American Cancer Society (ACS), lung cancer in both men and women is the second most prevalent cancer in the United States [6]. Approximately 228,820 new cases and approximately 135,720 lung cancer deaths, based on ACS figures, will occur [5]. Figure 1 shows the statistical information for the lead cancers and their death number in 2019 based on the American Cancer Society of lung cancer screening [7].

Air pollution due to the industrialization of cities, tobacco use, and genetic factors are the main causes of these diseases [8]. Early diagnosis of lung disease will have a major

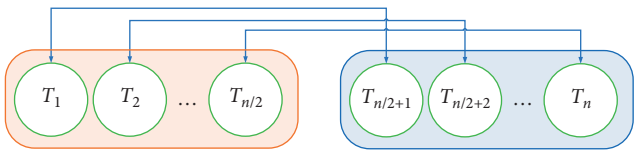


Figure 1: The heat transfer groups and the environment and cooling objects pairs.

impact on the possibility of definitive treatment of the disease. Major diagnostic methods for lung cancer include radiographic imaging and computed tomography, biopsy, bronchoscopy, and examination of cells in the sputum. Meanwhile, the CT scan imaging method is widely used as a superior diagnostic method. In this diagnostic method, the doctor examines possible nodules on the images. A pulmonary nodule is a small, round, opaque mass that forms inside the lung tissue [9, 10]. In other words, nodules are spherical radiographic opacities of less than three centimeters in diameter.

Formerly, lung diseases were diagnosed based on the help of experts' eye ability with no use for computer science. However, recently based on the different imaging techniques based on computer science and artificial intelligence, the diagnosis can be more precise. In most of these methods, after capturing the images from the patient, different image processing methods have been performed for tumor diagnosis [11].

2. Image Preprocessing

The analyzed images for validation in this study are collected from the *Lung CT-Diagnosis database* which has been provided by the Cancer Imaging Archive [12]. This database contains a collection of publicly available medical images for different cancers [13] with contrast-enhanced CT scan images stored in the Dicom format. This database is collected from 61 patients such that 4682 different images have been acquired from them.

After image acquisition, the min-max normalization method is established for them to scale the acquired images between 0 and 1. This study uses 250×250 scale normalization for this purpose. By considering a grayscale image with n dimension which has the following limitation: $I_n = [X \subseteq R^n] \longrightarrow [a, \ldots, b]$, the normalized image, I^*, can be achieved by the following formulation:

$$I^* = a_{\text{new}} + \frac{b_{\text{new}} - a_{\text{new}}}{b - a} \times (I - a), \tag{1}$$

where $I^* = [X \subseteq R^n] \longrightarrow [a_{\text{new}}, \ldots, b_{\text{new}}]$, a and b describe the intensity values in the grayscale image, and a_{new} and b_{new} describe the intensity values for the normalized image [14].

Also, since the reason that all CT scan images have a kind of visual noise, they need a denoising filter to resolve the problem. A CT scan image faces different factors and includes different types, such as Gaussian noise, Shot Noise, Poisson Noise, Speckle Noise, and Salt and Pepper Noise [15]. Noise hides slight details of the CT scan image. This shows that we need a tool for noise removal before starting the CT scan images.

One of the popular noise reduction techniques for CT scan images is mean filtering. The definition of the mean (average) filter works on averaging any aspect of the picture to the neighbors [16]. The average filter calculates and divides the sum of all the pixels in the filter window by the total number of pixels [4]. It then replaces the value of the center pixel with the calculated average. The result value for each indexed pixel (i, j) is determined as follows:

$$R(i,j) = \frac{[I(i-1,j-1) + I(i-1,j) + I(i-1,j+1) + I(i,j-1) + I(i,j) + I(i,j+1) + I(i+1,j-1) + I(i+1,j) + I(i+1,j+1)]}{9}. \tag{2}$$

3. Image Segmentation

One of the most important issues in image processing is the identification and separation of the image into its main components. Image segmentation determines the success or eventual failure of image analysis methods. And yet, due to its wide application, it has suitable research fields. The accuracy of this study in fields such as medicine is very important to preserve and protect human life. Thresholding is one of the most convenient methods for image segmentation [17]. By applying a thresholding method to a grayscale image, a binary image is obtained that delineates the boundaries of the objects in the image with appropriate accuracy.

The lower the threshold is, the more errors are detected and the more sensitive the results are to noise and unrelated image features [18]. On the other hand, a high threshold may miss weak errors or parts of errors. Since the main purpose of the fabric fault detection system is to find all possible faults, the threshold must be chosen to achieve this goal [19]. There are several methods for this practice in image processing science, and some separation methods are used for specific images [20]. One way to select the appropriate threshold is to use the trial-and-error method, in which different values of the threshold are selected and the image resulting from the application of this threshold is judged by the viewer. The simplest type of separation is called general separation, which is based on the image histogram. The input of this function is a gray image or a color image [21]. Its output is also a black and white image (binary). In this method, the threshold value at any point in the image is defined based on the local properties of the image in the neighborhood of that pixel. In this paper, Kapur thresholding has been used. Assume that the gray levels in an

image with N pixels and L gray levels are in the range $[0, 1, \ldots, L-1]$. Kapur method obtains the thresholds based on Kapur entropy maximization and based on the information obtained from the histogram of gray surfaces, which is defined as follows in the case of two-level segmentation (thresholding):

$$\text{Maximize } f_1(t) = H(0,t) + H(t,L), \tag{3}$$

such that

$$\begin{aligned} H(0,t) &= -\sum_{i=0}^{t-1} \frac{p_i}{w_o} \ln \frac{p_i}{w_o}, \quad w_o = \sum_{i=0}^{t-1} p_i, \\ H(t,L) &= -\sum_{i=0}^{t-1} \frac{p_i}{w_i} \ln \frac{p_i}{w_i}, \quad w_i = \sum_{i=0}^{t-1} p_i. \end{aligned} \tag{4}$$

The value of the optimal threshold point is the amount of gray area t that maximizes the function. For correct diagnosis of the tumors, the cancerous region should be detected with high precision. As mentioned before, image thresholding has been used for identifying the cancerous region. However, mathematical morphology is needed to better filter the detected areas based on image thresholding [22]. Mathematical morphology algorithm is a new technique for processing and analyzing signals and images. The basic idea of this technique is based on the analysis of geometric information by exploring an image with a small geometric pattern called a structuring element. This study uses three popular techniques including opening, closing, and filling holes. The first operator in the region filling is based on complementation, intersections, and dilation operators and is achieved by the following equation:

$$X_k = \left(X_{k-1} \oplus B\right) \cap A^c, \quad k = 1, 2, 3, \ldots, \tag{5}$$

where A indicates a group of boundaries and B defines the organizing element. This operator will end if $X_k = X_{k-1}$.

The mathematical opening is the second utilized operator. The opening of element A with element B has been obtained with A erosion by structure element B, followed by a dilation of the resulted image by structure element B, that is,

$$A \circ B = (A \ominus B) \oplus B. \tag{6}$$

The key goal of the mathematical opening is to remove the minor blemishes in the area that can be missed during the diagnosis of lung cancer.

Finally, the mathematical closing operator is used to smooth the counters, fuses narrow breaks, eliminates small holes, and fills gaps in the contour and long thin gulfs. This operation is formulated as follows:

$$A \cdot B = (A \oplus B) \ominus B. \tag{7}$$

4. Features Extraction

The purpose of feature extraction is to make raw data more usable for future statistical processing. Feature extraction is a very common process in different types of data processing such as image processing and audio processing. Feature extraction means selecting a feature that can describe the image with little information. These features must have properties so that a set of these features is described uniquely in each image. If a set of these attributes are the same for two samples, then you will not be able to distinguish two samples with any classifier in the classification section.

The main reasons for extracting features from images are image simplification, reduced processing time and memory, and increased accuracy and efficiency. So, feature extraction is a process in which data is mapped in a high-dimension space to a lower dimension space. This mapping can be linear (such as principal component analysis) or nonlinear. How to select these features requires data properties to be examined, and to extract it, preprocessing operations and various filters must be applied to the image to turn the image into the desired information. In this study, GLCM features have been used for extracting the lung cancer images information.

The GLCM method is one of the most efficient techniques for extracting tissue from medical images. This matrix is a square matrix with dimensions $N \times N$ where N is the number of degrees of gray in the image. Each element of this matrix represents the number of pairs of pixels that have degrees of gray on the surface of the image and are spaced in a certain direction from each other and to a certain pixel distance. After calculating the matrix, different parameters of the image texture can be extracted from it. In this study, the mentioned technique was used to extract tissue in lung tumor images. In the following, the utilized features have been explained.

4.1. Contrast. The contrast regulates the intensity value of the pixel and its neighbor in the image. This feature is achieved by the following equation:

$$\text{Contrast} = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (i-j)^2 f(i,j). \tag{8}$$

4.2. Correlation. The correlation feature describes the dependency on spatial features among the pixels. This feature can be mathematically given as follows:

$$\text{Cor} = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (i,j) f(i,j) - \mu_i \mu_j}{\sigma_i \sigma_j}. \tag{9}$$

4.3. Homogeneity. Homogeneity is a local uniformity feature that makes single/multiple intervals govern the textured and nontextured characteristics. This feature is achieved by the following:

$$H = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \frac{1}{1 + (i-j)^2} f(i,j). \tag{10}$$

4.4. Energy. The energy feature regulates the number of repetitive pixel pairs. This feature is mathematically obtained by the following equation:

$$\text{Energy} = \sqrt{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} f^2(i,j)}. \tag{11}$$

4.5. Entropy. The entropy is a feature that indicates the image selected interference based on the following equation:

$$\text{Ent} = -\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} \log_2 f(i,j). \tag{12}$$

5. Features Selection

Some of the different extraction characteristics in the images are so important and crucial for classification. In the meantime, those features that contain information which is not notable can have high potential when combined with other features. Any of these attributes could also have no useful data at all. This shortcoming can be resolved by different works. In this paper, an optimization-based methodology has been proposed for this purpose. The optimization method for the features selection can be achieved by the following equation:

$$\text{FF} = \frac{(\text{TP}\times\text{TN}) - (\text{FP}\times\text{FN})}{((\text{TN}+\text{FP})\times(\text{TP}+\text{FP})\times(\text{TP}+\text{FN})\times(\text{TN}+\text{FN}))^{(1/2)}}, \tag{13}$$

where TP signifies the true positive, TN describes the true negative, FN represents the false negative, and FP defines the false positive.

6. Improved Thermal Exchange Optimization Algorithm

In Thermal Exchange Optimization (TEO), the temperature of the objects indicates the individual position and with objects grouping, it is started to be exchanged. Therefore, new temperatures indicate their updated positions [23].

6.1. Newton's Law of Cooling. In the seventeenth century, the English scientist Isaac Newton studied the cooling of objects. The experiments showed that the cooling rate was approximately proportional to the temperature difference between the heated object and the environment. This fact is written as a differential relation:

$$\frac{dQ}{dt} = \alpha \times A \times (T_s - T_b), \tag{14}$$

where Q describes the heat, A signifies the area of the body surface that transmits heat, T_b defines the body temperature, T_s represents the ambient temperature, and α determines the heat transfer coefficient which is dependent on the geometry of the object, surface state, heat transfer mode, and other factors.

The heat loss in time dt is $\alpha \times A \times (T_s - T)dt$ which defines the alteration in stored heat as the temperature falls dT, i.e.,

$$V \times \rho \times c \times dT = -\alpha \times A \times (T - T_b)dt, \tag{15}$$

where V defines the volume (m^3), ρ describes the density (kg/m^3), and c signifies the specific heat (J/kg/K).

Therefore,

$$\frac{T - T_b}{T_M - T_b} = \exp\left(\frac{-\alpha \times A \times t}{V \times \rho \times c}\right), \tag{16}$$

where T_M is the starting high temperature.

The above equation is valid when $(\alpha \times A \times t)/(V \times \rho \times c)$ is not a function of T, i.e.,

$$\gamma = \frac{\alpha \times A}{V \times \rho \times c}, \tag{17}$$

where γ is constant. Therefore,

$$\frac{T - T_b}{T_M - T_b} = \exp(-\gamma t). \tag{18}$$

And finally, the equation can be rewritten as follows:

$$T = (T_M - T_b) \times \exp(-\gamma t) + T_b. \tag{19}$$

6.2. Inspiration. During the TEO algorithm, some individuals have been considered as the cooling objects and the residual individuals have been considered as the environment; then it has been in reverse. The method of simulation of the TEO algorithm is given in the following.

The first step is initialization. The initial temperature for all of the objects has been defined in an m-dimensional solution space as follows:

$$T_i^0 = T_{\min} + \text{rand}(T_{\max} - T_{\min}), \tag{20}$$

where T_i^0 signifies the initial solution vector of the object number i, rand describes a random vector with components in the range $[0, 1]$, $T_{\min}$ and $T_{\max}$ describe the minimum and the maximum limitations for the decision variables.

After initializing the objects, the value of the objective function for all of the individuals is evaluated. During the process, some historically best T vectors have been stored in a memory called Thermal Memory (TM) to use their position to develop the algorithm efficiency with no extra computational cost. After selecting some best values in TM, they have been added to the population and for the same numbers of them, worst individuals have been eliminated. Individuals have been divided into two equal groups. Figure 1 shows this division. For example, T_1 defines an environment object for $T_{(n/2)+1}$ cooling object and contrariwise.

During the process, if an object has lower γ, the temperature exchanging has been established slowly. Therefore, the value of γ for the objects has been established based on the following equation:

$$\gamma = \frac{\text{Cos}(\text{object})}{\text{Cos}(\text{worst object})}. \tag{21}$$

Time is another parameter that is important in this algorithm. This parameter is related to the iteration number. This parameter can be formulated by the following equation:

$$t = \frac{\text{iteration}}{\text{Max} \cdot \text{iteration}}. \tag{22}$$

Generally, an important ability of the metaheuristics is to escape from the local optimum. During this process, the environmental temperature has been altered by the following equation:

$$T_i^e = \left(1 - \left(c_1 + c_2 \times (1 - t) \times \text{rand}\right)\right) \times T_i^{'e}, \tag{23}$$

where c_1 and c_2 represent the control variables. $T_i^{'e}$ describes the object earlier temperature that has been modified to T_i^e.

Based on the previous models, the new temperatures of the objects have been updated by the following equation:

$$T_i^+ = T_i^e + \left(T_i^{\text{old}} - T_i^e\right)\exp(-\gamma t). \tag{24}$$

Another parameter in the algorithm is Pr with (0, 1) that stated whether a component in the cooling objects should be changed or not.

All of the Pr agents have been compared with $R(i)$ $(i = 1, 2, \ldots, n)$ which is a randomly distributed value in the range [0, 1]. If $R(i) < \text{Pr}$, one dimension of the agent number i has been randomly chosen and the value is redeveloped by the following:

$$T_{i,j} = T_j^{\min} + \text{rnd}\left(T_j^{\max} - T_j^{\min}\right)\exp(-\gamma t), \tag{25}$$

where $T_{i,j}$ signifies the j^{th} variable of the i^{th} agent and $T_j^{\min}$ and $T_j^{\max}$ represent the lower and the upper bounds of the j^{th} variable, respectively. To keep the structures of the agents unchanged, just one dimension has been altered.

Finally, the stopping criteria are checked to terminate the algorithm in the considered criteria.

6.3. Improved Thermal Exchange Optimization (ITEO). The basic Thermal Exchange Optimization algorithm suffers some disadvantages like stability and premature convergence problems. This case leads us to design a modified version of the TEO algorithm to refine these drawbacks as possible.

The first modification is to use Lévy flight (LF) as a proper mechanism. This mechanism has been commonly employed in metaheuristic algorithms to solve premature convergence shortcomings [24]. During this mechanism, a random walk policy has been utilized for proper adjusting of the local search that is mathematically represented as follows:

$$\begin{aligned} \text{Le}(x) &\approx \frac{1}{x^{1+\tau}}, \\ x &= \frac{A}{|B|^{(1/\tau)}}, \\ \sigma^2 &= \left\{\frac{\Gamma(1+\tau)}{\tau\Gamma((1+\tau)/2)} \times \frac{\sin(\pi\tau/2)}{2^{(1+\tau)/2}}\right\}^{(2/\tau)}, \end{aligned} \tag{26}$$

where $A \sim N(0, \sigma^2)$, $B \sim N(0, \sigma^2)$, τ signifies the Lévy index which is located in the range [0, 2] (here, $\tau = 1.5$ [25]), $\Gamma(\cdot)$ represents Gamma function, and ω defines the step size.

By assuming the above equations, the updating formulation for the TEO algorithm is as follows:

$$T_{i,j^{\text{new}}} = T_j^{\min} + \text{Le}\left(T_j^{\max} - T_j^{\min}\right)\exp(-\gamma t). \tag{27}$$

The second modification is to use the chaos mechanism for improving the system convergence speed. Here, we used Singer function for chaos modification [26, 27]. By considering this mechanism, rnd can be updated as follows:

$$\text{rnd}_{i+1} = 1.07\left(7.9 \times \text{rnd}_i - 23.3 \times \text{rnd}_i^2 + 28.7 \times \text{rnd}_i^3 - 13.3 \times \text{rnd}_i^4\right), \tag{28}$$

where $\text{rnd}_0 \in [0, 1]$.

6.4. Algorithm Authentication. The simulations are applied to a Core (TM) i7-4720HQ 1.60 GHz with 8 GB RAM under and simulations are applied to the MATLAB 2017b environment. To prove the effectiveness of the suggested ITEO algorithm, it is performed on some different benchmark functions, i.e., Ackley, Rastrigin, Sphere, and Rosenbrock, and the results have been compared with some well-known and new optimization techniques, i.e. Locust Swarm Optimization Algorithm (LSO) [28], Crow Search Algorithm (CSA) [29], Multi-Verse Optimizer (MVO) [30], search and rescue (SAR) algorithm [31], and the basic Thermal Exchange Optimization (TEO) [10]. Table 1 tabulates the parameter settings of the studied algorithms.

To clarify the studied benchmark functions, the equations and the boundaries have been given in Table 2. During the simulation, all of the algorithms run 40 times independently for all of the benchmark functions and the maximum iteration for all of the algorithms is considered 500.

Table 3 tabulates the results of the analysis of the algorithms on the benchmark functions based on four measurement indicators including minimum, maximum, mean, and standard deviation (std).

As can be observed from the results, the results of all four indicators based on the suggested ITEO have the minimum value. This indicates that the suggested algorithm has the minimum error ratio for all of the metrics. The minimum value of the indicator "minimum" shows that the proposed ITEO algorithm has the highest precision among the others. Plus, the minimum value of the "std" for the ITEO algorithm shows its higher consistency toward the other methods.

Table 1: The parameter settings of the studied algorithms.

Algorithm	Parameter	Value	Algorithm	Parameter	Value
SAR [31]	SE	0.5	LSO [28]	F	0.6
	MU	20		L	1
TEO [10]	c_1	0		g	20
	c_2	1	MVO [30]	Traveling distance rate	[0.6, 1]
	pro	0.3		Wormhole existence prob.	[0.2, 1]
	TM	28.5	CSA [29]	fl	2

Table 2: The equations and the boundaries of the studied benchmarks in the analysis.

Function	Equation	Constraint
Rastrigin	$f_1(x) = 10\ D + \sum_{i=1}^{D}(x_i^2 - 10\cos(2\pi x_i))$	$[30, 50]^D$
Rosenbrock	$f_2(x) = \sum_{i=1}^{D-1}(100(x_i^2 - x_{i+1}) + (x_i - 1)^2)$	$[-2.045, 2.045]^D$
Ackley	$f_3(x) = -20\exp(-0.2\sqrt{(1/D)\sum_{i=1}^{D}(x_i^2)}) - \exp((1/D)\sum_{i=1}^{D}(\cos(2\pi x_i))) + 20 + e$	$[-10, 10]^D$
Sphere	$f_4(x) = \sum_{i=1}^{D} x_i^2$	$[-512, 512]^D$

Table 3: The performance analysis of the studied algorithm.

Algorithm		f_1	f_2	f_3	f_4
LSO [28]	Min	11.428	0.035	0.022	1.327
	Max	4.284*e*3	6.534*e*2	0.679*e*5	6.192
	Mean	2.168*e*3	1.295*e*2	2.195*e*4	3.084
	Std	5.153*e*2	21.931	3.645*e*4	4.339
CSA [29]	Min	8.3161	0.028	7.65*e* − 1	1.446
	Max	492.57	2.193	4.15*e* − 1	11.28
	Mean	120.75	1.137	5.17*e* − 1	8.294
	Std	69.253	0.066	1.01*e* − 1	4.495
MVO [30]	Min	9.192	4.168*e* − 3	8.195*e* − 5	0.957
	Max	1.15 *e*2	4.279*e* − 2	7.086*e* − 4	1.492
	Mean	22.49	9.153*e* − 3	1.517*e* − 4	1.137
	Std	38.76	1.097*e* − 3	3.349*e* − 5	0.846
SAR [31]	Min	0.716	4.82*e* − 16	8.16*e* − 7	1.651
	Max	11.05	5.61*e* − 14	2.94*e* − 6	3.287
	Mean	4.13	5.29*e* − 15	1.19*e* − 6	0.201
	Std	3.62	2.46*e* − 14	4.36*e* − 7	0.146
TEO [32]	Min	0.356	9.208*e* − 21	4.192*e* − 9	3.062*e* − 9
	Max	15.349	11.193*e* − 19	5.930*e* − 8	9.836*e* − 8
	Mean	9.3462	34.255*e* − 19	3.591*e* − 8	3.491*e* − 8
	Std	2.896	3.673*e* − 19	5.086*e* − 9	2.038*e* − 9
ITEO	Min	0.00	8.376*e* − 22	6.956*e* − 10	6.763*e* − 13
	Max	4.207	3.205*e* − 21	3.907*e* − 9	6.358*e* − 12
	Mean	3.164	1.343*e* − 21	1.2084*e* − 9	1.928*e* − 12
	Std	1.237	2.934*e* − 21	3.666*e* − 10	1.117*e* − 13

7. Classification

For the final diagnosis of the features, a classifier is needed. In this study, a new optimized version of Artificial Neural Network (ANN) has been used [33]. The ANN is an almost new methodology with high efficiency for different classification applications. The ANN makes a proper relationship between input and output data. The sensitivity of the ANN is so low toward errors [34]. This method can be used for the accurate classification of medical images by correct training with no mathematical modeling [35, 36]. A popular and simple technique among different types of ANN is the multilayer perceptron (MLP) neural network. An MLP is a mathematical model of a natural brain [34]. An MLP is a model with several numbers of weights and biases that are connected for brain performance modeling. The popular method for error minimization in MLPs is backpropagation (BP). In BP, the values of the weights and biases have been adjusted for minimizing the error between the output value and the desired value [37]. This method uses Gradient Descent (GD) algorithm for minimization. One significant problem of the GD algorithm is that they have been easily trapped into the local minimum.

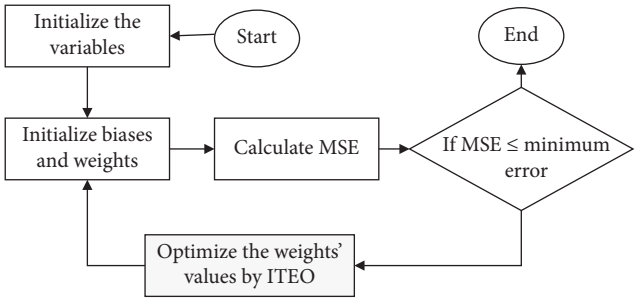


Figure 2: The flowchart of classification based on optimized ANN/ITEO.

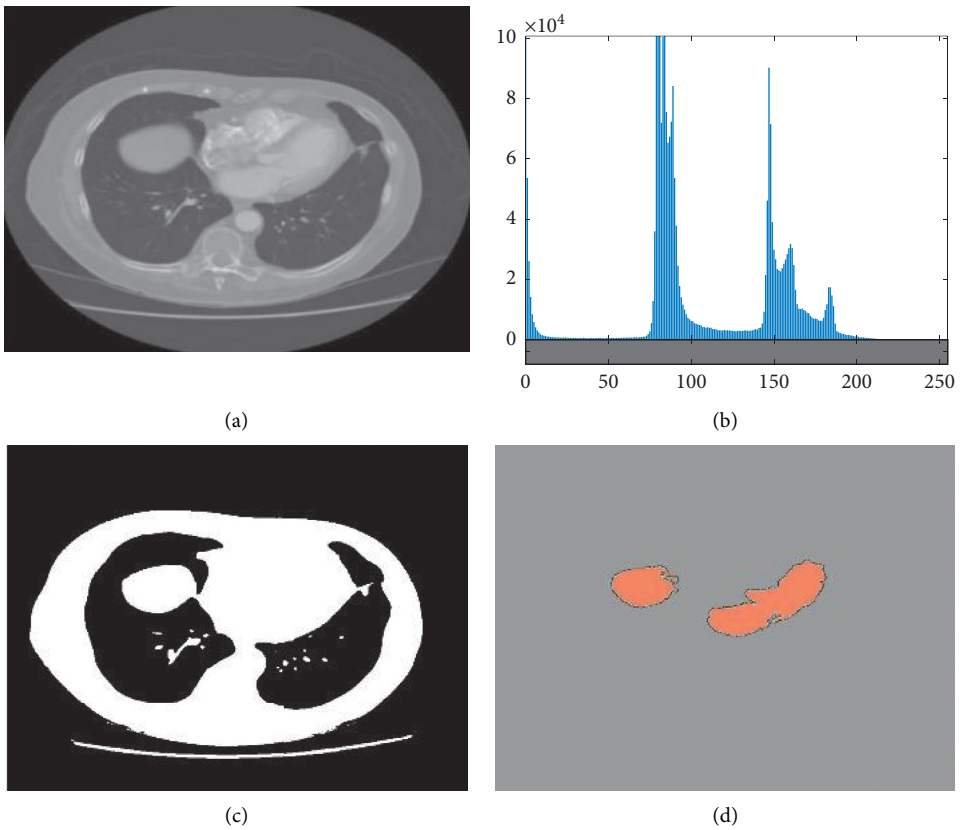


Figure 3: An example of the image segmentation: (a) Original image, (b) histogram of the image, (c) image segmentation based on Kapur technique, and (d) image (c) after mathematical morphology.

The output of each layer in the network is as follows:

$$z_i = \sum_{j=1}^{n} w_{ij}\alpha_i + b_j, \tag{29}$$

where α_i indicates the i^{th} input variable, b_j signifies the j^{th} bias for the neuron, and w_{ij} describes the relation weight between α_i and the j^{th} invisible neuron.

After that, the activation function has been performed to trigger the output of the neurons. This research employs sigmoid function for this purpose:

$$f_j(x) = \frac{1}{1 + e^{-z_j}}. \tag{30}$$

And the output layer gives

$$y_i = \sum_{j=1}^{m} w_{kj}\alpha_i + \beta_k. \tag{31}$$

The ANN calculates mean square error (MSE) between the desired and the observed output, i.e.,

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} y_i - d_i, \tag{32}$$

where n describes the number of the steps in the training data collection and y_i and d_i represent the observed value and the desired value, respectively.

Table 4: The GLCM data results of 20 first images from the lung CT-diagnosis database.

Image #	Homogeneity	IMC 1	IMC 2	Inverse difference	Highest probability	Sum average	Sum entropy	Sum of square variance	Sum variance
1	−0.916	−0.914	0.654	0.997	0.717	2.372	0.407	0.553	0.662
2	−0.900	−0.879	0.212	0.992	0.531	2.357	0.552	0.506	0.711
3	−0.884	−0.88	0.625	0.991	0.891	2.68	0.69	0.147	0.374
4	−0.891	−0.929	0.476	0.992	0.529	2.472	0.601	0.147	0.21
5	−0.911	−0.903	0.476	0.992	0.692	2.517	0.558	0.211	0.091
6	−0.895	−0.915	0.274	0.993	0.865	2.532	0.23	0.123	0.075
7	−0.915	−0.875	0.713	0.993	0.674	2.308	0.384	0.397	0.43
8	−0.882	−0.914	0.405	0.991	0.771	2.41	0.495	0.659	0.554
9	−0.881	−0.859	0.355	0.995	0.861	2.434	0.423	0.517	0.078
10	−0.888	−0.866	0.416	0.992	0.891	2.194	0.198	0.453	0.483
11	−0.917	−0.905	0.437	0.992	0.911	2.437	0.296	0.68	0.382
12	−0.913	−0.895	0.544	0.995	0.584	2.517	0.629	0.174	0.14
13	−0.925	−0.877	0.554	0.997	0.614	2.356	0.179	0.709	0.677
14	−0.893	−0.865	0.449	0.998	0.896	2.602	0.645	0.181	0.433
15	−0.887	−0.885	0.451	0.995	0.762	2.539	0.213	0.451	0.328
16	−0.885	−0.888	0.525	0.998	0.722	2.316	0.538	0.204	0.679
17	−0.907	−0.911	0.614	0.995	0.77	2.373	0.445	0.479	0.479
18	−0.904	−0.899	0.799	0.995	0.861	2.332	0.29	0.113	0.336
19	−0.931	−0.876	0.655	0.993	0.735	2.565	0.484	0.616	0.068
20	−0.886	−0.86	0.452	0.994	0.589	2.541	0.237	0.67	0.393

Table 5: The GLCM data results of 20 first images from the lung CT-diagnosis database.

Image #	Autocorrelation	Cluster prominence	Cluster shade	Contrast	Correlation	Difference entropy	Difference variance	Dissimilarity	Energy	Entropy
1	1.5292	0.9281	0.2484	0.0031	0.951	0.0101	0.001	0.001	0.8911	0.2308
2	1.7732	1.4425	0.5614	0.0049	0.9617	0.017	0.003	0.0029	0.749	0.424
3	1.5332	0.9471	0.2584	0.0024	0.9637	0.0023	0.0005	0.0004	0.8903	0.2274
4	2.0352	1.5474	0.6374	0.0086	0.957	0.0357	0.0063	0.0066	0.6421	0.5764
5	1.4202	0.5069	0.0204	0.0031	0.9247	0.0045	0.001	0.002	0.9595	0.101
6	1.3862	0.3493	0.1646	0.0015	0.9515	0.0133	0.0014	0.0014	0.984	0.0444
7	0.7428	0.7435	0.3716	0.0035	0.9413	0.029	0.0033	0.0035	0.9226	0.1792
8	0.7528	0.8409	0.4256	0.0024	0.975	0.0352	0.0046	0.0044	0.9064	0.2137
9	0.7668	0.8365	0.4216	0.0007	0.994	0.0257	0.003	0.0029	0.9105	0.2015
10	1.4782	0.7423	0.3716	0.0029	0.9603	0.0381	0.0051	0.005	0.921	0.1889
11	1.5302	0.9288	0.4746	0.0024	0.9826	0.0353	0.0045	0.0046	0.891	0.2391
12	1.5432	0.9597	0.2674	0.0039	0.9666	0.0466	0.006	0.005	0.881	0.2612
13	1.7702	1.4266	0.5524	0.0057	0.9891	0.0522	0.0077	0.0058	0.7585	0.4381
14	1.7072	0.8725	0.5024	0.0071	0.9858	0.0315	0.005	0.004	0.784	0.3995
15	1.0348	0.9531	0.7806	0.0072	0.9916	0.0294	0.0052	0.0051	0.756	0.4349
16	0.8828	0.6969	0.6166	0.0038	0.9594	0.0109	0.0018	0.0017	0.8405	0.3082
17	0.8948	0.7077	0.6256	0.0091	0.9283	0.0383	0.009	0.0069	0.8264	0.3453
18	2.0578	1.2224	0.1826	0.0204	0.94	0.105	0.0199	0.02	0.4803	0.7899
19	0.6988	0.5715	0.2796	0.002	0.955	0.0184	0.002	0.0019	0.9541	0.1186
20	1.7448	1.3354	0.5706	0.0141	0.9508	0.0811	0.0141	0.014	0.5192	0.7325

As mentioned earlier, GD algorithm for minimizing the MSE has some problems. Therefore, here for minimizing the MSE and optimizing the classifier, the suggested Improved Thermal Exchange Optimization has been used. Figure 2 shows the method of this idea.

8. Results and Discussion

As mentioned before, the main idea of this study is to propose a pipeline methodology for lung cancer diagnosis. The method starts with a preprocessing method for enhancing the quality of the original image. After image preprocessing, a simple image thresholding has been done based on Kapur for lung areas. Afterward, optimal features have been selected from GLCM features based on an improved version of the Thermal Exchange Optimization algorithm. Finally, an optimized MLP system was based on the introduced Improved Thermal Exchange Optimization algorithm. The method has been validated based on the Lung CT-Diagnosis dataset collected by the Cancer Imaging Archive [12]. Figure 3 shows an example of the image segmentation.

Table 6: The optimum features using the proposed ITEO applied to GLCM.

Image #	Autocorrelation	Correlation	Energy	Homogeneity	Inverse difference	Highest probability	Sum average	Class
1	1.163	0.97	0.8911	0.9972	0.9971	0.9417	2.1095	Yes
2	1.407	0.9807	0.749	0.9966	0.9965	0.8594	2.2728	Yes
3	1.167	0.9827	0.8903	0.9979	0.9978	0.9411	2.1119	Yes
4	1.669	0.976	0.6421	0.9948	0.9947	0.769	2.4488	Yes
5	1.054	0.9437	0.9595	0.997	0.997	0.969	2.036	Yes
6	1.02	0.9705	0.986	0.9985	0.9984	0.9931	2.0131	Yes
7	1.111	0.9603	0.9246	0.9993	0.9972	0.9611	2.0748	Yes
8	1.121	0.954	0.9064	0.9988	0.9987	0.9512	2.0955	Yes
9	1.135	0.975	0.9105	0.9996	0.9995	0.9534	2.0928	Yes
10	1.112	0.9413	0.921	0.9985	0.9984	0.959	2.0782	Yes
11	1.164	0.9636	0.891	0.997	0.998	0.94	2.1127	Yes
12	1.177	0.9476	0.879	0.996	0.997	0.9351	2.121	Yes
13	1.404	0.9701	0.7565	0.9971	0.995	0.859	2.2727	Yes
14	1.341	0.9648	0.784	0.9952	0.9951	0.8774	2.2339	No
15	1.403	0.9706	0.756	0.9954	0.9953	0.8585	2.2723	No
16	1.251	0.9784	0.8405	0.9971	0.997	0.9142	2.1682	No
17	1.263	0.9473	0.8284	0.9945	0.9944	0.9062	2.1787	No
18	2.426	0.961	0.4823	0.9907	0.9907	0.5114	2.9592	No
19	1.067	0.976	0.9541	1.0001	1	0.9767	2.0473	No
20	2.113	0.9718	0.5192	0.994	0.9939	0.6199	2.7485	No

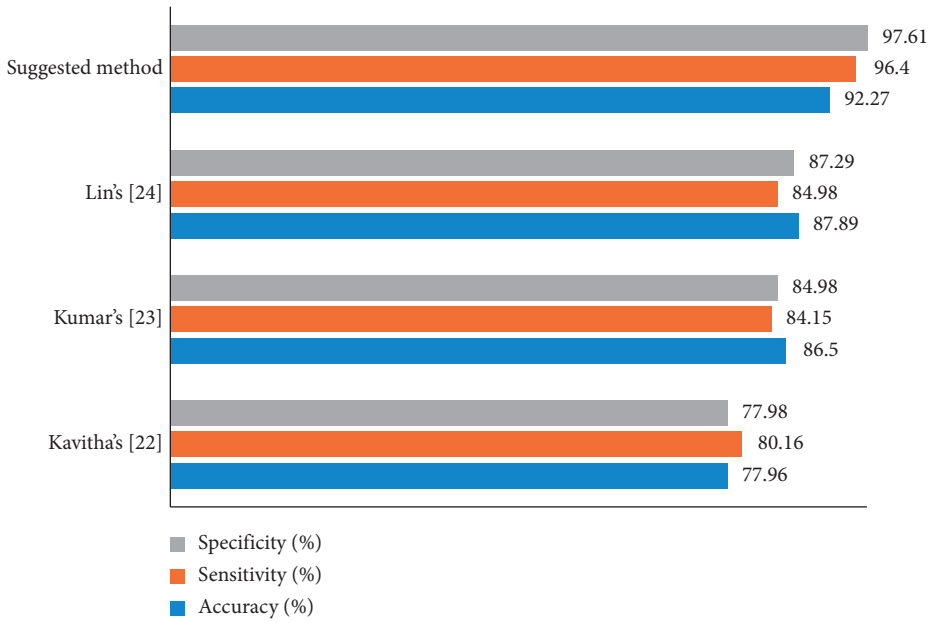


Figure 4: The efficiency validation of the suggested method compared with some well-known methods based on three mentioned indicators.

8.1. Dataset Description. In this study, the Lung CT-Diagnosis database has been utilized. This database contains several images in Dicom format, employed from capturing 61 patients. The number of the total images is 4682. The method has been simulated based on MATLAB 2017b environment and performed to the database based on the following configuration: Corei7 laptop with 16 GB RAM and CPU@2.6 GHz processor.

8.2. Simulation Results. Table 4 indicates the GLCM data results of 20 first images from the Lung CT-Diagnosis database. As can be observed from Table 4, nineteen numbers of GLCM features are employed for the feature extraction (Table 5).

After feature selection based on the Improved Thermal Exchange Optimization, the optimum features based on the suggested ITEO methodology are evaluated and shown in Table 6.

The optimum threshold achieved by ITEO has been used to select the features. For the cost function, the best optimal value earned is 0.75. To test the final efficacy of the suggested technique, three measurement indicators including sensitivity, precision, and accuracy have been analyzed. The mathematical formulations of the indicators are given below:

$$\text{Accuracy}(\%) = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{FN} + \text{TN})},$$
$$\text{Specificity}(\%) = \frac{\text{TN}}{(\text{FP} + \text{TN})}, \quad (33)$$
$$\text{Sensitivity}(\%) = \frac{\text{TP}}{(\text{TP} + \text{FN})},$$

where TN is truly negative, TP is truly positive, FN is false negative, and FP is false positive.

To verify the higher efficiency of the proposed method, a comparison analysis of the method has been applied toward some state-of-the-art algorithms, including Kavitha's [38], Kumar's [39], and Lin's [40], applied to the Lung CT-Diagnosis database. The comparison results are illustrated in a bar chart in Figure 4.

As can be observed from Figure 4, the proposed method has the best precision for the Lung CT-Diagnosis database and Lin's, Kumar's, and Kavitha's are placed in the next ranks. The results show also the proposed method.

9. Conclusions

The main purpose of this study is to propose an optimal pipeline for precise lung cancer diagnosis based on different approaches. The method started with applying a preprocessing process based on a min-max normalization for the input data and an average filter for denoising the input image. Afterward, Kapur entropy maximization along with mathematical morphology was used for segmentation of the lung area. Then, 19 numbers of the GLCM features have been extracted from the segmented images and the features with higher priority were selected based on a new optimization design. The new design, called Improved Thermal Exchange Optimization (ITEO) algorithm, was designed and employed to optimize the feature selection step by considering more accuracy and convergence ability as was shown in the validation stage. Finally, the images were classified into healthy or cancerous cases by using an artificial neural network optimized by ITEO. The simulation was compared with some different state-of-the-art methods including Lin's, Kumar's, and Kavitha's, and the results showed that the proposed method with 92.27% accuracy, 96.4% sensitivity, and 97.61% specificity has the highest efficiency toward the other state-of-the-art methods. In future work, we will work on using convolutional features of the lung cancer images to provide a method with higher accuracy in the system.

Data Availability

The data that support the findings will be available in the Lung CT-Diagnosis database of lung cancer images at https://doi.org/10.7937/K9/TCIA.2015.A6V7JIWX.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] R. Navid, M. Ashourian, M. Karimifard et al., "Computer-aided diagnosis of skin cancer: a review," *Current Medical Imaging*, vol. 16, no. 7, pp. 781–793, 2020.

[2] S. C. Satapathy, N. Sri Madhava Raja, V. Rajinikanth, A. S. Ashour, and N. Dey, "Multi-level image thresholding using Otsu and chaotic bat algorithm," *Neural Computing and Applications*, vol. 29, no. 12, pp. 1285–1307, 2018.

[3] Q. Liu, Z. Liu, S. Yong, K. Jia, and N. Razmjooy, "Computer-aided breast cancer diagnosis based on image segmentation and interval analysis," *Automatika*, vol. 61, no. 3, pp. 496–506, 2020.

[4] A. Hu and R. Navid, "Brain tumor diagnosis based on metaheuristics and deep learning," *International Journal of Imaging Systems and Technology*, vol. 31, no. 2, pp. 657–669, 2020.

[5] R. Wender, E. T. H. Fontham, E. Barrera et al., "American Cancer Society lung cancer screening guidelines," *CA: A Cancer Journal for Clinicians*, vol. 63, no. 2, pp. 106–117, 2013.

[6] N. Ghadimi, "An adaptive neuro-fuzzy inference system for islanding detection in wind turbine as distributed generation," *Complexity*, vol. 21, no. 1, pp. 10–20, 2015.

[7] A. Costa, Y. Kieffer, A. Scholer-Dahirel et al., "Fibroblast heterogeneity and immunosuppressive environment in human breast cancer," *Cancer Cell*, vol. 33, no. 3, pp. 463–479, 2018.

[8] N. Razmjooy, V. V. Estrela, and H. J. Loschi, "Entropy-based breast cancer detection in digital mammograms using world cup optimization algorithm," *International Journal of Swarm Intelligence Research*, vol. 11, no. 3, pp. 1–18, 2020.

[9] N. Dey, V. Rajinikanth, A. Ashour, and J. M. Tavares, "Social group optimization supported segmentation and evaluation of skin melanoma images," *Symmetry*, vol. 10, no. 2, p. 51, 2018.

[10] V. Rajinikanth and S. C. Satapathy, "Segmentation of ischemic stroke lesion in brain MRI based on social group optimization and Fuzzy-Tsallis entropy," *Arabian Journal for Science and Engineering*, vol. 43, no. 8, pp. 4365–4378, 2018.

[11] WHO, *Cancer: Early Diagnosis and Screening*, WHO, Geneva, Switzerland, 2020, https://www.who.int/cancer/prevention/diagnosis-screening/en.

[12] O. Grove, A. E. Berglund, M. B. Schabath et al., "Quantitative computed tomographic descriptors associate tumor shape complexity and intratumor heterogeneity with prognosis in lung adenocarcinoma," *PLoS One*, vol. 10, no. 3, Article ID e0118261, 2015.

[13] M. Ghiasi, N. Ghadimi, and E. Ahmadinia, "An analytical methodology for reliability assessment and failure analysis in distributed power system," *SN Applied Sciences*, vol. 1, no. 1, p. 44, 2019.

[14] G. Guo and N. Razmjooy, "A new interval differential equation for edge detection and determining breast cancer regions in mammography images," *Systems Science & Control Engineering*, vol. 7, no. 1, pp. 346–356, 2019.

[15] A. Arshaghi, N. Razmjooy, V. V. Estrela et al., *Image Transmission in UAV MIMO UWB-OSTBC System over Rayleigh Channel Using Multiple Description Coding (MDC)*, IET, London, UK, 2020.

[16] U. R. Acharya, S. L. Fernandes, J. E. WeiKoh et al., "Automated detection of Alzheimer's disease using brain MRI images: a study with various feature extraction techniques," *Journal of Medical Systems*, vol. 43, no. 9, p. 302, 2019.

[17] R. Navid, V. V. Estrela, and H. J. Loschi, "Chapter 1: a survey of potatoes image segmentation based on machine vision," in *Applications of Image Processing and Soft Computing Systems in Agriculture*, pp. 1–38, IGI Global, Hershey, PA, USA, 2019.

[18] V. V. Estrela, A. C. B. Monteiro, R. P. França, Y. Iano, A. Khelassi, and N. Razmjooy, "Health 4.0 as an application of industry 4.0 in healthcare services and management," *Medical Technologies Journal*, vol. 2, no. 4, pp. 262–276, 2018.

[19] R. Sheykhahmad, N. Razmjooy, and M. Ramezani, "A novel method for skin lesion segmentation," *International Journal of Information, Security and Systems Management*, vol. 4, no. 2, pp. 458–466, 2015.

[20] C. E. V. Marinho, V. V. Estrela, C. Marinho et al., "A model for medical staff idleness minimization," in *Proceedings of the Brazilian Technology Symposium*, Springer, Cham, Switzerland, 2018.

[21] R. Navid, M. Ramezani, and N. Ghadimi, "Imperialist competitive algorithm-based optimization of neuro-fuzzy system parameters for automatic red-eye removal," *International Journal of Fuzzy Systems*, vol. 19, no. 4, pp. 1144–1156, 2017.

[22] N. S. M. Raja, S. L. Fernandes, N. Dey, S. C. Satapathy, and V. Rajinikanth, "Contrast enhanced medical MRI evaluation using Tsallis entropy and region growing segmentation," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–12, 2018.

[23] R. Navid, V. V. Estrela, H. J. Loschi, and W. S. Farfan, "A comprehensive survey of new meta-heuristic algorithms," in *Recent Advances in Hybrid Metaheuristics for Data Clustering*Wiley Publishing, Hoboken, NJ, USA, 2019.

[24] C. Choi and J.-J. Lee, "Chaotic local search algorithm," *Artificial Life and Robotics*, vol. 2, no. 1, pp. 41–47, 1998.

[25] X. Li, P. Niu, and J. Liu, "Combustion optimization of a boiler based on the chaos and Lévy flight vortex search algorithm," *Applied Mathematical Modelling*, vol. 58, pp. 3–18, 2018.

[26] D. Yang, G. Li, and G. Cheng, "On the efficiency of chaos optimization algorithms for global optimization," *Chaos, Solitons & Fractals*, vol. 34, no. 4, pp. 1366–1375, 2007.

[27] C. Rim, S. Piao, G. Li, and U. Pak, "A niching chaos optimization algorithm for multimodal optimization," *Soft Computing*, vol. 22, no. 2, pp. 621–633, 2018.

[28] E. Cuevas, F. Fausto, and A. González, "The locust Swarm optimization algorithm," in *New Advancements in Swarm Algorithms: Operators and Applications*, pp. 139–159, Springer, Cham, Switzerland, 2020.

[29] A. Askarzadeh, "A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm," *Computers & Structures*, vol. 169, pp. 1–12, 2016.

[30] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-verse optimizer: a nature-inspired algorithm for global optimization," *Neural Computing and Applications*, vol. 27, no. 2, pp. 495–513, 2016.

[31] A. Shabani, B. Asgarian, M. Salido, and S. Asil Gharebaghi, "Search and rescue optimization algorithm: a new optimization method for solving constrained engineering optimization problems," *Expert Systems with Applications*, vol. 161, Article ID 113698, 2020.

[32] A. Kaveh and A. Dadras, "A novel meta-heuristic optimization algorithm: thermal exchange optimization," *Advances in Engineering Software*, vol. 110, pp. 69–84, 2017.

[33] R. Navid, V. V. Estrela, and H. J. Loschi, "A study on metaheuristic-based neural networks for image segmentation purposes," in *Data Science*, pp. 25–49, CRC Press, Boca Raton, FL, USA, 2019.

[34] D. Yu, Y. Wang, H. Liu, K. Jermsittiparsert, and N. Razmjooy, "System identification of PEM fuel cells using an improved Elman neural network and a new hybrid optimization algorithm," *Energy Reports*, vol. 5, pp. 1365–1374, 2019.

[35] Y. Cao, Y. Li, G. Zhang, K. Jermsittiparsert, and N. Razmjooy, "Experimental modeling of PEM fuel cells using a new improved seagull optimization algorithm," *Energy Reports*, vol. 5, pp. 1616–1625, 2019.

[36] Y. Cao, Y. Wu, L. Fu, K. Jermsittiparsert, and N. Razmjooy, "Multi-objective optimization of a PEMFC based CCHP system by meta-heuristics," *Energy Reports*, vol. 5, pp. 1551–1559, 2019.

[37] R. Navid, F. R. Sheykhahmad, and N. Ghadimi, "A hybrid neural network–world cup optimization algorithm for melanoma detection," *Open Medicine*, vol. 13, no. 1, pp. 9–16, 2018.

[38] M. S. Kavitha, J. Shanthini, and R. Sabitha, "ECM-CSD: an efficient classification model for cancer stage diagnosis in CT lung images using FCM and SVM techniques," *Journal of Medical Systems*, vol. 43, no. 3, p. 73, 2019.

[39] K. Senthil Kumar, K. Venkatalakshmi, and K. Karthikeyan, "Lung cancer detection using image segmentation by means of various evolutionary algorithms," *Computational and Mathematical Methods in Medicine*, vol. 2019, Article ID 4909846, 16 pages, 2019.

[40] C.-J. Lin, S.-Y. Jeng, and M.-K. Chen, "Using 2D CNN with taguchi parametric optimization for lung cancer recognition from CT images," *Applied Sciences*, vol. 10, no. 7, p. 2591, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 9523039, 22 pages
https://doi.org/10.1155/2021/9523039

Review Article

Single and Combined Neuroimaging Techniques for Alzheimer's Disease Detection

Morteza Amini,[1] **Mir Mohsen Pedram**,[2,3] **Alireza Moradi**,[4,5] **Mahdieh Jamshidi**,[6] **and Mahshad Ouchani**[7]

[1]*Department of Cognitive Modeling, Institute for Cognitive Science Studies, Shahid Beheshti University, Tehran, Iran*
[2]*Department of Electrical and Computer Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran*
[3]*Department of Cognitive Modeling, Institute for Cognitive Science Studies, Tehran, Iran*
[4]*Department of Clinical Psychology, Faculty of Psychology and Educational Science, Kharazmi University, Tehran, Iran*
[5]*Department of Cognitive Psychology, Institute for Cognitive Science Studies, Tehran, Iran*
[6]*Department of Mathematical Sciences, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran*
[7]*Institute for Cognitive and Brain Science, Shahid Beheshti University, Tehran, Iran*

Correspondence should be addressed to Morteza Amini; mor_amini@sbu.ac.ir

Received 21 April 2021; Revised 4 June 2021; Accepted 30 June 2021; Published 13 July 2021

Academic Editor: V. Rajinikanth

Alzheimer's disease (AD) consists of the gradual process of decreasing volume and quality of neuron connection in the brain, which consists of gradual synaptic integrity and loss of cognitive functions. In recent years, there has been significant attention in AD classification and early detection with machine learning algorithms. There are different neuroimaging techniques for capturing data and using it for the classification task. Input data as images will help machine learning models to detect different biomarkers for AD classification. This marker has a more critical role for AD detection than other diseases because beta-amyloid can extract complex structures with some metal ions. Most researchers have focused on using 3D and 4D convolutional neural networks for AD classification due to reasonable amounts of data. Also, combination neuroimaging techniques like functional magnetic resonance imaging and positron emission tomography for AD detection have recently gathered much attention. However, gathering a combination of data can be expensive, complex, and tedious. For time consumption reasons, most patients prefer to throw one of the neuroimaging techniques. So, in this review article, we have surveyed different research studies with various neuroimaging techniques and ML methods to see the effect of using combined data as input. The result has shown that the use of the combination method would increase the accuracy of AD detection. Also, according to the sensitivity metrics from different machine learning methods, MRI and fMRI showed promising results.

1. Introduction

Alzheimer's disease (AD) can be considered a gradually progressive neurodegenerative disease process that involves gradual synaptic integrity and loss of cognitive functions [1]. Early detection of AD will help to prevent catastrophic damage to the brain. One of the most significant signs and biomarkers for AD are beta-amyloid, tau protein, and miRNA [2]. This marker has a more critical role for AD detection than other diseases because beta-amyloid can extract complex structure form [3]. Detecting biomarkers deposition and brain structure examination with neuroimaging techniques like functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) approaches have been widely exploited nowadays. Amyloid PET was used to determine brain amyloid plaque load scores as a biomarker [4]. Observation of fMRI techniques will help detect dementia and change in neuron connections, determining the change in brain function. On the other hand, the level of amyloid deposition in certain parts of the brain, which can be seen by amyloid PET biomarker, will help to survey AD severity for patients [5]. However, especially in

cases of dementia diagnosis, mentioned biomarkers could not help accurately identify or predict cognitive deterioration due to individual threshold differences in each subject [6]. The use of PET for AD detection will consume more time, and because of isotope injection, it is considered an invasive technique. MRI and its branches like fMRI for AD detection and classification are other areas of research in neuroimaging. For the change in structure detection use of the fMRI technique is more convenient than other neuroimaging techniques.

The better change in brain detection structure using structural magnetic resonance imaging (sMRI) can help. Suppose that, with the survey of sMRI, enough structural features have not been extracted. In that case, resting-state fMRI (rs-fMRI) can provide more valuable and complementary information to distinguish early-stage dementia and AD detection in each patient [7]. Combining fMRI and PET into one unique scanner has allowed the researcher to explore the underlying neurochemistry of brain function in more detail [8]. A combination of PET and fMRI techniques resulted in spatial and quantifiable inconsistencies as active data acquisition. By the combination of PET and fMRI, Wehrl et al. [9] extracted functional connectivity of the subjected rat brain. In a nutshell, the fMRI technique will help extract nine well-known biological neural networks in brain structure.

In contrast, the PET technique identified seven glucose metabolism-related biological networks. Different studies have shown comprehensive and complementary information using combination techniques to decode brain function and brain networks further, so that the question of which brain neuroimaging technique will be more helpful and practical for Alzheimer-related disease (ARD) detection and classification arises. In this research, we focus on different branches of AI (artificial intelligence) like machine learning (ML) and deep learning (DL) algorithms for ARD prediction with PET, fMRI, sMRI, and combination methods. Increasing the number of patients with AD-related problems in the future is inevitable. It has been estimated that 1 of 85 individuals in 2050 would suffer from AD-related disease, so, with growth in the number of patients with AD-related problems [10] and with new corona virus pandemics emergence in 2019, different studies have categorized patients with neurodegeneration problems like AD, mild cognitive impairment (MCI), and other ARD at high risk for COVID-19 [11]. Different stages of AD and its complications will be associated with high morbidity and mortality rate. ARD will cause memory capacity loss. Most Alzheimer's patients will forget how to correctly conduct the recommendations from public health authorities or World Health Organization (WHO) to reduce the COVID-19 spread in a high-risk community.

For example, WHO's known recommendations are constant hand washing, covering mouth and nose when coughing or sneezing, monitoring physical and temperature conditions for reporting symptoms of COVID-19, and preserving at least 6 ft physical distance from other people, especially elderly peoples [11]. This specific order is crucial to AD patients, especially when considering these people's age groups, which usually are more than 65 years. Even patients with MCI or milder dementia may forget to conduct these procedures due to oblivion or depression. Those with more severe dementia cannot correctly comprehend or remember most of these orders due to the strictness of their short-term memory.

Another complication of ARD consists of financial problems. Single Medicare beneficiaries diagnosed with ARD have a higher probability of missing payments on credit accounts. These negative financial consequences continued after the progress of ARD and will cause 10% to 15% of missed payments. These financial complications from AD were common in none of the educational college groups [12]. With this significant complication, proper and fast detection of Alzheimer in the early stages will help patients prevent financial and physical complications. ML and DL algorithms for ARD classification and detection have gained much attention in the past decade. With the growth in computational power and emergence of more sophisticated and supervised algorithms like convolutional neural network (CNN) development of artificial intelligence (AI) application in healthcare has increased rapidly [13, 14]. All artificial intelligence models will use some training data such as pictures from neuroimaging techniques and other electronic healthcare data to extract full features or direct samples to classify, detect, and recognize ARD. Numerous ML applications involve tasks that can be set up as supervised and semisupervised learning [15]. ML algorithms often have reached more than 96% to classify AD [16, 17]. This state-of-the-art result for binary classification of AD is no surprise due to the structure of DL algorithms and the usage of neuroimaging data. Proper preprocessing of input data into the balance group of data mentioned result is no longer a surprise. Nevertheless, the vacancy of specific research on the effect of single and combined neuroimaging techniques for reaching the mentioned results and comparing them is increasing. This paper reviewed AD detection, early recognition, and classification using different machine learning applications with fMRI, PET, sMRI, and combined neuroimaging techniques.

2. Data Set

Brain imaging gathering procedure can be categorized as noninvasive like fMRI and sMRI techniques or invasive techniques like PET. For example, rs-fMRI is a neuroimaging technique commonly used to study the progressive and pathogenic procedure of neurodegeneration diseases like AD. Other techniques we have surveyed in this article impacted the brain's specific marker or specific area. All images presented in this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) repository, which can be found at https://www.loni.ucla.edu/ADNI and https://www.adni-info.org. The ADNI database contains 1.5 T and 3.0 T t1w MRI scans for AD, MCI, and cognitively normal conditions in various patients with different ethnic groups. This repository offers data as the image for three categories of patients: NC, AD, and MCI. These three categories will comprise the whole condition of each patient in this research.

2.1. MRI. This data gathering technique will use radio waves and magnetic fields to acquire compelling images with high discriminative features and high-resolution images. This technique can return 2D and 3D images of brain structures with high quality. This technique would not use any harmful radiations from X-rays or other radioactive tracers, so it will be categorized as a noninvasive and nonharmful neuroimaging technique. The most used MRI techniques for AD detection and classification comprise sMRI, which helps technicians evaluate brain volumes in pictures to detect brain loss of tissue, cells, neurons, and other destructive changes. Brain degeneration is an unavoidable component of ARD that causes memory loss and self-unawareness in patients [18]. Figure 1 shows an example of an MRI picture which has been used to detect brain decrease in tissue. In a nutshell, the MRI technique will exploit the nucleus of hydrogen atoms as small magnets as a tracer [19]. Then the vibration of this nucleus of hydrogen atoms can be manipulated to become a tracer and generate a signal that will turn into an image.

2.2. fMRI. fMRI can be categorized as noninvasive technique because this technique focuses on measuring and mapping brain activities without any injected tracer to patients' bodies. With the change in the body's activity, neuron activity in the brain constantly fluctuates [20]. This technique is also using the effect of magnetic fields on brains in order to gather data. The cylindrical tube of an MRI scanner houses a potent electromagnet field, which will be used to gather information whereas patients are conducting some activities. As it has been told before, this strong field will affect the magnetic behavior of atoms. Typically the atomic nuclei orientation is stochastic and does not create any specific pattern. However, under the encouragement of a magnetic field, the pattern of nuclei becomes allied with the direction of the field and makes a specific pattern. As stronger magnetic field becomes significant, the effect will be on the nuclei and degree of alignment. At last, with these powerful aligned signals from these individual nuclei, measurements of these pales will be possible. In fMRI neuroimaging techniques, the behavior of magnetic signal from the core of hydrogen atoms in a water molecule is detected as a tracer. The different neural activity of the body of different parts of the brain will increase; therefore, an improvement in demand for oxygen will be generated (see Figure 2).

With high blood pressure and movement of the capillary blood cells, the local response in a different part of the brain will increase biological neural activity. Red blood cells have hemoglobin. Hemoglobin is diamagnetic when it conveys oxygen in blood and paramagnetic when it conveys carbon dioxide [21], for gathering MRI picture. This difference in magnetic properties will be practical to use. As it has been said before, since the amounts of oxygen that each blood cell will transform will be different according to the different levels of biological neural activity in the brain, these can be used as markers to detect brain activity. A sample of fMRI pictures containing 64 different images of a patient was surveyed in this research. The fMRI picture of the brain is shown in Figure 2. So, with gathering different parts of the brain, brain structure will be revealed, and with different actions in the body like the movement of arms or feet, different parts of the brain will be activated and can be seen in the fMRI scans.

The use of fMRI scans for AD detection was widely used last decade [22, 23]. A different part of brain activity can be discriminated even from a different view with this image. Some pictures from a different region of the brain are shown in Figure 3.

2.3. rs-fMRI. rs-fMRI focuses on low-frequency that is defined as less than 0.1HZ impulsive fluctuations in the level of blood cells that carry oxygen in the brain section. The whole procedure of data gathering rs-fMRI occurs when a patient is at rest situation [24]. The main key characteristic in rs-fMRI was synchronism of low-frequency fluctuations in the tracer signal arising from the right and left primary motor regions of cortices at rest condition of each patient for surveying. With this neuroimaging technique, most researchers have reached a connectivity pattern that is sufficiently close to the activation pattern from a two-sided finger-tapping task [25]. It extracted abnormal patterns in the resting situation of each patient. It disrupted connectivity in various parts of biological neural networks so the brain can help clarify some of the motor and nonmotor shortfalls seen in patients with ARD [26]. A sample of the rs-MRI picture is shown in Figure 4. This picture contains 54 stages of patients in resting condition. The size of samples in this figure is $256 * 64$.

So, with different stages of the brain, even in resting conditions, different patterns can be exploited to classify AD patients.

2.4. PET. PET is the most exact and delicate neuroimaging technique for capturing the molecular image, which helps extract communications and pathways within humans' brains [27]. The specificity rises from the variety of positron-emitting radionuclides choice, which can be defined as specific biomarkers for pathway recognition, biochemicals, and pharmaceuticals without disturbing their biological function [28]. Furthermore, the radiation used as a radiolabeled tracer can be sensed when the reflect wave has been reported above the low natural radiation background. The initial provocation for PET use was for human brain studies. Researchers have used PET for gathering anatomical and biological complexity from brain organs. Because of the use of injected tracer in this procedure, this technique will be considered an invasive method [29]. The compassion and chronological resolution of PET scanners provided to researcher scans with kinetical features that constants of neurotransmitter pathways and binding could be extracted from correct input data [30]. The main concern of the PET technique is to gather more high-quality pictures with better discriminative spatial resolution features and increased axial coverage. With the success of this technique, most researchers worldwide have used this picture for regional brain activation detection. PET pictures are sometimes helping to

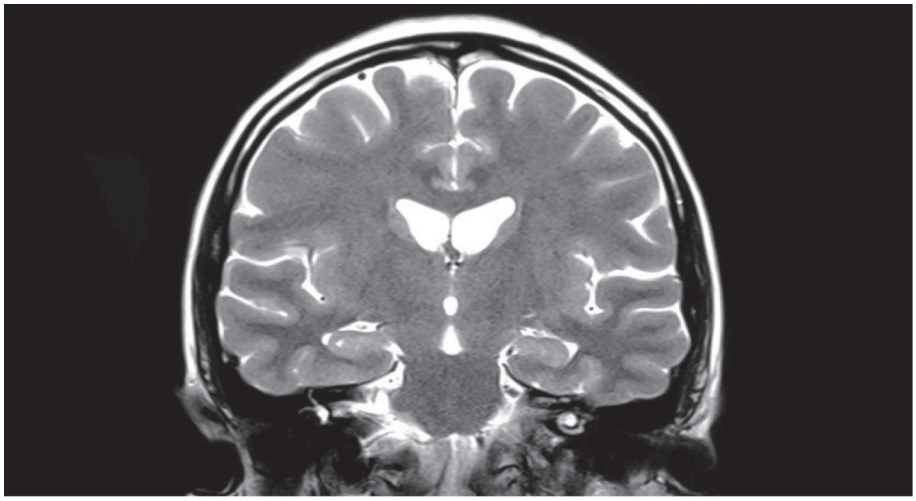
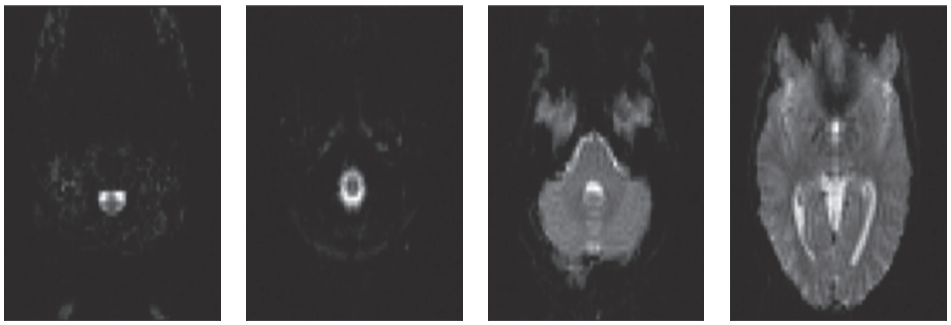


Figure 1: Example of magnetic resonance imaging (MRI).



(a) (b) (c) (d)

Figure 2: Example of functional magnetic resonance imaging (fMRI) in different stages. (a) Slice number: 0. (b) Slice number: 10. (c) Slice number: 20. (d) Slice number: 30.

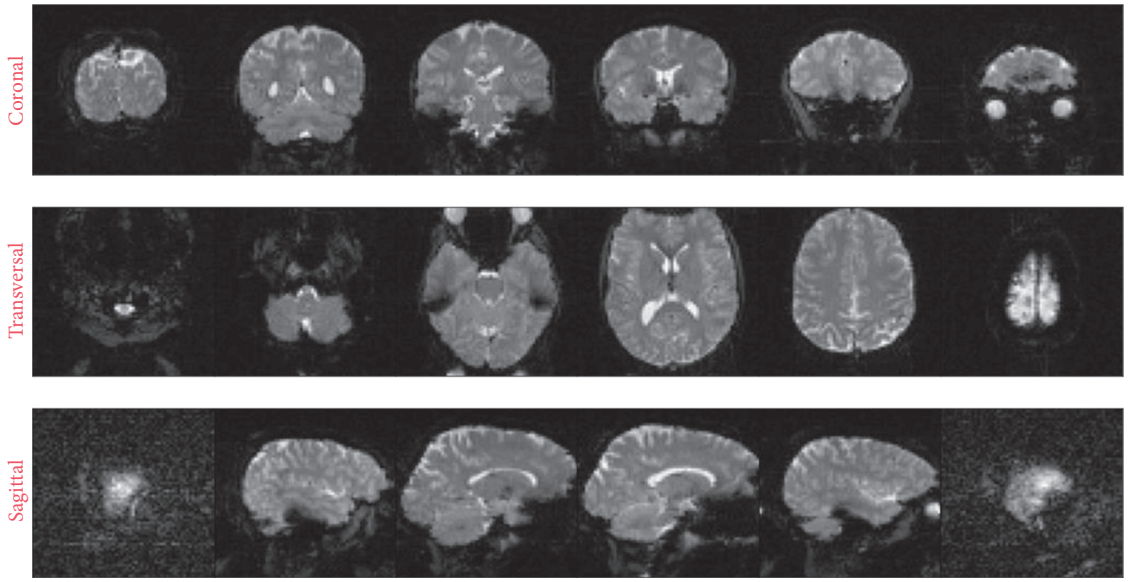
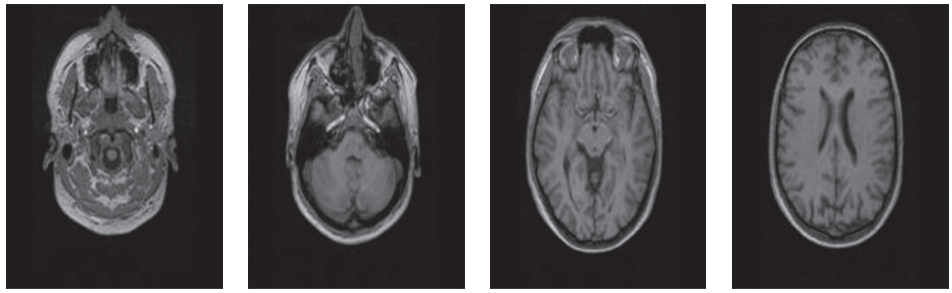


Figure 3: Example of fMRI in different stages.



(a) (b) (c) (d)

Figure 4: Example of rs-fMRI in different stages. (a) Slice number: 0. (b) Slice number: 10. (c) Slice number: 20. (d) Slice number: 30.

detect diseases like Alzheimer's faster than other imaging procedures [31]. Some samples of PET with $336 * 336$ are shown in Figure 5. As it has been shown in Figure 5, these PET pictures have more different stages than fMRI or rs-fMRI.

Here the number of different pictures has reached 654. This increase in pictures will cause better observation and PET's image will become more conventional. However, as said before, capturing all of these pictures can lead to 4 to 6 hours of tedious procedure. Also, this neuroimaging technique is invasive [32]. The most used case of PET is to survey the chemical activity in the brain tissue of each subject. These neuroimaging techniques will help to determine the different conditions of each patient, which include brain disorders. The pictures from a PET scan provide diverse data which are uncovered by other kinds of neuroimaging techniques, like computerized tomography (CT) or MRI. A PET scan or a combined CT-PET scan enables neuroscientists to identify illness and measure the condition of the brain of each patient more conveniently. PET images nowadays will be recorded by about 1% of the equivalent couples of photons emitted from the brain of each subject and the learned coincidences data from each patient. This data will be stored in list mode as separate events with a time brand or sorted into arrays, sinogram. Then, using this data cluster of tracers 3D distribution of the tracer can be recreated [33] (see Figure 6).

3. Combination of PET and MRI Techniques

MRI and PET techniques have opposite natures for gathering data from the brain. By combining these neuroimaging techniques, more accurate AD diagnosis or classification tasks can be conducted [34]. In Figure 6, a combination of PET and rs-fMRI has been shown. New researches show that a mixture of one or more biomarkers may deliver complementary material for ARD diagnosis; also, combination data can help to increase the classification accuracy. This combination of biomarkers can be presented as fluorodeoxyglucose positron emission tomography (FDG-PET), sMRI, cerebrospinal fluid (CSF) protein levels, and apolipoprotein-E (APOE) genotype [35]. Although published approaches based on the combination of techniques have applied dissimilar biomarkers to develop a new neuroimaging biomarker for AD, this usage of combination methods may be limited [36]. The usage of combination techniques has led to early detection in some approaches [36]. Based on the performance of ensemble learning models and multikernel learning success on combination neuroimaging techniques in the last decade, these techniques are popular for AD detection and classification [37, 38]. The combining method showed promising results and can be considered the future of AD detection and classification input data, especially in early detection cases [39, 40].

3.1. Combination Method Preprocessing. After gathering different data set use of various preprocessing procedures is necessary for better prediction of results. As it has been said before use of a combination of neuroimaging techniques has resulted in better classification and detection of ARD. Methods like Dartel are considered preprocessing procedures for combination input data. Dartel is a proper tool for increasing intersubject recording or three-dimensional normalization of functional and structural scans, providing less flattening and improving MRI-PET combination data [41]. Another approach has worked on a framework based on an early union procedure that uses different combination rules to combine opposite data from different biomarker modalities into a single feature vector [42]. In another research, they focused on region of interest (ROI) for gathering complementary information of each neuroimaging method. Many researchers used data from the ADNI database and divided brains based on two atlases: LONI Probabilistic Brain Atlas and Automated Anatomical Labeling. Then baseline images of sMRI and 18F-fluorodeoxyglucose PET were used to calculate average gray-matter density and average relative cerebral metabolic rate for glucose in each region [43]. In 2008 compatible PET detector tools for gathering synchronized PET/MR images of the human brain were conducted. With these new tools and new studies, researchers successfully achieved brain glucose consumption images in two subjected patients using 18F-FDG-PET, MR imaging, and MR spectroscopy [44]. With the combination of data, they established that PET/MR imaging combination is possible in humans' brains. With this combination of data for the first time, a field of new possibilities in molecular imaging areas and ARD detection have been unlocked. The essential step of combination

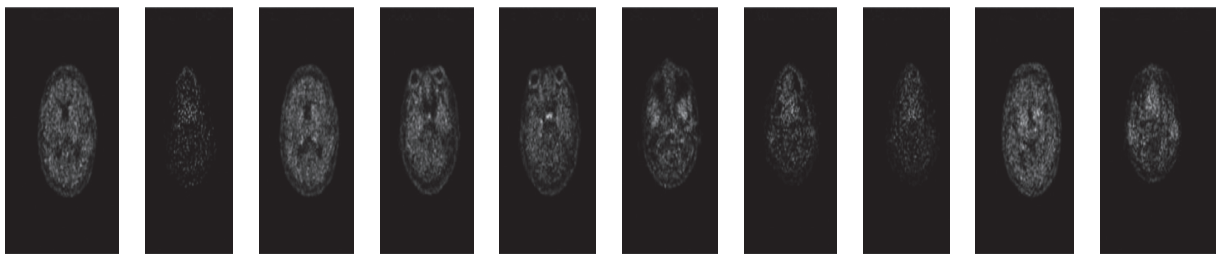


Figure 5: Example of PET in different stages. (a) Slice number: 0. (b) Slice number: 20. (c) Slice number: 40. (d) Slice number: 60. (e) Slice number: 80. (f) Slice number: 100. (g) Slice number: 120. (h) Slice number: 140. (i) Slice number: 160. (j) Slice number: 180.

techniques and preprocessing methods is shown in Figure 7. A typical pipeline for the combination of PET with rs-fMRI is shown in Figure 7.

In the first step, frontal commissure and subsequent commissure correction for all subject images can be extracted with this combination of data. After data gathering use of N4 bias field correction using ANT's toolbox will help correct the intensity of nonhomogeneity for each patient's image [45]. In some cases, elimination of skull has been conducted, which was unnecessary if images were already preprocessed. For the MRI images, aligning them to the MNI152 T1-weighted standard image using a standard procedure will be done next. In normalization steps, the extracted features will throw a standard scalar function, which transforms the array of input matrix data sets into a standard distribution with minimum and maximum of each column vector, which can help to reduce the redundancy and dependency of the data [46]. For structural and functional segmentation of brain into the anatomical area and enumerating these extracted features from each specific ROI from each sMRI image, toolbox with a conventional procedure like NiftyReg with 2 mm Brainnetome Atlas template has been used [47]. After gathering ROI from labeled sMRI images, computed volume of gray-matter tissues in that ROI will be used as an input feature for the detection task.

3.2. Single Method Preprocessing. For processing fMRI pictures into using robust data for early MCI detecting, a handful of researchers have used Data Processing Assistant for Resting-State (or in brief DPARSF) [48–51]. To process fMRI pictures in this platform, users need to arrange their DICOM files and specify their intended parameters. DPARSF then will deliver all the preprocessed data as different variety of desirable data for classification. This desirable data consists of slice timing images, normalized images, smooth data, functional connectivity with specific data, ReHo, ALFF/fALFF, degree centrality, and voxel-mirrored homotopic connectivity (VMHC) results [52]. For PET image processing, using a standard CL pipeline is a conventional method [53]. PET images were intensity normalized using the whole cerebellum as a reference region. In simple image preprocessing, the fusion parameters of combination methods have been eliminated.

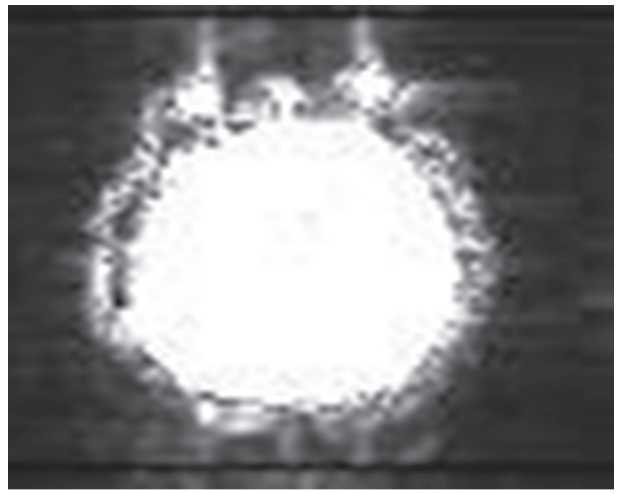


Figure 6: Combination of PET with rs-fMRI pictures.

However, the use of feature selection and normalization part of the procedure remains the same. All of the mentioned algorithms help decrease noise in the picture and use the whole part of the brain, which help model to classify, detect, or recognize ARD. After the acquisition of using complete information, some models will convert pictures into 2D features. With the use of these procedures, the number of features increases significantly. Also, even small shape images like $64 * 64$ or $128 * 128$ have about 4096 or 16384 features. With this amount of information, even power full computational GPU would not conduct the code and render results [54]. So, dimension reduction is a preferable procedure for dealing with image processing tasks. One of the main steps of dimensionality reduction is using correct data and eliminating undesirable data from each picture. Dealing with this task using different norms like l1 and l2 and hybrid classification will help reach a better result. Other studies have used features computed from MRI images to discriminate between different cognitive states related to AD [55]. With the rise in prostate cancer, patient's attention to prostate image segmentation has been conducted. With the use of this data, a semiautomatic method has been designated [56]. This procedure consists of two new algorithms for better feature selection. Experiments on prostate CT images have shown the effect of this method for segmentation and regression tasks [57].

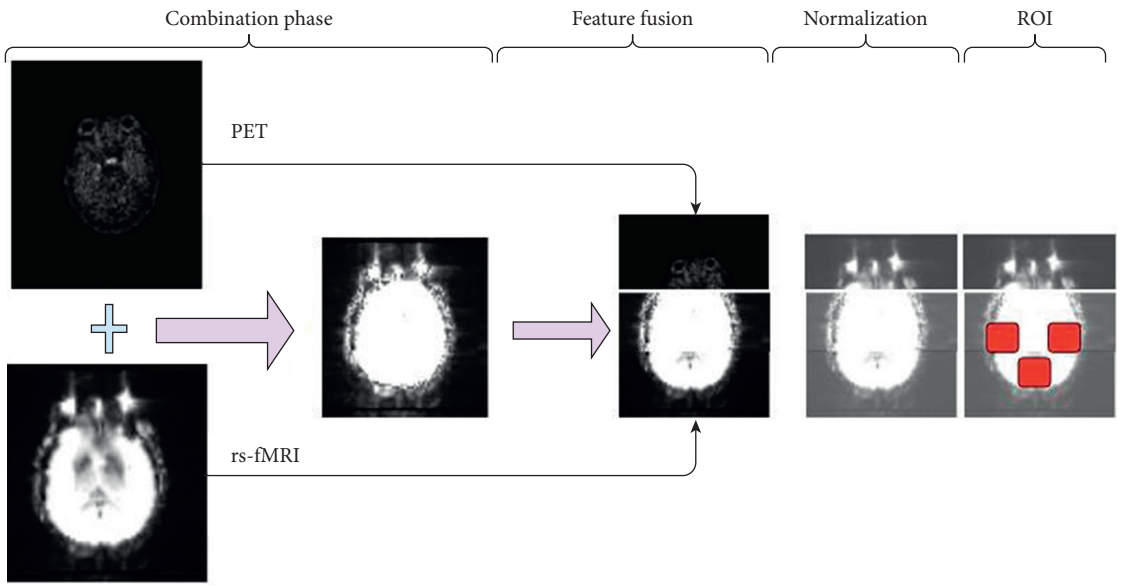


Figure 7: Combination of PET with rs-fMRI pictures in the processing of data preprocessing.

Principal Component Analysis (PCA) is another feature selection technique. PCA, in a nutshell, determines the alliance or axis responsible for the most significant amount of variance in the input data set. Respectively, PCA will determine a second axis that must be orthogonal to the first axis responsible for the largest total of remaining data variance. A standard matrix factorization method is called singular value decomposition (SVD) for dealing with this task. This technique will decompose the training data matrix into three different matrices. PCA will be the result of the multiplication of three multiplications. These three matrices will be A, B, and C, where C contains all the principal components used as a principal feature [58]. The whole structure of feature selection and PCA is shown in Figure 8. Effect of proper feature selection and dimension reduction for reaching the better performance of classifier algorithms is necessary. Fewer features will lead to less time consumption in order to train the different models. Using proper features and tuned model state-of-the-art result will be achieved in AD detection and classification.

4. Methods

In the last two decades, due to vast improvement of computational power, more conventional GPU and online platforms for the implementation of artificial intelligence (AI) systems have been developed. So, interest in the use of AI, ML, and DL to synthesize the applications for studying mental health has increased rapidly [59]. Different ML algorithms have been used for the prediction and classification of ARD. The main goal of these different algorithms was to separate different patients into AD, MCI, and NC classes [60]. In ML and DL fields, algorithms have been separated into three categories:

(i) Unsupervised algorithms

(ii) Semisupervised algorithms

(iii) Supervised algorithms

Supervised algorithms refer to algorithms targeting specific targets, and all samples have their target [61]. Most varieties of ML and DL algorithms belong to this category. In this category, gathering samples and labeling them would be tedious. DL can be categorized as a subfield of ML, usually used on big data as input. Different data structures like pictures, time-dependent data, and images can be used for AD classification tasks. It has attracted massive attention in the last few years, especially in image analysis [62]. Several DL architectures such as Convolutional Neural Network (CNN), Deep Neural Network (DNN), Recurrent Neural Network (RNN), and autoencoder (AE) are some examples of these fields which have been used for ARD and classification. Semisupervised learning refers to algorithms that work with a data set that most of its labels are unclear. Most data will be labeled by knowledge about already known labels from data sets [63]. Unsupervised learning refers to algorithms in which labels of the whole data set are not clear. In most cases, most data do not have any labels, making this category very important [64].

4.1. Semisupervised Methods

4.1.1. K Nearest Neighbor (KNN). Most used cases of semisupervised learning algorithms rely on Euclidean distances. The famous semisupervised learning algorithms are K Nearest Neighbor (KNN) and its branch [65]. The work structure here is simply finding a cluster of labeled data, computing the average of this cluster, and then computing Euclidean distances of unlabeled data from this mean of labeled data. Finally, labeling data set based on their nearest known average cluster of data will be done. After labeling the data set, supervised algorithms such as CNN or DNN will classify the data set [66]. How semisupervised algorithms will deal with semilabeled data has been shown in

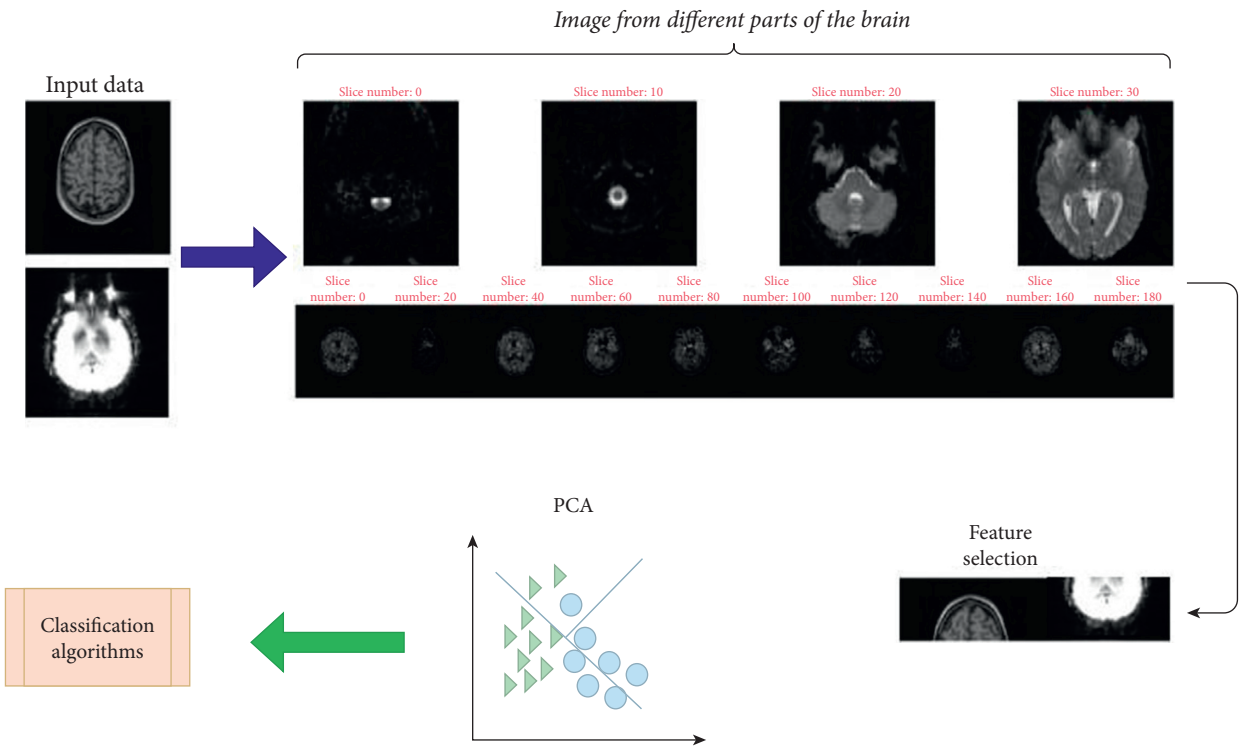


Figure 8: Diagram of feature selection and PCA.

Figure 9. Most of the different classification methods will repeat some examples from nature. Semisupervised algorithms will use the procedure of learning in humans. With different years of education, we humans will learn a little labeled information and solve unseen challenges. A variety of different models have been used as a semisupervised algorithm for classification. El-Yacoubi et al. [67] have proposed an approach based on generating each subject cluster and analyzing the correlation of these clusters with NC, AD, and MCI profiles. The main aim of their work was to find the optimal number of clusters and a subset of valuable features that help reach an excellent discriminative algorithm. They used a semisupervised algorithm based on normalized mutual information feature selection, which guides a clustering algorithm to optimize the choice for the number of optimal clusters and the discriminative power of each three-output class. Pohl and Davatzikos [68] used semisupervised algorithms for classification tasks which use both labeled and unlabeled data for training as all semisupervised algorithms do.

They used clustering methods to deal with unlabeled data; then, for training the labeled data, they used the linear Laplacian support vector machine (LapSVM) [69]. Gorriz et al. [70] Proposed a novel case-based model selection method at their time, which syndicates hypothesis testing from a separate set of expected results and feature extraction. For the training and validation part, they have used a cross-validation strategy for avoiding overfitting. This proposed model will take advantage of proper feature selection. Using good features, this model tries to improve the network's performance on validation and test sets.

4.1.2. Generative Adversarial Network (GANs). GANs were first developed and pioneered in 2014 and, from then until now, have gathered much attention on image generation tasks [71]. The use of GANs as semisupervised methods is one of the most capable areas of real-world application of GANs. In a nutshell, semisupervised GAN (SGAN) is a subset of GANs in which discriminator is a multiclass classifier, and its generator is an expanded CNN. Instead of distinguishing between only two classes, it learns to distinguish three or more classes with the production of fake images. Generator in SGAN is not the essential part, unlike other conventional GANs which have aimed to produce new high-quality data from the useless noisy data set [72]. The structure of SGAN has been shown in Figure 10. As shown in Figure 10, the generative part of the network will expand the dimension of noise to create fake images. The SGAN generator's aim is the same as in the original GAN. The generator of ordinal GANs will take a vector of random like Gaussian noise. It will produce fake examples or samples where the goal is very similar to natural images as input data of the training data set. The goal of the generator is the same in SGANs and GANs. The SGAN discriminator, on the other hand, differs obviously from the original GAN procedures. Discriminators of SGANs will get three sorts of inputs: fake examples produced by the generator model, real examples without any labels from the train set, and real examples with labels from the training data. Instead of binary classification, the SGAN discriminator's goal is to distinguish between real and fake examples and then use the labeled data and fake image to classify the input into different classes. In

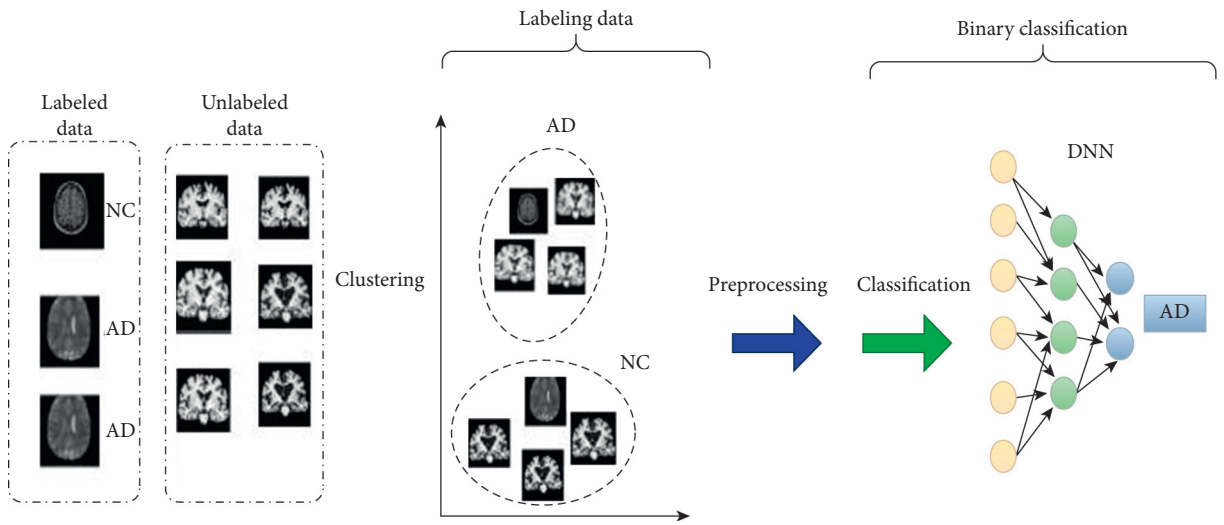


Figure 9: Semisupervised learning method for binary AD classification.

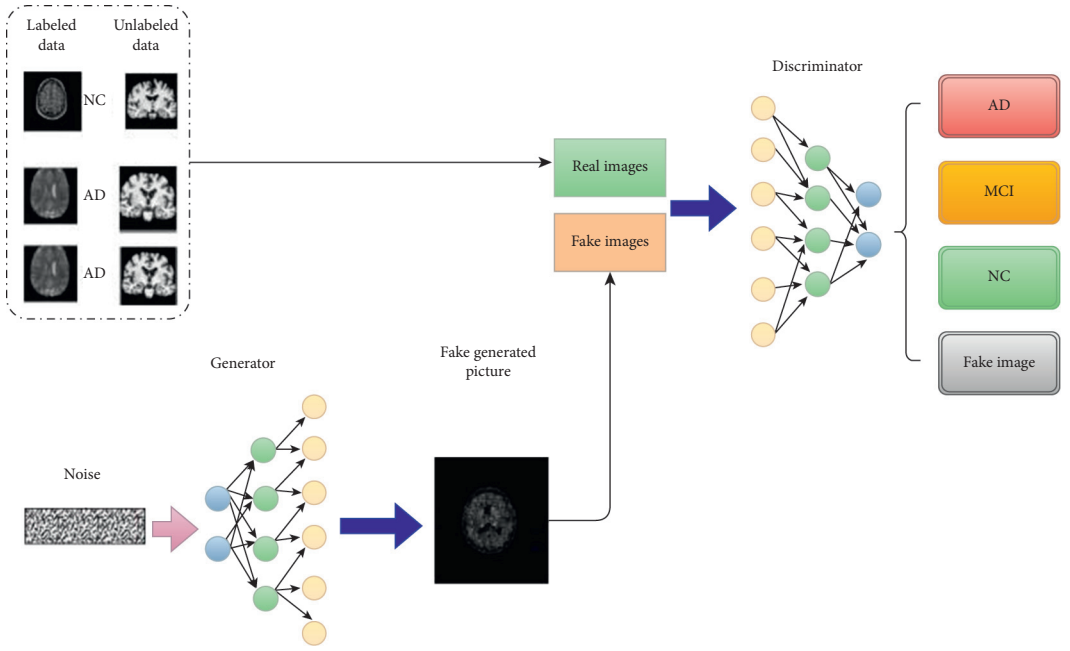


Figure 10: Semisupervised GAN learning method for binary AD classification.

research by Yu et al. [73], the authors used SGAN to predict MCI and AD. They proposed that THS-GAN is designed for semisupervised classification. They have used partially labeled data set input.

Then used the distribution of labels to predict the label for both labeled and unlabeled data and the newly generated samples. Their model can profit from the mechanical information of the brain. Also, they introduced high-order pooling, which helps to exploit more essential features by using the second-order statistics of the holistic MRI images. The result of THS-GAN demonstrates that the classification of MCI vs. NC has been done with 89.29% accuracy. AD vs. NC classification has been done with 95.92% accuracy.

4.2. Supervised Learning. Supervised methods have higher popularity because of their performance. In this section, supervised methods for AD detection have been reviewed. In real-world data sets, much of the data are not correctly labeled. So, the use of supervised methods will be restricted only to the labeled data set. This review first focused on DL algorithms and then used ML algorithms such as support vector machine (SVM) [74].

4.2.1. DNN. The DNN structure is similar to the traditional multilayer perceptron network structure, and with more perceptron layers, the structure of models will get deeper. This deep model can learn the more sophisticated pattern and

relations from input data [75]. The model with a deep layer can determine the best features for classification. DNNs have been used only as supervised methods. We separate the DL method into Conventional Neural Network (CNN), Recurrent Neural Network (RNN), and MLP (Vanilla or Residual Model). 2D CNN is the most used type of CNN for the classification and detection of ARD. The new research focused on using 3D and 4D convolutional layers to extract information from videos [76, 77]. First, we describe CNN because of the common use of this method for AD classification.

4.2.2. Convolutional Neural Network (CNN). The difference between DNN and CNN methods can be described in the connectivity of neurons in different CNN network layers. The connection in the primarily convolutional layer is not connected to all connections of the second layer. The first layer of convolutional layers extracts simple structure from images like orthogonal and diagonal lines. As the CNN structure goes further and gets deeper complex shapes like face and trapezoid shape can be extracted from images [78]. In each layer of CNN, CNN's top layers are connected to a restricted number of neurons in the next or precious layer located within a specific rectangle shape. This building structure of CNN permits the proposed model to focus on a small sublevel of features in the first hidden layer. Model uses them into more extensive higher-level features in the second hidden layer, and the same structure will be repeated until the last layer. This hierarchical structure is typical in real-world images, like ARD. Complete structural work of CNN consists of the following:

(i) Specifying the convolutional kernels which are defined by a width and heights

(ii) Specifying the number of input and output channels of each convolutional layer

(iii) Specifying the depth of each convolutional layer must be the same as the number of RGB colors in input data

The structure of CNN consists of a pooling layer that helps whole algorithms work with the smaller size of pictures and have the same performance as full pictures [79]. At the end of the convolutional layer, a flatten layer and stack of multilayer perceptron layers complete the whole structure of CNN. The structure of some CNN is shown in Figure 11.

As shown in Figure 11, CNN will increase the number of channels or depth of the input data while decreasing the widths and height of the input picture. The exact location of a feature is less important than its irregular location comparative to other features in the convolutional layer. It is the idea behind the use of a pooling layer in convolutional neural networks. The pooling layer will extract the essential features from the output of convolutional layers. With the pooling layer, the height and width of the input layer will decrease by a factor of the pooling layer's window. The use of the pooling layer will help to reduce overfitting and computational power, which is needed to train the CNN model. After the convolutional layer, a fully connected layer will be used. This model uses MLP as a feedforward network and tries to classify input features [80]. Sarraf and Tofighi [81] used the fundamental type of ordinal CNN called vanilla CNN to exploit different patterns from input data to develop a proper model for AD diagnosis among elderly patients. They proposed a state-of-the-art DL-based procedure to distinguish AD from NC using MRI and fMRI. The use of proposed pipelines was performed on a GPU-based powerful device as the computing platform. They have categorized their input data into three parts of train, test, and validation. Their research use of fMRI data has been used for the first time in the DL model as an application to distinguish between AD, MCI, and NC. Spasov et al. [82] have proposed a DL to classify input data, combining sMRI, demographic, neuropsychological, and APOe4 genotyping pictures as input data for classification tasks. The innovation of their work consists of the DL model, which learns to distinguish between MCI vs. AD and AD vs. NC. All the analyses of this work were performed on a subset of the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The data set used in their research consists of 785 participants subcategorized as 192 AD, 409 MCI, and 184 NC. Their research found that the most helpful combination of input data included the sMRI images and the demographic, neuropsychological, and APOe4 data. More and more CNN algorithms have been developed for AD classification. For better understanding, the convolutional layer and the effect of each layer at its input are shown in Figure 12. As shown in Figure 12, as the network goes deeper and deeper, more structure of each data set will get extracted and difference between two input pictures will get more precise.

4.2.3. RNN with CNN. RNN is a subtype of DNN that remembers earlier time-series data and uses this information with present input for predicting the future. RNN is a structure of time-variant algorithms with repeated information along with its layers.

The min different part of RNN which separates them from other DNN models is the structure of hidden states. This hidden state will help to extract useful information in the sequence of data. The same recurrent layer will be used as a stacked layer after each other. The use of this structure will be helpful in input data as video for the AD classification task. It results in the reduction of the complexity parameters, unlike the other DNN [83]. Two main subcategories of RNN algorithms are Long-Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). The main purpose of LSTM is to help to use and maintain the error signals through the structure of network models on the long sequence of data.

This maintenance of signal error will be done with short-term memory and long-term memory. The common activation functions used in RNN structure are sigmoid and tanh, which can help backpropagate error signal through different layers. In LSTM, different gates will be used for the preservation of information and ignoring redundant information. These gates consist of learning gates, forget gates, use gates, and remember gates. LSTM model is similar to a computer's memory. This cell's structure can be used

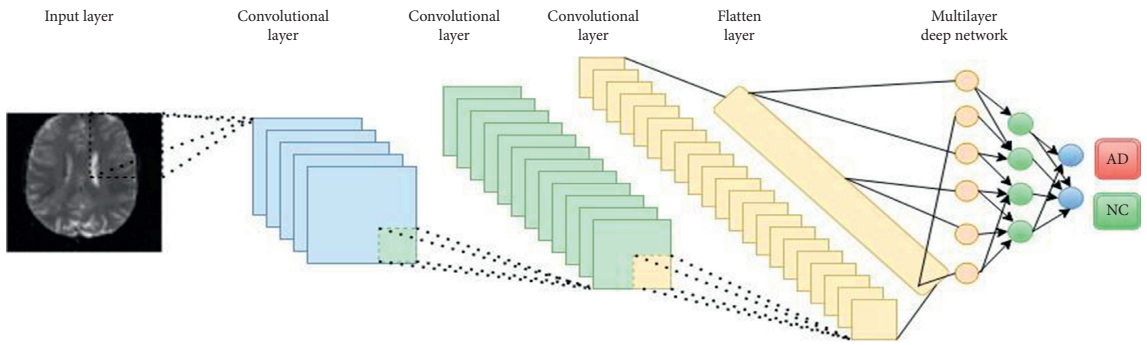
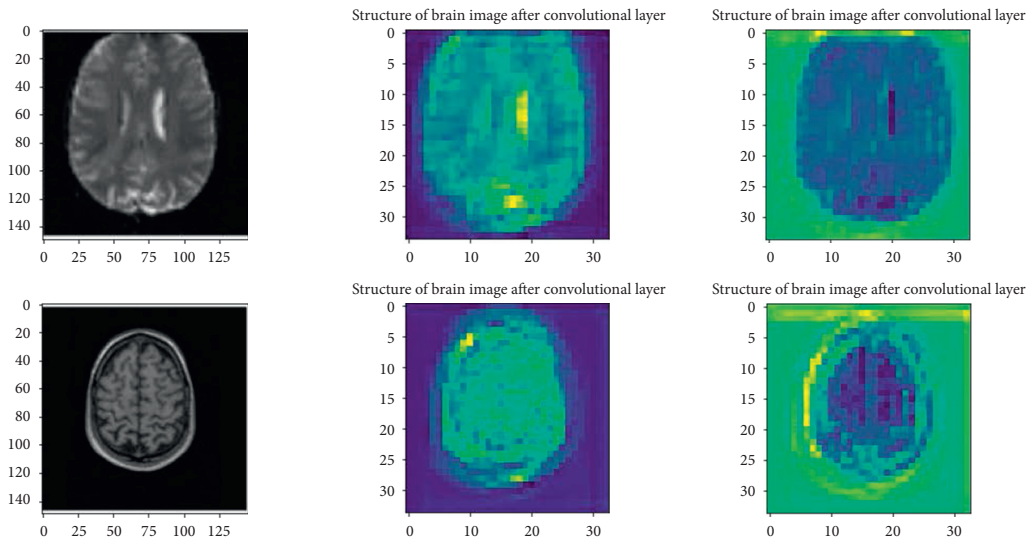
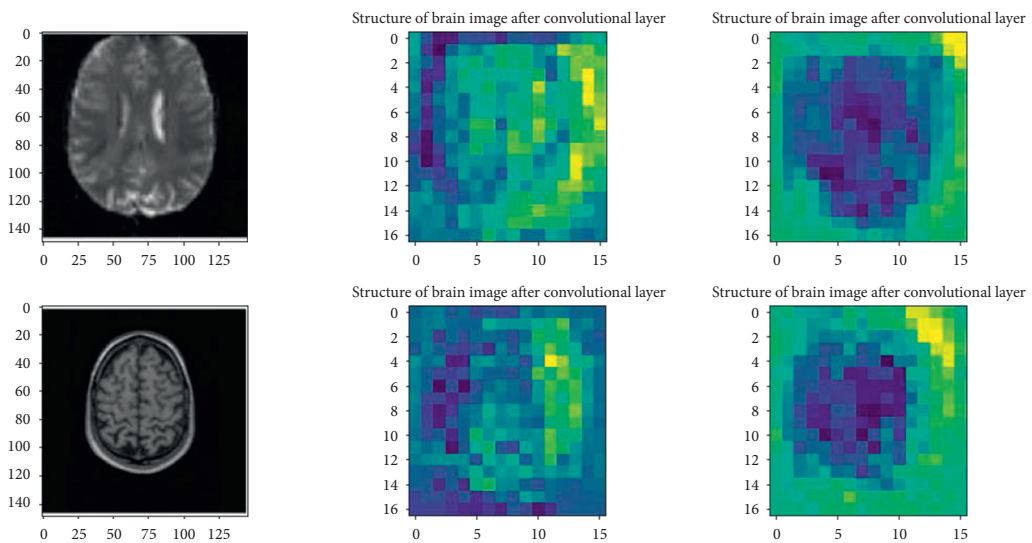


Figure 11: Structure of CNN.



(a)



(b)

Figure 12: Continued.

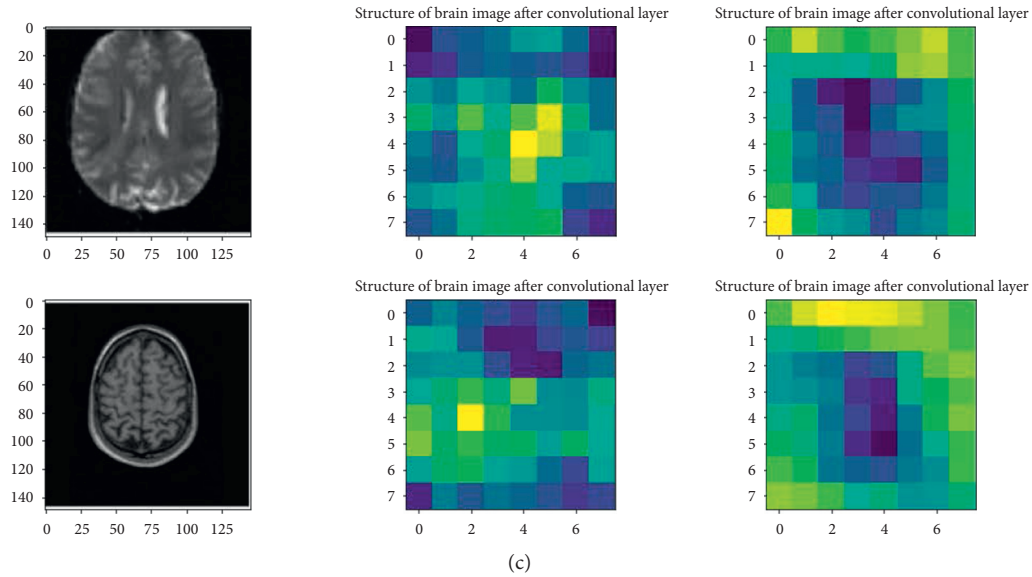


Figure 12: Result of convolutional layer on input image for (a) first convolutional layer, (b) second convolutional layer, and (c) last convolutional layer.

independently to decide which information to store and which information to forget [84]. The structure of the RNN and LSTM network is shown in Figure 13. As shown in Figure 13, the LSTM structure is more complex, but the flow of loss throws each layer for a better optimization process. Karlekar et al. [85] had used language changes as a sign of a patient's cognitive functionality for early diagnosis of AD. In their work, they have used Natural Language Processing (NLP) for AD classification tasks. In the proposed work, they have used CNN, LSTM\RNN to distinguish between language samples from AD and other stages of AD. They have reached 91.1% accuracy with this newly conducted procedure.

4.2.4. Machine Learning Algorithm. The use of supervised ML algorithms for classification is not more different than a semisupervised algorithm for classification. The difference between these two algorithms appears in part of the labeling data set. For a supervised classifier, all data labels are specified, but in semisupervised one, all data labels are not clear. A method that has been used for a handful of researches is SVM. In research by Kloppel et al. [86], they used SVM with the linear kernel to classify MRI scans from proven AD patients and MCI in elderly cases with two different scanning equipment and neuroimaging techniques. Finally, they used these methods to differentiate between patients suffering from AD and patients with frontotemporal lobar degeneration. The result of the classification models consists of 89% accuracy. In another research by Montagne et al. [87], they proposed a model based on a noninvasive neuroimaging technique for early diagnosis of ARD. They have used SVM to classify Alzheimer's disease versus NC group of patients. They have reached better classification rates by focusing on parietal and temporal lobes of brains with SVM. Other ML methods used for AD classification comprise using second-order derivation or Hessian of loss function for updating weight. Extreme Learning Machine (ELM) is another popular algorithm for classification too.

ELM is a learning algorithm conducted without using multiple different stacked layers and tuning this vast majority of hidden layer and input so that the computation time will decrease [87]. In ELM, unlike other DNN algorithms and SVM, the hidden layer parameters consist of weights and biases. One hidden layer does not need to be tuned after importing the data set and can be generated randomly before the training samples are acquired. This modification will help the network for faster learning processes at the cost of higher loss [88]. In an article by Lama et al. [89], they proposed an AD diagnosis approach using sMRI images to discriminate AD, mild MCI, and NC. They have used SVM and a regularized ELM for prediction. Lama et al. experimented on the ADNI data sets. They showed that regularized ELM with the feature selection techniques could significantly improve the classification accuracy of AD from MCI and NC subjects.

5. Comparison Based on the Different Types of Data

5.1. fMRI, sMRI, and rs-fMRI. In this section use of different algorithms based on fMRI and sMRI pictures has been surveyed. In research by Duc et al. [7], different variants of MRI like sMRI and rs-MRI scans of 331 subject patients for Mini Mental State Examination prediction have been used. In their work, 3-dimensional CNN, a method, has been developed for the mentioned classification task. Task linear regression, support vector regression, bagging-based ensemble regression, and tree regression were used. They have reached a test accuracy of 85.27% for the classification of AD versus NC. Also, it was mentioned that the SVM method with desired features has reached the lowest root mean

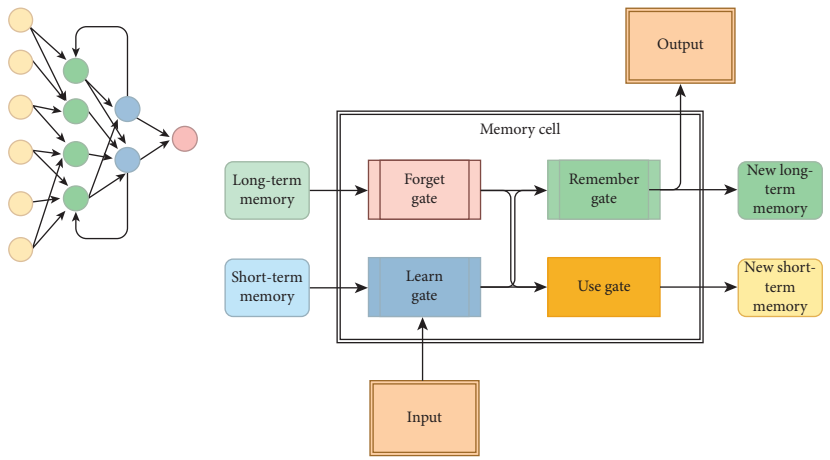


Figure 13: Structure of (a) RNN versus (b) LSTM.

square error of 3.27 and the highest R2 value of 0.63. As seen from this work structure, a perfect relation between input data and regression task had not been reported.

In another research, Ramzan et al. [90] have studied the effect of rs-fMRI for multiclass classification of AD and different stages of AD-related disease. They have used one of the famous structures of CNN, which is ResNet-18 [91]. The mentioned structure consists of skip connection for better use of error signal in backpropagation technique for optimization. Skip connection will help the model to get deeper with the vanishing of gradient signals. They have conducted model training from scratch with single-channel input. On the other hand, they also have performed transfer learning with and without an extended network architecture. Transfer learning helped them reach a state-of-the-art result with an average accuracy of 97.92% and 97.88% for all the AD stages prediction. Altinkaya et al. [92] have used superresolution on the low quality of input MRI picture and then used CNN for AD prediction. With superresolution, image processing time has been shortened, and images with high-quality features have been obtained. The result of their study concluded that the performance of proposed methods increased day by day. The resulting accuracy with CNN for AD detection using 302 MRI and fMRI instances was 99.9%.

Korolev et al. [93] have used 3DCNN and achieved better performance without including preprocessing steps like feature extraction in their proposed architecture. These proposed methods consist of CNN and the residual NN. In terms of performance metric Area under Receiver Operating Characteristic (AROC), receiver operating characteristics (ROC) curves, and accuracy have been evaluated. They have proposed a branch of CNN which is called Vox CNN and ordinal ResNet. AD vs. NC achieved the best result with AUC 0.88 and acc 0.79 using VoxCNN and AUC 0.87 with acc 0.80 using ResNet.

Li et al. [94] have used hippocampal MRI as input data. They surveyed 2146 subjects for prediction of MCI in each subject and how this stage will progress and lead to dementia in a time-to-event analysis setup. This study focused on the hippocampus region of the brain. They have reached 0.813 AUROC. For better results use of whole-brain MRI data can help different DL models to reach better classification tasks.

Yang et al. [95] have used SGANs with clustering as the novel semisupervised deep-clustering method. They have surveyed 8,146 scans which included NC, those with MCI, and dementia cognitively. One-fourth of their input data consists of sMRI data. Their proposed method has separated patients into four types of peoples: NC, MCI, relatively more significant memory impairment, and advanced dementia. Results of their work confirmed that the Smile-GAN model was able to cluster participants with 99.9% accuracy even with very severe confusing patterns as inputs. Zeng et al. [88] have proposed a total baseline for the diagnosis of AD. Their proposed model consists of MRI image preprocessing, feature extraction, PCA, and SVM algorithm developments at the end. For optimization, a particle swarm optimization algorithm was proposed to optimize the SVM parameters instead of traditional optimization methods. With their proposed model, they successfully conducted a classification of AD and MCI using MRI scans from the ADNI data set. The proposed algorithm in their research has a state-of-the-art performance compared to other presented methods [87]. Dua et al. [96] used MRI scans from different online repositories to create a diverse data set. They have deployed CNN, RNN, and LSTM individually and as ensemble methods together. The results of the proposed work show that, with the combination of CNN with RNN and CNN with LSTM, an accuracy of 89.75% has been achieved. Meanwhile, they have also used ensemble method with the bagging strategy of the first models. They achieved an accuracy of 92.22%. According to the structure of video and 4D, CNN's use of RNN with this specific data structure has shown promising results [97]. It has been shown that use of 3DCNN and stacked bidirectional LSTM could help the researcher to reach state-of-the-art performance based on accuracy and loss [97]. In research by Kruthika and Maheshappa [76], a DNN has used capsule networks, a branch of CNN. This method will use the benefit of not using

any max-pooling layers. The structure of the capsule net has been developed to use spatial information in the pictures. They have proposed a 3DCNN, which works with video structure too. Their work proposed a method based on CNN and capsule networks for AD prediction using MRI data as input. In the end, they have used the proposed model for creating an application for the AD detection task. They found that both the 3D capsule network method and CNN with pretrained 3D autoencoder improved the predictive performance compared to other structures of CNN trained from scratch. In another research, multiclass classification between AD, MCI, and NC has seen a state-of-the-art CNN called MCADNNet [98]. Using the mentioned method, the researcher reached 92.6% accuracy with distinguishable accuracy of 97% for MCI classification versus AD subjects. Furthermore, even after applying the decision-making algorithm, accuracy rates of 99.77% and 97.5% were achieved for MRI and fMRI pipelines to classify AD versus NC. Amini et al. [99] used CNN and ML models for finding the severity of AD. They have used fMRI pictures as input data set. Also, they have used a sophisticated procedure for converting raw fMRI input data into a valuable data set for training CNN and ML models. The performance of the proposed CNN model for different stages of AD classification was 98.1%, 92.4%, 97.0%, and 100% precision for the low, mild, moderate, and severe status of Alzheimer's patients. The absolute accuracy for their work was reported at 96.7%. They have used only vanilla CNN with one convolutional layer and three fully connected layers. Mentioned structure without any pooling layer is a variant of CNN, which helps to extract features from low dimension feature maps [100].

5.2. PET. The use of PET as a biomarker for AD-related detection is a relatively new procedure. Ozsahin et al. [101] have proposed amyloid-beta plaques combination as a marker. They have implied that mentioned marker should be taken for granted as a "start" of the degenerative process in the brain of most cases. This symptom can be seen earlier than other clinical symptoms, which will appear later in AD subjects. They have used DNN with 18F-florbetapir PET data for automatic classification of four patient groups into four different classes of people, which consist of the following:

(i) AD
(ii) Late MCI
(iii) Early MCI
(iv) Significant memory concern
(v) NC

With this specification, even early stages of AD disease can be detected. Their work on 30% of data as test cases based on sensitivity, specificity, and accuracy was measured as 92.4%, 84.3%, and 87.9%, respectively. Other parts of the second experiment, which consist of the classification of NC versus late MCI images, resulted in sensitivity, specificity, and accuracy of 62.9%, 70.0%, and 66.4%. Other experiments in the classification of NC versus early MCI images have shown the sensitivity of 60.0%, specificity of 60.0%, and accuracy of 60.0%. Finally, NC has significant memory concern as other classes have been used for classification, and the result of their work showed sensitivity, specificity, and accuracy values of 60.0%, 45.7%, and 52.9%.

The use of ML for improving AD detection in the analysis of digital biomarkers within PET imaging has emerged recently. Islam et al. [102] have proposed a 3DCNN for AD diagnosis using only PET scans. They have reached the classification accuracy of 88.76% for NC versus AD categories. Their experiment also developed a regular CNN model using axial, coronal, and sagittal segments from each subject's brain from PET data; in the end, they have achieved 71.45% accuracy for NC versus AD classification.

In another research by Vasan et al. [103], F-FDG-PET brain images from the ADNI repository have been used to train the models for testing the result of the training model; they have used an independent test set from individual patients. They have used a branch of CNN which is called InceptionV3 [104]. Using this model, they reached AUC to predict AD, MCI, and NC, of 92%, 0.63%, and 0.73%, respectively. The AUROC for this classification task was 0.98, 0.52, and 0.84 to predict AD, MCI, and NC, respectively. The reported sensitivity consists of 100%, 43%, and 35% for AD, MCI, and NC prediction, respectively. As it is shown, the best result has been achieved for AD classification.

Lu et al. [105] used Fluoro-deoxy-glucose positron emission tomography (FDG-PET) to detect the brain's metabolic activity in the different subjects as data set. They proposed a novel DL method at the time for analyzing the FDG-PET. They used this information to classify MCI subjects with symptomatic AD and distinguish them from other subjects with MCI stages. The result of their work shows 82.51% accuracy of classification just using measures from a single modality. Because of the similarity between these two stages, their work can distinguish between two similar stages rather than AD versus NC.

Adeli et al. [106] proposed a semisupervised algorithm for first dealing with little labeled data and the vast majority of unable data. The discriminative classification method based on the least-squares formulation of linear discriminant analysis has been used in work. The use of a linear discriminant model has helped them to deal with noisy and unwanted data. They have surveyed Parkinson's and AD diseases as neurodegenerative diseases. They have used their framework to create an application for neurodegenerative disease diagnosis. They have reached 92.1% accuracy for AD versus NC classifications. In another work by Gamberger et al. [105], they have analyzed 5-year longitudinal outcomes and biomarker data from 562 subjects with MCI from ADNI. The mentioned algorithm identified homogenous clusters of MCI subjects with evidently diverse predictive cognitive courses. In the end, they have reported high sensitivity and specificity for the classification of AD versus NC.

In another research by Liu et al. [107], they used FDG-PET data as input for training. They have used DL models with the combination of GRU and convolutional layers. Their proposed model has abled them to use intrasegment and intersegment to classify different stages of AD disease. They have used the different frames of video and then

developed a CNN model on these images. In the structure of the proposed model, the convolutions layer has been used only to capture the valuable features from input image data. After reaching good data, they used GRU and RNN to learn and integrate the intersegment features for classification tasks. They reached 95.3% for AD versus NC classification and 83.9% for MCI versus NC classification based on AUC metrics. Their proposed method reached 91.2% accuracy, 91.4%sensitivity, and 91.0 specificity for AD versus NC condition, respectively.

5.3. Combination of Data. Combining PET and different branches of MRI data with the aim of reaching a diverse data set with complementary effects has gained attention. Gupta et al. [35] have used a combination of four different biomarkers: FDG-PET, sMRI, the level of protein in cerebrospinal fluid (CSF) data, and apolipoprotein (APOE) genotype. They have used data set from ADNI as their baseline data set. In total, they have surveyed 158 patients whose all of the mentioned input data are available for each of these subjects. In their study, patients were divided into 38 subjects of AD, 82 subjects of MCI groups, and the remaining 38 subjects in the NC group. With these categories of data set, their data set was imbalanced. They used a kernel-based multiclass SVM classifier with a grid-search method and truncated PCA to determine the best feature to use for training. They have reached AUROC 98.33%, 93.59%, 96.83%, 94.64%, 96.43%, and 95.24% for AD versus NC, MCIs versus MCIc, AD versus MCIs, AD versus MCIc, NC versus MCIc, and NC versus MCIs classification. The accuracy, sensitivity, and F1-score of AD and NC classification were 98.42%, 100%, and 98.42%. The result of their work has shown some state-of-the-art results for AD classifications. The result of their work has shown a significant rise in accuracy in different stages from previous works.

Youssofzadeh et al. [34] used MRI and PET in combination due to their complementary nature as the input data set. They have used a multimodal imaging ML to enhance AD classification performance metrics. They have used 58 AD subjects, 108 MCI subjects, and 120 NC subjects from the Australian imaging, biomarkers, and lifestyle data set. For classification of AD versus NC, MCI versus NC, and AD versus NC, they reached 95.7%, 95.8%, and 95.1% accuracy, respectively. Also, they have found multikernel learning regression analysis for excellent predictions of diagnosis of AD in subjected samples with a relation factor of 0.86. In addition, they have reached significant correlations between developed methods and delayed memory recall scores with a relation factor of 0.62.

In another research, Li et al. [108] used whole-brain images as input and designed a disease-image-specific neural network for the classification task of AD subjects. They have used MRI and PET scans as input data for classification tasks. Also, they have used feature-consistent GANs to produce some images for better results of classification. Using this branch of GAN, they have encouraged the proposed model to use these produced images and authentic images as a consistent data set for final prediction. Their work conducted a state-of-the-art performance in both AD identification and MCI conversion prediction tasks at times. They have reached 34.18% in terms of the PR-AUC score.

In another study, Dukart et al. [109] used MRI and fluorodeoxyglucose FDG-PET to improve the detection accuracy of differentiation subjects with AD complication and frontotemporal lobar neuron connection failure. They have used an SVM classifier for this task. SVM classification has used combined information from different ROI of each subject from FDG-PET and MRI based on comprehensive quantitative meta-analyses. For the ADNI data set, accuracy rates of 88% have been achieved. In another work by Triebkorn et al. [110], they have used an amyloid-beta marker from each PET data from each patient. Also, they have used MRI, specifically amyloid-beta binding tracer PET, and tau protein (Tau) binding PET from 33 participants of ADNI3. Their work aims to classify AD MCI and NC using SVM and Random Forest classifiers together. They have reached 90.5% accuracy for NC versus AD and MCI classification. For AD versus NC and MCI classification, they have reached 78.3% accuracy. The sensitivity of this classifier method for NC versus AD and MCI classification was 90.5%. They also have reached AD versus HC and MCI classification of 78.3%.

In research by Kim and Lee [111], they proposed an autoencoder and sparse ELM to classify AD versus NC and MCI. They have used MRI, PET, and CSF pictures from 93 individual subjects. They have extracted volume and means ROIs as input features. At last, they have used a stacked sparse ELM autoencoder for the classification task. The use of an autoencoder for changing the input space into smaller later space has been done. For evaluating the proposed model, they have used 10-fold cross-validation. The classification result has shown more than 96% and 86.44% accuracy for classifying AD versus NC and MCI versus NC subjects.

5.4. Comparison between Different Modalities. In this section, we compared different neuroimaging modalities according to their performance. Comparison has been made based on different metrics such as accuracy, precision, recall, F1-score, and area under the receiver AROC. Accuracy is given in the following:

$$\begin{aligned}
\text{sensitivity} &= \left[\frac{\text{TP}}{\text{TP} + \text{FN}}\right] \times 100,\\
\text{specificity} &= \left[\frac{\text{TN}}{\text{TN} + \text{FP}}\right] \times 100,\\
\text{positive predictive value (PPV)} &= \left[\frac{\text{TP}}{\text{TP} + \text{FP}}\right] \times 100,\\
\text{negative predictive value (NPV)} &= \left[\frac{\text{TN}}{\text{TN} + \text{FN}}\right] \times 100,\\
\text{accuracy (ACC)} &= \left[\frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}\right] \times 100.
\end{aligned} \tag{1}$$

Table 1: Summary of methods and neuroimaging techniques for binary classification of AD versus NC results.

Num.	Reference	Year	Method data	Neuroimaging	Accuracy	Sensitivity	AUROC
1	Duc et al. [7]	2020	3DCNN	rs-fMRI	85.27%	—	—
2	Ramzan et al. [90]	2020	CNN (ResNet-18)	fMRI	97.88%	—	—
3	Altinkaya et al. [92]	2020	CNN	fMRI	99.9 (on 303 samples)	—	—
4	Korolev et al. [93]	2017	CNN (ResNet)	MRI	80%	—	0.87
5	Li et al. [101]	2019	CNN	MRI	—	—	0.82
6	Yang et al. [95]	2021	GANs	MRI	99%	—	—
7	Dua et al. [96]	2020	CNN-LSTM (RNN)	MRI	92.22%	91.92%	
8	Zeng et al. [88]	2018	CNN (PSO)	MRI	76.85%	—	—
9	Kruthika and Maheshappa [76]	2019	3DCNN (capsule net)	sMRI	92.98%	—	0.98
10	Sarraf et al. [112]	2019	CNN (MCADNNet)	fMRI + MRI	99%	95%	—
11	Ozsahin et al. [102]	2019	ANN	PET	87.9%	92.4%	—
12	Islam and Zhang [113]	2019	3DCNN	PET (FDG)	88.76%	—	—
13	Vasan et al. [103]	2020	CNN (Inceptionv3)	PET	92%	100%	0.98
14	Lu et al. [105]	2017	ANN	PET (FDG)	82.51%	—	—
15	Adeli et al. [106]	2018	RF	PET	92%	—	0.94
16	Liu et al. [110]	2021	CNN-RNN (GRU)	PET	91.2%	92.4%	0.94
17	Gupta et al. [34]	2017	SVM	PET + MRI	92%	98.42%	0.98
18	Youssofzadeh et al. [109]	2013	MKML	PET + MRI	95.7%	—	—
19	Li et al. [108]	2018	GANs	PET (FDG) + MRI	—	—	0.32
20	Dukart et al. [107]	2018	SVM	PET (FDG) + MRI	78.3%	90.5%	—
21	Kim and Lee [111]	2018	ANN (AE)	PET (FDG) + MRI	96%	—	—
22	Billones et al. [114]	2016	CNN (DemNET)	MRI	98.33%	98.99%	—
23	Talo et al. [115]	2019	CNN (DeepResNet-50)	MRI	95.23%	97.16%	—
24	Yu et al. [73]	2020	THS-GAN	MRI	95.92%	—	—
25	Moradi et al. [116]	2015	Semisupervised model	MRI	87%	82%	
26	Kloppel et al. [86]	2004	SVM	MRI	88%	—	—
27	Montagne et al. [87]	2013	SVM	PET	82%	81%	—
28	Lama et al. [89]	2017	ELM	Smri	83.38%	93.01%	0.85
29	Lin et al. [117]	2021	GANs + 3D VGG	PET (FDG) + MRI	74.1%	75.00%	0.92
30	Zhou et al. [118]	2021	GANs + DNN	MRI	88.73%	63.09%	0.932
31	Baydargil et al. [119]	2021	GANs	PET	96.03%	—	0.7521
32	Venugopalan et al. [120]	2021	DNN + RF	PET + MRI	89%	96%	—
33	Zhang et al. [121]	2021	3DCNN	MRI	97.35%	97.10%	0.9970
34	Mehmood et al. [122]	2021	CNN	MRI	98.73%	—	—
35	Raju et al. [123]	2021	CNN (VGG16 based model)	MRI	99.2%	98.5%	—
36	Subramoniam [124]	2021	CNN (ResNet 101)	MRI	99.71%	0.99	1
37	Lella et al. [125]	2021	ELM	MRI	89.25%	0.78	—
38	Acharya et al. [126]	2019	KNN	MRI	94.54%	0.96	—
39	Amini et al. [99]	2021	SVM + PCA	fMRI	85.8%	0.95	0.92
40	Amini et al. [99]	2021	KNN + PCA	fMRI	77.5%	0.98	0.93
41	Amini et al. [99]	2021	CNN	fMRI	96.7%	0.98	1

Based on mentioned metrics, an evaluation of binary and multiclass classification can be done. As shown in Table 1, different methods used MRI as input data. Most of them have reached a good accuracy too. Using PET as data set for the sake of AD detection only based on accuracy is lower than MRI as input. Combined modalities have shown better performance based on accuracy than single data. The best algorithms for prediction consist of a different branch of CNN.

6. Results

Alzheimer's disease is the main cause of memory loss and dementia in people older than 65. AD consists of a gradually progressive neurodegenerative disease process consisting of gradual synaptic integrity and loss of cognitive functions. For detecting AD, a marker as a sign of progressive neurodegenerative is needed. This marker has a more critical role for AD detection than other diseases because beta-

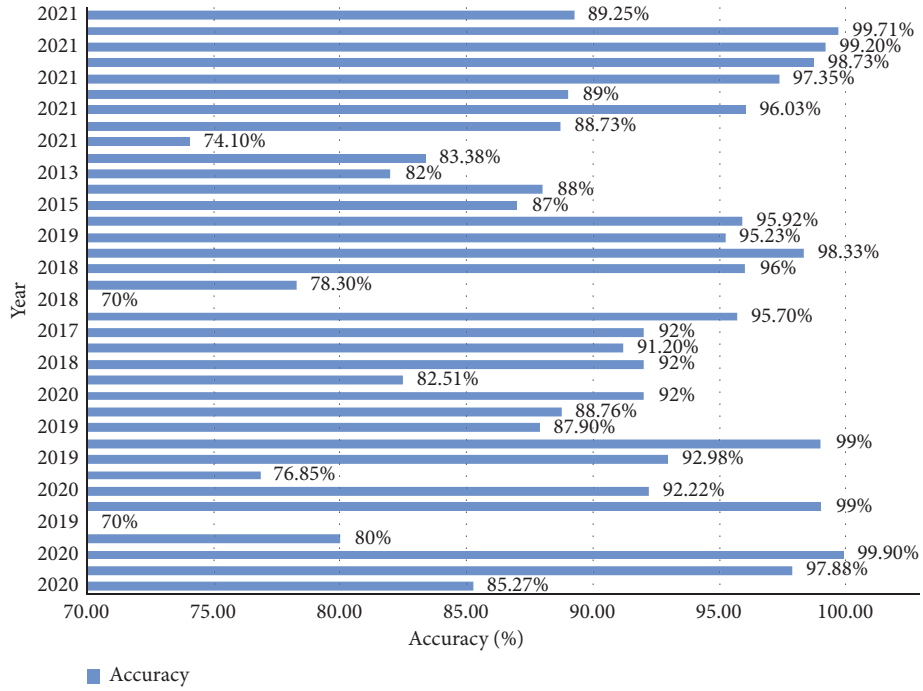


Figure 14: Different accuracy based on different years.

amyloid can form a complex structure with some metal ions. For assessing this marker, some neuroimaging techniques like fMRI and sMRI, PET, PET (FDG), and combination of these data have been used (see Table 1).

AI and its branches have helped with AD detection in patients with various methods. Some of these algorithms consist of ANN, CNN, and GANs as supervised and semisupervised algorithms. Other famous ML algorithms such as KNN, SVM, and RF have been used too. This review article discussed different neuroimaging techniques and has shown some effects of some of the DL methods on this input data for AD detection. In the end, we have to summarize these different techniques and compare them based on accuracy, sensitivity, and AUROC. The results showed that combined neuroimaging techniques are a newly open field, and DL methods for detecting AD on average are of high accuracy. For people older than 65, if we want a model to be sensitive and its prediction considers low false-negative outcomes, MRI and CNN base methods have shown better results. In full use of the combination, methods have much more unspecific areas for research. As shown in Table 1 more than half of the reviewed article was published in 2020 and 2021. This year's focus on GANs and using VGG and resent structure with combination data from different repositories have gained more attention. Some articles have reached 99% accuracy with old popular CNN for AD versus NC classification task only. With proceeding years, using different methods for reaching a good accuracy has paid off, and most of the articles have reached a state-of-the-art accuracy. The result of accuracy versus years is shown in Figure 14. Time-series signals related to specific parts of the brain as input data and use of CNN or ML models for AD classification are another center of attention for researchers [100, 127]. As discussed, we focused on AD versus NC classification improvements as the main priority. The reputation of authors or the number of citations for each research has not been considered in this article.

7. Conclusion and Discussion

Different ML techniques for AD classification and detection have been reviewed in this article. For gathering different images, data time consumption and the financial issue should be considered. Also, the use of combined data has shown promising performance. So, this is a trade-off that should solve this. Most of the research has worked on the ADNI data set due to its convenient access and variety of stages of the data set. AD versus NC and MCI classification has been surveyed in most of the researches. The use of different data set will add more controversy and variety to the data set. This variety is an essential part of each data set. So, for better and more comprehensive results use of a different variety of data must be considered. Also, for better early AD detection, different AD stages as a data set should be considered.

Data Availability

Data used in this paper's preparation were obtained from the ADNI database (http://adni.loni.usc.edu/).

Disclosure

The funding sources have no involvement support in the study design, collection, analysis, or interpretation of data,

writing of the manuscript, or in the decision to submit the manuscript for publication.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] K. Blennow, M. J. de Leon, and H. Zetterberg, "Alzheimer's disease," *The Lancet*, vol. 368, pp. 387–403, 2006.

[2] J. A. Weber, D. H. Baxter, S. Zhang et al., "The microRNA spectrum in 12 body fluids," *Clinical Chemistry*, vol. 56, no. 11, pp. 1733–1741, 2010.

[3] M. Brzyska, A. Bacia, and D. Elbaum, "Oxidative and hydrolytic properties of β-amyloid," *European Journal of Biochemistry*, vol. 268, no. 12, pp. 3443–3454, 2001.

[4] Y. Y. Syed and E. Deeks, "[18F]Florbetaben: a review in β-amyloid PET imaging in cognitive impairment," *CNS Drugs*, vol. 29, no. 7, pp. 605–613, 2015.

[5] M. E. Schmidt, L. Janssens, D. Moechars et al., "Clinical evaluation of [18F] JNJ-64326067, a novel candidate PET tracer for the detection of tau pathology in Alzheimer's disease," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 47, no. 13, pp. 3176–3185, 2020.

[6] T. A. Pascoal, S. Mathotaarachchi, S. Mathotaarachchi et al., "Amyloid and tau signatures of brain metabolic decline in preclinical Alzheimer's disease," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 45, no. 6, pp. 1021–1030, 2018.

[7] N. T. Duc, S. Ryu, M. N. I. Qureshi, M. Choi, K. H. Lee, and B. Lee, "3D-deep learning based automatic diagnosis of Alzheimer's disease with joint MMSE prediction using resting-state fMRI," *Neuroinformatics*, vol. 18, no. 1, pp. 71–86, 2020.

[8] C. Y. Sander, "Simultaneous PET/fMRI for imaging neuroreceptor dynamics," Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2014.

[9] H. F. Wehrl, M. Hossain, K. Lankes et al., "Simultaneous PET-MRI reveals brain function in activated and resting state on metabolic, hemodynamic and multiple temporal scales," *Nature Medicine*, vol. 19, no. 9, pp. 1184–1189, 2013.

[10] D. P. Veitch, M. W. Weiner, P. S. Aisen et al., "Understanding disease progression and improving Alzheimer's disease clinical trials: recent highlights from the Alzheimer's disease neuroimaging initiative," *Alzheimer's Dementia*, vol. 15, no. 1, pp. 106–152, 2019.

[11] E. E. Brown, S. Kumar, T. K. Rajji, B. G. Pollock, and B. H. Mulsant, "Anticipating and mitigating the impact of the COVID-19 pandemic on Alzheimer's disease and related dementias," *The American Journal of Geriatric Psychiatry*, vol. 28, no. 7, Article ID 712721, 2020.

[12] L. H. Nicholas, K. M. Langa, J. P. Bynum, and J. W. Hsu, "Financial presentationof Alzheimer disease and related dementias," *JAMA Internal Medicine*, vol. 181, no. 2, Article ID 220227, 2021.

[13] N. Wang, M. Chen, and K. P. Subbalakshmi, "Explainable CNN-attention networks (C-attention network) for automated detection of Alzheimer's disease," 2020, https://arxiv.org/abs/2006.14135.

[14] G. Folego, M. Weiler, R. F. Casseb, R. Pires, and A. Rocha, "Alzheimer's disease detection through whole-brain 3D-CNN MRI," *Frontiers in Bioengineering and Biotechnology*, vol. 8, 2020.

[15] J. Bagherzadeh and H. Asil, "A review of various semi-supervised learning modelswith a deep learning and memory approach," *Iran Journal of Computer Science*, vol. 2, no. 2, pp. 65–80, 2019.

[16] H.-m. Lv, D. Zhao, and X. b. Chi, "Deep learning for early diagnosis of Alzheimer's disease based on intensive AlexNet," *Journal of Computational Science*, vol. 44, pp. 50–60, 2017.

[17] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. Ying, "Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 173–183, 2017.

[18] K. Trojachanec, I. Kitanovski, I. Dimitrovski, and S. Loshkovska, "Longitudinal brain MRI retrieval for Alzheimer's disease using different temporal information," *IEEE Access*, vol. 6, pp. 9703–9712, 2018.

[19] S. Kumar, C. Dabas, and S. Godara, "Classification of brain MRI tumor images: a hybrid approach," *Procedia Computer Science*, vol. 122, pp. 510–517, 2017.

[20] J. Sun, B. Wang, Y. Niu et al., "Complexity analysis of EEG, MEG, and fMRI in mild cognitive impairment and Alzheimer's disease: a review," *Entropy*, vol. 22, no. 2, p. 239, 2020.

[21] W. Li and P. C. van Zijl, "Quantitative theory for the transverse relaxation time of blood water," *NMR in Biomedicine*, vol. 33, no. 5, p. 427, 2020.

[22] K. Sakurai, T. Shintani, N. Jomura, T. Matsuda, A. Sumiyoshi, and T. Hisatsune, "Hyper bold activation in dorsal raphe nucleus of APP/PS1 Alzheimer's disease mouse during rewARD -oriented drinking test under thirsty conditions," *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020.

[23] G. Castellazzi, M. G. Cuzzoni, M. Cotta Ramusino et al., "A machine learning approach for the differential diagnosis of alzheimer and vascular dementia fed by MRI selected features," *Frontiers in Neuroinformatics*, vol. 14, p. 25, 2020.

[24] M. D. Fox and M. E. Raichle, "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging," *Nature Reviews Neuroscience*, vol. 8, no. 9, pp. 700–711, 2007.

[25] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI," *Magnetic Resonance in Medicine*, vol. 34, no. 4, pp. 537–541, 1995.

[26] T. Wu, L. Wang, Y. Chen, C. Zhao, K. Li, and P. Chan, "Changes of functional connectivity of the motor network in the resting state in Parkinson's disease," *Neuroscience Letters*, vol. 460, no. 1, pp. 6–10, 2009.

[27] T. Jones and D. W. Townsend, "History and future technical innovation in positronemission tomography," *Journal of Medical Imaging*, vol. 4, no. 1, pp. 110–113, 2017.

[28] G. Muehllehner and J. S. Karp, "Positron emission tomography," *Physics in Medicine Biology*, vol. 51, no. 13, Article ID R117, 2006.

[29] A. Gustavsson, M. Svensson, F. Jacobi et al., "Cost of disorders of the brain in europe 2010," *European Neuropsychopharmacology*, vol. 21, no. 10, pp. 718–779, 2011.

[30] R. E. Carson, "Tracer kinetic modeling in PET," in *Positron Emission Tomography*, pp. 127–159, 2005.

[31] A. Drzezga, D. Altomare, C. Festari et al., "Diagnostic utility of 18F-fluorodeoxyglucose positron emission tomography (FDG-PET) in asymptomatic subjects at increased risk for Alzheimer's disease," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 45, no. 9, pp. 1487–1496, 2018.

[32] M. Colom, B. Vidal, and L. Zimmer, "Is there a role for GPCR agonist radiotracers in PET neuroimaging?" *Frontiers in Molecular Neuroscience*, vol. 12, p. 255, 2019.

[33] A. J. Reader, S. Ally, F. Bakatselos et al., "One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays," *IEEE Transactions on Nuclear Science*, vol. 49, no. 3, pp. 693–699, 2002.

[34] V. Youssofzadeh, B. McGuinness, L. P. Maguire, and K. Wong-Lin, "Multi-kernellearning with dartel improves combined MRI-PET classification of Alzheimer's disease in AIBL data: group and individual analyses," *Frontiers in Human Neuroscience*, vol. 11, p. 380, 2017.

[35] Y. Gupta, R. K. Lama, and G. R. Kwon, "Prediction and classification of Alzheimer's disease based on combined features from apolipoprotein-E genotype, cerebrospinal fluid, MR, and FDG-PET imaging biomarkers," *Frontiers in Computational Neuroscience*, vol. 13, p. 72, 2019.

[36] Y. Fan, S. M. Resnick, X. Wu, and C. Davatzikos, "Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study," *Neuroimage*, vol. 41, no. 2, pp. 277–285, 2008.

[37] D. Zhang and D. Shen, "Alzheimer's disease neuroimaging initiative, multi-modalmulti-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease," *NeuroImage*, vol. 59, no. 2, pp. 895–907, 2012.

[38] D. Zhang, Y. Wang, L. Zhou, H. Yuan, and D. Shen, "Alzheimer's disease neuroimaging initiative, multimodal classification of Alzheimer's disease and mild cognitive impairment," *Neuroimage*, vol. 55, no. 3, pp. 856–867, 2011.

[39] K. Ritter, J. Schumacher, M. Weygandt, R. Buchert, C. Allefeld, and J. D. Haynes, "Alzheimer's disease neuroimaging initiative, multimodal prediction of conversion to Alzheimer's disease based on incomplete biomarkers," *Alzheimer's Dementia: Diagnosis, Assessment Disease Monitoring*, vol. 1, no. 2, pp. 206–215, 2015.

[40] A. Drzezga, H. Barthel, S. Minoshima, and O. Sabri, "Potential clinical applications of PET/MR imaging in neurodegenerative diseases," *Journal of Nuclear Medicine*, vol. 55, no. 2, pp. 47S–55S, 2014.

[41] J. Ashburner, "A fast diffeomorphic image registration algorithm," *NeuroImage*, vol. 38, no. 1, p. 95, 2007.

[42] J. Young, M. Modat, M. J. Cardoso, A. Mendelson, D. Cash, and S. Ourselin, "Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment," *NeuroImage: Clinical*, vol. 2, pp. 735–745, 2013.

[43] Y. Asim, B. Raza, A. K. Malik, S. Rathore, L. Hussain, and M. A. Iftikhar, "A multimodal, multi-atlas-based approach for Alzheimer detection via machine learning," *International Journal of Imaging Systems and Technology*, vol. 28, no. 2, pp. 113–123, 2018.

[44] H. P. W. Schlemmer, B. J. Pichler, M. Schmand et al., "Simultaneous MR/PET imaging of the human brain: feasibility study," *Radiology*, vol. 248, no. 3, pp. 1028–1035, 2008.

[45] N. J. Tustison, B. B. Avants, P. A. Cook et al., "N4ITK: improved N3 bias correction," *IEEE Transactions on Medical Imaging*, vol. 29, no. 6, pp. 1310–1320, 2010.

[46] T. Hara, K. Inagaki, N. Kosaka, and T. Morita, "Sensitive detection of mediastinal lymph node metastasis of lung cancer with 11C-choline PET," *Journal of Nuclear Medicine*, vol. 41, no. 9, pp. 1507–1513, 2000.

[47] L. Fan, H. Li, J. Zhuo et al., "The human brain netome atlas: a new brain atlas based on connectional architecture," *Cerebral Cortex*, vol. 26, no. 8, pp. 3508–3526, 2016.

[48] T.-E. Kam, H. Zhang, Z. Jiao, and D. Shen, "Deep learning of static and dynamic brain functional networks for early MCI detection," *IEEE Transactions on Medical Imaging*, vol. 39, no. 2, pp. 478–487, 2019.

[49] L. Kang, J. Jiang, J. Huang, and T. Zhang, "Identifying early mild cognitive impairment by multi-modality MRI-based deep learning," *Frontiers in Aging Neuroscience*, vol. 12, p. 206, 2020.

[50] T. E. Kam, H. Zhang, and D. Shen, "A novel deep learning framework on brain functional networks for early MCI diagnosis," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 293–301, Granada, Spain, September 2018.

[51] S. H. Hojjati, A. Ebrahimzadeh, A. Khazaee, and A. Babajani-Feremi, "Predicting conversion from MCI to AD by integrating rs-fMRI and structural MRI," *Computers in Biology and Medicine*, vol. 102, pp. 30–39, 2018.

[52] Y. Du and Y. Fan, "Group information guided ICA for fMRI data analysis," *NeuroImage*, vol. 69, pp. 157–197, 2013.

[53] W. E. Klunk, R. A. Koeppe, J. C. Price et al., "The centiloid project: standardizing quantitative amyloid plaque estimation by PET," *Alzheimer's Dementia*, vol. 11, no. 1, pp. 1–15, 2015.

[54] R. Haase, L. A. Royer, P. Steinbach et al., "CLIJ: GPU-accelerated image processing for everyone," *Nature Methods*, vol. 17, no. 1, pp. 5-6, 2020.

[55] S. I. Dimitriadis, D. Liparas, and M. N. Tsolaki, "Random forest feature selection, fusion and ensemble strategy: combining multiple morphological MRI measures to discriminate among healhy elderly, MCI, cMCI and Alzheimer's disease patients: from the Alzheimer's disease neuroimaging initiative (ADNI) database," *Journal of Neuroscience Methods*, vol. 302, pp. 14–23, 2018.

[56] Y. Shi, Y. Gao, S. Liao, D. Zhang, Y. Gao, and D. Shen, "A learning-based CT prostatesegmentation method via joint transductive feature selection and regression," *Neurocomputing*, vol. 173, pp. 317–331, 2016.

[57] M. Guinin, S. Ruan, B. Dubray, L. Massoptier, and I. Gardin, "Notice of removal: feature selection and patch-based segmentation in MRI for prostate radiotherapy," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 2663–2667, Phoenix, AZ, USA, August 2016.

[58] J. Lever, M. Krzywinski, and N. Altman, "Points of significance: principal component analysis," *Nature Methods*, vol. 14, no. 7, pp. 641-642, 2017.

[59] S. Graham, C. Depp, E. E. Lee et al., "Artificial intelligence for mental health and mental illnesses: an overview," *Current Psychiatry Reports*, vol. 21, no. 11, pp. 1–18, 2019.

[60] C. J. Brown and G. Hamarneh, "Machine learning on human connectome data fromMRI," 2016, https://arxiv.org/abs/1611.08699.

[61] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, pp. 1–16, 2019.

[62] E. Barbará-Morales, J. Pérez-González, K. C. Rojas-Saavedra, and V. Medina-Bañuelos, "Evaluation of brain tortuosity measurement for the automatic multimodal classification of subjects with Alzheimer's disease," *Computational Intelligence and Neuroscience*, vol. 2020, Article ID 4041832, 11 pages, 2020.

[63] H. Huo, Z. Rong, O. Kononova et al., "Semi-supervised machine-learning classification of materials synthesis procedures," *Npj Computational Materials*, vol. 5, no. 1, pp. 1–7, 2019.

[64] J. F. Rodriguez-Nieva and M. Scheurer, "Identifying topological order through unsupervised machine learning," *Nature Physics*, vol. 15, no. 8, pp. 790–795, 2019.

[65] K. Shankar, S. K. Lakshmanaprabu, A. Khanna, S. Tanwar, J. J. Rodrigues, and N. R. Roy, "Alzheimer detection using group grey wolf optimization based features with convolutional classifier," *Computers Electrical Engineering*, vol. 77, pp. 230–243, 2019.

[66] K. R. Kruthika and H. D. Maheshappa, "Multistage classifier-based approach for Alzheimer's disease prediction and retrieval," *Informatics in Medicine Unlocked*, vol. 14, pp. 34–42, 2019.

[67] M. A. El-Yacoubi, S. Garcia-Salicetti, C. Kahindo, A. S. Rigaud, and V. Cristancho Lacroix, "From aging to early-stage Alzheimer's: uncovering handwriting multimodal behaviors by semi-supervised learning and sequential representation learning," *Pattern Recognition*, vol. 86, pp. 112–133, 2019.

[68] D. H. Ye, K. M. Pohl, and C. Davatzikos, "Semi-supervised pattern classification: application to structural MRI of Alzheimer's disease," in *Proceedings of the 2011 International Workshop on Pattern Recognition in NeuroImaging*, pp. 1–4, Seoul, Korea, May 2011.

[69] V. Sindhwani, P. Niyogi, M. Belkin, and S. Keerthi, "Linear manifold regularization for large scale semi-supervised learning," in *Proceedings of the 22nd ICML Workshop on Learning with Partially Classified Training Data*, vol. 28, Bonn, Germany, August 2005.

[70] J. M. Gorriz, J. Ramirez, J. Suckling et al., "A semi-supervised learning approach for model selection based on class-hypothesis testing," *Expert Systems with Applications*, vol. 90, pp. 40–49, 2017.

[71] Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan, and Y. Zheng, "Recent progress on generative adversusrial networks (GANs): a survey," *IEEE Access*, vol. 7, pp. 36322–36333, 2019.

[72] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, "A review on generative adversusrialnetworks: algorithms, theory, and applications," 2020, https://arxiv.org/abs/2001.06937.

[73] W. Yu, B. Lei, M. K. Ng, A. C. Cheung, Y. Shen, and S. Wang, "TensorizingGAN with high-order pooling for Alzheimer's disease assessment," 2020, https://arxiv.org/abs/2008.00748.

[74] H. Bhavsar and M. H. Panchal, "A review on support vector machine for data classification," *International Journal of Advanced Research in Computer Engineering Technology*, vol. 1, p. 10, 2012.

[75] F. Li, L. Tran, K. H. Thung, S. Ji, D. Shen, and J. Li, "Robust deep learning for improved classification of AD/MCI patients," in *Proceedings of the International Workshop on Machine Learning in Medical Imaging*, pp. 240–247, Boston, MA, USA, September 2014.

[76] K. R. Kruthika and H. D. Maheshappa, "CBIR system using capsule networks and 3D CNN for Alzheimer's disease diagnosis," *Informatics in Medicine Unlocked*, vol. 14, pp. 59–68, 2019.

[77] A. Khvostikov, K. Aderghal, J. Benois-Pineau, A. Krylov, and G. Catheline, "3DCNN-based classification using sMRI and MD-DTI images for Alzheimer disease studies," 2018, https://arxiv.org/abs/1801.05968.

[78] Y. D. Li, Z. B. Hao, and H. Lei, "Survey of convolutional neural network," *Journal of Computer Applications*, vol. 36, no. 9, pp. 2508–2515, 2016.

[79] M. Simonovsky and N. Komodakis, "Dynamic edge-conditioned filters in convolutional neural networks on graphs," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3693–3702, Honolulu, HI, USA, July 2017.

[80] D. Yu, H. Wang, P. Chen, and Z. Wei, "Mixed pooling for convolutional neuralnetworks," in *Proceedings of the International Conference on Rough Sets and Knowledge Technology*, pp. 364–375, Shanghai, China, October 2014.

[81] S. Sarraf and G. Tofighi, "DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI," *BioRxiv*, Article ID 070441, 2016.

[82] S. Spasov, L. Passamonti, A. Duggento, P. Lio, and N. Toschi, "A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease," *Neuroimage*, vol. 189, pp. 276–287, 2019.

[83] J. Dakka, P. Bashivan, M. Gheiratmand, I. Rish, S. Jha, and R. Greiner, "Learning neural markers of schizophrenia disorder using recurrent neural networks," 2017, https://arxiv.org/abs/1712.00512.

[84] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTMcells and network architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 2019.

[85] S. Karlekar, T. Niu, and M. Bansal, "Detecting linguistic characteristics of Alzheimer's dementia by interpreting neural models," 2018, https://arxiv.org/abs/1804.06440.

[86] S. Kloppel, C. M. Stonnington, C. Chu et al., "Automatic classification of MR scans in Alzheimer's disease," *Brain*, vol. 131, no. 3, pp. 681–689, 2008.

[87] C. Montagne, A. Kodewitz, V. Vigneron, V. Giraud, and S. Lelandais, "3D local binary pattern for PET image classification by SVM, Application to early Alzheimer disease diagnosis," in *Proceedings of the 6th International Conference on Bio-Inspired Systems and Signal Processing (BIOSIGNALS 2013)*, Barcelona, Spain, February 2013.

[88] N. Zeng, H. Qiu, Z. Wang, W. Liu, H. Zhang, and Y. Li, "A new switching delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 320, pp. 195–202, 2018.

[89] R. K. Lama, J. Gwak, J. S. Park, and S. W. Lee, "Diagnosis of Alzheimer's diseasebased on structural MRI images using a regularized extreme learning machine and PCA features," *Journal of Healthcare Engineering*, vol. 2017, Article ID 5485080, 2017.

[90] F. Ramzan, M. U. G. Khan, A. Rehmat et al., "A deep learning approach for automated diagnosis and multiclass classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks," *Journal of Medical Systems*, vol. 44, no. 2, pp. 1–16, 2020.

[91] M. Pak and S. Kim, "A review of deep learning in image recognition," in *Proceedings of the 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT)*, pp. 1–3, IEEE, Kuta Bali, Indonesia, August 2017.

[92] E. Altinkaya, K. Polat, and B. Barakli, "Detection of Alzheimer's disease and dementia states based on deep learning from MRI images: a comprehensive review," *Journal of the Institute of Electronics and Computer*, vol. 1, no. 1, pp. 39–53, 2020.

[93] S. Korolev, A. Safiullin, M. Belyaev, and Y. Dodonova, "Residual and plain convolutional neural networks for 3D

brain MRI classification," in *Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging*, pp. 835–838, Melbourne, Australia, April 2017.

[94] H. Li, M. Habes, D. A. Wolk, and Y. Fan, "A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data," *Alzheimer's Dementia*, vol. 15, no. 8, pp. 1059–1070, 2019.

[95] Z. Yang, I. M. Nasrallah, H. Shou et al., "Disentangling brain heterogeneity via semi-supervised deep-learning and MRI: dimensional representations of Alzheimer's disease," 2021, https://arxiv.org/abs/2102.12582.

[96] M. Dua, D. Makhija, P. Y. L. Manasa, and P. Mishra, "A CNN–RNN–LSTM based amalgamation for Alzheimer's disease detection," *Journal of Medical and Biological Engineering*, vol. 40, no. 5, pp. 688–706, 2020.

[97] C. Feng, A. Elazab, P. Yang et al., "Deep learning framework for Alzheimer's disease diagnosis via 3D-CNN and FSBi-LSTM," *IEEE Access*, vol. 7, pp. 63605–63618, 2019.

[98] A. Sarraf, M. Azhdari, and S. Sarraf, "A comprehensive review of deep learning architectures for computer vision applications," *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, vol. 77, no. 1, pp. 1–29, 2021.

[99] M. Amini, M. Pedram, A. Moradi, and M. Ouchani, "Diagnosis of Alzheimer's disease severity with fMRI images using robust multitask feature extraction method and convolutional neural network (CNN)," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 5514839, 15 pages, 2021.

[100] M. Amini, M. Pedram, A. Moradi, and M. Ouchani, "Diagnosis of Alzheimer's disease by time-dependent power spectrum descriptors and convolutional neural network using EEG signal," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 5511922, 17 pages, 2021.

[101] I. Ozsahin, B. Sekeroglu, and G. S. Mok, "The use of back propagation neural networks and 18F-florbetapir PET for early detection of Alzheimer's disease using Alzheimer's disease neuroimaging initiative database," *PLoS One*, vol. 14, no. 12, Article ID e0226577, 2019.

[102] J. Islam and Y. Zhang, "Understanding 3D CNN behavior for Alzheimer's disease diagnosis from brain PET scan," 2019, https://arxiv.org/abs/1912.04563.

[103] D. Vasan, M. Alazab, S. Wassan, H. Naeem, B. Safaei, and Q. Zheng, "IMCFN: image-based malware classification using fine-tuned convolutional neural network architecture," *Computer Networks*, vol. 171, Article ID 107138, 2020.

[104] D. Lu, K. Popuri, G. W. Ding, R. Balachandar, and M. F. Beg, "Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease," *Medical Image Analysis*, vol. 46, pp. 26–34, 2018.

[105] D. Gamberger, N. Lavrac, S. Srivatsa, R. E. Tanzi, and P. M. Doraiswamy, "Identification of clusters of rapid and slow decliners among subjects at risk for Alzheimer's disease," *Scientific Reports*, vol. 7, no. 1, pp. 1–12, 2017.

[106] E. Adeli, K. H. Thung, L. An et al., "Semi-supervised discriminative classification robust to sample-outliers and feature noises," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 2, pp. 515–522, 2018.

[107] M. Liu, D. Cheng, and W. Yan, "Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images," *Frontiers in Neuroinformatics*, vol. 12, p. 35, 2018.

[108] W. Li, Y. Wang, Y. Cai, C. Arnold, E. Zhao, and Y. Yuan, "Semi-supervised rare disease detection using generative adversusrial network," 2018, https://arxiv.org/abs/1812.00547.

[109] J. Dukart, K. Mueller, H. Barthel, A. Villringer, O. Sabri, and M. L. Schroeter, "Meta-analysis based SVM classification enables accurate detection of Alzheimer's disease across different clinical centers using FDG-PET and MRI," *Psychiatry Research: Neuroimaging*, vol. 212, no. 3, pp. 230–236, 2013.

[110] P. Triebkorn, L. Stefanovski, K. Dhindsa et al., "Multi-scale brain simulation with integrated positron emission tomography yields hidden local field potential activity that augments machine-learning classification of Alzheimer's disease," *BioRxiv*, 2021.

[111] J. Kim and B. Lee, "Identification of Alzheimer's disease and mild cognitive impairment using multimodal sparse hierarchical extreme learning machine," *Human Brain Mapping*, vol. 39, no. 9, pp. 3728–3741, 2018.

[112] S. Sarraf, D. D. Desouza, J. A. Anderson, and C. Saverino, "MCADNNet: recognizing stages of cognitive impairment through efficient convolutional fMRI and MRI neural network topology models," *IEEE Access*, vol. 7, pp. 155584–155600, 2019.

[113] Y. Ding, J. H. Sohn, M. G. Kawczynski et al., "A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain," *Radiology*, vol. 290, no. 2, pp. 456–464, 2019.

[114] C. D. Billones, O. J. L. D. Demetria, D. E. D. Hostallero, and P. C. Naval, "DemNet: a convolutional neural network for the detection of Alzheimer's disease and mild cognitive impairment," in *2016 IEEE Region 10 Conference*, pp. 3724–3727, Singapore, November 2016.

[115] M. Talo, O. Yildirim, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Convolutional neural networks for multi-class brain disease detection using MRI images," *Computerized Medical Imaging and Graphics*, vol. 78, Article ID 101673, 2019.

[116] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, and J. Tohka, "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects," *Neuroimage*, vol. 104, pp. 398–412, 2015.

[117] W. Lin, W. Lin, G. Chen, H. Zhang, Q. Gao, and Y. Huang, "Bidirectional mapping of brain MRI and PET with 3D reversible GAN for the diagnosis of Alzheimer's disease," *Frontiers in Neuroscience*, vol. 15, p. 357, 2021.

[118] X. Zhou, S. Qiu, P. S. Joshi et al., "Enhancing magnetic resonance imaging-driven Alzheimer's disease classification performance using generative adversusrial learning," *Alzheimer's Research & Therapy*, vol. 13, no. 1, pp. 1–11, 2021.

[119] H. B. Baydargil, J. S. Park, and D. Y. Kang, "Anomaly analysis of Alzheimer's disease in PET images using an unsupervised adversusrial deep learning model," *Applied Sciences*, vol. 11, no. 5, Article ID 2187, 2021.

[120] J. Venugopalan, L. Tong, H. R. Hassanzadeh, and M. D. Wang, "Multimodal deep learning models for early detection of Alzheimer's disease stage," *Scientific Reports*, vol. 11, no. 1, pp. 1–13, 2021.

[121] J. Zhang, B. Zheng, A. Gao, X. Feng, D. Liang, and X. Long, "A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification," *Magnetic Resonance Imaging*, vol. 78, pp. 119–126, 2021.

[122] A. Mehmood, S. Yang, Z. Feng et al., "A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images," *Neuroscience*, vol. 460, pp. 43–52, 2021.

[123] M. Raju, M. Thirupalani, S. Vidhyabharathi, and S. Thilagavathi, "Deep learning based multilevel classification of Alzheimer's disease using MRI scans," in *IOP Conference Series: Materials Science and Engineering*, vol. 1084, no. 1, Bristol, UK, IOP Publishing, Article ID 012017, 2021.

[124] M. Subramoniam, "Deep learning based prediction of Alzheimer's disease from magnetic resonance images," 2021, https://arxiv.org/abs/2101.04961.

[125] E. Lella, A. Pazienza, D. Lofù, R. Anglani, and F. Vitulano, "An ensemble learning approach based on diffusion tensor imaging measures for Alzheimer's disease classification," *Electronics*, vol. 10, no. 3, p. 249, 2021.

[126] U. R. Acharya, S. L. Fernandes, J. E. Wei Koh et al., "Automated detection of Alzheimer's disease using brain MRI images–a study with various feature extraction techniques," *Journal of Medical Systems*, vol. 43, no. 9, pp. 1–14, 2019.

[127] Y. Höller, K. H. Butz, A. C. Thomschewski et al., "Prediction of cognitive decline in temporal lobe epilepsy and mild cognitive impairment by EEG, MRI, and neuropsychology," *Computational Intelligence and Neuroscience*, vol. 2020, Article ID 8915961, 16 pages, 2020.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 5863496, 14 pages
https://doi.org/10.1155/2021/5863496

Research Article

Presentation of Novel Hybrid Algorithm for Detection and Classification of Breast Cancer Using Growth Region Method and Probabilistic Neural Network

Zeynab Nasr Isfahani,[1] Iman Jannat-Dastjerdi,[2] Fatemeh Eskandari,[1] Saeid Jafarzadeh Ghoushchi,[3] and Yaghoub Pourasad[4]

[1]*Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran*
[2]*Department of Radiology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran*
[3]*Department of Industrial Engineering, Urmia University of Technology, Urmia, Iran*
[4]*Department of Electrical Engineering, Urmia University of Technology, Urmia, Iran*

Correspondence should be addressed to Saeid Jafarzadeh Ghoushchi; s.jafarzadeh@uut.ac.ir

Received 9 April 2021; Accepted 10 June 2021; Published 21 June 2021

Academic Editor: Suresh Manic

Mammography is a significant screening test for early detection of breast cancer, which increases the patient's chances of complete recovery. In this paper, a clustering method is presented for the detection of breast cancer tumor locations and areas. To implement the clustering method, we used the growth region approach. This method detects similar pixels nearby. To find the best initial point for detection, it is essential to remove human interaction in clustering. Therefore, in this paper, the FCM-GA algorithm is used to find the best point for starting growth. Their results are compared with the manual selection method and Gaussian Mixture Model method for verification. The classification is performed to diagnose breast cancer type in two primary datasets of MIAS and BI-RADS using features of GLCM and probabilistic neural network (PNN). Results of clustering show that the presented FCM-GA method outperforms other methods. Moreover, the accuracy of the clustering method for FCM-GA is 94%, as the best approach used in this paper. Furthermore, the result shows that the PNN methods have high accuracy and sensitivity with the MIAS dataset.

1. Introduction

Breast cancer is a deadly and frequent illness that affects people all over the world. In the next 20 years, the number of new breast cancer patients is expected to increase by 75 percent. Consequently, according to the WHO in 2019, precise and early detection plays a critical role in developing the diagnostic and increasing the patients' survival rate with breast cancer from 20% to 60%. Tumors come in various forms that must be identified independently since each might lead to different treatment options and prognoses [1]. To aid oncologic decision-making, cancer categorization strives to give an accurate diagnosis of the illness and a prognosis of tumor activity. Traditional breast cancer categorization, which is mainly focused on clinicopathologic aspects and the use of routine biomarkers, may not represent the wide range of clinical outcomes experienced by individual breast cancers. The biology that underpins cancer genesis and progression is complex. Recent high-throughput technology results have added to our understanding of breast cancer's underlying genetic changes and biological processes [2].

Mammography is the most effective method for the early detection of breast cancer [3]. However, a look back reveals that many lesions that can be seen on a mammogram are overlooked by radiologists, which can have various causes, such as poor quality of the mammogram image, benign appearance of the lesions, and eye fatigue or neglect by

radiologists. Utilizing diagnostic approaches in the early stages of cancer development can be very effective and essential for patient treatment so that this early diagnosis can help doctors treat patients and significantly reduce the mortality of patients. Examination of breast tumors has a special place in the initial diagnosis of breast cancer [4]. Due to this, diagnosis with the eye can be prone to error, and the radiologist may not identify the tumor and cancer. Therefore, having an image processing system with the power to extract features that the human eye cannot detect or recognize with low accuracy can be very useful. A tumor is an abnormal mass of cells. Tumor cells grow for reasons that are still unknown, and they grow regardless of the body's needs [5]. Moreover, because nutrients absorb normal cells from the blood, they are often harmful to the body. Tumors are often called neoplasms or neoplasms. Body tissues are permanently repaired and replaced with new cells following damage or damage caused by natural cell depletion. Therefore, in general, growth and repair depend on the body's needs. Specific organs can grow in size (hypertrophy) or increase the number of cells (hyperplasia) if the organ is required to do more than its capacity [6].

A breast cancer diagnosis can help physicians to treat patients and significantly reduce mortality. Also, it increases the 5-year survival rate of patients with this cancer from 14% to 49% [7]. It is very important to check for breast cancer and to diagnose the tumor quickly and accurately. It is because eye diagnosis can be prone to error, and the radiologist may not detect the tumor. Therefore, there is an image processing system with high extraction power for detecting tumors, which can be very useful. The reduction in breast cancer mortality during screening may be even more significant. Since investigations have revealed that radiologists have failed to identify a remarkable number of breast cancer cases, the reasons for these cases are the failure of mammography screening. It is often unclear that disorders that are not visible in the images should be ignored. However, cancer may not be detected due to the absence of symptoms. Computer-aided diagnosis (CAD) is being developed. These methods utilized pattern recognition approaches to find features in the image that characterize breast cancer tumor location. Therefore, CAD systems assist the radiologist during the examination [8]. The image of suspicious areas is used. Most CAD systems also have diagnostic errors. However, there is also evidence that the CAD system can enhance the radiologist's ability to interpret detection lesions. Although the results of a small number of recent studies indicate that the performance of existing commercial CAD systems still needs to be developed, they can meet the needs of imaging centers and clinics. Therefore, improving the performance of CAD systems is a crucial issue for investigation, and future developments remain [9].

The convolutional neural network (CNN) is a multilayer system that recovers features from raw input. It is a symbol for hierarchical structuring [10]. Convolution layers, fully connected layers, pooling, and an output layer are among the layers that make up a deep neural network (DNN) [11]. A convolution layer is one of these layers that is beneficial for learning high-level characteristics such as the edges of an image. FC layers are used to learn pixel-by-pixel characteristics. A pooling layer can minimize the quantity of convolved features, lowering the amount of computing power required. Max pooling and average pooling are two operations that this layer may execute [12, 13]. There are two types of CNNs utilized for breast image or data classification: de novo trained models and transfer learning-based models. The term "de novo model" refers to CNN-based models created and trained from the ground up [14]. On the contrary, transfer learning networks are CNN models that use previously trained neural network models such as AlexNet, visual geometry group, and residual neural network [15, 16].

This study aims to cluster the breast cancer area using the region growth method in combination with the FCM-GA approach. These results are compared with the manual selection method and Gaussian Mixture Model method for verification. In the second part of the paper, the classification is performed to diagnose breast cancer type in two datasets of MIAS and BI-RADS using features of GLCM and probabilistic neural network (PNN).

2. Literature Review

Veena and Padma [17] preprocessed the input image by reducing the noise coefficient using the intermediate filter method, reducing the image noise. Then, they use Gaussian mixed model (GMM), one of the well-known clustering algorithms for image segmentation, and finally, by applying probabilistic neural network (PNN) classifier on the features extracted with coincident matrix algorithm. Gray area (GLCM) is classified into three categories of benign, malignant, and normal cases. Punitha et al. [18] used intelligent artificial bee colonization and Improved Monarch Butterfly Optimization Technique (IABC-EMBOT) to detect breast cancer. The method used is of good speed and accuracy. Classification accuracy is 97.53%, sensitivity is up to 96.75%, specificity is up to 97.04%, and the average processing time is 113.42. Sakri et al. [19] presented a feature selection method for predicting the recurrence of breast cancer. The selection method is the Particle Swarm Optimization (PSO-RM) feature which uses three different classifiers KNN, NB, and the fast decision tree. Among the 34 features, the proposed method chooses the best quadratic method and improves the accuracy of all three classifiers. KNN accuracy improved from 70% to 81%, NB from 76% to 80%, and the fast decision tree from 66% to 75%.

Karthik et al. [20] used a deep neural network to learn data characteristics (DNNs). They categorized breast cancer data using multiple-layer DNNs. Experimental results show that the accuracy obtained from this system is 97.66%, and the sensitivity is slightly less than 0.98. The deep network designed in this study is for breast cancer datasets only. Unni et al. [21] used a general thresholding method to estimate the basal chest muscle boundary and then applied morphological methods to correct the extracted area boundaries and the mean filter to eliminate noise. The GLCM algorithm is

used to extract the property. Then, a subset of these features that provides the best classification rate is selected using a genetic algorithm. Finally, Support Vector Machine Classification (SVM) is used to classify benign and malignant cancers.

Selvathi and Poornila [22] propose a general thresholding method for extracting chest bounds in images in which images are converted to binary using a fixed threshold value of 18. Each component with a significant number of pixels connected is considered as the chest area. The region boundary is then smoothed using a disk with a radius of 5 pixels using morphologically based filtering operations. Sasikala and Ezhilarasi [23] also proposed a general thresholding method for the extraction of breast boundaries. The 8-bit image noise used by the mean filter is reduced, and the image contrast is improved by the Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm. The images are then converted to binary images with a fixed threshold value followed by morphology-based filtering operations to eliminate small background objects. Results reported for this study included a maximum accuracy of 97.1%, a sensitivity of 98.8%, and a specificity of 95.4%. Heidari et al. [24] took mammographic image features and created an optimal classification model to estimate the risk of breast cancer. The data analyzed are 500 and divided into 50% high risk and 50% low risk. To anticipate the risk of cancer diagnosis, they proposed an LPP model based on the combination of several features to reduce the dimensions of the feature space. Unlike typical feature selection techniques that select a set of optimal features from the primary feature, LPP creates a new optimum feature array that includes features different from the main features in the feature pool, which ultimately created a 9.7% rise in risk anticipation accuracy.

Tariq et al. [25] conducted a study to classify mammographic images of breast cancer. GLCM algorithm was used to extract texture-type features from images. Then, a smaller set was made by the individual, in addition to the whole set of features. 60% of the data is used for training, 20% for validation, and 20% for testing. Using ANN as a classifier, the results of this study achieved 99% accuracy in the image recognition process. Kashyap et al. [26] used a partial differential equation adjustment process to extract the chest area, mammogram images, dark masking, and moderate filtering and map suspicious anomalies. Fuzzy c-mean clustering is applied. To calculate the texture characteristics of suspected fragmented masses, the local binary pattern was rotated, and local binary patterns were calculated. At the end of the support vector machine, polynomial kernel and radial basis function and multilayer and linear perceptron were used to classify areas of suspicion of abnormal and ordinary clades.

Chowdhary et al. [43] used the fuzzy c-mean bounded probability (IPFCM) method to overcome the bugs in traditional methods, namely, noise and random clustering, and used a fuzzy histogram algorithm in initial preprocessing in mammographic images. Finally, extraction, classification, and validation are performed to assist the radiologist in tumor diagnosis.

Rampun et al. [44] used a pretrained modified version of AlexNet with detailed adjustments in the database of CBIS-DDSM mammography images. The AlexNet network architecture used different parameters with more advanced functions such as PReLu instead of ReLu. The experimental results of this study are classification accuracy (ACC) of 80.4% and area under the curve (AUC) of 0.84. Kim et al. [45] evaluated the feasibility of using data-driven imaging biomarkers (DIB-MG) to characterize the deep CNN algorithm in mammographic images, including normal and benign classes. Using the algorithm, they have achieved a sensitivity of 76% and a specificity of 89%.

Zhang et al. [46] used a neural network algorithm to classify mammographic images. They evaluated ten different CNN architectures and concluded that combining both data addition methods (each original image into eight images) and a circular neural network would improve classification performance. The area under the ROC curve is 0.73. Salama et al. [47] developed a new computer-aided diagnosis (CAD) system to diagnose breast cancer in digital mammography. They used the WBCT algorithm to extract the feature. GA-SVM-MI optimization technique was used to select the optimal set of features. In this algorithm, the classification accuracy is between normal-abnormal (97.5%) and benign (96%) cases (see Table 1).

3. Methods and Materials

3.1. Data Collection. MIAS (Mammography Image Analysis Society), a UK research community engaged in understanding mammography, is a database for digital mammograms that have been developed. Films taken from the UK National Breast Screening Program were digitized with a Joyce-Loebl scanning microdensitometer to the 50-micron pixel tip, a linear unit in the optical density range 0–3.2, and an 8 bit term for each pixel. The list is split into film pairs, where each pair reflects a single patient's left (even filename numbers) and right (odd filename numbers) mammograms. The file resolution is 1024 pixels x 1024 pixels for ALL images. In the matrix, the images have been centered. Center positions and radii, rather than individual calcification, refer to clusters when calcification is present. The bottom-left corner is the root of the coordination system. In some cases, calcification is widely dispersed rather than concentrated at a single position throughout the image. Center positions and radii are unsuitable in these situations and removed [48].

The second dataset is a BI-RADS data collection intended to standardize reporting on breast imaging and minimize uncertainty in the understanding of breast imaging. It also promotes monitoring of outcomes and evaluation of quality. It provides a lexicon for systematic terminology and chapters on report organizations and guidance for use in everyday practice, for Mammograms, US breast, and MRI. A database of clinical mammograms comprising 60 patient photos was taken from mammogram screening centers. A broad range of cases that are difficult to classify by radiologists is included in the real-time database. Both clinical mammograms obtained from screening clinics were positive for abnormalities in their presence. Initially, as

Table 1: Literature review summary.

	Authors	Method	Objective	Results
1	Ribli et al. [27]	Faster R–CNN used mammography to identify tumors and showed that this method was quite time-efficient; however, the faster R–CNN is generally weak, meaning that the training set must contain a large set of ROIs yet complete enough to include all possible waste changes	Diagnosis and classification of masses in breast tissue	The system can detect 90% of malignant lesions in the INbreast dataset with only 0.3 false-positive marks per image
2	Gao [28]	They used a low-energy image (LE) similar to full-field digital mammography (FFDM) and a recombination image; in the proposed algorithm, the shallow CNN has the task of "image reconstruction," and the deep CNN has "extracting features"	Diagnosis and classification of masses in breast tissue	Experimental results on 89 FFDM datasets are obtained using recombinant "virtual" imaging features with an accuracy of 0.90 and AUC = 0.92
3	Jung et al. [29]	The model used a single-stage detector and a two-stage detector; one-step detectors such as RetinaNet generate a fixed number of projections over a network to cover possible positive sample spaces; unlike RetinaNet, the R–CNN mask can classify finite boxes in any range of scales and aspect ratios in any situation by segmenting the pixel surface	Diagnosis and classification of masses in breast tissue	The best result of this algorithm for TPR @ FPPI in the GURO dataset is (0.99 ± 0.01@ 3.0) (1.00 ± 0.00@1.3) (0.94 ± 0.03@0.5)
4	Shams et al. [30]	A deep generative multitasking network (DiaGRAM) has been used to solve data loss and limited training data to interpret the lesions because it is a costly and time-consuming task; this multifunctional network is built on a combination of Concentration neural networks and a generational opposite network	Classification of masses in breast tissue	The results of this algorithm for the INbreast dataset are as follows: Accuracy equal to $9.2\% \pm 5/93$ And AUC = $92.5 \pm 2.4\%$
5	Al-masni et al. [31]	A CAD system based on deep learning ROI (area of interest) techniques using CNN called YOLO was developed to diagnose and classify breast masses into benign and malignant states in DDSM mammographic images	Diagnosis and classification of masses in breast tissue	The results of this algorithm are as follows Sensitivity = 93.2% Feature = 78% AUC = 87.7%
6	Chougrad et al. [32]	CAD describes a system based on deep CNN that distinguishes malignant and benign breast mass in high-resolution mammographic images; the models used are VGG16, ResNet50, and Inceptionv3	Classification of masses in breast tissue	With the ResNet50 network in the DDSM dataset, they achieved an accuracy of 27.27%
7	Ragab et al. [33]	The region-based method was used to determine the threshold of 76 and determine the most significant area; at the feature extraction stage, DCNN was used; AlexNet network retraining was used to distinguish between the two classes; for better accuracy, the last layer of DCNN was replaced with SVM	Diagnosis of masses in breast tissue	Accuracy, AUC, sensitivity, specificity, accuracy, and F1 score were 80.5%, 88.8%, 88.4%, 84.2%, 86%, and 81.5%, respectively
8	Hazarika and Mahanta [34]	The method of extraction of breast border area using threshold-based zoning with morphological operations was proposed; for this purpose, the two-stage image contrast enhancement method was used; in the first phase, a two-step histogram correction technique is used to improve the image at the general level, and in the second phase, a nonlinear filter based on local mean and the local standard deviation for each pixel is applied to the image with the modified histogram	Diagnosis of cancer in breast tissue	The result of the proposed algorithm with 322 images is 98% accuracy

Table 1: Continued.

	Authors	Method	Objective	Results
9	Rajendra et al. [35]	Four texture extraction algorithms were used to extract features from mammographic images, and the SVM classification method was used to classify mammographic images into normal and abnormal categories	Classification of masses in breast tissue	The best results are related to the GLCM feature Accuracy = 92%, sensitivity = 94%, specificity = 93%, accuracy = 95%
10	Eltoukhy et al. [36]	The CAD system, based on feature extraction, was applied using the Gauss-Hermitage method, and the features were classified into four different categories: K-NN, random forest, and AdaBoost	Diagnosis of masses in breast tissue	The accuracy of the method used on the two sets of IRMA and MIAS images is 93.27 and 90.56, respectively
11	Padmavathy et al. [37]	A practical method was used by NSST + ANFIS to diagnose breast cancer; NSST was used to parse original images in multiple directions and several scales, and ANFIS-compliant clustering was used to classify input images	Diagnosis of masses in breast tissue	The results of the proposed algorithm Accuracy = 98.2%, sensitivity = 90.4%, specificity = 90.6%
12	Tahmooresi et al. [38]	A method for the early diagnosis of breast cancer has been proposed. In which it combines different machine learning methods, Support Vector Machine (SVM), ANN, KNN, Decision Tree (DT)	Classification of masses in breast tissue	Accuracy of 99.8% was achieved for the diagnosis of breast cancer
13	Amrane et al. [39]	They studied different machine learning methods such as support vector machine, Naive Bayes classifier, and nearest neighbor (KNN) to classify images of breast cancer and claimed that the KNN classification method performed better than vector machine; gain support and Naïve Bayes	Classification of masses in breast tissue	The KNN method has a higher accuracy of 97.57. However, the NB method also has a good accuracy of 96.99%
14	Anjaiah et al. [35]	Multi-ROI segmentation is one of the ways these authors have used mammography images; also, images were extracted using statistical criteria to measure the texture characteristics of mammographic images	Classification of masses in breast tissue	This method helps better detect the texture and shape of suspicious mammography images and better diagnose breast cancer
15	Vijayarajeswari et al. [40]	They efficiently categorized normal and abnormal classes into mammographic images using Huff transforms; improve results using other features such as mean, variance, and entropy; finally, the SVM cluster was used for classification	Classification of masses in breast tissue	The diagnosis accuracy in standard images is 65%, and in nonnormal images, 71%
16	Chowdhary et al. [41]	The goal of this study is to use intuitionistic possibilistic fuzzy c-mean clustering to segment medical photos	Segmentation of breast cancer	For MIAS pictures with varying noise levels of 5%, 7%, and 9% of the presented method, the average segmentation accuracy is 91.25 percent, 87.50 percent, and 85.30 percent
17	Chowdhary et al. [42]	Deep convolution neural networks are used to classify breast cancer using computer vision and image processing	Classification of breast cancer	For benign and malignant pictures, the traditional computer vision and image processing paradigm has a classification accuracy of 85 percent and 82 percent, respectively

an input, we take a 2D mammogram image of size M×N and add the average filter to it. The pictures are 20 benign, 20 malignant, and 20 regular images of the breast tumor.

3.2. Growth Region Algorithm. This approach focuses on the definition employed to divide the image into distinctive regions to analyze the homogeneity and centered on the resemblance or the homogeneity of the neighboring pixels. The pixels in each region are comparable to particular parameters, such as color and strength. These histogram-based image clustering approaches concentrate only on the dispersion of image pixels at the grayscale, whereas district

growth techniques note that near-gray levels are often present in the surrounding pixels.

Area-based methods are performed as follows:

(1) The beginning of the algorithm is regarded to be the number of initial seeds.

(2) The region starts to develop from these seeds. The pixels are inserted in this area similar to the initial pixels.

(3) The next grain is considered when the region's growth ends, and the following area's growth continues.

(4) These steps keep going until one region belongs to all the pixels in the image.

The following measures refer to the growth method of the region (Figure 1).

Step 1: *select initial seeds*. It must manually insert the initial points to start the algorithm. In the handy process, the algorithm begins by choosing the initial points for the user. Several techniques automatically take out the initial points in the sector: the utilization of a random step algorithm to identify the first points, for instance. To pick the initial points, this study proposes an algorithm that focuses on FCM-GA methods. The purpose of the algorithm is to use the fuzzy clustering algorithm to implement clustering initially. Moreover, it is distinguished by membership grade M and cluster centers C. Then, through the genetic algorithm, the appropriate values of these parameters are gained based on reducing target function. In the following approach, the error performance criterion E is

$$E(C) = \sum_{j=1}^{n} \sum_{i=1}^{K} m_{ij} x_j - c_{iA}^2, \tag{1}$$

where the m relation is as follows:

$$\begin{aligned} 0 < \sum_{j=1}^{n} m_{ij} < n, \quad & i = 1, 2, \ldots, K, \\ \sum_{i=1}^{K} m_{ij} = 1, \quad & j = 1, 2, \ldots, n, \\ \sum_{i=1}^{K} \sum_{j=1}^{n} m_{ij} = n. & \end{aligned} \tag{2}$$

Furthermore, from the following relationship, m and C are computed:

$$\begin{aligned} m_{ij} &= \left[\sum_{k=1}^{K} \left(\frac{x_j - c_{iA}}{x_k - c_i} \right) \sum \right]^{-1}, \quad 1 \le i \le K, \ 1 \le j \le n, \\ c_i &= \frac{\sum_{j=1}^{n} m_{ij} x_j}{\sum_{j=1}^{n} m_{ij}}, \quad 1 \le i \le K. \end{aligned} \tag{3}$$

Step 2: *determine the similarity of regions*. The selection of the similarity criterion between the regions is achieved. Then, the initial points have been defined in the previous stage. The criterion of similarity is utilized to evaluate the resemblance of the new pixels. The pixels of the regions define the assignment of the new pixel to the corresponding area.

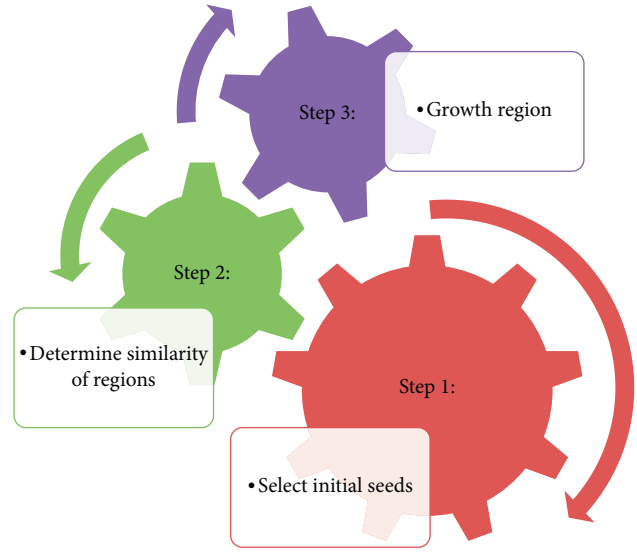


Figure 1: Conceptual diagram of presented growth region algorithm.

The standard deviation criterion is one of these criteria for similarity. The new pixel $I_{n+1}(x, y)$ relates to the region when it is used under the following condition and is used in the mean and standard deviation areas:

$$\mu_n - X\sigma_n < I_{n+1}(x + y) < \mu_n + X\sigma_n, \tag{4}$$

where X is a variable that describes the difference in the area by how many pixels, usually, 99.7% of the actual seeds in the field, are placed in an isolated area where X is considered to be 3. The smaller X is, the fewer points the region contains, and the picture splits into more regions. The threshold criterion is another deemed criterion. In this process, the average of the regions is determined, and the new pixel $I_{n+1}(x, y)$ relates to the area as the following condition:

$$\mu_n - X < I_{n+1}(x + y) < \mu_n + X. \tag{5}$$

In this context, when the gap is less than the defined limit between the pixels and the average area, the area covers that pixel. For color images (red, green, and blue), this situation must also extend to all three layers to connect these pixels to this region.

Step 3: *the third step is growth region*. The region's growth is carried out after choosing the initial seeds for beginning the algorithm and the criterion of similarity for pixels with zones. The area's development is such that the adjacent regions are chosen, beginning from the initial seeds.

3.3. Fuzzy C-Means (FCM). FCM Bezdek et al. (1984) developed a classification algorithm regarding the reductions of the objective function:

$$J_q = \sum_{i=1}^{n} \sum_{j=1}^{m} u_{ij}^q d\left(x_i, \theta_j\right), \tag{6}$$

where q manages the nonresolution degree and the classification spike.

U seems to be the inextensible membership in the center class of the x_i data. The interval between the x_i data and the middle of class j and d should be the interval.

U is subject to the following situations:

$$u_{ij} \in [0,1], \sum_{j=1}^{n} u_{ij} = 1 \& 0 < \sum_{j=1}^{n} u_{ij} < n. \tag{7}$$

For each group of the following relationships, the membership function and center membership are gained:

$$u_{ij} = \frac{1}{\sum^{m} \left(d\left(x_i, \theta_j\right)/d\left(x_i, \theta\right)_i\right)^{2/(q-1)}},$$
$$\theta_j = \frac{\sum_{k=1}^{N} u_{ij}^{q} x_i}{\sum_{k=1}^{N} u_{ij}^{q}}, \tag{8}$$

3.4. Genetic Algorithm (GA). The genetic algorithm (GA) is an optimization tool based on the Darwinian evolutionary rule. In each step of applying the GA, a set of search points is applied to arbitrary processing. Every point has appointed a sequence of characters in this method, and genetic operators carry out these sequences. To obtain new results in the search space, the resulting sequences are then diced. In the end, the likelihood of their presence in the next level is calculated based on the premise that each point has an objective function.

The fitness function was described in this research as follows. This function is determined based on the difference between the degraded database image and the image acquired by the growth method of the field, beginning at the initial random point, as follows:

$$\text{cost} = \left(G - I_1\right)^2 + \left(W - I_2\right)^2 + \left(B - I_3\right)^2, \tag{9}$$

where gray, white, and black layers are I_1 and I_3 are images derived from fragmentation, respectively.

3.5. Adaptive Median Filter. The adaptive median filter classifies each pixel of the image as noise with its surrounding pixels. The size of the neighborhood, as well as the reference threshold, are adjustable. As impulse noise, a pixel that varies from most of its neighbors and is not functionally compatible with those pixels to which it is identical is labeled. The objects are (1) deleteing noise from impulse, (2) smoothing other noises, and (3) diminishing distortion, such as extreme thinning or object boundary thickening [49].

3.6. Gaussian Mixture Model. As a random parameter, the pixel's value in an image (i.e., the intensity or color) can be taken. Since each random variable has a probability distribution, then pixel values have a distribution of probability [50]. A reasonable probability distribution for pixel values of an image is the Gaussian mixture distribution. The form is the Gaussian mixture distribution formula:

$$f\left(x_s\right) = \sum_{i=1}^{k} P_i N\left(x_s | \mu_i, \sigma_i\right), \tag{10}$$

where

$$N\left(x_s | \mu_i, \sigma_i\right) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_i^2}\left(x_s - \mu_i\right)^2\right]. \tag{11}$$

We presume that an image is categorized into a $C_i = 1, 2, \ldots, k$ class known to the class k. μ_i, σ_i^2, and p_i parameters are of C_i-class in terms of mean, variance, and likelihood.

It means the following:

$$p_i = p\left(x_s \in C_i\right), \quad 0 < p_i < 1,$$
$$\sum_{i=1}^{k} p_i = 1. \tag{12}$$

3.7. Expectation-Maximization Algorithm. In setup to the K-means, the EM algorithm is very identical. Similarly, choosing the input partitions is the first part. In this case, to compare findings more realistic, the same initial partitions as utilized in the color segmentation with K-Means were used. Here is the comparison parameter, RGB color was again chosen. The EM cycle starts with the following equation described by an Expectation phase [51]:

$$E\left[z_{ij}\right] = \frac{p\left(x_i | \mu_i\right)}{\sum_{n=1}^{k} p\left(x_i | \mu_i\right)} = \frac{\exp\left[-1/2\sigma_i^2\left(x_s - \mu_i\right)^2\right]}{\sum_{n=1}^{k} \exp\left[-1/2\sigma_i^2\left(x_s - \mu_i\right)^2\right]}. \tag{13}$$

This equation asserts that concerning partition j, the assumptions or weight for pixel z equals the possibility that x is pixel x_i provided that μ is partition μ_i divided by the sum over all components k of the same probability subsequently defined. For weights, this contributes to the lower expression. The sigma expresses the covariance of pixel data squared shown in the second expression. The M step or maximization step starts once the E step has been implemented, and every pixel has a weight of expectation for each partition. The following equation defines this step:

$$\mu_j \leftarrow \frac{1}{m} \sum_{i=1}^{m} E\left[z_{ij}\right] x_i. \tag{14}$$

This equation indicates that the partition value j is modified to the weighted average of the pixel values for this specific partition, where the weights are the weights of the E phase. For each new set of partitions, this EM loop is replicated until, as in the K-Means algorithm, the partition values do not shift by a significant amount anymore.

3.8. Hidden Markov Random Field Model (HMRFM). HMRFM is defined by a sequence of observations as accidental processes generated by a sequence of Markov whose state sequence cannot be explicitly tracked. Each observation of the state series is presumed to be a stochastic function.

According to the $l \times l$ transition probability matrix, in which l represents the number of states, the underlying Markov chain alters its state. HMMs have been effectively used in speech recognition and handwritten script recognition [52]. Since the original HMMs were created as 1D Markov chains, 2D/3D problems like image segmentation cannot be used directly with first-order neighborhood structures. Here, we consider an HMM case that contains a Markov random field rather than a Markov chain and thus is not confined to 1D as the underlying stochastic process. We call this particular case an HMRF model. The following characterizes an HMRF model mathematically:

(i) *Hidden Random Field (MRF).* Let ℓ be a finite status space with the distribution of probability; the arbitrary vector $x = \{x_i\, i \in s\}$ is an underlying MRF (5). The condition of X is not observable.

(ii) *Objective Random Field.* $y = \{y_i, i \in s\}$ is a random field with space D as a finite state. Each Y_i is the conditional probability distribution $p\ (y_i \mid x_i)$ for function $f\,(y_i; \theta_{xi})$ given any unique configuration$x \in X$, where the parameters associated areθ_{xi}. The emission likelihood function is termed this distribution, and Y is often pointed to as the arbitrary vector released.

(iii) *Conditional Dependency.* For any $x \in X$, the arbitrary parameters Y_i are conditional independent:

$$P\,(Y - X) = \prod_{i \in s} p\,(y_i | x_i). \tag{15}$$

3.9. Probabilistic Neural Network (PNN). The PNN is a network for neural feed-forward, widely employed in diagnosis and pattern recognition algorithm. According to the discriminant analysis of the Kernel Fisher, this form of ANN was extracted. The operations are categorized into four layers of a multilayered feed-forward network within a PNN as input, pattern, summation, and output layers (see Figure 2).

In classification problems, PNN is also utilized [53]. The first layer measures the difference when input is available, from the input vector to the training input vectors. The pattern layer calculates the relationship between each type of inputs and produces the net output as a likelihood vector. Eventually, a competitive transfer feature on the output of the second layer chooses the sum of those possibilities. For that class and nontarget classes, binary identification is created, respectively. Each neuron in the input layer defines a predictor variable. In categorical variables, when there are N category numbers, N-1 neurons are considered. The pattern layer includes a neuron for every case in the training process. It keeps the values of the input variables for the scenario, along with the output value. In the summation layer, the values for the class they represent are added by the pattern neurons. In the output layer, for each target category, it contrasts the weighted collected in the pattern layer and uses the maximum values to predict the targets.

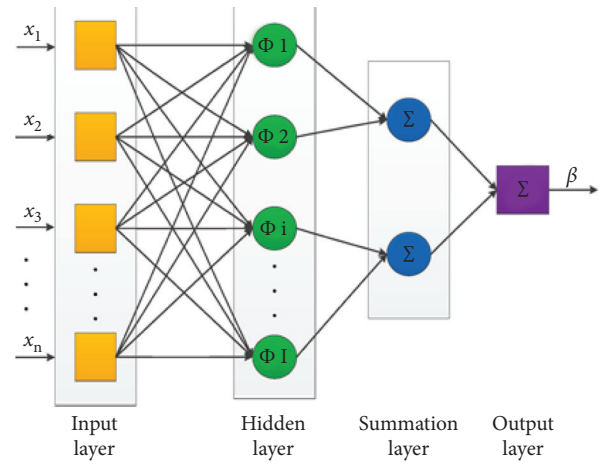


Figure 2: The diagram of the PNN method.

3.10. ROC Curve. To determine the outcome of binary classification (duality), sensitivity and specificity are used in the statistics of both measures. The consistency of the outcomes of a test that separates the information into these two categories is observable and descriptive using sensitivity and attribute metrics when the data can be separated into positive and negative classes. Sensitivity means the percentage of positive cases that would be correctly checked as positive. Specificity means the percentage of negative cases that correctly label them as negative. The sensitivity is the division of true-positive cases in statistical terms into the sum of true-positive cases and false-negative cases:

$$\begin{aligned} \text{Sensitivity or recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{Specificity} &= \frac{\text{TN}}{\text{TN} + \text{FP}}. \end{aligned} \tag{16}$$

The sensitivity of the test and its specificity depend on the quality of the test and the type of test utilized. However, it is not possible to describe the outcome of a test using sensitivity and specificity alone:

$$\begin{aligned} \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \\ \text{Fall} - \text{out} &= \frac{\text{FP}}{\text{FP} + \text{FN}}. \end{aligned} \tag{17}$$

The ROC curve is a plot displaying the diagnostic capabilities of the binary classifier system as its discrimination threshold is different. The ROC curve is formed by plotting the true-positive rate against fall-out or the false-positive rate at different threshold settings. Sensitivity, recall, or likelihood of identification are also recognized as the true-positive rate. Accuracy, specificity, and precision are other performance analysis parameters.

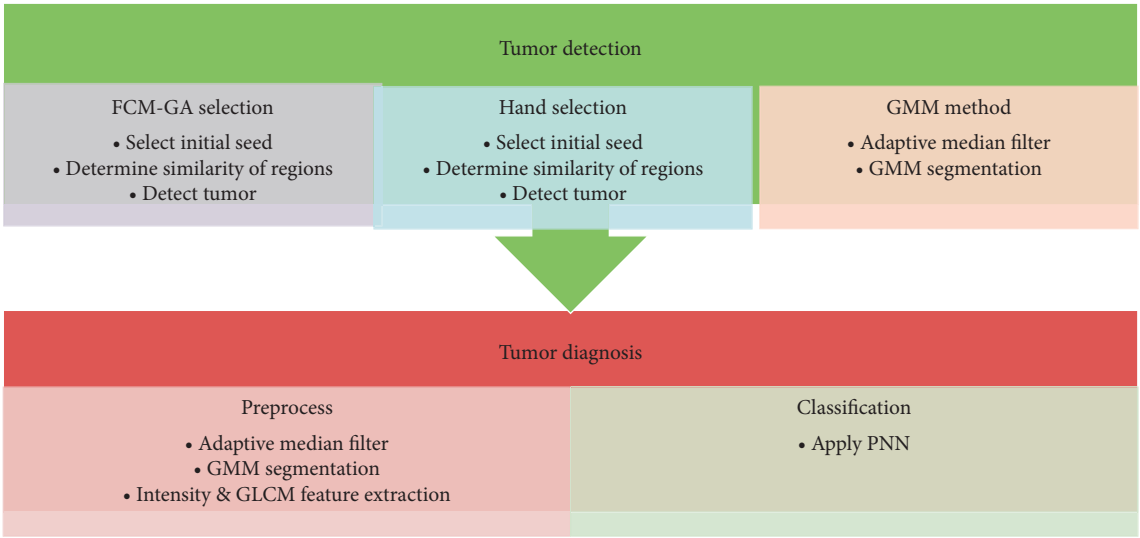


Figure 3: The diagram of the proposed model.

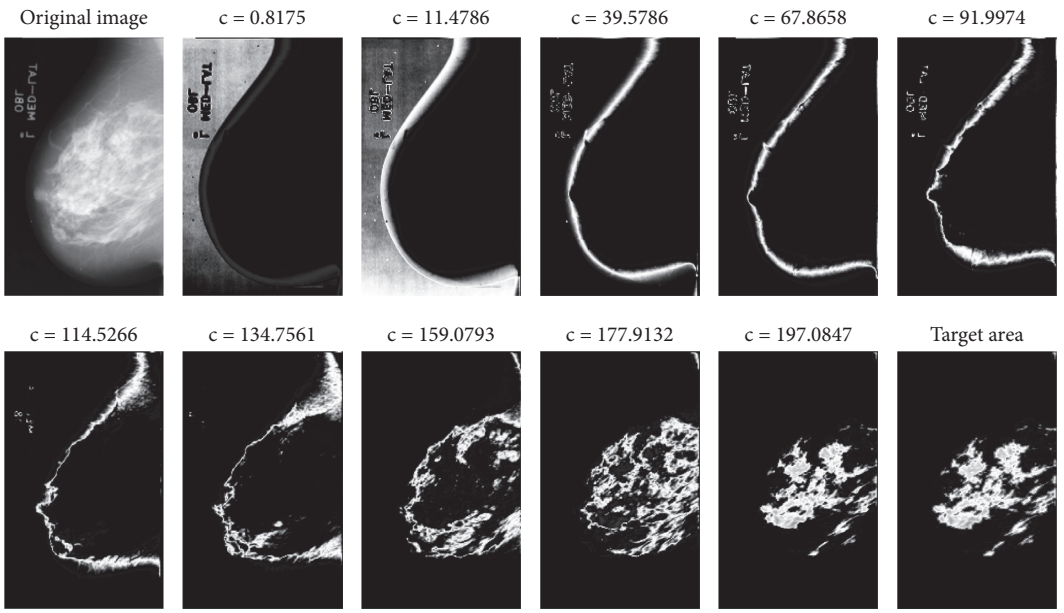


Figure 4: The results of the method provided for breast cancer detection.

4. Results and Discussion

4.1. Presented Approach. The diagram of the approach presented is shown in Figure 3. We used the growth region technique for tumor detection and adaptive median filters for classification to eliminate noise from the image since it is best for all spatial filtering and recognizes noise from fine information. The Adaptive Median Filter implies space processing to evaluate the impulse noise pixels in an image. By comparing each pixel with its surrounding pixels in the image, the adaptive media filters classify pixels as noise. The size of the neighborhood, as well as the comparison threshold, is adjustable.

4.2. The Tumor Detection Stage

4.2.1. Growth Region Method. In the growth region algorithm, the average value of the initial seeds is the average of the tumor area in this work, and the initial STD is equal by zero for region growth beginning from the initial seeds in the input image. Neighborhoods are referred to for their initial point 8. The analysis is done so that the pixels around it are checked, starting with the initial seed, and if they relate to that class, they are attached to it.

By accumulating points to that class as follows, the average and STD of the old class are recursively updated by using the mean and STD of the previous stage:

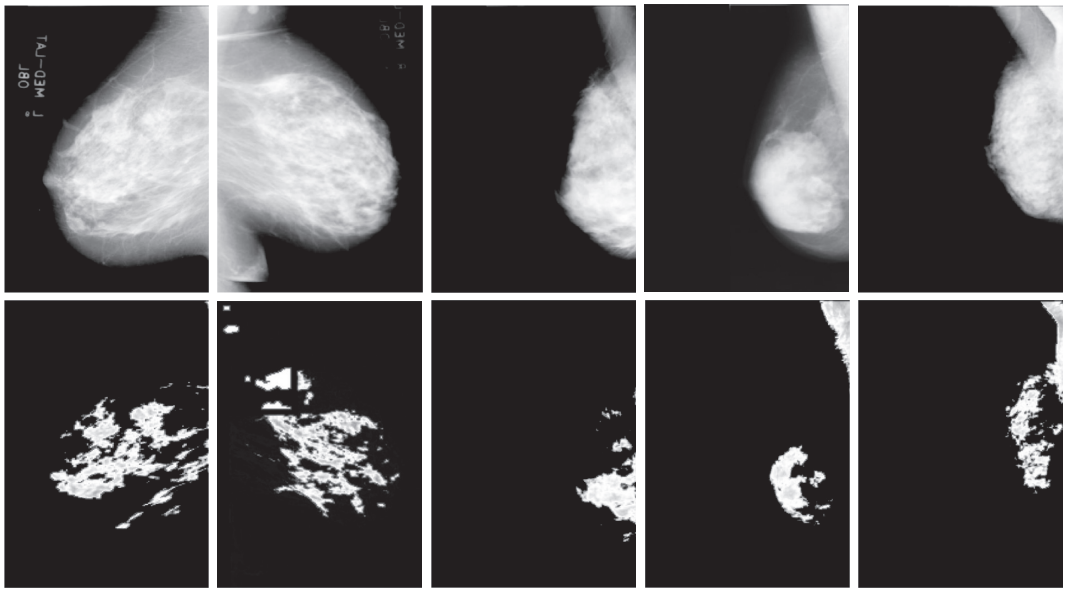


Figure 5: Breast cancer detection. Top: input image; bottom: output image.

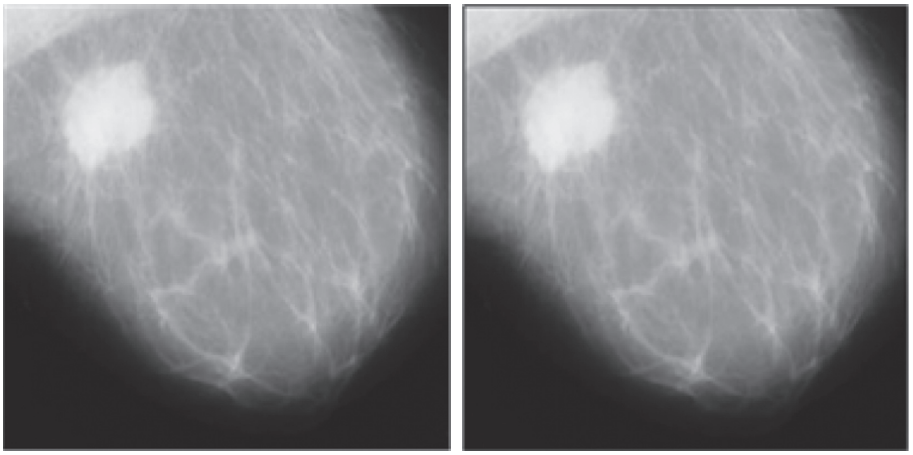


Figure 6: Results of the adaptive median filtration.

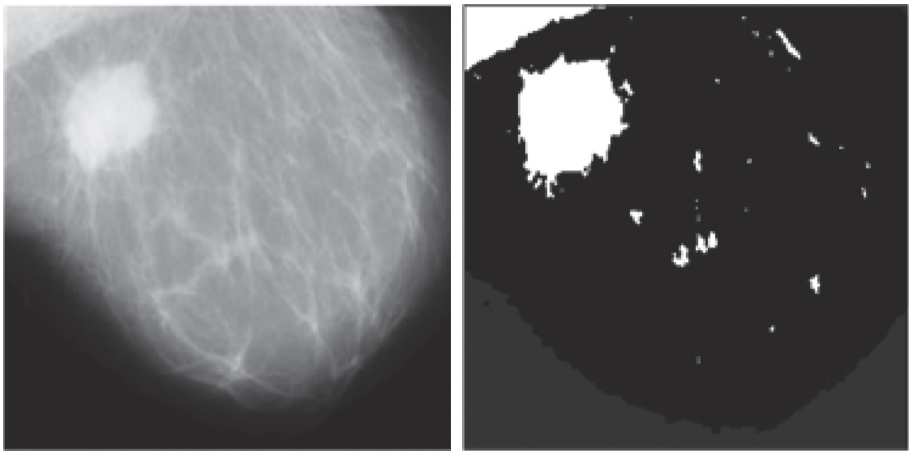


Figure 7: Results of the GMM method after filtration for a malignant tumor.

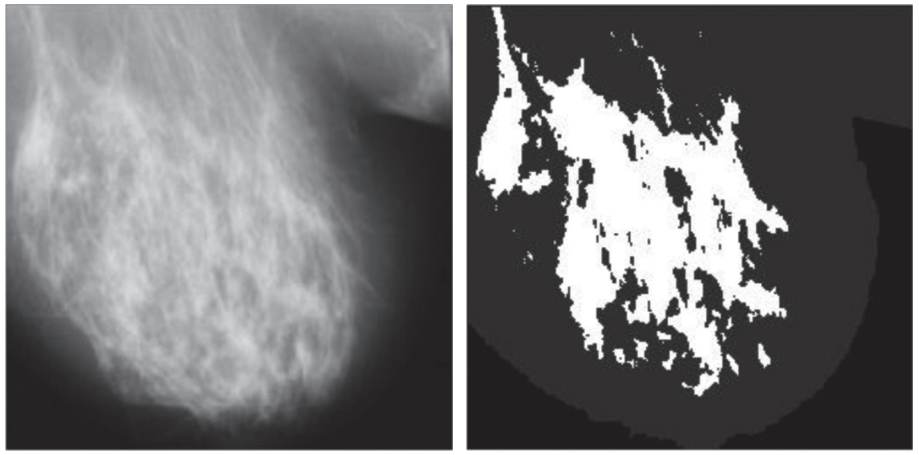


Figure 8: Results of the presented approach for a benign tumor.

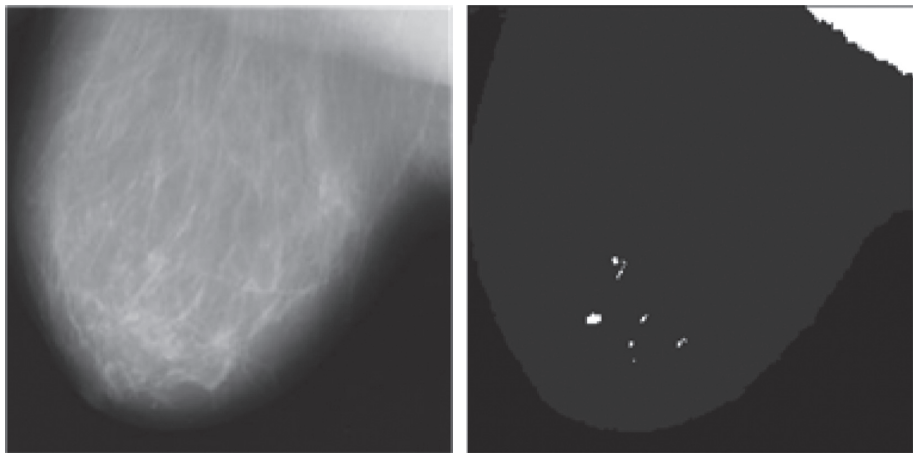


Figure 9: Results of the presented approach for a normal breast.

$$\mu_N = \frac{(N-1)(\mu_{N-1}) + I_N}{N},$$
$$\sigma_N = \sqrt{\frac{(N-2)(\sigma_{N-1})^2 + N/N - 1(I_N - \mu_N)^2}{N-1}}. \quad (18)$$

Then, they will also be located in the same neighborhood, and the criteria for the neighboring point applied to this class will be changed. This search will continue until ultimately detecting the first class and not add any other point. The cluster size of the FCM is ten, and the population size of GA is 100. Furthermore, 0.2 is the rate of mutations. Figure 4 shows the clustering findings using the FCM-GA approach mentioned. The precise location of the target tumor area was detected in the final image (see Figure 5).

4.2.2. Segmentation Using GMM. The outcomes of Adaptive Median Filtration are also shown in Figure 6. The image is clarified, and the noise is reduced. These noise pixels would then substitute the value of the median pixel in the adjacency. Initially, the image is transformed into a grayscale image, adaptive median filtering is then applied to the output image, and the image is transformed to an integer eight that is unsigned. Then, with two regions, two GMM components, and 100 iterations, the GMM clustering is performed on the preprocessed image.

We used the k-means clustering method ($k = 2$) and applied HMRF-EM. Figures 7 and 8 display the effects of the GMM process. Malignant and benign tumors are illustrated in Figures 7 and 8, respectively, and we can segment the tumors using the presented procedure.

Figure 9 demonstrates the effects of the approach provided for normal breasts. The results do not reveal the critical section of the tumor in the model. Also, the chest muscle is present in the white position in the top corner of the figure.

4.2.3. The Performances of Approaches. Table 2 shows the fuzzy fitness value with three presented techniques in the genetic algorithm and is compared by the maximum values of Jaccard and the minimum Jaccard distance. In the genetic algorithm, the FCM based fitness function works well due to similarities between the techniques and the smallest Jaccard distance.

Table 2: Comparison of the approach presented and selective seed methods from the approach to the growth area.

Method	Jaccard	Similarity
Growth region hand selection	0.82	0.63
Growth region FCM-GA selection	0.95	0.71
GMM method	0.93	0.65

Table 3: Comparison of the proposed algorithm with handy selection FCM-GA and GMM.

Method	RMS
Growth region hand selection	0.5711
Growth region FCM-GA selection	0.3681
GMM method	0.4216

Table 4: The results of performance analysis.

	Sensitivity	Specificity	Precision	Fall-out
Hand selection	0.72	0.65	0.76	0.35
FCM-GA selection	0.83	0.79	0.94	0.21
GMM method	0.76	0.63	0.92	0.37

The clustering methods are presented based on the growth region method. Then, the analysis was used in this work on a variety of performance criteria. The suggested approach can diminish the RMS error, as shown from the results (Table 3). It indicates that the division of the picture was performed more precisely. In this analysis, FCM-GA was employed to separate the initial points using the area growth process. As mentioned, the appropriate initial points have been selected using a genetic algorithm and evaluating the proper fitness function for image clustering using fuzzy logic. With this hybrid method, we get the required initial seeds for starting the growth process.

Consequently, the presented method was implemented on mammography MRI breast cancer images, and the outcomes were recorded. The finding revealed that the suggested algorithm could reduce the clustering error. We utilized 212 healthy breasts and 110 breast cancer images for the implementation of the algorithm. The performance criteria are shown in Table 4. The findings indicate that for presented techniques, maximum sensitivity is shown. Besides, the GMM approach offers minimal fall-out and optimum sensitivity with better detection performance using the GA to collect initial seeds.

We demonstrate that this method is appropriate according to the receiver operating characteristic (ROC) curve. Since it has a sensitivity above the line of guess, the optimal outcome should have minimal fall-out and optimum sensitivity concerning Figure 10.

4.3. The Classification Stages. We used two BI-RADS and MIAS datasets in the classification stage. We utilized 60 breast cancer images in BI-RADS, including 100 malignant and 100 benign breast cancer images. Table 5 displays the sensitivity, fall-out, precision, accuracy, and specificity parameters of classification using the PNN method. The result

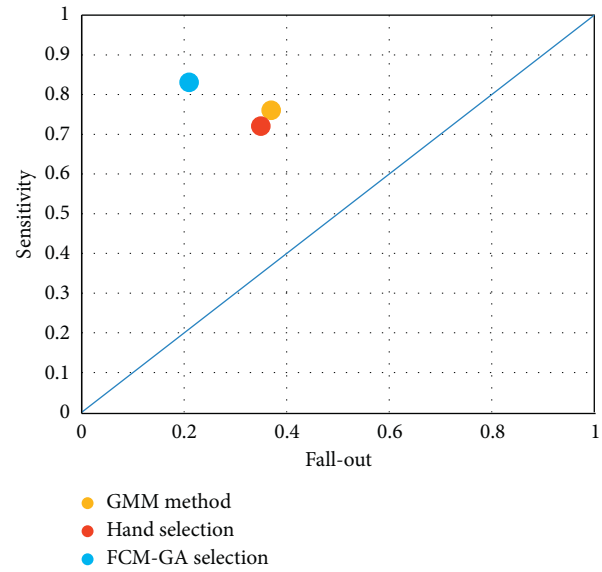


Figure 10: The ROC curve of the presented method.

Table 5: The results of detection parameters.

	Sensitivity	Specificity	Precision	Accuracy	Fall-out
MIAS	0.91	0.92	0.99	0.92	0.08
BI-RADS	0.88	0.72	0.95	0.80	0.28

shows that the PNN methods have high accuracy and sensitivity with the MIAS dataset. Also, the fall-out for the MIAS dataset is 0.08, which is lower than BI-RADS as 0.8, so the number of datasets will trigger classification outcomes.

5. Conclusion

In this study, an automated clustering method is proposed to detect breast cancer from mammographic MRI images based on the growth region method. Also, FCM-GA was utilized to find initial seeds. MRI images are proposed for target area detection and extraction utilizing genetic algorithm-based dynamic image analysis. Comparison results of RMSE error reveal that the suggested algorithm has the lowest error rates relative to the other method. Moreover, the suggested FCM-GA method has a higher Jaccard index and the minimum Jaccard distance than other methods.

Moreover, the suggested GMM solution is close to the method of FCM-GA. As shown in the findings, the RMS error could be minimized by the suggested approach. For the automatic selection of the appropriate initial seeds, a genetic algorithm has been used based on fuzzy logic clustering techniques to start the growth process. Also, we used the adaptive median filter for classification to eliminate noise from the picture of breast cancer, to distinguish between fine details and noise. We then performed the GMM clustering on the preprocessed image. Moreover, the PNN method is used to classify or diagnose tumor types based on GLCM features extracted from GMM resulted in images. We utilized two datasets MIAS and BI-RADS for this target. With minimal fall-out and optimum sensitivity, we can see perfect outcomes.

Data Availability

The mammography data of this study are available in the MIAS and BI-RADS repositories (the Mammographic Image Analysis Society Digital Mammogram Database, ACR BI-RADS® Mammography).

Conflicts of Interest

The authors declare that they have conflicts of interest.

References

[1] S. Bharati, P. Podder, and M. Mondal, "Artificial neural network based breast cancer screening: a comprehensive review," 2020, http://arxiv.org/abs/2006.01767.

[2] J. Tsang and G. M. Tse, "Molecular classification of breast cancer," *Advances in Anatomic Pathology*, vol. 27, no. 1, pp. 27–35, 2020.

[3] R. T. Chlebowski, G. L. Anderson, A. K. Aragaki et al., "Association of menopausal hormone therapy with breast cancer incidence and mortality during long-term follow-up of the women's health initiative randomized clinical trials," *Jama*, vol. 324, no. 4, pp. 369–380, 2020.

[4] F. M. Sutter and A. Ye, "Radiation therapy in the management of breast cancer and the impact of BC cancer center for the north on patient choice of treatment," *British Columbia Medical Journal*, vol. 60, 2018.

[5] S. Dorosti, S. J. Ghoushchi, E. Sobhrakhshankhah, M. Ahmadi, and A. Sharifi, "Application of gene expression programming and sensitivity analyses in analyzing effective parameters in gastric cancer tumor size and location," *Soft Computing*, vol. 24, no. 13, pp. 9943–9964, 2020.

[6] C. I. Ullrich, R. Aloni, M. E. Saeed, W. Ullrich, and T. Efferth, "Comparison between tumors in plants and human beings: mechanisms of tumor development and therapy with secondary plant metabolites," *Phytomedicine*, vol. 64, Article ID 153081, 2019.

[7] C. Allemani, T. Matsuda, V. Di Carlo et al., "Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries," *The Lancet*, vol. 391, no. 10125, pp. 1023–1075, 2018.

[8] T. Syeda-Mahmood, "Role of big data and machine learning in diagnostic decision support in radiology," *Journal of the American College of Radiology*, vol. 15, no. 3, pp. 569–576, 2018.

[9] M. Ahmadi, A. Sharifi, S. Hassantabar, and S. Enayati, "QAIS-DSNN: tumor area segmentation of MRI image with optimized quantum matched-filter technique and deep spiking neural network," *BioMed Research International*, vol. 2021, Article ID 6653879, 16 pages, 2021.

[10] Y. Liu, S. Stojadinovic, B. Hrycushko, Z. Wardak, S. Lau, W. Lu et al., "A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery," *PloS One*, vol. 12, no. 10, Article ID e0185844, 2017.

[11] S. P. RM, P. K. R. Maddikunta, M. Parimala et al., "An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture," *Computer Communications*, vol. 160, pp. 139–149, 2020.

[12] O. Hadad, R. Bakalo, R. Ben-Ari, S. Hashoul, and G. Amit, "Classification of breast lesions using cross-modal deep learning," in *Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pp. 109–112, Melbourne, Australia, 2017, April.

[13] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Into Imaging*, vol. 9, no. 4, pp. 611–629, 2018.

[14] A. Abdel Rahman, S. Belhaouari, A. Bouzerdoum, H. Baali, T. Alam, and A. Eldaraa, "breast mass tumor classification using deep learning," in *Proceedings of the 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, Doha, Qatar, February 2020.

[15] M. Ahmadi, A. Sharifi, M. Jafarian Fard, and N. Soleimani, "Detection of brain lesion location in MRI images using convolutional neural network and robust PCA," *International Journal of Neuroscience*, vol. 12, pp. 1–12, 2021 b.

[16] S. Charan, M. J. Khan, and K. Khurshid, "Breast cancer detection in mammograms using convolutional neural network," in *Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1–5, IEEE, Sukkur, Pakistan, 2018, March.

[17] M. Veena and M. C. Padma, "Detection of breast cancer using digital breast tomosynthesis," in *Emerging Research in Electronics, Computer Science and Technology*, pp. 721–730, Springer, Singapore, 2019.

[18] S. Punitha, A. Amuthan, and K. Suresh Joseph, "Enhanced monarchy butterfly optimization technique for effective breast cancer diagnosis," *Journal of Medical Systems*, vol. 43, no. 7, 206 pages, 2019.

[19] S. Sakri, N. B. Abdul Rashid, and Z. Muhammad Zain, "Particle swarm optimization feature selection for breast cancer recurrence prediction," *IEEE Access*, vol. 6, pp. 29637–29647, 2018.

[20] S. Karthik, R. Srinivasa Perumal, and C. Mouli, "Breast cancer classification using deep neural networks," in *Knowledge Computing and its Applications*, pp. 227–241, Springer, Singapore, 2018.

[21] A. Unni, E. Nidheep, S. Vinod, and S. Lekha, "Tumour detection in double threshold segmented mammograms using optimized GLCM features fed SVM," in *Proceedings of the 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 554–559, IEEE, Bangalore, India, September 2018.

[22] D. Selvathi and A. Aarthy Poornila, "Deep learning techniques for breast cancer detection using medical image analysis," in *Biologically Rationalized Computing Techniques for Image Processing Applications*, pp. 159–186, Springer, Berlin, Germany, 2018.

[23] S. Sasikala and M. Ezhilarasi, "Fusion of k-Gabor features from medio-lateral-oblique and craniocaudal view mammograms for improved breast cancer diagnosis," *Journal of Cancer Research and Therapeutics*, vol. 14, no. 5, 1036 pages, 2018.

[24] M. Heidari, A. B. Hollingsworth, G. Danala et al., "Prediction of breast cancer risk using a machine learning approach embedded with a locality preserving projection algorithm," *Physics in Medicine & Biology*, vol. 63, no. 3, Article ID 035020, 2018.

[25] N. Tariq, B. Abid, K. Ali Qadeer, I. Hashim, Z. Ali, and I. Khosa, "Breast cancer classification using global discriminate features in mammographic images," *Breast Cancer*, vol. 102 pages, 2019.

[26] K. L. Kashyap, M. K. Bajpai, P. Khanna, and G. George, "Mesh-free based variational level set evolution for breast

region segmentation and abnormality detection using mammograms," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 34, no. 1, Article ID e2907, 2018.

[27] D. Ribli, H. Anna, Z. Unger, P. . Pollner, and I. Csabai, "Detecting and classifying lesions in mammograms with deep learning," *Scientific Reports*, vol. 8, no. 1, 4165 pages, 2018.

[28] F. Gao, T. Wu, L. Jing et al., "A shallow-deep CNN for improved breast cancer diagnosis," *Computerized Medical Imaging and Graphics*, vol. 70, pp. 53–62, 2018.

[29] H. Jung, B. Kim, I. Lee et al., "Detection of masses in mammograms using a one-stage object detector based on a deep convolutional neural network," *PLoS One*, vol. 13, no. 9, Article ID e0203355, 2018.

[30] S. Shams, R. Platania, J. Zhang, J. Kim, K. Lee, and S.-J. Park, "Deep generative breast cancer screening and diagnosis," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 859–867, Springer, Granada, Spain, September 2018.

[31] M. A. Al-masni, M. A. Al-antari, J.-M. Park et al., ""Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system," *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 85–94, 2018.

[32] H. Chougrad, Z. Hamid, and A. Omar, "Deep convolutional neural networks for breast cancer screening," *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 19–30, 2018.

[33] D. A. Ragab, M. Sharkas, S. Marshall, and J. Ren, "Breast cancer detection using deep convolutional neural networks and support vector machines," *Peer Reviewed Journal*, vol. 7, Article ID e6201, 2019.

[34] M. Hazarika and L. B. Mahanta, "A new breast border extraction and contrast enhancement technique with digital mammogram images for improved detection of breast cancer," *Asian Pacific Journal of Cancer Prevention: APJCP*, vol. 192141 pages, 2018.

[35] P. Anjaiah, K. Rajendra Prasad, and C. Raghavendra, "Effective texture features for segmented mammogram images," *International Journal of Engineering &Technology*, vol. 7, no. 3, pp. 666–669, 2018.

[36] M. M. Eltoukhy, M. Elhoseny, K. M. Hosny, and K. Amit, "Computer aided detection of mammographic mass using exact Gaussian–Hermite moments," *Journal of Ambient Intelligence and Humanized Computing*, vol. 167, pp. 1–9, 2018.

[37] T. V. Padmavathy, M. N. Vimalkumar, and D. S. Bhargava, "Adaptive clustering based breast cancer detection with ANFIS classifier using mammographic images," *Cluster Computing*, vol. 22, no. 6, pp. 13975–13984, 2019.

[38] M. Tahmooresi, A. Afshar, B. Bashari Rad, K. B. Nowshath, and M. A. Bamiah, "Early detection of breast cancer using machine learning techniques," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 10, no. 3-2, pp. 21–27, 2018.

[39] M. Amrane, S. Oukid, I. Gagaoua, and T. Ensarİ, "Breast cancer classification using machine learning," in *Proceedings of the 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)*, pp. 1–4, IEEE, Istanbul, Turkey, July 2018.

[40] R. Vijayarajeswari, P. Parthasarathy, S. Vivekanandan, and A. Alavudeen Basha, "Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform," *Measurement*, vol. 146, 2019.

[41] C. L. Chowdhary, M. Mittal, P. A. Pattanaik, and Z. Marszalek, "An efficient segmentation and classification system in medical images using intuitionist possibilistic fuzzy C-mean clustering and fuzzy SVM algorithm," *Sensors*, vol. 20, no. 14, 3903 pages, 2020.

[42] C. L. Chowdhary, P. G. Shynu, and V. K. Gurani, "Exploring breast cancer classification of histopathology images from computer vision and image processing algorithms to deep learning," *International Journal of Advanced Science and Technology*, vol. 29, pp. 43–48, 2020.

[43] Chowdhary, C. Lal, and D. P. Acharjya, "Segmentation of mammograms using a novel intuitionistic possibilistic fuzzy c-mean clustering algorithm," in *Nature Inspired Computing*, pp. 75–82, Springer, Singapore, 2018.

[44] A. Rampun, B. W. Scotney, P. J. Morrow, and H. Wang, "Breast mass classification in mammograms using ensemble convolutional neural networks," in *Proceedings of the 2018 IEEE 20th International Conference on E-Health Networking, Applications and Services (Healthcom)*, pp. 1–6, IEEE, Ostrava, Czech Republic, September 2018.

[45] E.-K. Kim, H.-E. Kim, K. Han et al., "Applying data-driven imaging biomarker in mammography for breast cancer screening: preliminary study," *Scientific Reports*, vol. 8, no. 1, 2762 pages, 2018.

[46] X. Zhang, Y. Zhang, Y. Erik et al., "Classification of whole mammogram and tomosynthesis images using deep convolutional neural networks," *IEEE Transactions on Nanobioscience*, vol. 17, no. 3, pp. 237–242, 2018.

[47] M. S. Salama, A. S. Eltrass, and H. M. Elkamchouchi, "An improved approach for computer-aided diagnosis of breast cancer in digital mammography," in *Proceedings of the 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 1–5, IEEE, Rome, Italy, June 2018.

[48] Suckling, J., Parker, J., Dance, D., et al. (2015). Mammographic Image Analysis Society (MIAS) Database V1.21 [dataset].

[49] M. Rakshit and S. Das, "An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter," *Biomedical Signal Processing and Control*, vol. 40, pp. 140–148, 2018.

[50] B. Zong, Q. Song, M. R. Min et al., "Deep autoencoding Gaussian mixture model for unsupervised anomaly detection," in *Proceedings of the International Conference on Learning Representations*, New Orleans, LA, USA, 2018, February.

[51] E. Tzoreff and A. J. Weiss, "Expectation-maximization algorithm for direct position determination," *Signal Processing*, vol. 133, pp. 32–39, 2017.

[52] R. Liu, C. Huang, T. Li, L. Yang, and H. Zhu, "Statistical disease mapping for heterogeneous neuroimaging studies," in *Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 1415–1418, IEEE, Washington, DC, USA, 2018 April.

[53] Y. Zeinali and B. A. Story, "Competitive probabilistic neural network," *Integrated Computer-Aided Engineering*, vol. 24, no. 2, pp. 105–118, 2017.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 9980326, 12 pages
https://doi.org/10.1155/2021/9980326

Research Article

Multi-Features-Based Automated Breast Tumor Diagnosis Using Ultrasound Image and Support Vector Machine

Zhemin Zhuang,[1] **Zengbiao Yang**,[1] **Shuxin Zhuang**,[1] **Alex Noel Joseph Raj**,[1] **Ye Yuan**,[1] **and Ruban Nersisson**[2]

[1]*Department of Electronic Engineering, Shantou University, Shantou 515063, China*
[2]*School of Electrical Engineering, Vellore Institute of Technology, Vellore 632014, India*

Correspondence should be addressed to Alex Noel Joseph Raj; jalexnoel@stu.edu.cn

Received 30 March 2021; Accepted 7 May 2021; Published 19 May 2021

Academic Editor: V. Rajinikanth

Breast ultrasound examination is a routine, fast, and safe method for clinical diagnosis of breast tumors. In this paper, a classification method based on multi-features and support vector machines was proposed for breast tumor diagnosis. Multi-features are composed of characteristic features and deep learning features of breast tumor images. Initially, an improved level set algorithm was used to segment the lesion in breast ultrasound images, which provided an accurate calculation of characteristic features, such as orientation, edge indistinctness, characteristics of posterior shadowing region, and shape complexity. Simultaneously, we used transfer learning to construct a pretrained model as a feature extractor to extract the deep learning features of breast ultrasound images. Finally, the multi-features were fused and fed to support vector machine for the further classification of breast ultrasound images. The proposed model, when tested on unknown samples, provided a classification accuracy of 92.5% for cancerous and noncancerous tumors.

1. Introduction

International Agency for Research on Cancer (IARC) reported that breast cancer accounts for about 24.2% of cancers diagnosed in women worldwide [1]. It is also the leading fatal cause in women, accounting for about 15%. With the development of modern medicine, if breast cancer is diagnosed early, the survival rate of patients will be significantly improved [2]. Breast tumors are usually examined by computerized tomography (CT), magnetic resonance imaging (MRI), molybdenum target X-ray, far infrared, ultrasound (US), and other methods. Among them, the US has become the preferred choice for early breast cancer screening due to its cost-effectiveness and more robust images [3]. However, the Breast Imaging-Reporting and Data System (BI-RADS) [4] grades diagnosed by different clinicians for the same patient are subjective and different since some features in the breast ultrasound (BUS) images are not typically visible to diagnose [5]. Besides, different breast lesions show different features in BUS images. Also, experience and the ability to understand the visual clues from BUS images are essential in reducing false negative detection. The count shows that the missed diagnosis of medical imaging in disease diagnosis can be between 10% and 30% [6].

Artificial intelligence (AI) can assist doctors in making a more accurate judgment because of its objectivity and versatility. AI diagnosis of benign and malignant BUS images can be divided into deep learning and feature extraction. Deep learning transforms the raw BUS images into much higher-dimension expression through convolutional neural networks (CNNs). Multi-level and multi-aspect features are obtained by training the network model, which makes the obtained features have a more robust generalization and expression. Deep learning is often used in the automatic classification of BUS images. For example, in reference [7], 166 malignant and 292 benign BUS images were trained and classified by using a neural network

composed of three convolution layers and two fully connected layers. Qi et al. [3] used Mt-Net (malignant tumors) and Sn-Net (solid nodules) to classify BUS images, where Mt-Net was used to detect malignant tumors and Sn-Net was used to detect solid nodules. Although deep learning has achieved good results, they are constrained by the need for a higher number of ground truth (GT).

Feature extraction techniques identify useful characteristic features (CFs) from the original images, where the original image is transformed into a group of features with obvious physical significance, to achieve the purpose of dimensionality reduction. For example, in reference [8], the region growing method was used to segment the lesion, and the histogram method was used to calculate six histograms from the posterior shadowing region (PS). Finally, BUS images were divided into PS enhancement and PS nonenhancement by using the six histograms and multilayer perceptron (MLP). However, PS is only one of the features to judge the benign and malignant BUS images and lacks accuracy to make a classification.

In this paper, (1) by analyzing the different manifestations of benign and malignant breast tumors in ultrasound images, combined with the clinical experience of experts, different and effective characteristic features were designed manually. (2) In order to assist the classification of breast tumor ultrasound images, we used transfer learning for extracting deep learning features. Finally, (3) SVM was used to integrate characteristic features and deep learning features and present an effective classification.

2. Materials and Methods

Since benign and malignant breast BUS images have different histological structures and biological characteristics, they exhibit different properties on BUS images. The malignant tumors such as ductal carcinoma in situ [9], due to their invasive nature, penetrate through the ducts and into adjacent fibrous and adipose tissues. It forms a blurred mixed zone between the tumor and the tissues and complex edge. Besides, the complex interstitial components and hyaline transformation of malignant lesions often lead to scattering of acoustic signals [10, 11]. During the decision process, the specialists often consider the orientation, posterior shadowing (PS), edge indistinctness (EI), and shape complexity (SC) of the tumors as essential features to identify them as benign or malignant. These characteristic features (CFs) of BUS images are not only a critical judgment for clinical diagnosis but also a significant basis for BUS image classification based on feature extraction. Hence, we proposed a method to classify breast lesions using multifeatures (MFs) and SVM. The proposed method firstly employed a level set technique to segment the tumor region of BUS images. From the contour of the segmented tumor, orientation and edge indistinctness (EI) scores were calculated. Next, using Hu moments, we determined the characteristic of the posterior shadowing (PS) region. Later, the fractal divider method was used to calculate the shape complexity (SC) score of tumor contour. Meanwhile, the pretrained VGG16 model was used as a feature extractor to get deep learning features (DLFs) of BUS images. Finally, we classified the BUS images based on the multi-features (MFs) obtained above and SVM. The above process is illustrated in Figure 1.

3. Lesion Segementation and Feature Calculation

Due to the complexity involved during the ultrasound examination, the acquired images contain speckle noise, image artifacts, and weak boundaries that hurt the segmentation process. Accurate segmentation can effectively improve the accuracy of classification [12]. Therefore, for accurate extraction of tumor regions, conventional segmentation techniques may not provide desired results. Literature suggests that level set techniques are useful for segmentation problems related to topological changes, and hence we used an advanced level set segmentation algorithm based on geometric active contour model and curve evolution theory to complete the lesion segmentation of BUS images [13]. The technique employed an iterative method to segment the tumor region within the BUS image accurately. The following paragraph briefly explains the use of a level set algorithm to segment the BUS images.

The level set algorithm that we used in this paper does not depend on the gradient information of the BUS image. Therefore, it is insensitive to noise and has a significant advantage in medical image processing [14, 15]. Here we employed the Distance Regularized Level Set Evolution (DRLSE) [16] model that eliminates the need for reinitialization but employs a distance regularization term and energy functions to propagate the zero-level set function (LSF) towards the desired locations. The energy function $E(\phi)$ can be defined as follows.

For ϕ: $\Omega \longrightarrow \mathfrak{R}$,

$$E(\phi) = \mu \int_{\Omega} p(|\nabla\phi|)\mathrm{d}x + \lambda \int_{\Omega} g\delta(\phi)|\nabla\phi|\mathrm{d}x + \alpha \int_{\Omega} gH(-\phi)\mathrm{d}x, \tag{1}$$

where $\int_{\Omega} p(|\nabla\phi|)\mathrm{d}x$ is the regularization term and $\int_{\Omega} g\delta(\phi)|\nabla\phi|\mathrm{d}x$ and $\int_{\Omega} gH(-\phi)\mathrm{d}x$ are the external energy terms. $\mu > 0$, $\lambda > 0$, $\alpha \in \mathfrak{R}$ are the coefficients, respectively, H and δ are the Heaviside function and Dirac delta function, respectively, and p is a potential function.

Due to the addition of the distance regularization term $p(|\nabla\phi|)$ in equation (1), the deviation between the level set and the signed distance function (SDF) is automatically corrected in each iteration of the level set function, thus maintaining stability and avoiding reinitialization.

The following gradient descent flow function can be obtained by differentiating equation (1), to realize the extraction of the tumor region in BUS image while minimizing the energy function.

$$\frac{\partial\phi}{\partial t} = \mu\left[\Delta\phi - \mathrm{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right] + \lambda\delta_{\varepsilon}(\phi)\mathrm{div}\left(g\frac{\nabla\phi}{|\nabla\phi|}\right) + \alpha g\delta_{\varepsilon}(\phi). \tag{2}$$

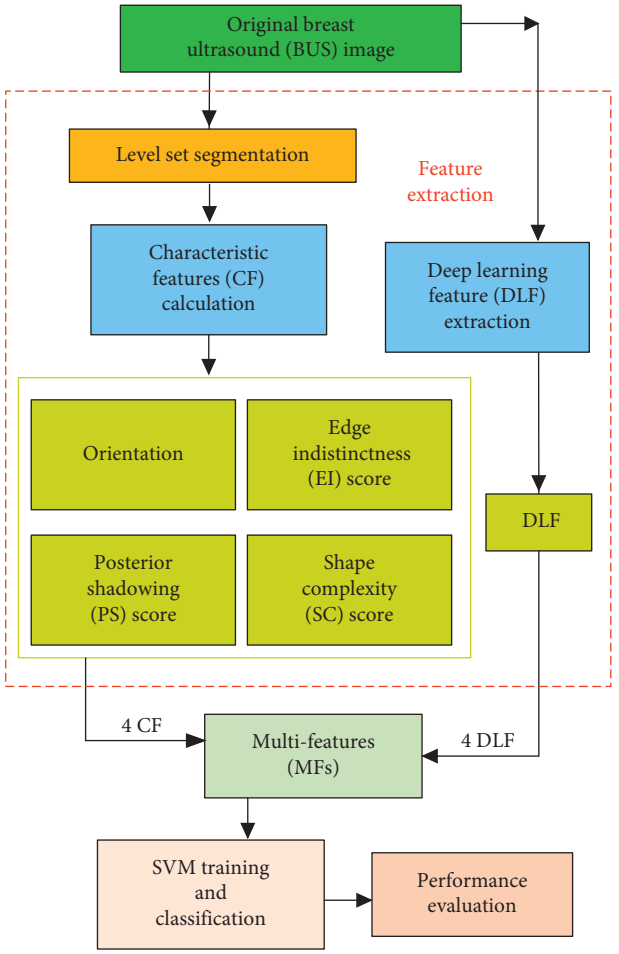


Figure 1: The proposed methodology.

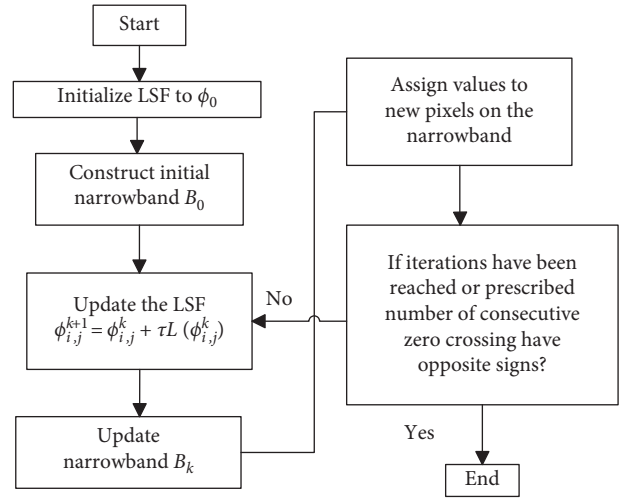


Figure 2: The evolution process of DRLSE.

The implementation of DRLSE for an application is based on the flowchart illustrated in Figure 2, which includes (a) initialization of LSF ϕ_0 and narrowband B_0; (b) updating the LSF and narrowband region; (c) updating the pixel values on the narrowband based on $\phi_{i,j}^{k+1}$, where (i, j) is called zero-crossing point, k is the number of iteration, and τ is the time step; and (d) termination of the process, if the prescribed number of consecutive zero-crossing has opposite signs or the expected number iterations is reached. The segmentation results of the lesions in breast ultrasound images are shown in Figure 3.

3.1. Computation of Orientation. The growth characteristics of benign and malignant tumors vary and therefore show different orientations. Here we used contour obtained from the segmentation process to facilitate the calculation of tumor orientation. First, we transversed the segmented contour in both horizontal and vertical directions and obtain four points: top (x_u, y_u), bottom (x_d, y_d), leftmost (x_l, y_l), and rightmost (x_r, y_r). These points are vertices obtained from the intersection of lines l_u, l_d, l_l, and l_r with upper, lower, leftmost, and rightmost extreme regions of contour, respectively, as shown in Figure 4.

Next, using equation (3), we computed the tumor orientation, which is expressed as a ratio between the height to the width of the tumor.

$$\text{Orientation} = \frac{y_u - y_d}{x_r - x_l}. \tag{3}$$

3.2. Computation of Edge Indistinctness (EI) Score of Lesions. Commonly malignant tumors penetrate deeper into the tissues causing indistinctive margins that are different from the benign ones. Therefore, to measure the edge indistinctness more comprehensively, we extracted a region of m x n pixels from the top and bottom vertices of the contour, i.e., around (x_u, y_u) and (x_d, y_d) points, as shown in Figure 5.

For the extracted $m \times n$ region, we separately calculated the standard deviation along the row and the column as given in equations (4) and (5). We defined EI score (equation (6)) as the maximum of the standard deviation computed along with the row and column directions:

$$x_{\text{std}} = \frac{\sum_{i=1}^{n} x_{\text{std}\_i}}{n}, \tag{4}$$

$$y_{\text{std}} = \frac{\sum_{j=1}^{m} y_{\text{std}\_j}}{m}, \tag{5}$$

$$\text{EI score} = \max\left(x_{\text{std}}, y_{\text{std}}\right). \tag{6}$$

For our experiments, we have selected two regions around (x_u, y_u) and (x_d, y_d) points and accordingly computed two EI scores: $\text{EI score}_{\text{up}}$ and $\text{EI score}_{\text{down}}$, respectively.

In cancer diagnosis, the edge strength (blurring) is an important index that is used to classify the tumor. However, there are differences in the degree of blurring across different sections along the tumor boundary. Thus, we defined the average values of $\text{EI score}_{\text{up}}$ and $\text{EI score}_{\text{down}}$ as EI score, as shown in the following equation:

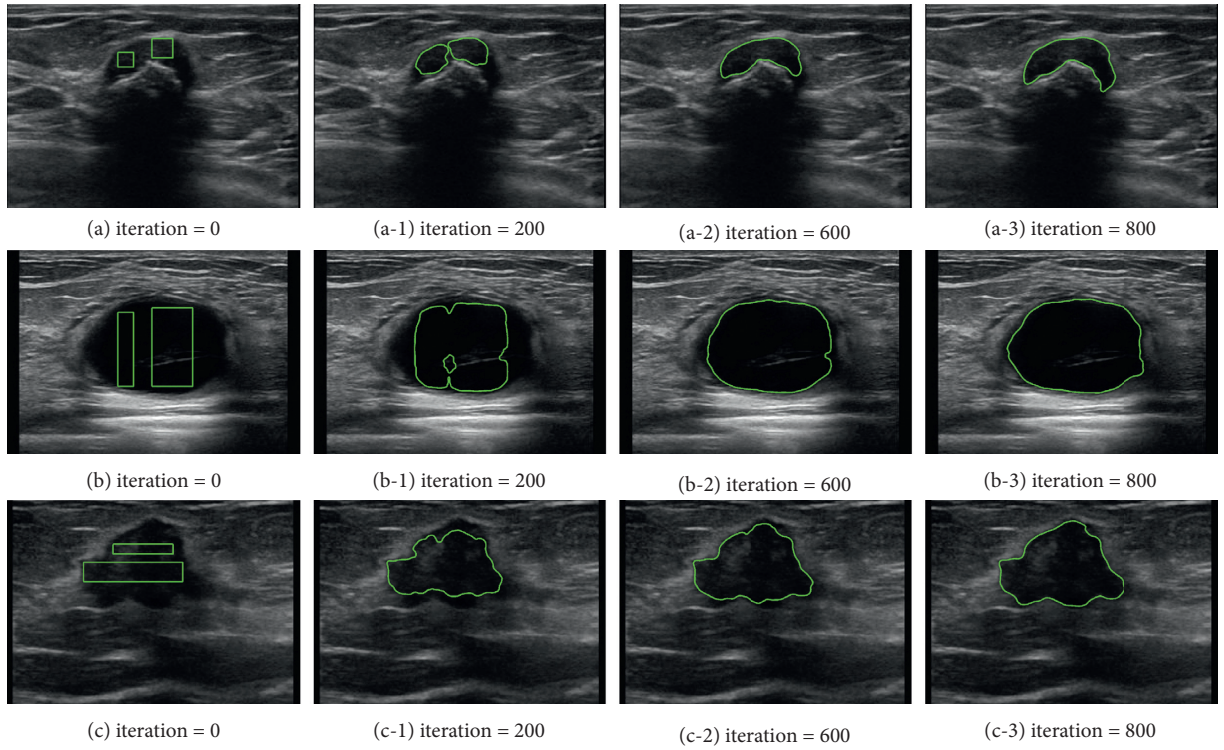


Figure 3: Segmentation results from the DRLSE model: (a), (b), and (c) correspond to zero LSF. (a-1) to (a-3), (b-1) to (b-3), and (c-1) to (c-3) illustrate the evolution of the LSF at 200, 600, and 800 iterations, respectively.

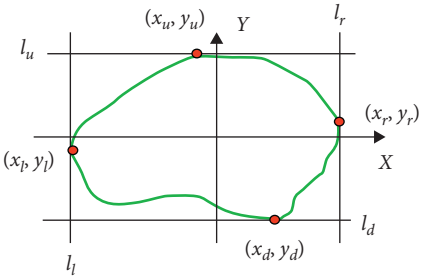


Figure 4: The contour of the segmented tumor region of the BUS image. (x_u, y_u), (x_d, y_d), (x_l, y_l), and $(x_r, y_r)$are the four vertices obtained from the intersection of lines l_u,l_d, l_l, and l_r with the upper, lower, leftmost, and rightmost extreme of contour, respectively.

$$\text{EI score} = \frac{\left(\text{EI score}_{\text{up}} + \text{EI score}_{\text{down}}\right)}{2}. \tag{7}$$

3.3. Computation of the PS Score of Posterior Shadowing Region by Using Hu Moments. The texture characteristics of the PS region are generally different for benign and malignant tumors. Literature suggests that moments can be used for analyzing texture characteristics [17], and therefore, we used moment invariants proposed by Hu [18] to compare the PS regions of different BUS images quantitatively. The PS region of the BUS image was extracted based on the bottom (x_d, y_d), the leftmost (x_l, y_l), and the rightmost (x_r, y_r) points of the contour, as shown in Figure 6.

Let $f(x, y)$ be the extracted PS region; then, its $(p + q)$ order of the geometric moment can be defined as follows:

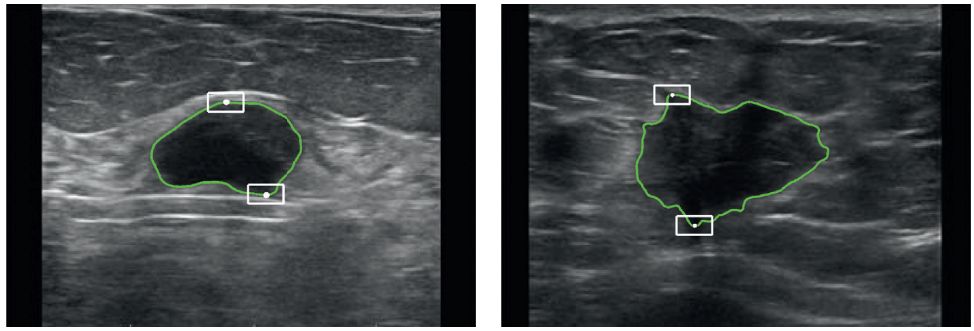
$$m_{pq} = \sum_{y=1}^{N} \sum_{x=1}^{M} x^p y^q f(x, y), \quad p, q = 0, 1, 2 \ldots. \tag{8}$$

Their central moments can be defined as

$$\mu_{pq} = \sum_{y=1}^{N} \sum_{x=1}^{M} (x - \overline{x})^p (y - \overline{y})^q f(x, y), \quad p, q = 0, 1, 2 \ldots, \tag{9}$$

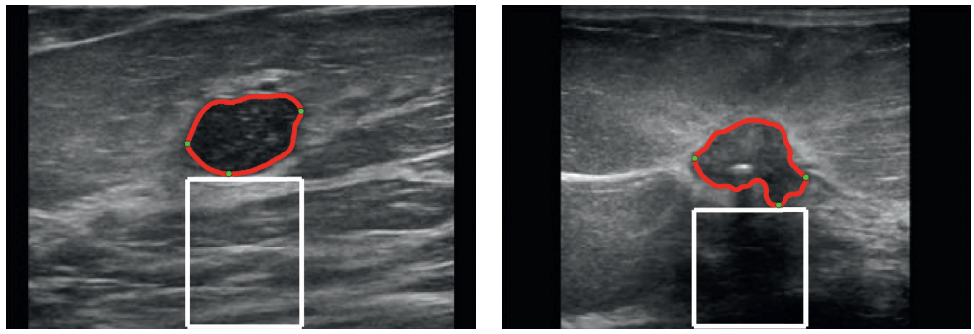
where $\overline{x}$ and $\overline{y}$ are the center of gravity coordinates of the image, given by

$$\overline{x} = \frac{m_{10}}{m_{00}}, \qquad \overline{y} = \frac{m_{01}}{m_{00}}. \tag{10}$$



(a) (b)

Figure 5: (a) and (b) illustrate segmented contour shown in green and selected $n \times m$ regions around (x_u, y_u) and (x_d, y_d) points in white, respectively. Here we set $n = 15$ and $m = 21$.



(a) (b)

Figure 6: The PS regions. (a) and (b) are the benign and malignant BUS images, respectively. The red line illustrates the segmented tumor contour; the white region represents the PS region considered for PS score computation.

The normalized central moment is defined as

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\rho}}, \tag{11}$$

where $\rho = ((p + q)/2) + 1$.

Normally, seven Hu moments can be computed using the second- and third-order normalized central moments. But here we only use the first moment as it is sufficient to provide a score that could differentiate PS regions of different BUS images effectively. Accordingly, we substituted $p = 0$ and $q = 2$ in equation (11) and defined the PS score as

$$\text{PS score} = \eta_{20} + \eta_{02}. \tag{12}$$

To have a clear distinction, the PS score is transformed as

$$\text{PS score} = -\text{sign}(\text{PS score}) * (\log 10(abs(\text{PS score}))). \tag{13}$$

3.4. Computation of Shape Complexity (SC) Score Based on Fractal Dimensions. The shape is one of the critical factors clinical experts use to classify tumors as benign or malignant. Cancerous tumors have complex contours, whereas benign ones have simpler structures. Therefore, we proposed a technique based on fractals to quantify the shape complexity of the segmented tumors. In image processing, fractals have been widely used in US image segmentation. Omiotek et al. [19] used fractals to quantify the texture of thyroid US images. Lin et al. [20] used fractals to determine the area of alveolar bone defect, and recently Zhuang et al. [21] collectively used fuzzy enhancement and fractal length to segment the US image of atherosclerosis successfully. Also, the fractal theory was successfully used to measure the irregular coastlines and the fault geometries propagated by earthquakes. For example, Mandelbrot [22] employed power law to relate the costal length to the different linear rulers, and Okubo and Aki [23] quantified complex fault geometries to large values of fractal dimension.

Here the divider method used by cartographers [23–25] was employed to quantify the shape complexity of the segmented tumor. To measure the complexity, we drew circles of different radii along the boundary of the segmented contour. The starting point was chosen as the top point (x_u, y_u) and the circles of different radii were drawn dividing the contours, as shown in Figure 7. Let $N(R)$ be the number of the circles and R be their corresponding radius; then, according to [26, 27], we can relate $N(R)$ and R as

$$\text{Log}_{10} N(R) = a + b\text{Log}_{10} R, \tag{14}$$

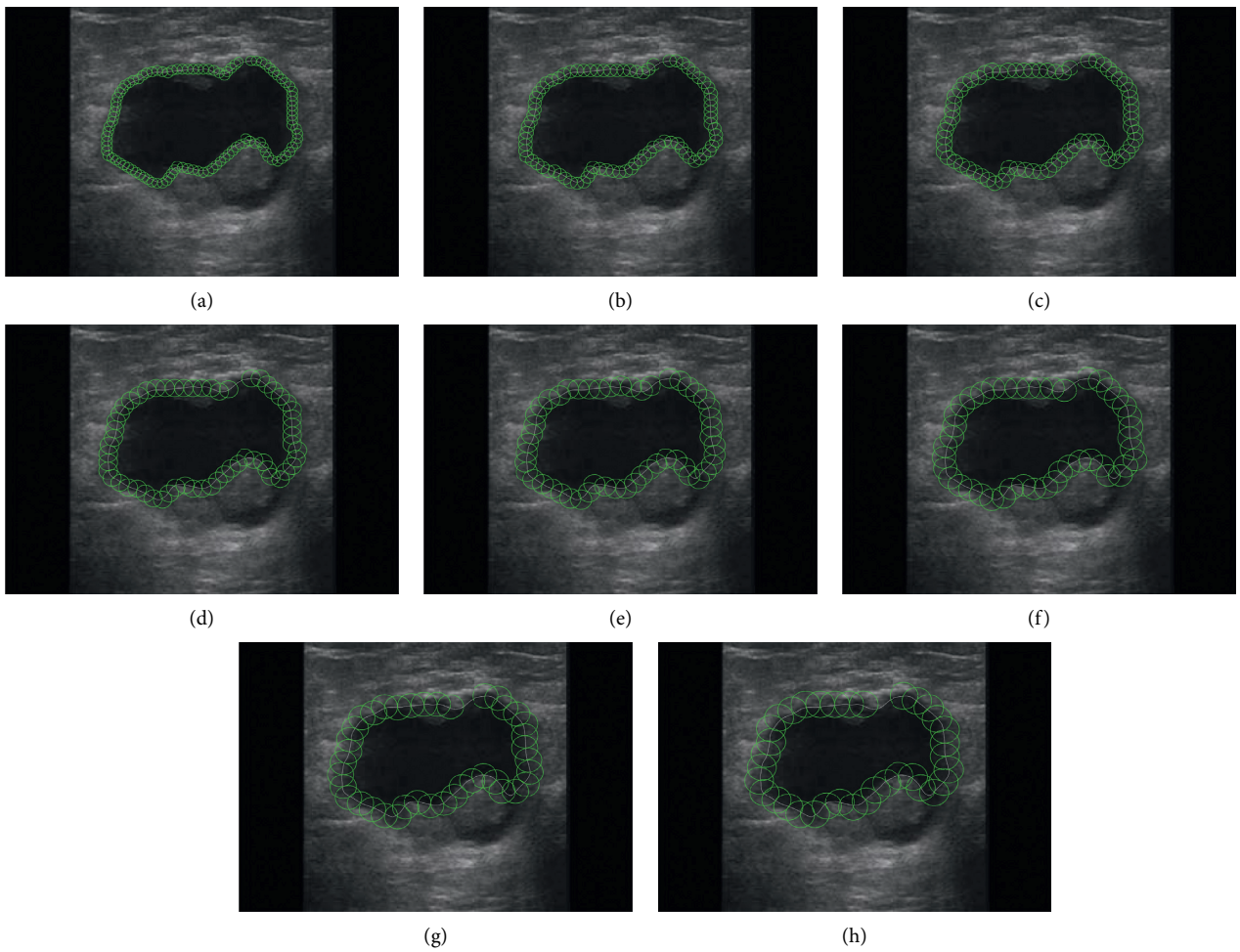


Figure 7: Illustration of divider method. (a)–(h) show the circles drawn with different radii along the contour of the segmentation tumor. (a) $R = 4$. (b) $R = 5$. (c) $R = 6$. (d) $R = 7$. (e) $R = 8$. (f) $R = 9$. (g) $R = 10$. (h) $R = 11$.

where "a" and "b" are constants that are obtained through least-squares fit between $\text{Log}_{10}N(R)$ and $\text{Log}_{10}R$ and "b" is the slope of the line that represents the fractal dimension, which determines the SC score of the contour.

The illustration of the divider method for a sample BUS image is shown in Figure 7, and Table 1 presents R and their corresponding $N(R)$ values. Further Figure 8 demonstrates the SC score obtained by the least-square fitting of $N(R)$ and R for the sample BUS image shown in Figure 7.

4. Deep Learning Features (DLF) Extraction and SVM-Based Classification

In the above, we identified handcrafted features such as shape score, PS score, and EI and orientation to identify the characteristic features (CFs) of the segmented lesion in a BUS image. Here we retrieved high-dimension features from the BUS image to assist the classification. A deep learning method can fix this problem by performing convolution operations on the input graphics multiple times. Therefore, by using deep learning, we extracted the DLF of the image and combined them with the manually extracted CF to distinguish benign and malignant BUS images.

Table 1: Number of circles $N(R)$ and their radius R for BUS image shown in Figure 7.

Figure 7	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
R	4	5	6	7	8	9	10	11
$N(R)$	109	85	72	61	53	46	43	37

Due to the limited dataset, our paper used transfer learning [28] to train the neural network as the DLF extractor. Initially, we compared the classification ability of VGG16 [29], VGG19 [29], ResNet50 [30], and Inception V3 [31] for BUS images. The experiment results show that VGG16 is better than other networks. Then, based on transfer learning, we froze the convolution layer of VGG16, added the global average pooling layer, and changed $7 \times 7 \times 512$ of the original VGG16 output into $1 \times 1 \times 512$. For the full connected layer, the nodes were set to 512, 128, 4, and 2, respectively. At the same time, the Relu activation function was added after each fully connected layer except the last one. The modified VGG16 is shown in Figure 9. For the output, the cross-entropy loss function was used to calculate the loss. After

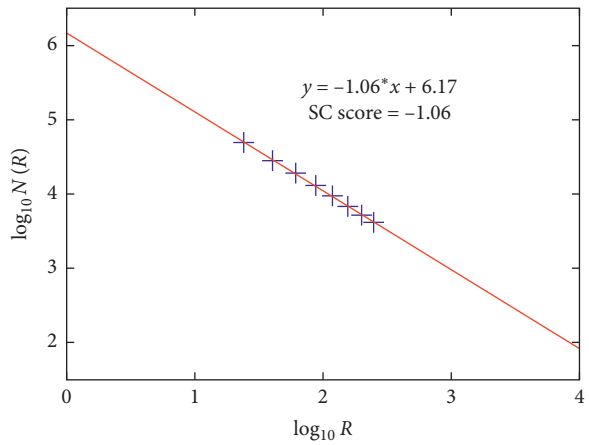


Figure 8: Computation of SC score for the measurements recorded in Table 1 on equation (15).

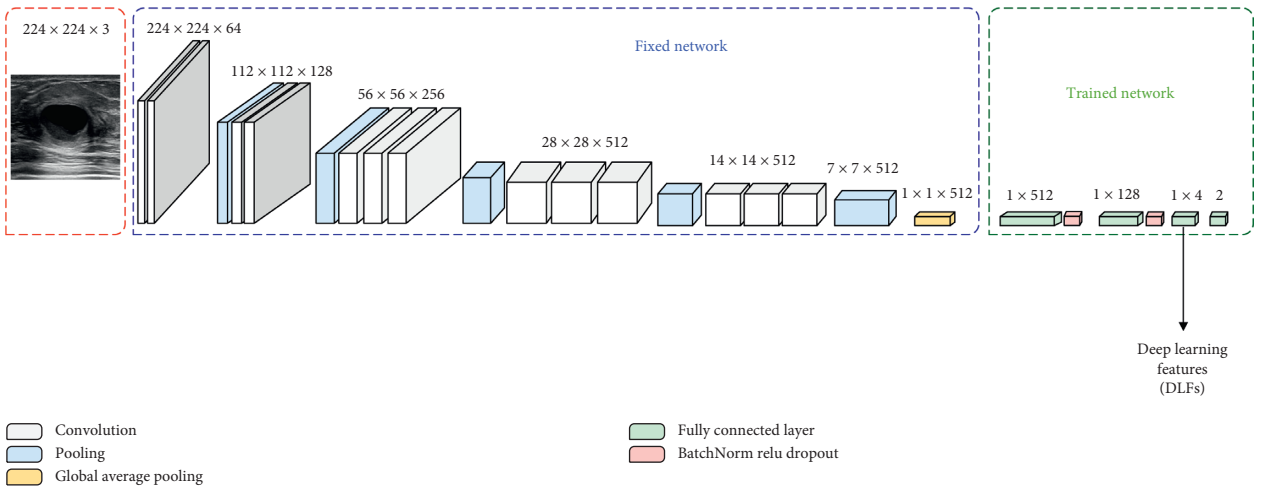


Figure 9: Modified VGG16.

training, for other input images, we could get four values (DLF) from the second fully connected layer of the trained model.

Once the features have been computed from the BUS images, we used the SVM classifier to classify them as benign or malignant. The SVM was chosen since (a) the availability of labeled large medical image data sets for training is not feasible. SVM could provide better classification accuracy for smaller training sets [32]; (b) SVM theory provides a way to avoid inseparable problem in low-dimensional space through the use of kernel functions [33]. This attribute simplifies the problems in higher-dimensional space providing a more generalized classification.

Normally, to solve the problem of linear inseparability in low-dimensional space, kernel functions are used to map the features from low-dimensional space to high-dimensional space, thus realizing the linear classification in higher-dimensional space. In the experiments, the radial basis function (RBF) [34] was used as the kernel function of SVM to classify the CF, the DLF, and the MF, respectively, and the classification results were evaluated by various evaluation indexes.

5. Results

For experimental results, we have used 1802 BUS images that include 787 benign and 1015 malignant BUS images. It contains two parts. The first part is provided by Ultrasoundcases.info, which is a professional breast cancer ultrasound website developed by Hitachi Medical Systems in Switzerland and Dr. Taco Geertsma, who works for Gelderse Vallei hospital in the Netherlands. It contains many ultrasound cases, which were collected and labeled by radiologists and ultrasound technicians of the hospital over the years. The other part is supported by the First Affiliated Hospital of Shantou University, Guangdong Province, China.

5.1. Evaluation Indexes. The following measures (equations (15)–(19)) were used as a metric to evaluate the performance of the SVM classifier model [35].

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{15}$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{16}$$

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{17}$$

$$F1 - \text{score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}, \tag{18}$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{19}$$

where true positive (TP): GT malignant and prediction malignant; false positive (FP): GT benign and prediction malignant; false negative (FN): GT malignant and prediction benign; and true negative (TN): GT benign and prediction benign.

5.2. Characteristic Feature (CF) Calculation. To illustrate the calculation of CF, we presented 6 BUS images as examples, as shown in Figure 10. The calculation results of CF are listed in Table 2. The original value is calculated by using the proposed methods, and the normalized value presents the normalized CF value.

It can be seen from Table 2 that the characteristic features (CFs) of Figures 10(a) and 10(d) are consistent with the biological properties of benign and malignant BUS images. For example, Figure 10(a) which was diagnosed as benign tumor has larger EI score compared with malignant tumor (Figure 10(d)). However, we can also find that benign and malignant BUS images may have the same properties, for example, the posterior shadowing (PS) shown in Figures 10(b) and 10(f) is relatively low. Therefore, only relying on a single feature to distinguish benign and malignant BUS images will present a higher probability of misjudgment.

5.3. Deep Learning Feature (DLF) Extractor. To select the best DLF extractor, based on transfer learning, we used VGG16, VGG19, ResNet50, and Inception V3 to train and test the dataset, which is composed of 955 malignant and 727 benign BUS images. In the experiment, the training set accounts for 80% of the total data and the remaining accounts for the test set. The experimental results are shown in Table 3.

As we can see from Table 3, the accuracy, sensitivity, specificity, and $F1$-score of VGG16 are 0.84, 0.86, 0.82, and 0.86, respectively, which are higher than those of VGG19, ResNet50, and Inception V3. From the table, it can be inferred that VGG16 has better classification ability; it can learn BUS image features better than other networks. Therefore, we use VGG16 to extract deep learning features (DLFs) to assist the SVM classification.

5.4. Classification of Breast Tumors Based on SVM. In this experiment, we used another 120 BUS images, which are totally different from the data used in the above section. Among them, 80 BUS images were used to train the SVM model, including 40 benign BUS images and 40 malignant BUS images. Also, we took another 40 BUS images as the test set. For those 120 BUS images, we extracted their characteristic features (CFs) and deep learning features (DLFs) and then concatenated them serially to form multi-features (MF), as shown in equation (20). Later, the MF was normalized and labeled for supervised learning. We use "1" to indicate malignant and "0" to label benign. After the preparation, we firstly used triple cross-validation on the training samples to get the best classifier and then used the classifier to test the training samples and test samples, respectively.

$$\text{MF} = \{\text{CF}, \text{DLF}\}, \tag{20}$$

where,

$$\text{CF} = \{\text{orientation}, \text{EI score}, \text{PS score}, \text{SC score}\}. \tag{21}$$

In Table 4, CF and a different number of deep learning features (DLFs) are used as multi-features (MFs) to classify BUS images based on SVM. The result shows that the MF composed of CF and 4 DLF could get better classification results. Therefore, we choose CF and 4 DLF as the final MF. Further, from Table 5, it can be seen that by using SVM, the classification accuracy, precision, sensitivity, specificity, and $F1$-score of MF are 0.925, 0.905, 0.95, 0.905, and 0.927, respectively, which are higher than those indicators obtained from other classification methods. In Figure 11, the ROC curve and AUC value of SVM classification based on characteristic features (CFs), four deep learning features (DLFs), and multi-features (MFs) are recorded. The results show that the AUC value of MF is 0.970, which is higher than that of CF (0.935) or 4 DLF (0.895). Therefore, the classification model based on MF is better than other classification techniques.

5.5. Triple Cross-Validation. Here, we used triple cross-validation to illustrate the advantages of using the multi-features (MF), that is, for different c, g values in SVM, the training dataset is randomly divided into three parts: two of them are considered as the training set and the rest is used for verification set. The average accuracy of the three validation sets is considered as the accuracy of the SVM with these c, g values. Here $c = 0.5$ is the regularization parameter of SVM, and $g = 0.25$ is the parameter of radial basis function (RBF).

The contour plot in Figure 12 represents different accuracy values obtained when different c and g were used during triple cross-validation. After the triple cross-validation, the c, g values that represent the highest accuracy among the triple cross-validation were considered as the parameters of the final SVM classifier. The accuracy among the triple cross-validation can reach over 88%, which is close to the final accuracy of 92.5% obtained by using the best c, g. This indeed illustrates that the model has good robustness.

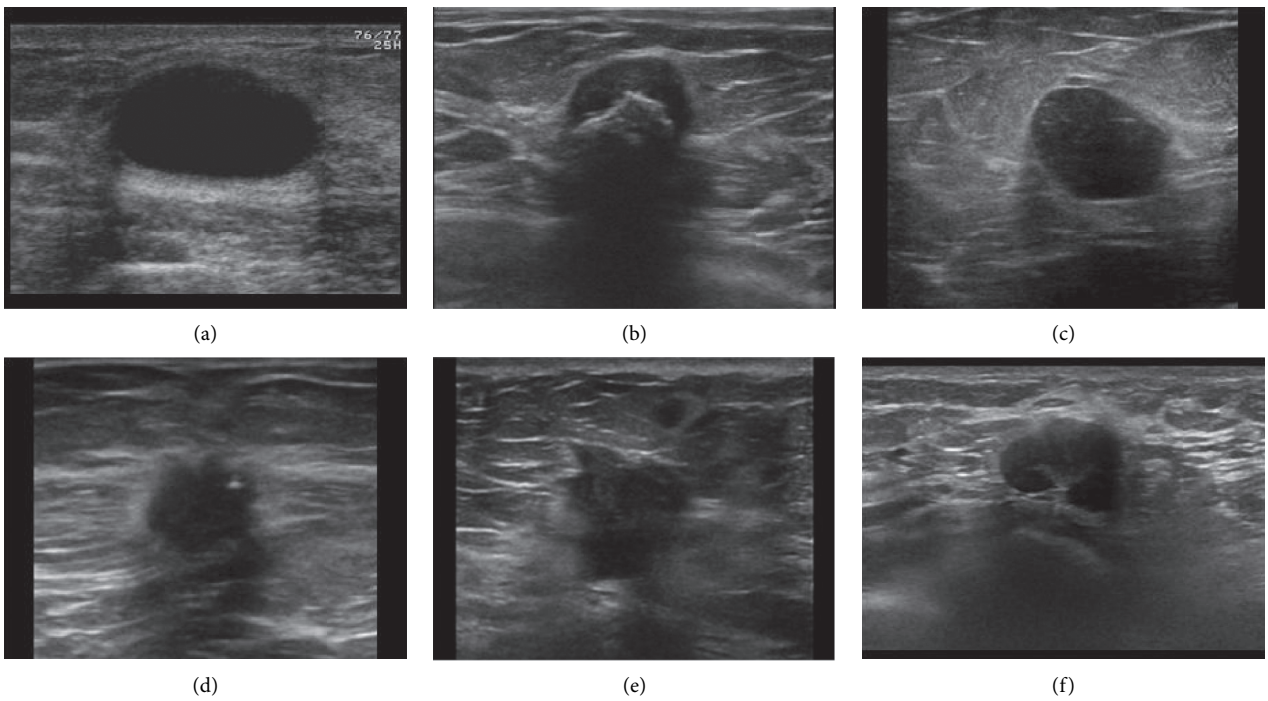


Figure 10: BUS images. (a), (b), and (c) are benign. (d), (e), and (f) are malignant.

Table 2: Calculation results of CF in Figure 10 (Figures 10(a)–10(c) are benign and Figures 10(d)–10(f) are malignant) (EI—edge indistinctness, PS—posterior shadowing, SC—shape complexity, OV—original value, and NV—normalized value).

Images	Orientation		EI		PS score		SC	
	OV	NV	OV	NV	OV	NV	OV	NV
(a)	0.544	0	27.530	0.572	2.751	1	−1.021	0
(b)	0.670	0.283	27.529	0.572	1.983	0	−1.025	0.05
(c)	0.835	0.654	35.151	1	2.248	0.371	−1.027	0.008
(d)	0.878	0.751	17.329	0	2.321	0.440	−1.077	0.7
(e)	0.989	1	30.194	0.722	2.474	0.693	−1.101	1
(f)	0.772	0.512	23.215	0.33	2.153	0.221	−1.034	0.163

Table 3: The classification result of using classical deep neural network.

Pretrained models	Accuracy	Sensitivity	Specificity	*F*1-score
VGG16	0.84	0.86	0.82	0.86
VGG19	0.81	0.81	0.77	0.82
ResNet50	0.78	0.77	0.77	0.85
InceptionV3	0.78	0.83	0.7	0.81

Table 4: The comparison of classification using different numbers of DLF and CF (CFs—characteristic features; DLFs—deep learning features).

Classification models	Accuracy	Sensitivity	Specificity	*F*1-score
2 DLF + CF + SVM	0.875	0.895	0.85	0.872
4 DLF + CF + SVM	0.925	0.905	0.95	0.927
6 DLF + CF + SVM	0.9	0.9	0.9	0.9
8 DLF + CF + SVM	0.85	0.889	0.8	0.84
10 DLF + CF + SVM	0.875	0.941	0.8	0.84
12 DLF + CF + SVM	0.85	0.889	0.8	0.84
16 DLF + CF + SVM	0.875	0.895	0.85	0.872

Table 5: The comparison of classification methods (CFs—characteristic features, DLFs—deep learning features, and MFs—multi-features).

Models	Accuracy	Precision	Sensitivity	Specificity	*F*1-score
MF + SVM	**0.925**	**0.905**	**0.95**	**0.905**	**0.927**
CF + SVM	0.875	0.895	0.85	0.895	0.827
DLF + SVM	0.8	0.875	0.7	0.875	0.778
[3]	0.901	NA	0.935	0.832	NA
[7]	0.83	NA	NA	0.824	NA
Inception V3 [36]	0.78	NA	0.77	0.78	NA
VGG19 [36]	0.82	0.70	0.70	0.78	NA
[37]	0.8667	NA	0.9245	0.7838	NA

The bold values indicate that the result of the proposed method is better than that of other classification methods.

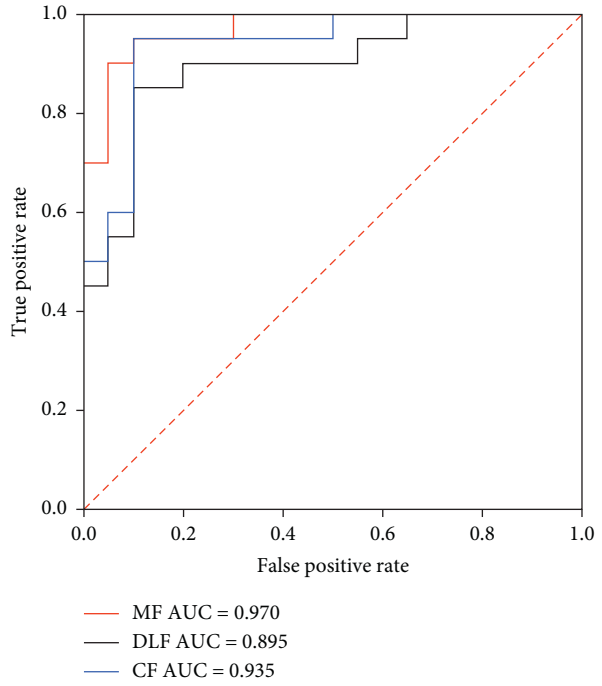


Figure 11: ROC curve and AUC value by using characteristic features (CFs), deep learning features (DLFs), and multi-features (MFs), respectively.

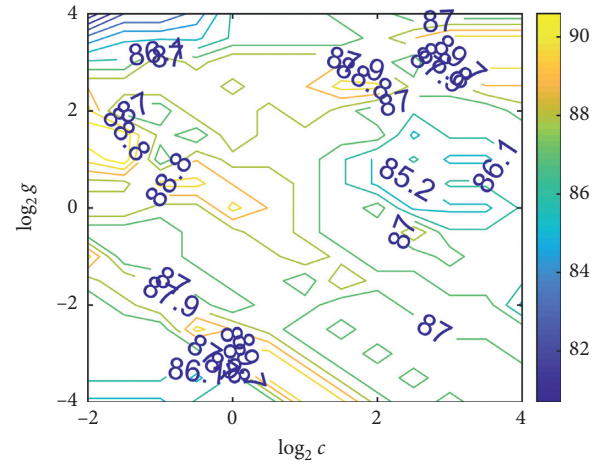


Figure 12: Triple cross-validation of multi-features (MFs).

6. Discussion

Recently, various models relied on deep learning that has been used in BUS image classification. For example, references [36, 38–40] used deep learning methods to extract BUS image features and present classification. Compared with the direct use of traditional deep learning models such as VGG16, VGG19, ResNet50, and Inception V3, they got better results. However, they have the following problems. First of all, because CAD system is mainly used to assist doctors in BI-RADS classification of breast tumors, rather than just relying on CAD system to determine benign and malignant breast tumors, the quantitative methods we proposed can better assist doctors in diagnosis. In addition, according to the clinical experience of doctors, on the basis of segmentation, we quantify the regions with different characteristics of benign and malignant breast tumors, such as posterior shadowing (PS), edge indistinctness (EI), and so on, so as to avoid the influence of irrelevant areas in BUS images on the experimental results and reduce the interference of inherent noise of BUS images.

7. Conclusions

In this work, firstly, with the help of the clinical experience of doctors, four characteristic features (CFs) of BUS images were obtained manually, including orientation, edge indistinctness (EI), characteristics of posterior shadowing region (PS), and shape complexity (SC). Based on the experiments, we compared the CF computed from different BUS images, which showed that the CF designed in this paper can characterize the different properties of benign and malignant BUS images. Meanwhile, the experiment showed that using a single feature to distinguish BUS images was prone to false interpretation. Also, this paper introduced deep learning features (DLFs) to improve the accuracy of classification further. In the experiment of DLF extraction, through comparing the classification results of several classical deep neural networks, it was found that the accuracy, sensitivity, specificity, and $F1$-score of VGG16 are 0.84, 0.86, 0.82, and 0.86, respectively, which were higher than those of other classical neural networks. Therefore, we employed a modified VGG16 as a deep learning feature extractor followed by SVM to classify the fused the CF and DLF. The results showed that the accuracy, sensitivity, specificity, and $F1$-score of this method are 92.5%, 90.5%, 95%, 90.5%, and 92.7% respectively, which were better than other methods. Finally, through the triple cross-validation of multi-features (MFs), the experiment results further indicated that the proposed method can be used to assist doctors to identify benign and malignant BUS images effectively.

Data Availability

Part of the datasets used in this study can be found on the following website: https://www.ultrasoundcases.info/cases/breast-and-axilla/, and the other part is the data provided by the hospital, which is not open to the public because it involves research privacy.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank Dr. Shunmin Qiu from the First Affiliated Hospital of Shantou University, China. She provided some datasets used in this paper and gave us some professional advice. This study was supported by the National Natural Science Foundation of China (grant no. 82071992), Basic and Applied Basic Research Foundation of Guangdong Province (grant no. 2020B1515120061), Guangdong Province University Priority Field (Artificial Intelligence) Project (grant no. 2019KZDZX1013), National Key R&D Program of China (grant no. 2020YFC0122103), Key Project of Guangdong Province Science & Technology Plan (grant no. 2015B020233018), and Research Project of Shantou University, China (grant no. NTF17016).

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] R. Vijayarajeswari, P. Parthasarathy, S. Vivekanandan, and A. A. Basha, "Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform," *Measurement*, vol. 146, pp. 800–805, 2019.

[3] X. Qi, L. Zhang, Y. Chen et al., "Automated diagnosis of breast ultrasonography images using deep neural networks," *Medical Image Analysis*, vol. 52, pp. 185–198, 2019.

[4] J. Y. An, K. M. L. Unsdorfer, and J. C. Weinreb, "BI-RADS, C-RADS, CAD-RADS, LI-RADS, lung-RADS, NI-RADS, O-RADS, PI-RADS, TI-RADS: reporting and data systems," *RadioGraphics*, vol. 39, no. 5, pp. 1435-1436, 2019.

[5] J. Chang, J. Yu, T. Han, H. J. Chang, and E. Park, "A method for classifying medical images using transfer learning a pilot study on histopathology of breast cancer," in *Proceedings of the 2017 IEEE 19th International Conference on E-Health Networking, Applications and Services (Healthcom)*, pp. 1–4, Dalian, China, October 2017.

[6] C. Dromain, B. Boyer, R. Ferre, S. Canale, S. Delaloge, and C. Balleyguier, "Computed-aided diagnosis (CAD) in the detection of breast cancer," *European Journal of Radiology*, vol. 82, no. 3, pp. 417–423, 2013.

[7] M. Byra, H. Piotrzkowska-Wróblewska, K. Dobruch-Sobczak, and A. Nowicki, "Combining Nakagami imaging and convolutional neural network for breast lesion classification," in *Proceedings of the 2017 IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, Washington, DC, USA, September 2017.

[8] H. A. Nugroho, M. Sahar, I. Ardiyanto, R. Indrastuti, and L. Choridah, "Classification of breast ultrasound images based on posterior feature," in *Proceedings of the 2016 International Conference on Information Technology Systems and Innovation (ICITSI)*, pp. 1–4, Bandung, Indonesia, October 2016.

[9] Y. Liu, K. Shou, J. Li et al., "Ductal carcinoma in situ of the breast: perspectives on tumor subtype and treatment," *BioMed Research International*, vol. 2020, Article ID 7251431, 9 pages, 2020.

[10] W. G. McCluggage, "Endometriosis-related pathology: a discussion of selected uncommon benign, premalignant and malignant lesions," *Histopathology*, vol. 76, no. 1, pp. 76–92, 2020.

[11] S. Shaikh and A. Rasheed, "Predicting molecular subtypes of breast cancer with mammography and ultrasound findings: introduction of sono-mammometry score," *Radiology Research and Practice*, vol. 2021, Article ID 6691958, 12 pages, 2021.

[12] X. Liang, J. Yu, J. Liao, and Z. Chen, "Convolutional neural network for breast and thyroid nodules diagnosis in ultrasound imaging," *BioMed Research International*, vol. 2020, Article ID 1763803, 9 pages, 2020.

[13] R. I. R. Thanaraj, B. Anand, J. A. Rahul, and V. Rajinikanth, "Appraisal of breast ultrasound image using Shannon's thresholding and level-set segmentation," *Progress in Computing, Analytics and Networking*, Springer, Berlin, Germany, 2020.

[14] M. Caroccia, A. Chambolle, and D. Slepčev, "Mumford-Shah functionals on graphs and their asymptotics," *Nonlinearity*, vol. 33, no. 8, pp. 3846–3888, 2020.

[15] T. A. Ngo, Z. Lu, and G. Carneiro, "Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance," *Medical Image Analysis*, vol. 35, pp. 159–171, 2017.

[16] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, vol. 19, no. 12, pp. 3243–3254, 2010.

[17] Y. Li, J. Zhang, M. Chen, H. Lei, G. Luo, and Y. Huang, "Shape based local affine invariant texture characteristics for fabric image retrieval," *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15433–15453, 2019.

[18] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, pp. 179–187, 1962.

[19] Z. Omiotek, "Fractal analysis of the grey and binary images in diagnosis of Hashimoto's thyroiditis," *Biocybernetics and Biomedical Engineering*, vol. 37, no. 4, pp. 655–665, 2017.

[20] P. L. Lin, P. W. Huang, P. Y. Huang, and H. C. Hsu, "Alveolar bone-loss area localization in periodontitis radiographs based on threshold segmentation with a hybrid feature fused of intensity and the H-value of fractional Brownian motion model," *Computer Methods and Programs in Biomedicine*, vol. 121, no. 3, pp. 117–126, 2015.

[21] Z. Zhuang, N. Lei, A. N. Joseph Raj, and S. Qiu, "Application of fractal theory and fuzzy enhancement in ultrasound image segmentation," *Medical & Biological Engineering & Computing*, vol. 57, no. 3, pp. 623–632, 2019.

[22] B. Mandelbrot, "How long is the coast of Britain? Statistical self-similarity and fractional dimension," *Science*, vol. 156, no. 3775, pp. 636–638, 1967.

[23] P. G. Okubo and K. Aki, "Fractal geometry in the San Andreas fault system," *Journal of Geophysical Research*, vol. 92, no. B1, pp. 345–355, 1987.

[24] A. Faghih and A. Nourbakhsh, "Implication of surface fractal analysis to evaluate the relative sensitivity of topography to active tectonics, Zagros mountains, Iran," *Journal of Mountain Science*, vol. 12, no. 1, pp. 177–185, 2015.

[25] J. Dong, Y. Ju, F. Gao, and H. Xie, "Estimation of the fractal dimension of Weierstrass-Mandelbrot function based on cuckoo search methods," *Fractals*, vol. 25, no. 6, Article ID 1750065, 2017.

[26] M. S. Taqqu, "Benoît Mandelbrot and fractional Brownian motion," *Statistical Science*, vol. 28, no. 1, pp. 131–134, 2013.

[27] J. Gatheral, T. Jaisson, and M. Rosenbaum, "Volatility is rough," *Quantitative Finance*, vol. 18, no. 6, pp. 933–949, 2018.

[28] E. S. Olivas, J. D. M. Guerrero, M. Martinez-Sober, J. R. Magdalena-Benedito, and L. Serrano, *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques: Algorithms, Methods, and Techniques*, IGI Global, Hershey, PA, USA, 2009.

[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https://arxiv.org/abs/1409.1556.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[31] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, Las Vegas, NV, USA, June 2016.

[32] A. Ortiz, J. Munilla, F. J. Martínez-Murcia, J. M. Górriz, and J. Ramírez, "Empirical functional PCA for 3D image feature extraction through fractal sampling," *International Journal of Neural Systems*, vol. 29, no. 2, Article ID 1850040, 2019.

[33] H. Chen, L. Xu, W. Ai, B. Lin, Q. Feng, and K. Cai, "Kernel functions embedded in support vector machine learning models for rapid water pollution assessment via near-infrared spectroscopy," *Science of the Total Environment*, vol. 714, Article ID 136765, 2020.

[34] V. Mehdipour and M. Memarianfard, "Ground-level O3 sensitivity analysis using support vector machine with radial basis function," *International Journal of Environmental Science and Technology*, vol. 16, no. 6, pp. 2745–2754, 2019.

[35] W. X. Liao, P. He, J. Hao et al., "Automatic identification of breast ultrasound image based on supervised block-based region segmentation algorithm and features combination migration deep learning model," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 4, pp. 984–993, 2019.

[36] N. Antropova, B. Q. Huynh, and M. L. Giger, "A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets," *Medical Physics*, vol. 44, no. 10, pp. 5162–5171, 2017.

[37] M. Wei, Y. Du, X. Wu et al., "A benign and malignant breast tumor classification method via efficiently combining texture and morphological features on ultrasound images," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 5894010, 12 pages, 2020.

[38] S. Han, H.-K. Kang, J.-Y. Jeong et al., "A deep learning framework for supporting the classification of breast lesions in ultrasound images," *Physics in Medicine & Biology*, vol. 62, no. 19, pp. 7714–7728, 2017.

[39] W. K. Moon, Y. W. Lee, H. H. Ke, S. H. Lee, C. S. Huang, and R. F. Chang, "Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks," *Computer Methods and Programs in Biomedicine*, vol. 190, Article ID 105361, 2020.

[40] T. Xiao, L. Liu, K. Li, W. Qin, S. Yu, and Z. Li, "Comparison of transferred deep neural networks in ultrasonic breast masses discrimination," *BioMed Research International*, vol. 2018, Article ID 4605191, 9 pages, 2018.

Hindawi
Computational Intelligence and Neuroscience
Volume 2021, Article ID 9950332, 12 pages
https://doi.org/10.1155/2021/9950332

Research Article

Performance Assessment of Certain Machine Learning Models for Predicting the Major Depressive Disorder among IT Professionals during Pandemic times

P. M. Durai Raj Vincent,[1] **Nivedhitha Mahendran**,[1] **Jamel Nebhen**,[2] **N. Deepa**,[1] **Kathiravan Srinivasan**,[1] **and Yuh-Chung Hu**[3]

[1]*School of Information Technology and Engineering, Vellore Institute of Technology (VIT), Vellore 632 014, Tamil Nadu, India*
[2]*Prince Sattam Bin Abdulaziz University, College of Computer Engineering and Sciences, P.O. Box: 151, Alkharj 11942, Saudi Arabia*
[3]*Department of Mechanical and Electromechanical Engineering, National ILan University, Shenlung Road, Yilan City 26047, Taiwan*

Correspondence should be addressed to Yuh-Chung Hu; ychu@niu.edu.tw

Received 11 March 2021; Revised 26 March 2021; Accepted 9 April 2021; Published 27 April 2021

Academic Editor: Navid Razmjooy

Major depressive disorder (MDD) is the most common mental disorder in the present day as all individuals' lives, irrespective of being employed or unemployed, is going through the depression phase at least once in their lifetime. In simple terms, it is a mood disturbance that can persist for an individual for more than a few weeks to months. In MDD, in most cases, the individuals do not consult a professional, and even if being consulted, the results are not significant as the individuals find it challenging to identify whether they are depressed or not. Depression, most of the time, cooccurs with anxiety and leads to suicide in few cases, among the employees, who are about to handle the pressure at work and home and mostly unnoticing such problems. This is why this work aims to analyze the IT employees who are mostly working with targets. The artificial neural network, which is modeled loosely like the brain, has proved in recent days that it can perform better than most of the classification algorithms. This study has implemented the multilayered neural perceptron and experimented with the backpropagation technique over the data samples collected from IT professionals. This study aims to develop a model that can classify depressed individuals from those who are not depressed effectively with the data collected from them manually and through sensors. The results show that deep-MLP with backpropagation outperforms other machine learning-based models for effective classification.

1. Introduction

In the present day pandemic scenario, where people always complain about stress, pressure, and anxiety, major depressive disorder is commonly seen as a leading mental disorder across the globe. When someone appears to have intense feelings such as sadness and distress for a considerable period, they might have major depressive disorder [1]. It has high impacts on mental and physical activities to the one suffering from it; also, there is a higher risk of suicide [2]. Those who have been suffering from MDD tend to feel uninterested in doing the activities they enjoyed doing once. Also, it affects their moods and behavior and finds difficulty in doing day-to-day activities. Most of those who die by killing themselves are found to have mental disorders that are treatable, mostly only due to depression they are doing so. The suicide rate is said to be around 15% among depressed people [3]. Major depressive disorder is a treatable mental disorder that appears when the individual is too stressed out because of various reasons of one's life including hormonal changes [4].

Major depressive disorder is termed as comorbid [5], that is, a medical condition that tends to occur, and it is a tedious task to identify whether the individual is suffering

from MDD or not. In many cases, the individual who is depressed would be reluctant to consult a professional because of the undertrained workforce and resources; it is difficult to diagnose and continue further treatment for it [6]. Therefore, in this paper, we have tried to apply machine learning techniques to distinguish between depressed and nondepressed individuals, mainly focused on IT professionals. IT professionals are majorly working on targets and failing to meet the target brings a lot of stress, leading to depression. On the other side, it will be difficult to diagnose those depression-affected people as the work style would not allow them to realize the reality. So, it is necessary to bring a system that would allow them to analyze themselves without much human intervention easily. That is why we are trying to bring out a machine learning-based model to help out the needy.

The introduction of smart bands has reduced the burden on the data collection process. Smart wearables can track the key factors that are required for measuring the health status of every individual. This would help handle this scenario as the usage of such devices is growing gradually. This work majorly depends on the data collected from the smart devices and the questionnaire collected from the employees. The data are handled in a way that the depressed person would be identified without much human intervention. The usage of sensors helps determine the status of the employees. This would be helpful among the IT employees as they would self-assess themselves without meeting a doctor.

When it comes to building a machine learning-based model, the necessity is to consider neural network-based approaches, which is also part of the machine learning paradigm. An artificial neural network is vaguely modeled from the human brain consisting of a functional unit called neurons or nodes, just as in the brain. Neurons or nodes are highly interconnected elements, which are the processing elements, and operate parallel [7]. The neural network's behavior is the capability to learn, recall, and implement them on unforeseen data [8]. Due to its ability to improve its performance over every iteration, it is possible to produce the result most accurately. Specifically, in the binary class classifications, this approach would produce the expected results.

Deep multilayer perceptrons are gaining its momentum due to the kind of structure it uses, and the backpropagation algorithm is very popular due to the availability of high-end computing facilities [9]. So, in this depression model, we will be using deep multilayer perceptrons with a backpropagation approach for producing the results more efficiently. Also, the comparison of results will be presented to show the proposed model's superiority over the other approaches [10]. Table 1 depicts the list of abbreviations used in the manuscript.

Among many previous works, it was aimed to study and understand the stress level manually. In the past, not many works targeted work-based stress and its implications on developing major depressive disorder, and among those, no works involved only IT professionals in their study. This work involves a complete model that will do prediction modeling with the most successful questionnaire-based method. Here, the whole process involves the data collection to preprocessing it and building a machine learning-based prediction model that is something new that we are trying to develop in this work. The main root cause for carrying out this research is to analyze its impact on IT employees during the pandemic, especially when they carry out their work from home. In addition to this, the highlights of this work are given as follows.

Table 1: List of abbreviations used in the manuscript along with their full form.

Abbreviations	Full form
MDD	Major depressive disorder
SVM	Support vector machine
HRN	Heart rate normalized
HRV	Heart rate variability
ARIMA	Autoregressive integrated moving average
kNN	K-nearest neighbor
CFS	Correlation-based feature selection
MLP	Multilayer perceptron
ReLU	Rectified linear unit
ROC curve	Receiver operating characteristic curve

The key contributions of this work are summarized as follows:

(a) In this paper, we are proposing a model to detect major depressive disorder among IT professionals

(b) The required data are collected from the questionnaire and sensors, including a pulse rate measuring sensor and a sleep pattern assessing sensor

(c) In this model, the collected data will be checked for anomalies, and the preprocessing steps with a data analysis approach would ensure the quality of the data used in this model

(d) The proposed binary classification model would be expected to produce the maximum possible effective results, which will be more than 98%

(e) Also, this model will be a noble initiation on addressing one of the important issues due to the lifestyle changes, especially among the IT professionals and this is going to be a very rare study that involves only IT employees during pandemic times

(f) This would be also helpful for an individual to assess themselves without any human intervention

The remaining portion of the paper is structured as related work, proposed methodology, results and discussion, and conclusion.

2. Related Works

Stress followed by depression is something very commonly happening problem in recent days. In a study [11], the authors have analyzed the problem of predicting major depressive disorder and generalized anxiety disorder using a novel machine learning pipeline to reanalyze data from an observational study. Another study was conducted to assess the possibility of anxiety and depression in the parents'

offspring having a history of anxiety and depression [12]. The results show that the progeny with parents having anxiety or depression are more likely to suffer from the same than the offspring with parents who do not have anxiety or depression [13]. A comparison was made between the healthy individuals and those who have MDD based on interpersonal and adaptive domains. The findings showed that individuals with MDD performed considerably lower than the health concerns in all the domains such as adaptive functioning, adaptive competence, perceptive competence, functional ability, and interpersonal functioning [14]. An investigation on the major depressive disorder and bipolar disorder was carried out [15], and on the other side, the major depressive disorder-based impact on sleep apnea patients was analyzed [16, 17].

The hospital Anxiety and Depression Scale is used in France to assess the level of stress, and basically, it is a questionnaire that is the most widely used approach to estimate the stress level. This kind of questionnaire approach was used to assess the stress level among the French employees where fourteen parameters were considered, which comprises both to test the anxiety and depression [18]. The participants were majorly involved in industries such as telecommunications, petroleum, and aeronautics. Stress analysis among banking employees was done across the countries from the Middle East to Africa [19]. It is also understood that job status is one of the prime reasons for getting stress. Among various factors, job security played a major role in that and was found in the study. More than four thousand employees from Iran participated in a study based on four questionnaires to analyze the somatic syndrome [20].

Data collected from smartphones are analyzed to predict depression among undergraduate students. The hourly mood throughout a week was analyzed when the participant visits various places. Data were collected through sensors available in the smart bands and smartphones. SVM-based approach was proposed to analyze the signals to analyze the MDD [21]. Inputs through EEG signals were collected for this MDD model, which was successfully deployed with machine learning-based approaches. In another attempt machine learning-based approach was proposed to predict panic disorder [22]. This model was mainly proposed to distinguish panic disorder from other types of anxiety-based disorders clearly. Again, SVM-based approach was used in this model to produce the appropriate results [23]. Ensemble-based classifier was proposed for analyzing the quality of life cycles [24]. Predictive modeling based on machine learning was proposed to analyze depression from health records [11]. How depressed people would return to work after prolonged treatment and its effectiveness was discussed [25]. A study on the impact on occupational-related stress was discussed from the country Ghana [26].

Machine learning plays a vital role in prediction or estimation, and among these, artificial neural networks play an important role in solving real-time problems. A weighted average ensemble model was proposed for handling MDD [27]. In this analysis [28], the authors have analyzed heart rate variability to distinguish between diastolic heart failure and systolic heart failure patients. They have implemented the nearest neighbor and deep multilayered perceptron classifiers in evaluating the performances of classification. The two classifiers were implemented with two measures, such as HRN (heart rate normalized) and HRV (heart rate variability), and the results show that a deep multilayered perceptron performs better than the nearest neighbor with higher accuracy [29].

In this experimental analysis [30], the authors have proposed a model to predict the coal prices by employing a deep multilayered perceptron with three hidden layers and having 3,11,3 neurons in each layer, respectively. The results show that the proposed method performs better than the autoregressive integrated moving average model (ARIMA). In this study [31], the authors have proposed an ensemble-based deep multilayered perceptron effective in analyzing stock market trends and predicting when to buy and sell the stocks. A deep-MLP-based approach was proposed [32] to handle the issue on roller bearing, and it was successfully addressed, and in another case [33], it was able to distinguish between COVID and non-COVID patients successfully.

Feature selection is used to play an important role in choosing the right parameters for effective model building. Pearson correlation-based feature selection approach was proposed for proper document classification problems [34]. A classical feature selection approach was proposed for blood cell disease recognition [35], and a robust feature selection approach was proposed for the application based on welding defects detection [36]. The importance of optimization was illustrated [37], and a hybrid approach of feature selection was proposed for the application related to agriculture [38].

Section 3 presents the proposed model, and the algorithm and the steps are presented in this section.

3. Proposed Model

As part of the data collection process, we have chosen 1032 IT professionals whose average age is 38. Followed by this, a questionnaire has been framed using the Hamilton rating scale for depression [39]. The data have been collected using that questionnaire form, and the questionnaire consists of 22 attributes, and using the range given by the Hamilton scale, the depression levels are calculated. Along with the questionnaire, the participants' data were collected from the smart band they were wearing. Heart rate monitoring is an important parameter that was considered during the data collection process. The participants were wearing the smart bands for about two weeks for the data collection process. Among the participants, about 60% are men and the remaining are women.

The data have been cleaned and reduced to relevant features using preprocessing and feature selection techniques. Then, deep multilayered neural network has been implemented to handle the collected data. The methodology includes three processes: preprocessing, feature selection, and applying artificial neural network-based deep multilayered perceptron with the backpropagation approach for the prediction process.

Preprocessing of data is essential as the data collected has its anomalies in missing information to wrong entries. Handling these is very important since the sample size considered in this work is optimal, and every information present is vital to building an effective model. So, in handling the missing information in the data collected from the IT professionals, kNN imputation techniques [40, 41] have been employed in this work. It consists of finding the closest K records and calculating the weights based on the distance computed using one of the distance calculation methods such as Manhattan, Euclidean, and Minkowski. Also, the input data are thoroughly analyzed, and all the categorical values are converted to numerical values with the help of a data label encoder. Since for handling data using machine learning approaches, it is desirable to have this conversion for effective data processing.

Usually, the source dataset consists of any number of attributes that may or may not be relevant to the classification process [42, 43]. The irrelevant attributes which depend on other attributes reduce the prediction accuracy. To overcome this and also to reduce the dimension of the feature, a feature selection technique must be implemented. A correlation-based feature selection technique is used in this work, which helps to find the features' subsets. CFS considers different attributes, and the correlation below the given threshold will be considered as part of this approach.

This study's data are collected from IT professionals; it consists of 22 features and 1032 samples. The data are cleaned for missing values with the help of K-nearest neighbor approach. Handling missing values is critical since it will affect the prediction power of the model. Also, features that are not related or dependent features will impact accuracy. Therefore, it is necessary to select only the required features to improve the developed model's accuracy further. The details of 22 features considered initially for this study are listed in Table 2.

We have applied the correlation-based feature selection approach to shortlist the required number of features. The initial number of features considered are 22, which is then reduced to 12. The chosen feature includes sleeping pattern, mood during work and other time, interest towards eating, weight, happiness quotient, level of concentration at work, and heart rate during the work and nonwork time. The duration of sleep is monitored with sensors and through questionnaires, whereas the heart rate inputs are completely dependent on the sensors available in the smart band and sleep monitoring. Even though these inputs are recorded, further inputs like in between wakeup details are recorded manually. The algorithm used for feature selection is presented in Algorithm 1 and the details of chosen features are listed in Table 3.

This approach helped remove some of the redundant features, which in turn helps to predict the outcome much more accurately. For classification of the individuals with a deep multilayered perceptron, this is employed in two phases. A deep multilayered perceptron consists of one or more hidden layers: the layers consisting of neurons between input and output. In the first phase, perceptron is trained without backpropagation, and in the second phase, the perceptron is trained with backpropagation. The functional unit of the neural network is called neurons or nodes. A perceptron is formed by combining several neurons into a layer. A perceptron has four components: input, bias, weights, activation function, and output [38]. Every neuron in the network is connected through a connection link, and each connection link consists of weight. The weight has information about the input signals. Bias indirectly impacts the output and helps in calculating the net input. There are two types of bias: one is a positive bias, which increases the network's net input, and the other one is a negative bias, which decreases the network's net input. Then, the activation function is applied on the net input to calculate the neural network's output, called as step function [44].

In phase one, the neural network is implemented without any backpropagation. The second phase is implemented with backpropagation. Backpropagation is also called a backward propagation of errors. The error which has been calculated in the output layer is again propagated

Table 2: Features of data samples.

Sl. No	Features
1	Age
2	Gender
3	Falling asleep during work
4	Reduction in food intake
5	Increase in food intake
6	Untime wakeup
7	Ability to enjoy with the team
8	Bad mood frequency
9	Irritated towards work
10	Uneasiness at work
11	Any relation between mood and time at work
12	Weight decreased recently
13	Weight increased recently
14	Sleeping more than the usual
15	Anxiousness during work
16	Concentration/decision-making during work
17	View of my future at my work place
18	Suicidal thoughts
19	Symptoms in body language
20	Alarming symptoms
21	Heart rate recorded average from sensors
22	Sleep pattern recorded average from sensors

Table 3: Features for the MDD model.

Sl. No	Chosen features
1	Decreased weight within last two weeks
2	Suicidal thoughts
3	Sleeping more than the usual
4	Alarming symptoms
5	Reduction in food intake
6	Anxiousness during work
7	Concentration/decision-making during work
8	Irritated towards work
9	Uneasiness at work
11	Heart rate recorded average from sensors
12	Sleep pattern recorded average from sensors

Input: X – training inputs, N – the list of inputs, D – number of features, F – feature finder
n – features 1 to n
Procedure
$S = 0$
$n = 1$
while $|S_n| < D$ do
if $|S_n| < D - 1$ then
S_n= CBF (S_{n-1}, X, F)
else
select the highest level feature S' to S_{n-1}
exit
$n = n+1$
exit while
Output:
S – selected feature set.

Algorithm 1

backward and distributed to all the neurons in the network so that new weights would be updated. The following equations show the calculation of net input, activation function, and the perceptron weight adjustments after applying backpropagation. Figure 1 shows the architectural diagram of the proposed deep-MLP model. Figure 2 represents the process flow of the deep-MLP network.

The net input or preactivation function is

$$\mathrm{PA} = \sum_{k=1}^{n} Wt_k I_k + b, \tag{1}$$

where PA is the preactivation function, Wt is the weight associated with the connection link, I is the inputs (I_1, I_2, I_3, . . ., I_n), and b is the bias. The activation function is

$$A(I) = \begin{cases} 1, & I \geq 0, \\ 0, & I < 0. \end{cases} \tag{2}$$

If the input from the neuron $I \geq 0$, then the output is 1, and if the input $I < 0$, then the output is 0. It can be written as

$$Z = A \sum_{k=1}^{n} Wt_k I_k + b. \tag{3}$$

To the perceptron weight adjustment, the weights are updated after applying the backpropagation algorithm. The new weight after adjusting is given by

$$\Delta Wt = L \times P \times I, \tag{4}$$

where ΔWt is the updated weight value, L is the learning rate, P is the predicted output, and I is the input data.

The process involved in the deep multilayer perceptron-based approach, which is to be used in this model-building exercise, is illustrated in the algorithm part. The whole process involved in this is the updation of weights after every iteration, which will tune the network to produce relevant results. Activation functions like the ReLU approach will help this algorithm normalize the output within the range (Algorithm 2).

In the supervised learning-based problem considered, the model is trained each time to calculate the error value deviation. According to the error value, the backpropagation approach will update each weight value, taking hundreds of iterations to prepare the network as part of the training process. Data preparation for building a model based on machine learning approaches requires the holdout method for model building and validation of results. The collected data after undergoing the preprocessing steps will be divided into 80 : 20 ratio for training and testing. Section 4 presents the results and discussion part in detail.

4. Results and Discussion

Predicting depression is a binary classification issue that classifies whether the person is depressed or not (0 or 1). For classifying the patients, we have used the deep multilayered perceptron. Also, to insist on the importance of the backpropagation, we have implemented deep-MLP without backpropagation in phase 1 and with backpropagation in the second phase and tabulated the results in Table 4. The phase 1 output of deep-MLP, without backpropagation, is shown in Figure 3, and the phase 2 deep-MLP, after applying the backpropagation, is shown in Figure 4. We can see the 12 important features given as the input for further processing in the input layer. The hyperparameters we chose in our deep-MLP are hidden layers, number of nodes per layer, activation function, and the learning rate. The values of the hyperparameters are determined using the Bayesian optimization technique. The Bayesian optimization technique iterates the data through 5 folds, and the average of each parameter is considered the final value for the model. Thus, our chosen hyperparameters' final values are hidden layers are 4 and the learning rate is 0.02, and number of nodes per layer 57. We used the ReLU as the activation function and the cross-entropy cost function to check the model's error rate.

In backpropagation, the gradient error function is estimated based on the weights. The total loss will then be

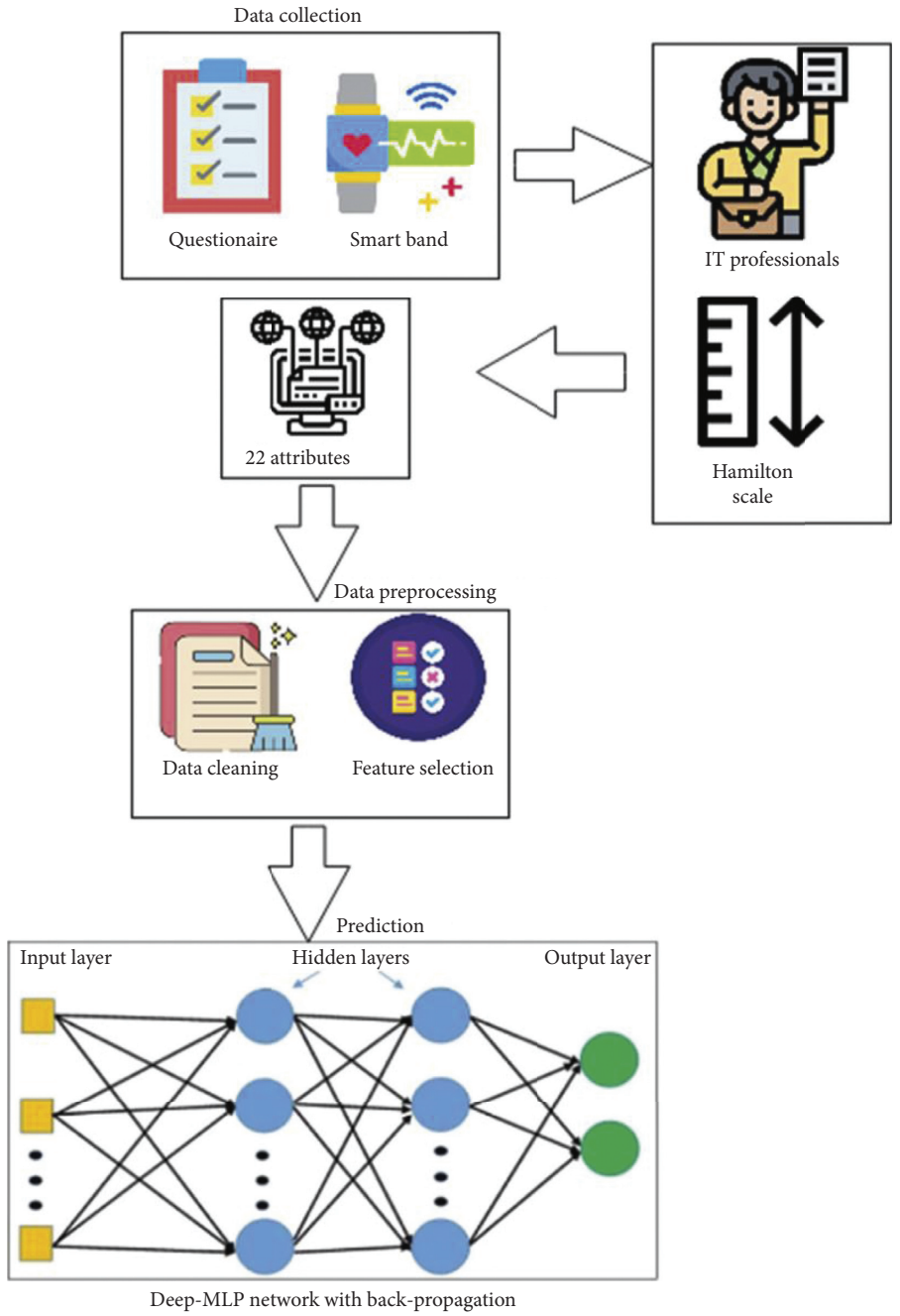


Figure 1: Architectural diagram of the proposed deep-MLP model.

propagated backward, and the weights will be updated. From the results, the importance of backpropagation is quite visible. The comparison of results between the deep-MLP with and without backpropagation is shown in Figures 5 and 6. The line in red color is the actual values, and the line in blue color is the predicted values. The figure clearly shows a considerable difference between the actual and predicted values. Figure 5 shows the actual vs. predicted value before applying the backpropagation algorithm, and Figure 6 shows the results after applying the backpropagation. The predicted values in the blue color overlap the actual values, which are red lines, which imply that the algorithm predicted almost all the outcomes correctly. The plot consists of only a few records out of 1032 records in the dataset.

We have first evaluated the model with the help of three performance metrics: error rate, steps, and accuracy. All three metrics show that the deep-MLP with backpropagation outperforms the deep-MLP without backpropagation. Figure 7 shows the number of steps involved to classify the depressed and not depressed before backpropagation, it was 16,395, and then it is 2253. Moreover, Figure 8 shows the error rate without backpropagation; it is noted that the error rate is 30.43 before applying backpropagation, and the error rate after backpropagation has been hugely decreased to

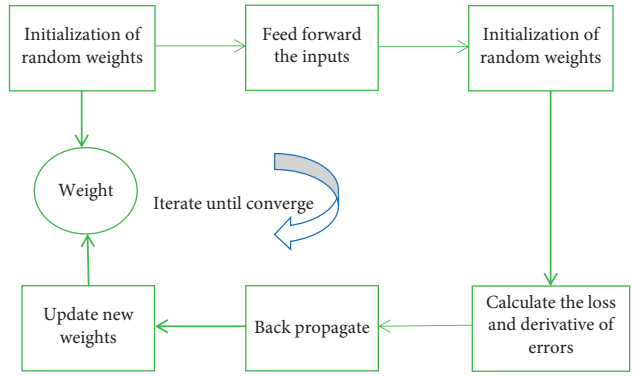


Figure 2: Process flow of deep-MLP network perceptron.

Input. The selected features from the correlation-based feature selection
Initialize the values of weights, biases, and the learning rate (α)
Do
For every input do Feedforward:
Process the inputs one by one: $a_i = x_i$, where *i value lies between* 1 *an d n*, 'n' stands for the total number of samples
Output: $b_i = \text{bias} + \sum_{i=1}^{n} a_i.w_i$
output calculation with the activation function:

$$A = \begin{pmatrix} 1 & \text{if } b_i > \theta \\ 0 & \text{if } -\theta \leq b_i \leq \theta \\ -1 & \text{if } b_i < -\theta \end{pmatrix}$$

Adjust weight and bias:
if $b \neq t$, if 'b' not equals to target 't' then update the weights
w_i new = $w_i$old + $\propto t_{ai}$
bias new = bias old + $\propto t$
if $b = t$, Stopping condition
Gradient descent calculation with respect to each error due to the selected weights w_i
$E(w) = 1/2\sum_i \sum_{j \in y} (t_i^j - b_i^j)^2$

Where t_i^j, b_i^j are the initial target value and the obtained output
Repeat the similar calculation and updation of new weights continues at the hidden layer
Output. Classification results and error rate.

Algorithm 2

Table 4: Performance analysis of Deep-MLP.

Model	Accuracy	Error rate	Steps
Deep-MLP: before backpropagation	0.928	30.43	16,395
Deep-MLP: after backpropagation	0.988	4.52	2253

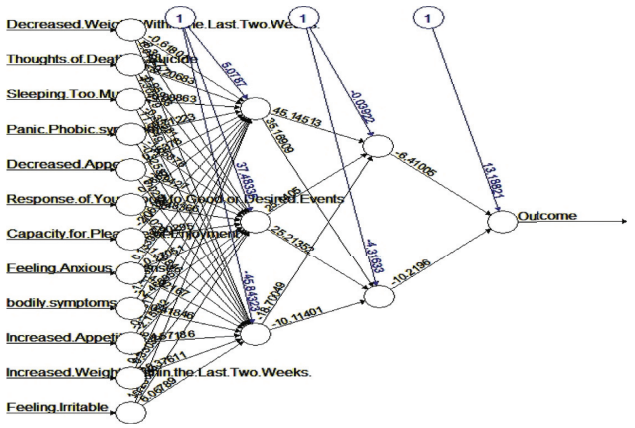


Figure 3: Deep-MLP network before BP.

4.52. Correspondingly, Figure 9 shows the accuracy of the deep multilayered perceptron before backpropagation learning, it was 0.92865, and after implementing backpropagation, it got higher to 0.987837, which is a considerable increase in the accuracy of classification. In all the three evaluation metrics used, the deep-MLP model with backpropagation outperforms the other. With this, we can understand the importance of backpropagation in the deep-MLP model. Table 3 shows the comparison between the two approaches.

These outcomes show that the necessity to have the backpropagation based approach to consider over the other. Furthermore, the proposed approach is considered for further evaluation with the other well-known approaches discussed in the literature. Here, we have considered the

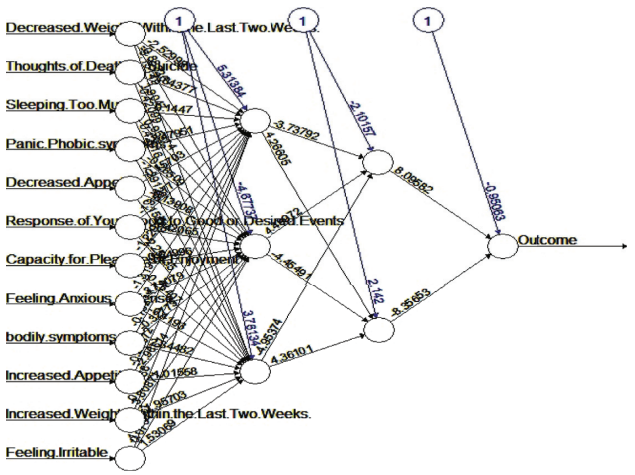


Figure 4: Deep-MLP network after BP.

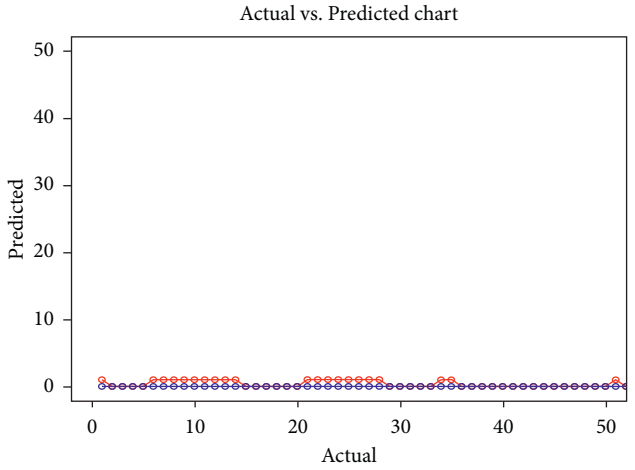


Figure 5: Actual vs. predicted before BP.

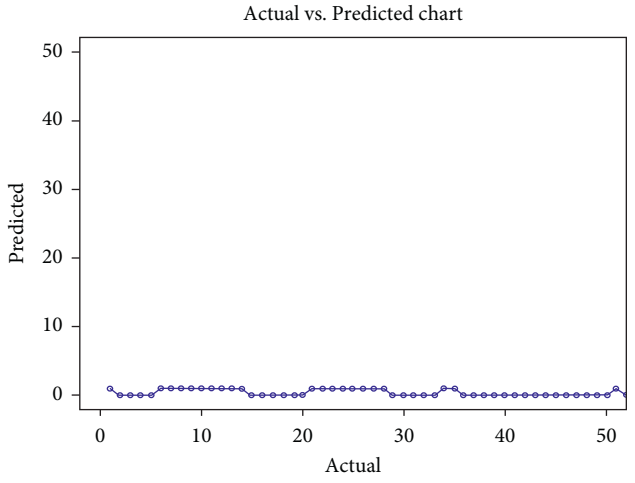


Figure 6: Actual vs. predicted after BP.

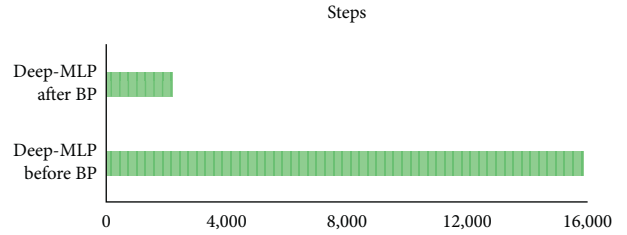


Figure 7: Steps before and after BP.

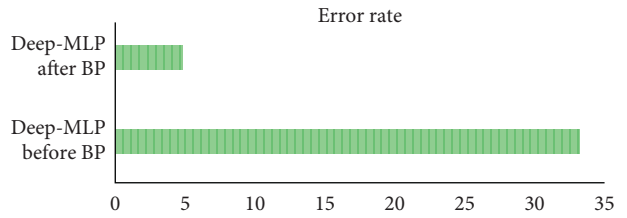


Figure 8: Error rate before and after BP.

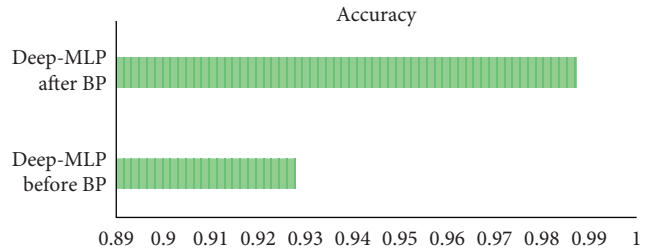


Figure 9: Accuracy before and after BP.

Table 5: Performance measurement analysis.

Performance measurement metric	Definition	Formula
Accuracy	Accuracy defines the capability of the proposed model to correctly classify the output	True negative + true positive/total samples
Sensitivity	Sensitivity is about verifying the identification of positive outcomes from the proposed model	True positive/true positive + false negative
Specificity	Sensitivity is about verifying the identification of negative outcomes from the proposed model	True negative/true negative + false positive
F-measure	F-measure is helpful in identifying the outcomes of precision with the recall of the proposed model	2 × (precision × recall)/ (recall + precision)

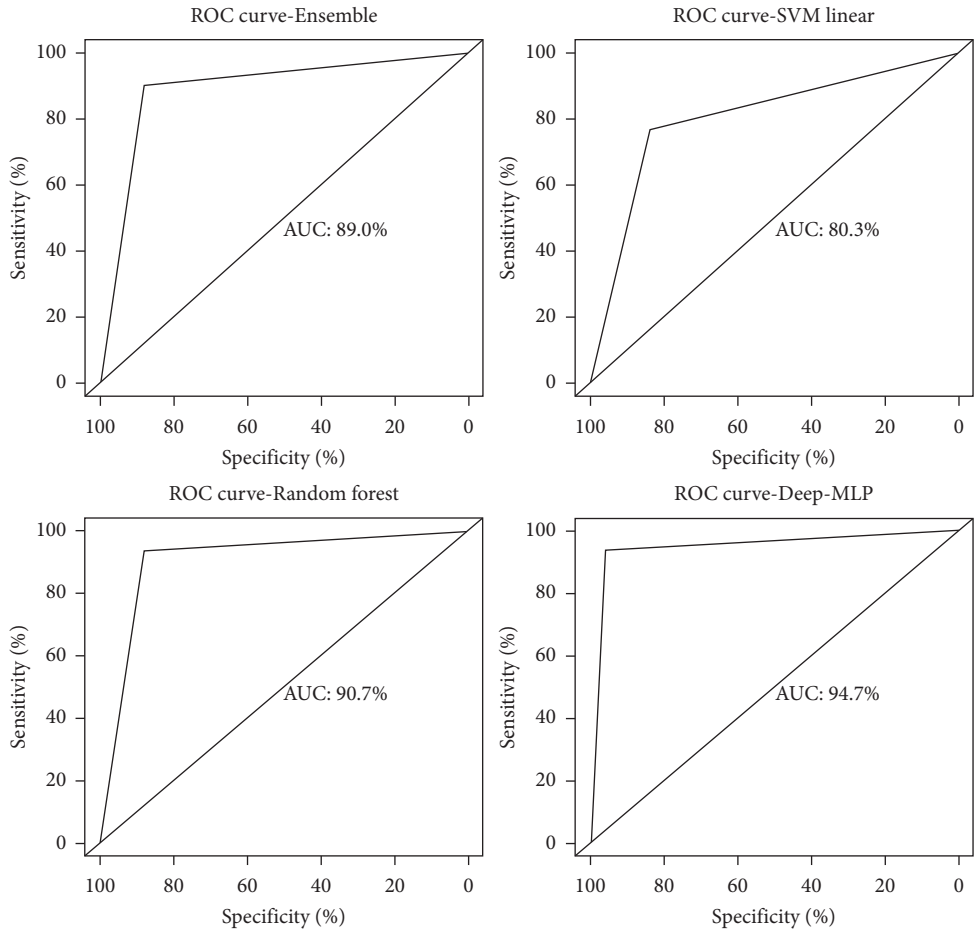


Figure 10: ROC curve of the different approaches including the proposed deep-MLP network.

performance metrics such as accuracy, sensitivity, specificity, and *F*-measure to compare with the other approaches. The definition of the metrics is given in Table 5.

There are three other approaches considered in this work for the purpose of comparison. These approaches are particularly considered for this comparison as these were proposed in other approaches that were considered during literature. The ROC curve shows the specificity measurement of each approach. Figure 10 shows that the proposed model with deep-MLP has better accuracy compared to the other approaches.

The different approaches considered for the comparison of results with the proposed model are the support vector machine-based approach and the random forest-based approach. With these, an ensemble model is also considered for the comparison of results. The results are tabulated in Table 6. The results show that the deep-MLP approach considered for this model performs better than the other proven approaches discussed in the literature.

From accuracy to *F*-measure, the results show that the outcomes are better for the proposed model, requiring the desired results for the proper classification. The detailed graphs are presented in Figures 11 to 14 for a better interpretation of the results.

Table 6: Performance analysis of deep-MLP with other machine learning models.

Models	Accuracy	Sensitivity	Specificity	*F*-measure
SVM (linear)	0.862	0.824	0.871	0.863
Random forest	0.849	0.807	0.878	0.881
Ensemble (SVM, random forest)	0.905	0.877	0.896	0.907
Deep-MLP	0.988	0.924	0.947	0.965

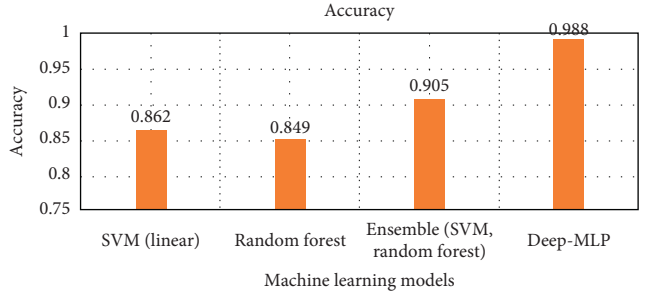


Figure 11: Comparison between deep-MLP and other models in terms of accuracy.

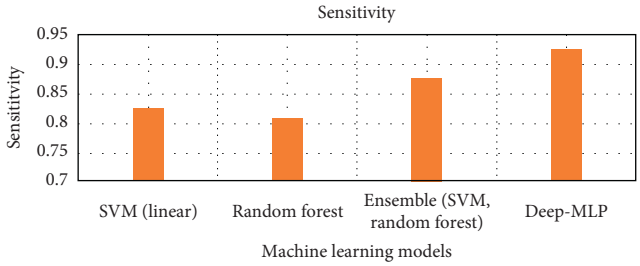


Figure 12: Comparison between deep-MLP and other models in terms of sensitivity.

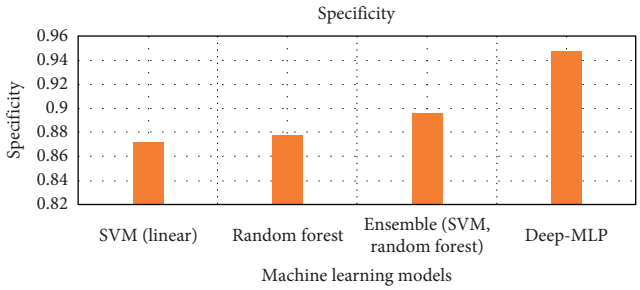


Figure 13: Comparison between deep-MLP and other models in terms of specificity.

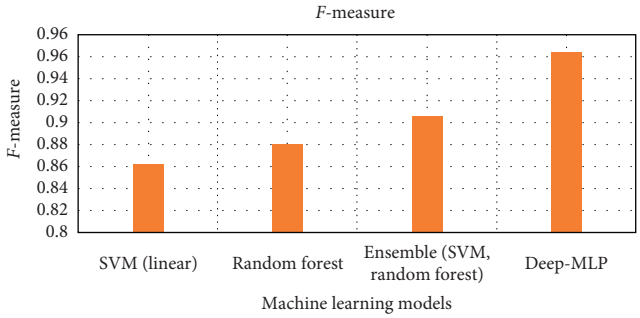


Figure 14: Comparison between deep-MLP and other models in terms of *F*-measure.

With the results obtained, the proposed model for assessing the IT employees has shown the better results, and the model can be considered for the real-time purposes. The considered sample size is sufficient for this model building exercises which also produces the results in a considerable manner.

5. Conclusions

MDD is one form of depression, which is comorbid. The diagnosis of any depression is tough for clinicians because of its subjectiveness. There are many methods to determine whether the individual is suffering from depression or not, using real-time brain images to questionnaire-based approaches. Recording brain images is expensive and a time-consuming one, which is also not a viable option for many. Thus, in this work, we have tried to build a classification-based model that can classify the IT employees into depressed and nondepressed categories. We have also compared the results of deep multilayered neural network perceptrons with backpropagation and without backpropagation in classifying the depressed and nondepressed individuals from the collected samples of IT professionals. The data are collected from IT professionals through a standard questionnaire and also with the sensors connected with them through smart bands during the pandemic times. The accuracy of the model is improved when developed with the necessary features. Therefore, a CFS technique on the cleaned data reduces the feature dimensions. The first phase was without backpropagation, just the feedforward deep-MLP, and in the second phase after feedforward, backpropagation is applied, and the weights are adjusted accordingly, and the process is iterated until convergence is reached. From the experimental results, we have also compared with the other proven approaches such as SVM and other ensemble-based models in terms of accuracy and other performance-oriented measures. This model's presented results are unique, which handled only the samples collected from the IT employees and did the appropriate classifications. In the future, we would experiment with this model that can be enhanced for the other professions, especially among the underprivileged who face many challenges comparatively with others.

Data Availability

The data used to support the findings of this study have not been made available so as to ensure the privacy and anonymity of the persons involved.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was funded by the Ministry of Science and Technology, Taiwan, (grant numbers MOST 109-2221-E-197-011 and MOST 109-2622-E-197-007-CC2). The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

[1] A. Karasz, C. Dowrick, R. Byng et al., "What we talk about when we talk about depression: doctor-patient conversations and treatment decision outcomes," *British Journal of General Practice*, vol. 62, no. 594, pp. e55–e63, 2012.

[2] G. I. Papakostas, T. Petersen, Y. Mahal, D. Mischoulon, A. A. Nierenberg, and M. Fava, "Quality of life assessments in major depressive disorder: a review of the literature," *General Hospital Psychiatry*, vol. 26, no. 1, pp. 13–17, 2004.

[3] M. D. Feldman, P. Franks, P. R. Duberstein, S. Vannoy, R. Epstein, and R. L. Kravitz, "Let's not talk about it: suicide inquiry in primary care," *The Annals of Family Medicine*, vol. 5, no. 5, pp. 412–418, 2007.

[4] J. Yang, Y. Lee, and U.-G. Kang, "Comparison of prediction models for coronary heart diseases in depression patients," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 10, no. 3, pp. 257–268, 2015.

[5] N. Kabra and A. Nadkarni, "Prevalence of depression and anxiety in irritable bowel syndrome: a clinic based study from India," *Indian Journal of Psychiatry*, vol. 55, no. 1, p. 77, 2013.

[6] N. Mahendran and D. R. Vincent, "Effective classification of major depressive disorder patients using machine learning techniques," *Recent Patents on Computer Science*, vol. 12, no. 1, pp. 41–48, 2019.

[7] K. Srinivasan, N. Mahendran, D. R. Vincent, C.-Y. Chang, and S. Syed-Abdul, "Realizing an integrated multistage support vector machine model for augmented recognition of unipolar depression," *Electronics*, vol. 9, no. 4, p. 647, 2020.

[8] G. Richter-Levin and L. Xu, "How could stress lead to major depressive disorder?" *IBRO Reports*, vol. 4, pp. 38–43, 2018.

[9] P. Lawrence, K. Murayama, and C. Creswell, "Systematic Review and Meta-Analysis: anxiety and depressive disorders in offspring of parents with anxiety disorders: a meta-analysis," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 58, no. 1, pp. 46–60, 2018.

[10] M. Milanovic, K. Holshausen, R. Milev, and C. R. Bowie, "Functional competence in major depressive disorder: objective performance and subjective perceptions," *Journal of Affective Disorders*, vol. 234, pp. 1–7, 2018.

[11] M. D. Nemesure, M. V. Heinz, R. Huang et al., "Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence," *Science Reports*, vol. 11, p. 1980, 2021.

[12] C. Bocéréan and E. Dupret, "A validation study of the Hospital Anxiety and Depression Scale (HADS) in a large sample of French employees," *BMC Psychiatry*, vol. 14, no. 1, p. 354, 2014.

[13] D. Volker, M. C. Zijlstra-Vlasveld, E. P. M. Brouwers, W. A. Homans, W. H. M. Emons, and C. M. van der Feltz-Cornelis, "Validation of the patient health questionnaire-9 for major depressive disorder in the occupational health setting," *Journal of Occupational Rehabilitation*, vol. 26, no. 2, pp. 237–244, 2016.

[14] F. Marcatto, L. Colautti, F. Larese Filon et al., "Work-related stress risk factors and health outcomes in public sector employees," *Safety Science*, vol. 89, pp. 274–278, 2016.

[15] C. W. Musket, N. S. Hansen, K. M. Welker, K. E. Gilbert, and J. Gruber, "A pilot investigation of emotional regulation difficulties and mindfulness-based strategies in manic and remitted bipolar I disorder and major depressive disorder," *International Journal of Bipolar Disorders*, vol. 9, no. 1, pp. 1–8, 2021.

[16] R. Geng and X. Huang, "Identification of major depressive disorder disease-related genes and functional pathways based on system dynamic changes of network connectivity," *BMC Medical Genomics*, vol. 14, no. 1, pp. 1–11, 2021.

[17] C. Chia-Min, K. Chia-Yu, W. Meng-Ni, H. Jen-Yu, H. Chung-Yao, and T. Ming-Ju, "Increased risk of major depressive disorder in sleep apnea patients in Taiwan," *Scientific Reports (Nature Publisher Group)*, vol. 11, no. 1, 2021.

[18] M. Ota, S. Takeda, S. Pu et al., "The relationship between cognitive distortion, depressive symptoms, and social adaptation: a survey in Japan," *Journal of Affective Disorders*, vol. 265, pp. 453–459, 2020.

[19] S. Mostafaei, K. Kabir, A. Kazemnejad et al., "Explanation of somatic symptoms by mental health and personality traits: application of Bayesian regularized quantile regression in a large population study," *BMC Psychiatry*, vol. 19, no. 1, p. 207, 2019.

[20] K. Schwarzkopf, D. Straus, H. Porschke, H. Znoj, and R. von Känel, "Is it all about depression? Determinants of burnout among employees referred for inpatient treatment of job-related stress disorders," *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, vol. 65, no. 2, pp. 183–197, 2019.

[21] N. C. Jacobson and Y. J. Chung, "Passive sensing of prediction of moment-to-moment depressed mood among undergraduates with clinical levels of depression sample using smartphones," *Sensors*, vol. 20, no. 12, p. 3572, 2020.

[22] S.-C. Liao, C.-T. Wu, H.-C. Huang, W.-T. Cheng, and Y.-H. Liu, "Major depression detection from EEG signals using kernel eigen-filter-bank common spatial patterns," *Sensors*, vol. 17, no. 6, p. 1385, 2017.

[23] K. S. Na, S. E. Cho, and S. J. Cho, "Machine learning-based discrimination of panic disorder from other anxiety disorders," *Journal of Affective Disorders*, vol. 278, pp. 1–4, 2020.

[24] X. Tao, O. Chi, P. J. Delaney, L. Li, and J. Huang, "Detecting depression using an ensemble classifier based on Quality of Life scales," *Brain Informatics*, vol. 8, no. 1, pp. 1–15, 2021.

[25] K. Nieuwenhuijsen, J. H. Verbeek, A. Neumeyer-Gromen, A. C. Verhoeven, U. Bültmann, and B. Faber, "Interventions to improve return to work in depressed people," *Cochrane Database of Systematic Reviews*, vol. 10, 2020.

[26] M. Dzodzomenyo, E. E. Kploanyi, and D. Dwomoh, "The effect of occupational stress on depression and insomnia: a cross-sectional study among employees in a ghanaian telecommunication company," *BMC Public Health*, vol. 20, 2019.

[27] N. Mahendran, D. R. Vincent, K. Srinivasan et al., "Sensor-assisted weighted average ensemble model for detecting major depressive disorder," *Sensors*, vol. 19, no. 22, p. 4822, 2019.

[28] E. Lin, P.-H. Kuo, Y.-L. Liu, Y. W.-Y. Yu, A. C. Yang, and S.-J. Tsai, "Prediction of antidepressant treatment response and remission using an ensemble machine learning framework," *Pharmaceuticals*, vol. 13, no. 10, p. 305, 2020.

[29] Y. Sler, "Discrimination of systolic and diastolic dysfunctions using multi-layer perceptron in heart rate variability analysis," *Computers in Biology and Medicine*, vol. 76, pp. 113–119, 2016.

[30] X. Fan, L. Wang, and S. Li, "Predicting chaotic coal prices using a multi-layer perceptron network model," *Resources Policy*, vol. 50, pp. 86–92, 2016.

[31] S. Mabu, M. Obayashi, and T. Kuremoto, "Ensemble learning of rule-based evolutionary algorithm using multi-layer perceptron for supporting decisions in stock trading problems," *Applied Soft Computing*, vol. 36, pp. 357–367, 2015.

[32] D. Zhang, E. Stewart, J. Ye, M. Entezami, and C. Roberts, "Roller bearing degradation assessment based on a deep MLP convolution neural network considering outlier regions," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 2996–3004, 2019.

[33] M. M. Ahsan, T. Alam, T. Trafalis, and P. Huebner, "Deep MLP-CNN model using mixed-data to distinguish between COVID-19 and Non-COVID-19 patients," *Symmetry*, vol. 12, no. 9, p. 1526, 2020.

[34] I. M. Nasir, M. A. Khan, M. Yasmin et al., "Pearson correlation-based feature selection for document classification using balanced training," *Sensors*, vol. 20, no. 23, p. 6793, 2020.

[35] M. A. Khan, M. Qasim, H. M. J. Lodhi et al., "Automated design for recognition of blood cells diseases from hematopathology using classical features selection and ELM," *Microscopy Research and Technique*, vol. 84, no. 2, pp. 202–216, 2021.

[36] K. Srinivasan, N. Deepa, P. M. Durai Raj Vincent, and S. Senthil Kumaran, "Realizing the resolution enhancement of tube-to-tube plate friction welding microstructure images via hybrid sparsity model for improved weld interface defects

diagnosis," *Journal of Internet Technology*, vol. 21, no. 1, pp. 61–72, 2020.

[37] K. Srinivasan, A. Kumar Cherukuri, D. R. Vincent, and A. Garg, "An efficient implementation of artificial neural networks with K-fold cross-validation for process optimization," *Journal of Internet Technology*, vol. 20, no. 4, pp. 1213–1225, 2019.

[38] D. Elavarasan, D. R. Vincent P M, K. Srinivasan, and C.-Y. Chang, "A hybrid CFS filter and RF-RFE wrapper-based feature extraction for enhanced agricultural crop yield prediction modeling," *Agriculture*, vol. 10, no. 9, p. 400, 2020.

[39] M. Hamilton, "The Hamilton rating scale for depression," in *Assessment of Depression*, pp. 143–152, Springer, Berlin, Heidelberg, 1986.

[40] S. Zhang, "Nearest neighbor selection for iteratively kNN imputation," *Journal of Systems and Software*, vol. 85, no. 11, pp. 2541–2552, 2012.

[41] S. Zhang, X. Li, M. Zong, X. Zhu, and D. Cheng, "Learning k for knn classification," *ACM Transactions on Intelligent Systems and Technology*, vol. 8, no. 3, pp. 1–19, 2017.

[42] N. Calik, M. A. Belen, and P. Mahouti, "Deep learning based modified MLP model for precise scattering parameter prediction of capacitive feed antenna," *International Journal of Numerical Modelling*, vol. 33, Article ID e2682, 2020.

[43] S. J. Yim, L. M. Lui, Y. Lee et al., "The utility of smartphone-based, ecological momentary assessment for depressive symptoms," *Journal of Affective Disorders*, vol. 274, 2020.

[44] Y. Meng, W. Speier, C. Shufelt et al., "A machine learning approach to classifying self-reported health status in a cohort of patients with heart disease using activity tracker data," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 3, pp. 878–884, 2019.