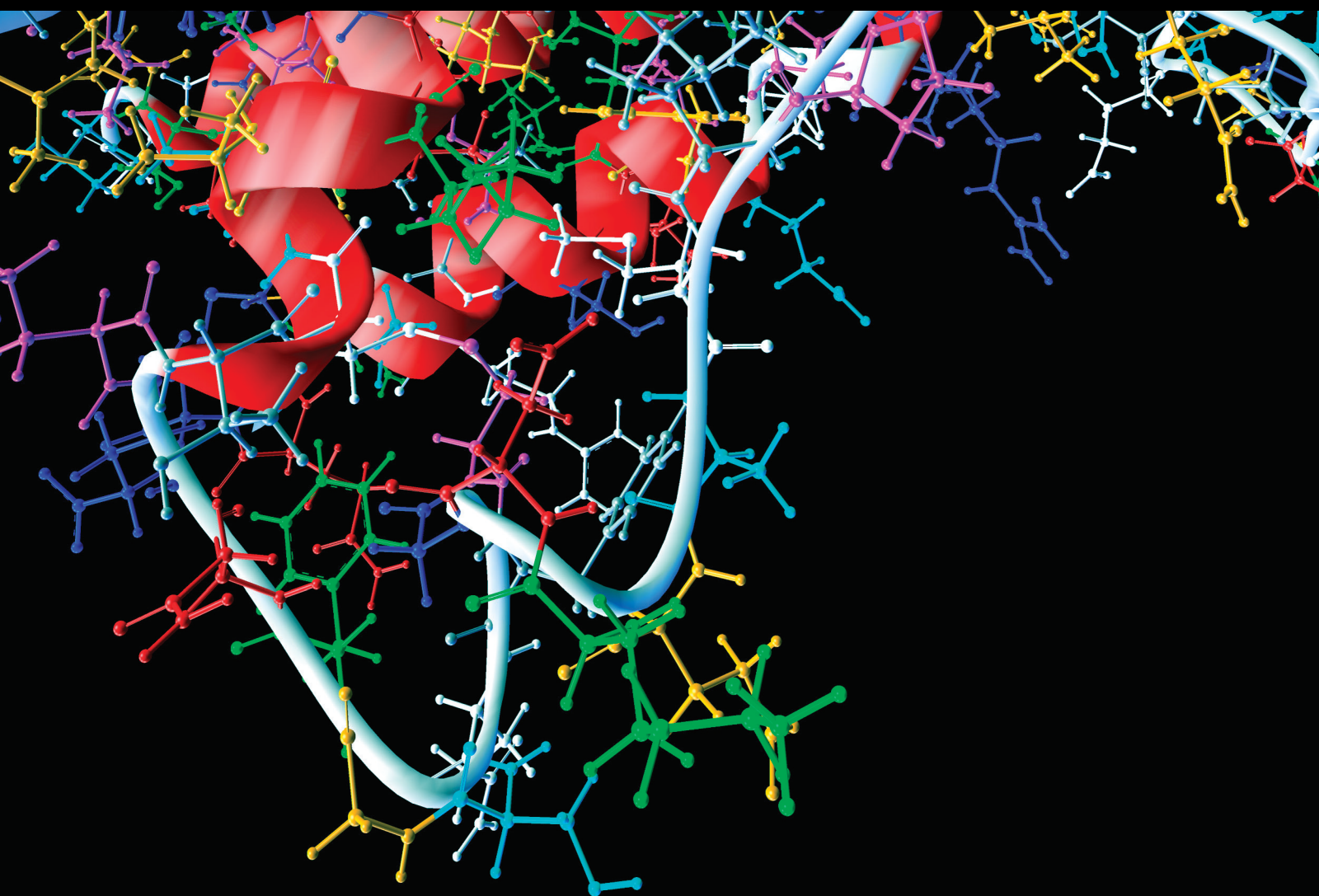# Automated and Semi-Automated Computational Intelligence Techniques for Medical Data Assessment

Lead Guest Editor: Venkatesan Rajinikanth
Guest Editors: Navid Razmjooy and S. Arunmozhi

Automated and Semi-Automated Computational Intelligence Techniques for Medical Data Assessment

Automated and Semi-Automated Computational Intelligence Techniques for Medical Data Assessment

Lead Guest Editor: Venkatesan Rajinikanth
Guest Editors: Navid Razmjooy and S. Arunmozhi

Hyuntae Park, Japan
Jovana Paunovic, Serbia
Manuel F. G. Penedo, Spain
Riccardo Pernice, Italy
Kemal Polat, Turkey
Alberto Policriti, Italy
Giuseppe Pontrelli, Italy
Jesús Poza, Spain
Maciej Przybyłek, Poland
Bhanwar Lal Puniya, USA
Mihai V. Putz, Romania
Suresh Rasappan, Oman
Jose Joaquin Rieta, Spain
Fathalla Rihan, United Arab Emirates
Sidheswar Routray, India
Sudipta Roy, India
Jan Rychtar, USA
Mario Sansone, Italy
Murat Sari, Turkey
Shahzad Sarwar, Saudi Arabia
Kamal Shah, Saudi Arabia
Bhisham Sharma, India
Simon A. Sherman, USA
Mingsong Shi, China
Mohammed Shuaib, Malaysia
Prabhishek Singh, India
Neelakandan Subramani, India
Junwei Sun, China
Yung-Shin Sun, Taiwan
Min Tang, China
Hongxun Tao, China
Alireza Tavakkoli, USA
João M. Tavares, Portugal
Jlenia Toppi, Italy
Anna Tsantili-Kakoulidou, Greece
Markos G. Tsipouras, North Macedonia
Po-Hsiang Tsui, Taiwan
Sathishkumar V E, Republic of Korea
Durai Raj Vincent P M, India
Gajendra Kumar Vishwakarma, India
Liangjiang Wang, USA
Ruisheng Wang, USA
Zhouchao Wei, China
Gabriel Wittum, Germany
Xiang Wu, China

KI Yanover, Israel
Xiaojun Yao, China
Kaan Yetilmezsoy, Turkey
Hiro Yoshida, USA
Yuhai Zhao, China

Contents

Retraction

Retracted: Optimization of Correlation Filters Using Extended Particle Swarm Optimization Technique

Computational and Mathematical Methods in Medicine

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Masood, A. Zafar, M. U. Ali et al., "Optimization of Correlation Filters Using Extended Particle Swarm Optimization Technique," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 6321860, 13 pages, 2021.

Research Article

Breast Cancer Diagnosis by Convolutional Neural Network and Advanced Thermal Exchange Optimization Algorithm

Xiuzhen Cai,[1] Xia Li [ID],[2] Navid Razmjooy [ID],[3] and Noradin Ghadimi [ID][4]

[1]*Community Health Service Center, Taibei Street, Jiangan District, Wuhan, 430010 Hubei Province, China*
[2]*Wuhan Vocational College of Software and Engineering, Wuhan, 430205 Hubei, China*
[3]*Department of Engineering, Tafresh University, Tafresh, Iran*
[4]*Young Researchers and Elite Club, Islamic Azad University, Ardabil Branch, Ardabil, Iran*

Correspondence should be addressed to Xia Li; lixia95272021@163.com, Navid Razmjooy; xnavid@gmail.com,
and Noradin Ghadimi; noradin.ghadimi@iauardabil.ac.ir

A common gynecological disease in the world is breast cancer that early diagnosis of this disease can be very effective in its treatment. The use of image processing methods and pattern recognition techniques in automatic breast detection from mammographic images decreases human errors and increments the rapidity of diagnosis. In this paper, mammographic images are analyzed using image processing techniques and a pipeline structure for the diagnosis of the cancerous masses. In the first stage, the quality of mammogram images and the contrast of abnormal areas in the image are improved by using image contrast improvement and a noise decline. A method based on color space is then used for image segmentation that is followed by mathematical morphology. Then, for feature image extraction, a combined gray-level cooccurrence matrix (GLCM) and discrete wavelet transform (DWT) method is used. At last, a new optimized version of convolutional neural network (CNN) and a new improved metaheuristic, called Advanced Thermal Exchange Optimizer, are used for the classification of the features. A comparison of the simulations of the proposed technique with three different techniques from the literature applied on the MIAS mammogram database is performed to show its superiority. Results show that the accuracy of diagnosing cancer cases for the proposed method and applied on the MIAS database is 93.79%, and sensitivity and specificity are obtained 96.89% and 67.7%, respectively.

1. Introduction

Breast cancer is common cancer and is the first cause of cancer mortality in women. Breast cancer is a malign tumor that develops from cells in the same organ. The disease usually begins in the lobules, or breast ducts, and can then penetrate the ducts and walls of the glands and attack the surrounding adipose tissue or even other parts of the body. There may be other lumps in the breast that are not cancerous, but in any case, the final diagnosis is up to the physicians. Science has proven that despite the existence of a way to prevent cancer and the lack of definitive treatment for this disease, early diagnosis of this disease helps physicians to at least prevent the progression of this disease. The number of cancer patients in the world is increasing. Based on the World

Health Organization (WHO), breast cancer has a great effect on about 2.1 million women annually. Based on this statistic, in 2018, 627,000 women have died of breast cancer which contains about 15% of all deaths among women cancer [1]. Figure 1 shows the statistical information of the cancer diagnosis and the cancer deaths in 2019 [2].

The best solution to decline breast cancer mortality is to diagnose it in the primary stage and treat it. Early diagnosis needs a precise and dependable diagnostic method. Among the various methods of diagnosing breast cancer, mammography is a very common and very popular method. Systematic screening of the female population with mammograms and early diagnosis of early stage breast cancer can also increase the patients' survival chances and reduce the negative side effects of necessary treatments. These results are

(a)

(b)
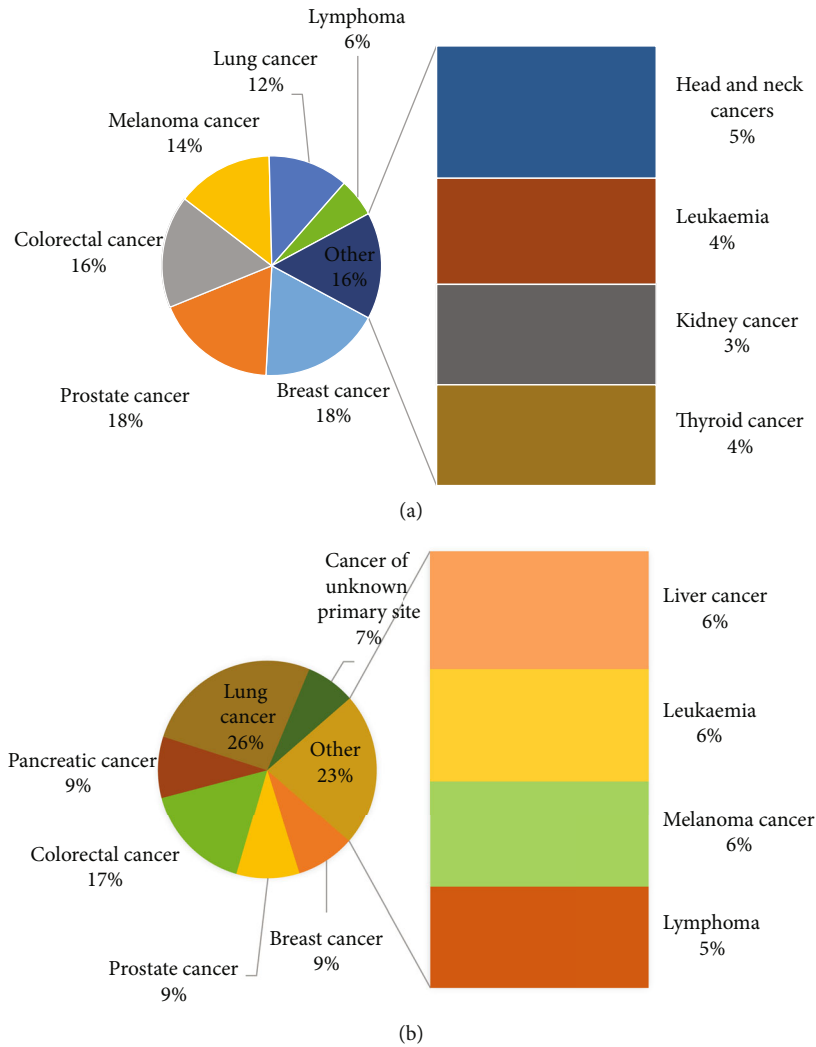
FIGURE 1: The statistical information of the cancers (a) and cancer deaths (b) in 2019 [2].

possible if the quality of services is provided in the best possible way.

On the other hand, the diagnosis of breast cancer based on mammography film has several problems. In some cases, there is a possibility that the film is damaged or the image is not suitable for diagnosis. Meanwhile, the film wears out over time, and the possibility of revision decreases. On the other hand, the only tool a doctor can use to diagnose a lesion is a visual video. The physician's visual observations to diagnose the lesion leads to two errors. The first is that a radiograph is shown twice to a physician or radiologist. If he does not know that both images are the same, his diagnosis may be different. Another error is showing an image to two physicians or two radiologists, each of whom has a different diagnosis. Although mammography is usually the best way to diagnose breast cancer, some classes of cancers are not diagnosed in this way. In this condition, providing the computer-aided systems can detect malignant lesions efficiently [1]. Based on the literature, the effectiveness of a computer diagnostic system is more precise than that of a physician. Recently, various kinds of research works were

performed in the area of the automatic early detection of breast cancers [3]. In other words, the higher efficiency of the computer-aided systems assists the physicians to diagnose cancer with lower complexity and higher speed. For instance, Liu et al. [4] proposed a proper image segmentation methodology for optimal breast cancer diagnosis regarding the interval uncertainties. For considering the indeterminacy, interval assessment was utilized. The method is guaranteed to providing suitable results in any changes in the imaging system. The main idea was to provide an interval-based Laplacian of Gaussian filter to simulate the intensity uncertainties. Final results have been performed to MIAS database, and a comparison of the results with some classic techniques was carried out to illustrate the method efficiency.

Toğaçar et al. [5] introduced a breast cancer diagnosis system by CNN. The method was improved by a technique called BreastNet. The image data was established by the expansion technique before implementing it into the model. Then, a precise classification method was accomplished based on the hypercolumn technique. Then, a comparison

of method results with some latest techniques was done to state the higher precision of the suggested system.

Carvalho et al. [6] used another method for breast cancer detection for utilizing in histopathological images. The authors presented a method for using phylogenetic diversity indexes to determine images to model creation and histopathological breast image classification into some classes. The method was then compared with several different latest techniques to demonstrate the technique's accuracy. The results showed significant robustness to the method to help experts at large medical centers.

It can be concluded from the previous studies that many works have been done for the automatic diagnosis of breast cancer. This paper proposes a new automated method for breast cancers diagnosed in mammogram images. Here, an optimized deep learning-based methodology based on a new improved metaheuristic, called Advanced Thermal Exchange Optimization algorithm, has been used for this purpose.

2. Image Preprocessing

The heterogeneity of light intensity in medical images has weakened the boundaries of medical images, especially the heterogeneity of light intensity in magnetic resonance images created by nonuniform magnetic fields by radiofrequency coils, which is why preprocessing is so important in medical research. Therefore, after obtaining the input information of the medical images, preprocessing operations should be performed, which are methods to eliminate noise and isolate and improve the differentiation of areas where there is a possibility of numerical information.

2.1. Image Contrast Improvement. Commonly, in images, several forms of contrast issues are existing, for instance, the inappropriate lighting and room conditions, the deficiency of enough applicant interface for imaging, and the inadequate quality of the measuring sensors and devices. These variables will disappear with some essential details, darkening or overexposure, and finally the image abnormalities. The presence of these variables increases their need for enhancement in numerous medical images. Lack of these parameters in most numbers of medical images increases their requirement for improvement. This improvement has been performed based on contrast enhancement.

Here, the contrast enhancement has been applied to the images to highlight the skin cancer areas with no changes on the other areas. A simple application to perform a piecewise linear contrast stretch operation on an image. The present study uses a 16-bit lookup table to improve the contrast of the images that are then stored on a disc. This is implemented based on the following formula:

$$y_{\text{hist}} = \frac{x_{\text{hist}} - \text{Min}_{\text{hist}}}{\text{Max}_{\text{hist}} - \text{Min}_{\text{hist}}}, \qquad (1)$$

where Max_{hist} and Min_{hist} stand for the highest and the lowest levels for the gray magnitudes of the main image histogram, respectively, and x_{hist} and y_{hist} represent the input

image before contrast enhancement and the output image after image contrast enhancement, respectively.

2.2. Noise Reduction. As aforementioned, due to different conditions in the medical imaging, there are some kinds of noises in them that should be removed before processing. This is done by using noise reduction. Noises can be white, random, or Gaussian (that contains a large part of medical images). Noise is usually in the high-frequency bands of the image; the important edges and details of the image are in the same bands. Therefore, noise removal along with preserving the edges and important image information is the main problem in the image noise removal process. Many noise removal techniques have been proposed in recent years. One of the proper methods for noise removal in medical images is to use the Wang-Mendel algorithm. This algorithm is a beneficial technique based on fuzzy theory [7]. Because of the simple conception of the fuzzy theory, the method of the Wang-Mendel algorithm is so easy to understand. Also, due to the higher speed of this method, it is too valuable for initial fuzzy model creation [8]. The method for the rule database is achieved by the following:

(1) Perform fuzzy separation from the input variable space which can be obtained based on the knowledge or using normalization technique. Afterward, it classifies it into two parts including equal or unequal by performing a fuzzy separation of the input variable space. Afterward, the membership function has been selected and the components have been given as a fuzzy package. Afterward, membership has been chosen and a fuzzy set is given to every part

(2) Creation candidate language rules which can be formed by choosing the most all-encompassing laws for the samples

(3) Level of validity assignment to the laws that are achieved based on multiplying the membership function values of the components and the membership function value of the result of the law

(4) Provide final database rules from the collection of candidate language which is done by classifying the candidate rules into various groups, where each of them contains candidate rules with similar assumptions. For obtaining the final rule base, the maximum degree of verification law is achieved in every set

3. Image Segmentation

3.1. Changing Color Space. However RGB has a good concept for the human, it is completely dependent on the three colors (red, green, and blue), as base colors of the RGB color space. This color space has also a high dependency on the ambient light intensity which limits its usages in a different application. To recover this issue, different color spaces have been introduced. In this study, the XYZ color space is utilized for the purpose after some trials and errors. The XYZ color space makes a link between the physiologically

(i) Compute the histogram and probabilities of each intensity level
(ii) Give primary values of $\omega_i(0)$ and $\mu_i(0)$ for all feasible threshold levels (highest intensity $(t = 1, 2, \cdots,)$
(iii) Renew ω_i and μ_i
(iv) Calculate $\sigma_b^2(t)$
(v) The optimal threshold here is the maximum of $\sigma_b^2(t)$

PSEUDOCODE 1: The pseudocode of the Otsu algorithm.

supposed colors in human color vision and distributions of wavelengths in the visible spectrum of the electromagnetic. In the XYZ color model, Y states the luminance, and X and Z indicate the color information. The formula for converting the RGB to XYZ is as follows:

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

(2)

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

3.2. Method of Segmentation. The red (R) color space provides the main dimension in RGB color space to give nearly the image intensity in medical images. As previously mentioned, both X and Z values provide similar color information for the XYZ color space. So, for segmentation, X dimension and red dimension are only normalized, i.e.,

$$
\widehat{R} = \frac{R}{\sqrt{R^2 + G^2 + B^2}},
$$
$$
\widehat{X} = \frac{X}{\sqrt{X^2 + Y^2 + Z^2}}.
$$

(3)

These normalization values are carried out on every pixel in the input images. The Otsu threshold is used after this normalization to provide a low-cost segmentation in the sense of time complexity.

By intergroup variance optimization and lessening pixels' intragroup variance, the Otsu process is an efficient way that is established to automatically pick the optimal threshold. There is an issue with the global threshold when the resolution of the image background is insufficient. To remove the heterogeneity effect, it is possible to use a local threshold. To remove inhomogeneity and add a global threshold to the processed image, this problem is solved by image preprocessing.

Based on the Otsu method, the threshold value has been searched which minimizes the class-in-between variance as follows:

$$
\sigma_\omega^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t),
$$

(4)

where ω_i describes the probability for two different groups with a threshold magnitude of t and σ_i^2 represents the vari-

ance amount of the groups. In other words, Otsu indicates that variance minimization of a class is like maximization of the variance in class-within, i.e.,

$$
\sigma_b^2(t) = \sigma^2 - \sigma_\omega^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2,
$$

(5)

where μ_i describes the mean value. The Otsu algorithm will be defined as the following pseudocode.

Subsequently, for better performance, postprocessing mathematical morphology, including filling, closing, and opening, was carried out on the images [9]. First, extra holes in the image are filled by applying the mathematical filling operator. The analytical model is as follows:

$$
X_k = (X_{k-1} \oplus B) \cap A^c, \quad k = 1, 2, 3 \cdots,
$$

(6)

where A and B describe the area that should be processed and the constructing element, respectively.

Afterward, the mathematical opening has been employed to the filled image to eliminate the ignitor information with no adjustments on other gray surfaces. The mathematical model of this operator is as follows:

$$
A \circ B = (A \ominus B) \oplus B.
$$

(7)

Then, the mathematical closing is executed based on the following equation for linking the narrow parts:

$$
A \cdot B = (A \oplus B) \ominus B.
$$

(8)

This study uses a 5×5 identity matrix as a structural element. Figure 2 gives some examples of the breast segmentation of images depending on the approach suggested.

4. Feature Extraction

Feature extraction is the process of reducing the dimensional of the images by dividing and reducing an initial set of images to more controllable groups. So, the next processing of the images can be simpler by this process. Of these large datasets, the most important characteristic is that they have numerous parameters. To process them, these parameters need many calculating origins. Therefore, extraction of these features helps to select and combine variables into features to get the best feature from those big datasets, thereby effectively decreasing the data volume. In this study, two popular features including DWT and GLCM were employed for feature extraction that is described subsequently.
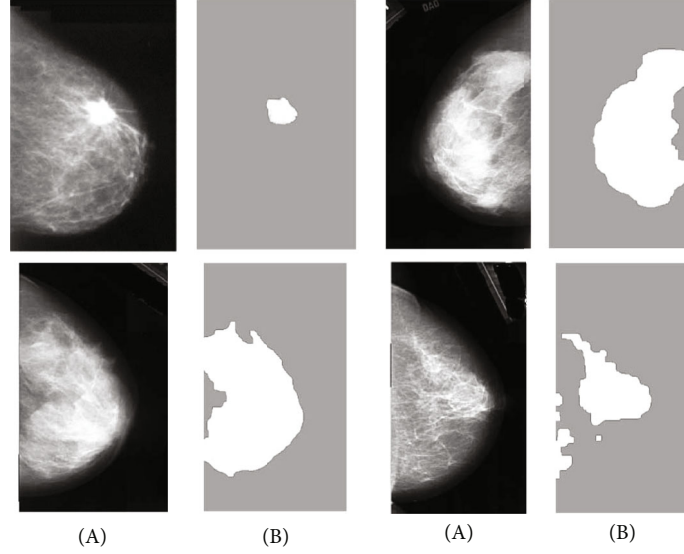
FIGURE 2: Several instances of the breast segmentation of images depending on the approach suggested: (a) basic image and (b) segmented image.

4.1. Discrete Wavelet Transform (DWT). Wavelet transform is an efficient tool for various applications in image processing and is utilized in numerous fields such as image noise removal, pattern recognition, coding, image compression, and feature extraction. Wavelet transform is a method in the frequency domain. In this method, instead of using sine and cosine functions (such as Fourier transform), a function called wavelet is used. When implementing a wavelet transform, the wavelet function retains its shape but is displaced along with the signal and compressed and opened during the displacement, thus encapsulating the entire signal. Unlike short-time Fourier transform, this method can create different resolutions for low- and high-frequency ranges.

There are various solutions for implementing DWT, the most common of which is the implementation of Multiresolution Analysis (MRA). In this method, the implementation of discrete wavelet transform is done with the help of a series of consecutive operations that each step of this operation includes signal filtering and downsampling. At each stage of the discrete wavelet process, the signal content is decomposed into two orthogonal subspaces including a low-pass filter (LL) and a high-pass (HH) filter [10], which is then split into four classifications: LH, LL, HH, and HL.

To increase the frequency resolution, this decomposition is rendered consecutively, such that the approximation signal is passed through a pair of high- and low-pass filters and decomposed into two new information and approximate signals. Afterward, the read rapidity was reduced by 50 percent. To give more information, the HL subbands with further efficiency are applied. This process is mathematically formulated as follows:

$$P_{dwt}(s) = \begin{cases} d_{i,j} = \sum f(s) \times H * i(s - 2 \times i \times j), \\ d_{i,j} = \sum f(s) \times L * i(s - 2 \times i \times j), \end{cases} \quad (9)$$

where $d_{i,j}$ signifies the feature of the component in signal $f(s)$, L and H describe the coefficients of low-pass and high-pass filters, respectively, and i and j stand for the wavelet and the translation factor scales, respectively.

4.2. Gray-Level Cooccurrence Matrix. To study the structure of different tissues, Haralick has proposed properties based on the GLCM, which is one of the most successful methods for studying the properties of different tissues [11]. In the gray surface cooccurrence matrix method, it is assumed that the texture image information is determined by a specific matrix. This method is relied on manipulating the gray surfaces of an image. In this method, in addition to examining the gray surface of the desired pixel, the gray surfaces of its neighboring pixels are also examined, and by creating a new matrix of gray surfaces of the pixel neighbors at different angles and distances, the image properties are identified and defined. The coevent matrix is a square matrix and its size is the amount to the gray surfaces' number.

The cooccurrence matrix of an image is defined using radius d and angle θ. Usually, d is selected in the range of 1 and 2. Since each pixel has 8 neighbors at θ equal to 0, 45, 90, 135, 180, 225, 270, and 315 degrees to define the cooccurrence matrix, so the angle selection may be up to 4 adjacent pixels at θ equal to 0, 45, 90, 135, and 180 degrees (horizontal, right diagonal, vertical, and left diameter). In addition to the radius and angle parameters used to define the cosmopolitan matrix, the gray number parameter can also be defined. In this study, to define the cooccurrence matrix, a radial distance of 1 with four zero angles and the number of 256 gray surfaces were used, for which a new matrix was extracted. Subsequently, the information about the utilized characteristics was explained. The first characteristic is *Contrast* that describes the intensity magnitude of the pixels and their neighborhood. The second feature is *Entropy* which defines the image selected interference. The third feature is *Energy* that describes the repetitive pixel pair

quantity. The fourth feature is *Correlation* that defines the spatial feature reliance among the pixels. Finally, *Homogeneity* as the last feature as a local uniformity feature creates multiple/single intervals for accusing the nontextured/textured characteristics. Table 1 indicates five gray-level cooccurrence matrix features extracted from the samples.

5. Convolutional Neural Networks

After feature extraction from the segmented images, they should be classified properly as the final step of diagnosis. In this study, convolutional neural network (CNN) was used for this purpose. CNNs are significant deep learning techniques where several layers are prepared strongly. This technique is very effective and is a usual technique in different applications of computer vision. An outline of convolutional neural network architecture is depicted in Figure 3.

Generally, a CNN is made of three major layers: the convolutional layer, the pooling layer, and the fully connected layer. Various layers do various tasks. Each convolution neural network includes two stages: feedforward and backward for preparation.

In the first step, the features enter the network, and this operation is the point multiplication between the input and the variables of each neuron, and finally, the application of convolution operations in every layer.

The output of the network is then computed. Here, to establish the variables related to network training, network output results are applied to compute the network error rate. To do this, a comparison of the network output to the correct solution (optimal solution) is carried out by an error function and the error rate is computed. In the later phase, by the computed error rate, the postrelease phase begins. The gradient of each variable is computed in this phase based on the chain rule, and all variables are altered by the influence they have on the error created in the network. Following parameters' updating, the feed-forward phase starts. Afterward, repeating a good number of these phases, the network preparing finishes. In this study, CNN is employed for local feature extraction in breast mammogram images. To offer optimum weighting among network connections, the backpropagation technique has been established. As the activation mechanism, a rectified linear unit (ReLU) is used.

With multiplying filter matrices by the images, feature maps are generated. To generate the feature map, the filter moves from left to right and up to down with a specific stride size to extract high-level features (like edges) until it finishes the full width. Here, the Max-pooling process uses the maximum value of the matrix in the feature maps to decrease the output neurons and the cross-entropy loss value based on backpropagation, which is formulated as follows:

$$
L = \sum_{j=1}^{N} \sum_{i=1}^{M} - d_j^{(i)} \log z_j^{(i)},
$$

$$
d_j = \left(0, \cdots, 0, \underbrace{1, \cdots, 1}_{k}, 0, \cdots, 0 \right),
\tag{10}
$$

TABLE 1: Five GLCM features extracted from the samples.

Feature name	Mathematical equation
Contrast	$\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (i-j)^2 f(i,j)$
Entropy	$-\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \log_2 f(i,j)$
Energy	$\sqrt{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} f^2(i,j)}$
Correlation	$\dfrac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (i,j) f(i,j) - \mu_i \mu_j}{\sigma_i \sigma_j}$
Homogeneity	$\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \dfrac{1}{1 + (i-j)^2} f(i,j)$

where d_j signifies the proper output vector and z_j determines the achieved output vector for the mth class. The softmax function is achieved as follows:

$$
z_j^{(i)} = \frac{e^{f_j}}{\sum_{i=1}^{M} e^{f_i}},
\tag{11}
$$

where M describes the sample number.

To adjust function followed by keeping higher values, a weighting penalty (ρ) is added that is illustrated in the following equation:

$$
L = \sum_{j=1}^{N} \sum_{i=1}^{M} - d_j^{(i)} \log z_j^{(i)} + \frac{1}{2} \rho \sum_{K} \sum_{L} W_{k,l}^2,
\tag{12}
$$

where W_k defines the weight of connections and L and K define the overall number of layers and the layer l connections, respectively.

The CNN layouts were usually used based on trials and errors, which yielded inaccurate results. Numerous automated and optimized works have been implemented to address this problem [12]. The use of metaheuristic algorithms is one of the normal approaches. A new, optimized metaheuristic was used in this study to present an effective CNN based on the previously described cases.

6. The Modified Thermal Exchange Optimizer

6.1. The Concept of Newton Law of Cooling. Heat transfer that occurs simultaneously with the movement of a fluid is called convection heat transfer. Depending on the process, heat transfer is divided into two categories: free and forced. In free movement, the energy transferred is due to natural factors such as Archimedes' force. But in forced displacement, external forces such as a pump or fan cause the fluid to move. The heat transfer analysis is complex due to the simultaneous process of thermal conductivity and fluid motion. The higher the fluid velocity, the higher the heat
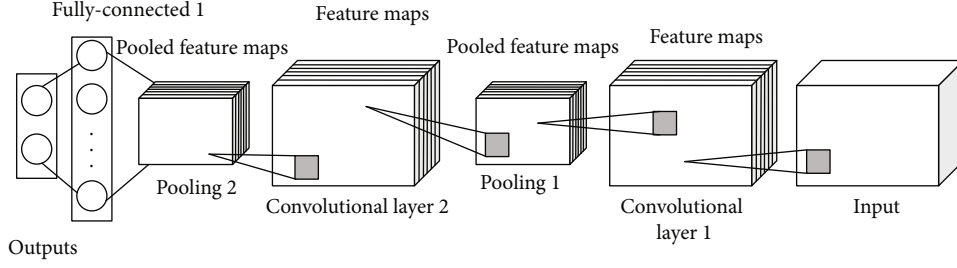
FIGURE 3: An overview of convolutional neural network architecture.

transfer rate. The transfer heat transfer velocity can also be expressed using Newton's law of cooling by the following formula:

$$\dot{Q} = \beta \times A \times (T_s - T_a), \tag{13}$$

where A describes the surface of the body that transfers heat, Q determines the heat, α signifies the coefficient of the heat transfer that relates to numerous cases like surface state, heat transfer mode, and object geometry, and T_b and T_a represent the body and the ambient temperatures.

According to the above equation, heat losing time is $\beta \times A \times (T_a - T) \, dt$ that defines reserved heat changing once the temperature dT falls, i.e.,

$$V \times \rho \times c \times dT = -\alpha \times A \times (T - T_b) dt, \tag{14}$$

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

$$\frac{T - T_b}{T_{eh} - T_b} = \exp\left(\frac{-\beta \times A \times t}{V \times \rho \times c}\right), \tag{15}$$

where T_{eh} describes the early high temperature. By considering the $(\alpha \times A \times t)/(V \times \rho \times c)$, a time-independent value, i.e.,

$$\zeta = \frac{\alpha \times A}{V \times \rho \times c}. \tag{16}$$

That ζ is a constant, the main equation can be reformulated as follows:

$$\frac{T - T_b}{T_{eh} - T_b} = \exp(-\zeta t). \tag{17}$$

Accordingly,

$$T = (T_{eh} - T_b) \times \exp(-\gamma t) + T_b. \tag{18}$$

6.2. Thermal Exchange Optimization Algorithm. After explanations about the concept of Newton's law of cooling, it is time to explain the concept of optimization and the relation of the Newton law of cooling and the optimization [13]. Generally, optimization contains all techniques that are used for finding the best solution for optimization problems. Several methods

of optimization techniques have been introduced for this aim. Classic methods give exact results for the optimization problems, but recently, by increasing the complexity of these problems, the ability to solve the problems with these algorithms is decreasing. Metaheuristics are intelligent algorithms that are used to find the optimal solution and resolve the before mentioned issues [14, 15]. Metaheuristic algorithms are approximation optimizers that have solutions to exit the local optimization and proper for a wide range of problems. Metaheuristic algorithms are an inspiration of various phenomena from the nature, behaviors of animals, breeding, to human societies and use these conceptions to simulate an approach for solving the optimization problem. Several kinds of metaheuristic algorithms have been proposed in recent years [16, 17], for example, biogeography-based optimization [18], elephant herding optimization [19], ant lion optimizer (ALO) algorithm [20], equilibrium optimizer [21], world cup optimizer [22], and Thermal Exchange Optimizer (TEO) [23].

Here, an enhanced design of the TEO algorithm is presented to provide more ability for this algorithm in terms of accuracy and consistency. The TEO algorithm is an inspiration of the temperature performance of the objects and their position which is exchanged between warm and cold portions to indicate the updated positions. In the TEO optimizer, the individual is split into two parts. One group is the candidates that are considered as cooling substances, and the other group is considered as the environment, and then, the reverse process has been made. Figure 4 shows the pairs of transfer objects.

The algorithm begins with a predefined number of random individuals as the initial solutions are as follows:

$$T_j^0 = \underline{T} + \theta \times (\bar{T} - \underline{T}), \\ j = 1, 2, \cdots, n, \tag{19}$$

where θ denotes a random magnitude in the range [0, 1], T_j^0 signifies the algorithm early population for the ith object, and $\underline{T}$ and $\bar{T}$ stand for the minimum and the maximum limitations.

After achieving the cost value of the generated candidates, T number of the best cost individual positions is saved as *Thermal Memory* (TM) to provide higher efficiency with lower complexity to the algorithm. The *TM* individuals are
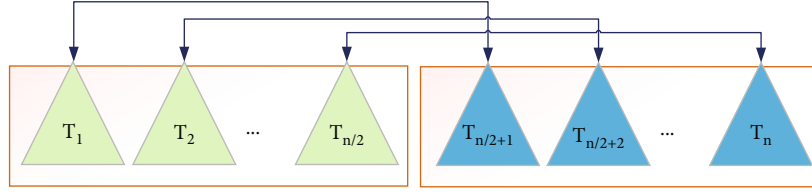
FIGURE 4: The pairs of transfer objects.

then combined to the individual, and the equal number of worst candidates is then taken out.

To provide more understanding, consider Figure 5. T_1 describes the environment object for $T_{(n/2)+1}$ cooling object, and contrariwise. If the object is less than ζ, the temperature exchanges gradually. ζ is formulated as follows:

$$\zeta = \frac{\text{Cos(object)}}{\text{Cos(worst object)}}. \tag{20}$$

Time is another term in the simulation of the optimizer that is related to the number of iteration. This term is obtained by the following equation:

$$t = \frac{\text{iteration}}{\text{Max.iteration}}. \tag{21}$$

To improve the global searching of the algorithm, the environmental temperature changing is considered that is formulated as follows:

$$T_i^e = \left(1 - \left(\alpha_1 + \alpha_2 \times (1 - t) \times \delta\right)\right) \times T_i^{\mathscr{E}}, \tag{22}$$

where δ describes a random number between 0 and 1, $T_i^{\mathscr{E}}$ represents the preceding temperature of the object modified by T_i^e, and α_1 and α_2 represent the control variables, respectively.

Finally, the new position for the object temperature is achieved by the following:

$$T_i^N = T_i^e + \left(T_i^{\text{old}} - T_i^e\right) \exp\left(-\zeta t\right). \tag{23}$$

The algorithm also defines whether a component changes in the cooling objects or not. This has been stimulated by a term, called Pr. The Pr contains some individuals that are compared with $R(i)$ which is a random value in the range [0, 1]. If $R(i)$ is less than Pr, one dimension of the ith candidate is randomly chosen and the magnitude is reformulated as given in the following:

$$T_{i,j} = \underline{T}_j + \delta\left(\bar{T}_j - \underline{T}_j\right) \exp\left(-\zeta t\right), \tag{24}$$

where $T_{i,j}$ describes the variable number j of the individual number i and $\underline{T}_j$ and $\bar{T}_j$ represent the lower and the higher limitations of the parameter number j, respectively. The
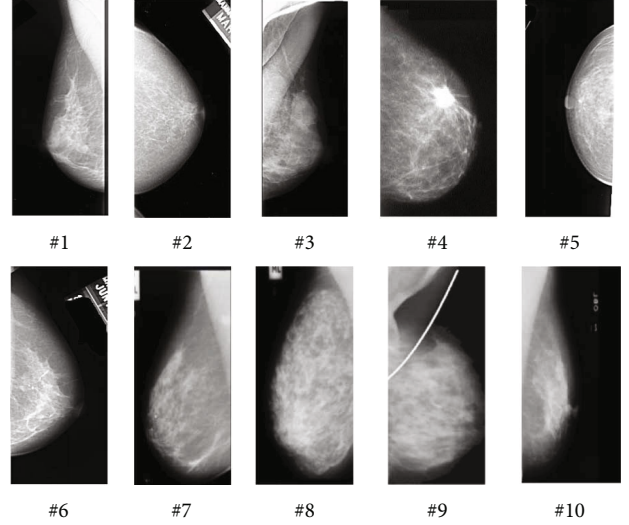
FIGURE 5: Some examples of the MIAS database mammography images.

algorithm is then terminated when the stopping criteria have been reached.

6.3. *Advanced Thermal Exchange Optimizer.* Although the Thermal Exchange Optimizer has a proper speed in solving the problems, it may be trapped in the local optimum point once solving complex and nonlinear optimization problems. Due to this problem, here, an advanced design of the TEO algorithm is designed and suggested to develop the search power of the original TEO algorithm and to resolve the mentioned issue. The movement of the worst individual in the groups (T_w) is improved in each iteration of the local search in the Advanced Thermal Exchange Optimization (ATEO) algorithm. First, an exchange vector is generated for the worst solution in each iteration:

$$T_i^a = T_b + \gamma \times \left(T_{r1} - T_{r2}\right), \tag{25}$$

where T_b signifies the best solution achieved by the current iteration, T_{r1} and T_{r2} represent two dissimilar agents that are randomly chosen from the population in each group, and γ describes the exchange coefficient to determine the differences range between T_{r1} and T_{r2}.

By considering the above condition, the value of the jth parameter of the vector T_i^{New} in the following iteration is achieved as follows:

$$T_i^{\text{New}} = \begin{cases} T_i^a & \text{if } \text{rand} < CG, \\ T_i^N & O.W., \end{cases} \qquad (26)$$

where CG signifies the general intersection constant between 0 and 1 and rand signifies a random constant in the range of [0, 1]. If the cost value of the new solution has proper value in comparison to the preceding solution, the new individual substitutes the former one; else, it will be kept with no changes.

To give a proper result with the TEO algorithm, the population size should be selected wisely. Indeed, population size is a term to define the number of individuals (candidates) that are randomly generated and tested on the objective function to get the best solution. However, this case is one of the difficult parts of all metaheuristics [24]. Here, a self-adaptive mechanism is used for adjusting this case in each iteration. The main characteristic of the self-adaptive mechanism is that it regulates the population size automatically in each iteration with no user intervention [25]. Based on this mechanism, the initial population size before starting the algorithm main loop is considered as follows:

$$PS = 10 \times D, \qquad (27)$$

where PS signifies the population size and D describes the problem dimensions. So, the new population size is achieved by the following:

$$PS^{\text{New}} = \text{round}(PS + \text{rnd} \times PS), \qquad (28)$$

where rnd defines a random magnitude in the range [-0.5, 0.5].

The population size will increase or decrease by up to half the current population size. If the population size obtained for the next iteration increases compared to the population size in the former iteration ($\text{PopSize}_{\text{new}} > \text{PopSize}$), then all members of the present individual are kept unchanged.

Once the population size in the former iteration decreases compared to the population size ($\text{PopSize}_{\text{new}} < \text{PopSize}$), the best members of the present population are kept and the weak members have been discarded. If the size of the population does not change ($\text{PopSize}_{\text{new}} = \text{PopSize}$), so there will be no population changing. Finally, if the new population size reduces from the problem dimensions ($PS^{\text{New}} < D$), then the population size becomes equal to the problem dimension.

6.4. Algorithm Verification. After designing and introducing the proposed Advanced Thermal Exchange Optimization (ATEO) algorithm, the performance of the method should be analyzed to ensure its ability in use in our purpose.

The present work uses the Single Objective Bound Constrained Numerical Optimization (CEC2020) benchmark standard benchmark in 20 dimensions to analyze the effectiveness of the method; in other words, each faction has 20 decision variables that should be optimally selected. The CEC2020 is known as one of the latterly introduced bench-

Table 2: The variable setting of the compared optimizers.

Algorithm	Parameter	Value
BH [27]	a	[0, 1]
	Number of stars	100
MVO [28]	Traveling distance rate	[0.6, 1]
	Wormhole existence prob.	[0.2, 1]
SHO [29]	$\vec{M}$	[0.5, 1]
	$\vec{h}$	[5, 0]

mark functions for analyzing optimization problems. The termination criteria, including the maximum number of the calculation of the fitness function and the minimum error value, are set $1e7$ and $1e-8$, respectively. The constraint of the decision variables is in the range [−100, 100], and 35 independent runs have been established for giving a reliable result. The formulation for each equation can be found in [26]. To provide an appropriate analysis for the suggested algorithm, a comparison of its results with some latest algorithms including blackhole (BH) [27], multiverse optimize (MVO) [28], spotted hyena optimize (SHO) [29], and original Thermal Exchange Optimization (TEO) [23] algorithm has been performed. Table 2 indicates the parameter setting of the compared algorithms.

During the simulations, for all of the compared algorithms, the population size is equal to 150. To analyze the algorithms' behavior, the mean magnitude and the (mean) and the standard deviation value (Std) have been extracted from the results. Table 3 discusses the achievements of the algorithms applied to the CEC2020 benchmark sets.

As seen in Table 3, the scores of the analyzed algorithms have been shown. It is clear from the results that in some test functions, the suggested ATEO algorithm escapes the local optimum and found the optimal value. Here, the mean value is employed to consider all of the runs for the algorithm, although the minimum value of the algorithms gives too better or even incomparable results. The minimum value of "Mean" for the proposed ATEO algorithm against the other compared algorithms indicates its higher accuracy to find the minimum value. On the other hand, the minimum value of "Std" for the proposed ATEO algorithm state's better reliability of the proposed algorithm than the comparative methods for the studied CEC2020 benchmark function.

7. Classification

With the advances made in the field of imaging and production of high-resolution digital images, the need for accurate image classification is felt so that one of the most basic parts of image processing is image classification. The important point in image classification is providing a method with high accuracy. Taking into account the above-mentioned reasons, backpropagation is in the form of preparing in the CNN mostly. We also clarified why different methods have been proposed to overcome it because of certain major drawbacks of the backpropagation method. Here, the proposed

TABLE 3: The comparison achievements between the suggested ATEO algorithm and the other compared algorithms on the CEC2020.

		ATEO	TEO [23]	BH [27]	MVO [28]	SHO [29]
F1	Mean	$7.38e8$	$5.83e11$	$4.22e15$	$8.37e13$	$9.07e15$
	Std	$1.29e8$	$6.19e10$	$5.13e13$	$4.38e11$	$5.46e11$
F2	Mean	$5.79e1$	$9.67e2$	$4.67e3$	$1.76e4$	$4.46e6$
	Std	$4.31e1$	$2.84e2$	$3.82e2$	$6.37e2$	$2.08e5$
F3	Mean	$2.08e1$	$6.92e2$	$9.37e2$	$5.17e2$	$4.83e5$
	Std	$1.46e0$	$3.27e0$	$4.28e1$	$8.09e1$	$6.17e4$
F4	Mean	0.00	$6.15e-10$	$5.80e-6$	$4.96e-7$	$7.67e-6$
	Std	0.00	$3.48e-11$	$9.37e-7$	$4.18e-8$	$4.08e-8$
F5	Mean	$1.76e2$	$4.53e2$	$6.37e4$	$6.55e3$	$9.86e7$
	Std	$3.82e1$	$1.27e2$	$5.19e3$	$2.41e2$	$8.19e3$
F6	Mean	$3.29e-1$	$6.12e-1$	$8.09e0$	$7.18e0$	$2.96e1$
	Std	$4.13e-1$	$2.73e-1$	$3.46e-1$	$4.82e-2$	$4.63e0$
F7	Mean	$3.18e0$	$4.16e0$	$8.09e1$	$5.33e2$	$4.29e3$
	Std	$1.24e0$	$1.08e0$	$6.17e1$	$6.81e1$	$2.82e2$
F8	Mean	$7.19e1$	$8.35e2$	$2.19e4$	$5.24e5$	$2.56e4$
	Std	$2.76e0$	$4.37e0$	$3.77e1$	$4.65e1$	$4.07e3$
F9	Mean	$1.96e2$	$3.17e2$	$6.51e3$	$2.85e3$	$5.11e4$
	Std	$1.07e1$	$2.03e1$	$8.09e2$	$6.19e1$	$6.97e1$
F10	Mean	$5.76e2$	$9.83e2$	$9.23e3$	$5.37e3$	$5.17e3$
	Std	$4.27e-1$	$5.94e-1$	$2.60e0$	$1.93e1$	$6.93e1$

TABLE 4: The feature extraction for training data.

#	H	CR	E	CN	ER
1	0.816	0.173	0.794	0.257	0.298
2	0.757	0.038	0.996	0.047	0.264
3	0.869	0.046	0.957	0.032	0.317
4	0.806	0.042	0.896	0.031	0.376
5	0.794	0.041	0.987	0.135	0.395
6	0.843	0.010	0.917	0.009	0.219
7	0.585	0.057	0.967	0.037	0.293
8	0.810	0.007	0.968	0.011	0.417
9	0.594	0.068	0.979	0.028	0.294
10	0.704	0.041	0.968	0.046	0.407

TABLE 5: The feature extraction for the testing data.

#	H	CR	E	CN	ER
1	0.794	0.072	0.794	0.047	0.272
2	0.758	0.053	0.856	0.012	0.311
3	0.786	0.032	0.851	0.046	0.347
4	0.865	0.029	0.783	0.019	0.215
5	0.749	0.028	0.764	0.018	0.337
6	0.708	0.029	0.886	0.029	0.318
7	0.819	0.017	0.895	0.053	0.319
8	0.693	0.022	0.851	0.031	0.420
9	0.649	0.034	0.817	0.050	0.433
10	0.684	0.069	0.963	0.079	0.351

Advanced Thermal Exchange Optimizer (ATEO) is developed and used to reduce the proper and output magnitude by the selection of suitable network weights replacing backpropagation in CNN for mean square error (MSE). The MSE can be mathematically described by the following equation:

$$MSE = \frac{1}{T} \sum_{j=1}^{N} \sum_{i=1}^{M} \left(y_j^i - d_j^i \right)^2, \tag{29}$$

where M and N represent the value of the output layers and the data, respectively, and y_j^i and d_j^i define the obtained and the proper magnitudes for j^{th} unit in the output layer of the CNN in time t, respectively.

8. Simulation Results

This study presents an efficient and automated method for brain tumor detection by combining deep neural networks and metaheuristics. The technique involves image preprocessing, image segmentation, extraction of features, and then classification. Based on digital mammogram images, the process is validated.
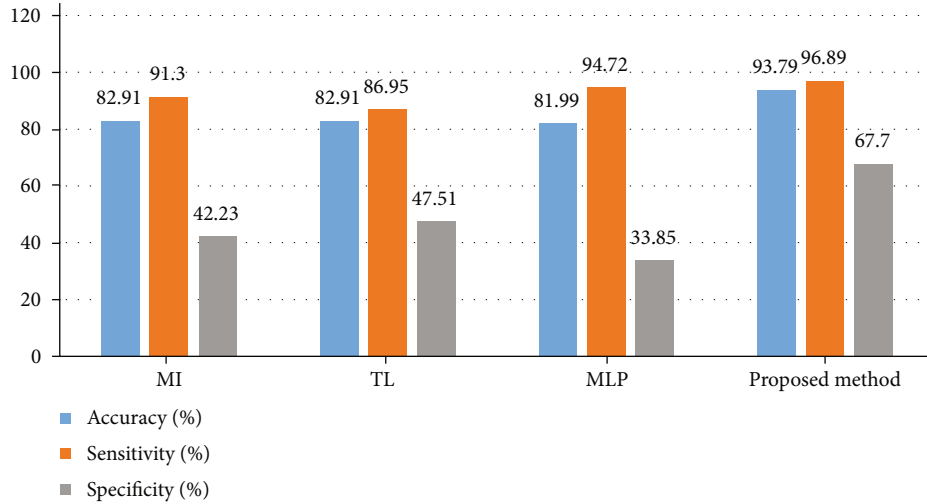
FIGURE 6: The comparison results between the suggested pipeline ATEO-based methodology and the mentioned methods applied to the MIAS database.

8.1. Dataset Description. To verify the accuracy and the ability of the suggested method, it is performed to a standard mammographic benchmark database, known Image xAnalysis Society Digital Mammogram Database (MIAS) [30]. The MIAS database is compiled by UK researchers to support researchers involved in working on mammogram images. The database includes 322 number of 1024×1024 digital mammography images that are taken from the UK National Breast Screening Program. The MIAS database also includes correct labels which are obtained with the help of experts. This database has been gotten available by the Pilot European Image Processing Archive (PEIPA) at the University of Essex. Figure 5 shows some examples of the MIAS database mammography images.

8.2. Simulations. By performing the discrete wavelet transform to the image and using decomposition of LL and HL characteristics from it, the GLCM characteristics are achieved by the wavelet decomposition's extracted levels. Then, the features are combined with the optimized CNN-based classifier that is arranged for the final detections. As mentioned before, five features including homogeneity (H), correlation (Cr), contrast (CN), energy (E), and entropy (ER) are employed to the LL and HL subband levels on the image. Table 4 illustrates the feature extraction for preparing data.

Also, Table 5 illustrates the feature extraction for the testing data.

For more clarification of the proposed automatic system, it is validated by three measurement indicators, precision, sensitivity, and specificity that are formulated in the following:

$$\text{Accuracy}(\%) = \frac{TP + TN}{TP + FP + FN + TN},$$

$$\text{Sensitivity}(\%) = \frac{TP}{TP + FN}, \quad (30)$$

$$\text{Specificity}(\%) = \frac{TN}{FP + TN},$$

where TN, TP, FN, and FP represent Truly Negative, Truly Positive, False Negative, and False Positive, respectively.

To proper validation of the suggested technique, it was compared with three latest techniques including Multilayer Perceptron (MLP) [31], Multiple Instance (MI) [32], and Transfer Learning (TL) [33]. Figure 6 shows the comparison results between the suggested pipeline Advanced Thermal Exchange Optimization algorithm-based methodology and the mentioned methods applied to the MIAS database.

Based on Figure 6, the suggested ATEO-based methodology with a 93.79% accuracy rate has the highest precision, and the method of MI, TL, and MLP with 82.91%, 82.91%, and 81.99% is placed in the later ranks. Furthermore, the specificity of the suggested method with 67.7% provides the best achievements than the other compared methods. Finally, the total achievements display optimal results for the suggested technique to automatic breast cancer diagnosis.

9. Conclusions

Breast cancer has been a cause of death in women in the last decade; the rate of breast cancer is increasing worldwide. This cancer is common cancer detected in women, and death from breast cancer is more common in women between the ages of 15 and 54. In recent years, much research was performed on mammographic images to be able to diagnose cancerous tumors ignoring the intervention of a person by image processing methods and computer programming. The present study presented a computer-aided diagnosis system for automatic detections of breast cancers. The mammogram images were first preprocessed based on image contrast enhancement and noise reduction to improve and prepare the image for the next steps. Afterward, a method based on color space was used for image segmentation that is followed by mathematical morphology. To achieve the main characteristics of the mammogram images, a combined gray-level cooccurrence matrix (GLCM) and discrete wavelet transform (DWT) was applied to the processed images. Finally, a new optimized version of convolutional

neural network (CNN) and a new improved metaheuristic, called Advanced Thermal Exchange Optimization algorithm, was applied for features' categorization. Simulation achievements of the suggested technique were finally compared with three other techniques including Multilayer Perceptron (MLP), Multiple Instances (MI), and Transfer Learning (TL) applied on the MIAS mammogram database to show its superiority.

Data Availability

The database for analysis is based on MIAS (the mammographic image analysis society digital mammogram database) which can be obtained as follows: http://peipa.essex.ac.uk/info/mias.html.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] N. Razmjooy, 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.

[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. 463–479.e10, 2018.

[3] N. Razmjooy, B. Somayeh Mousavi, F. Soleymani, and M. Hosseini Khotbesara, "A computer-aided diagnosis system for malignant melanomas," *Neural Computing and Applications*, vol. 23, no. 7-8, pp. 2059–2071, 2013.

[4] 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.

[5] M. Toğaçar, K. B. Özkurt, B. Ergen, and Z. Cömert, "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 123592, 2020.

[6] E. D. Carvalho, A. O. C. Filho, R. R. V. Silva et al., "Breast cancer diagnosis from histopathological images using textural features and CBIR," *Artificial Intelligence in Medicine*, vol. 105, article 101845, 2020.

[7] 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.

[8] 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.

[9] N. Razmjooy, B. S. Mousavi, B. Sadeghi, and M. Khalilpour, "Image thresholding optimization based on imperialist competitive algorithm," in *3rd Iranian Conference on Electrical and Electronics Engineering (ICEEE2011)*, Iran: Islamic Azad University of Gonabad, 2011.

[10] M. Tarafdar 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.

[11] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.

[12] 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, 2019.

[13] N. Ghadimi, "A method for placement of distributed generation (DG) units using particle swarm optimization," *International Journal of Physical Sciences*, vol. 8, no. 27, pp. 1417–1423, 2013.

[14] N. Ghadimi, A. Afkousi-Paqaleh, and A. Emamhosseini, "A PSO-based fuzzy long-term multi-objective optimization approach for placement and parameter setting of UPFC," *Arabian Journal for Science and Engineering*, vol. 39, no. 4, pp. 2953–2963, 2014.

[15] M. Saeedi, M. Moradi, M. Hosseini, A. Emamifar, and N. Ghadimi, "Robust optimization based optimal chiller loading under cooling demand uncertainty," *Applied Thermal Engineering*, vol. 148, pp. 1081–1091, 2019.

[16] 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.

[17] 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.

[18] D. Simon, "Biogeography-based optimization," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 6, pp. 702–713, 2008.

[19] G.-G. Wang, S. Deb, and L. D. S. Coelho, "Elephant herding optimization," in *2015 3rd international symposium on computational and business intelligence (ISCBI)*, Bali, Indonesia, 2015.

[20] M. Mani, O. Bozorg-Haddad, and X. Chu, "Ant lion optimizer (ALO) algorithm," in *Advanced Optimization by Nature-Inspired Algorithms*, O. Bozorg-Haddad, Ed., vol. 720 of Studies in Computational Intelligence, pp. 105–116, Springer, Singapore, 2018.

[21] A. Faramarzi, M. Heidarinejad, B. Stephens, and S. Mirjalili, "Equilibrium optimizer: a novel optimization algorithm," *Knowledge-Based Systems*, vol. 191, article 105190, 2020.

[22] R. Navid, V. V. Estrela, R. Padilha, and A. C. B. Monteiro, "World Cup Optimization Algorithm: An Application for Optimal Control of Pitch Angle in Hybrid Renewable PV/Wind Energy System," in *Metaheuristics and Optimization in Computer and Electrical Engineering*, pp. 25–47, Springer, Cham, 2021.

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

[24] 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 Reports*, vol. 6, pp. 885–894, 2020.

[25] R. Rao and K. More, "Conception optimale et analyse d'une tour de refroidissement a convection mecanique grace a un

algorithme Jaya ameliore," *International Journal of Refrigeration*, vol. 82, pp. 312–324, 2017.

[26] J. J. Liang, B. Y. Qu, and P. N. Suganthan, *Problem Definitions and Evaluation Criteria for the CEC 2014 Special Session and Competition on Single Objective Real-Parameter Numerical Optimization*, Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou China and Technical Report, Nanyang Technological University, Singapore, 2013.

[27] A. Hatamlou, "Black hole: a new heuristic optimization approach for data clustering," *Information Sciences*, vol. 222, pp. 175–184, 2013.

[28] 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.

[29] 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.

[30] J. Suckling, "The mammographic image analysis society digital mammogram database," *Digital Mammo*, pp. 375–386, 1994.

[31] M. Mehdy, P. Y. Ng, E. F. Shair, N. I. M. Saleh, and C. Gomes, "Artificial neural networks in image processing for early detection of breast cancer," *Computational and Mathematical Methods in Medicine*, vol. 2017, Article ID 2610628, 15 pages, 2017.

[32] P. Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte, and P. Honeine, "Multiple instance learning for histopathological breast cancer image classification," *Expert Systems with Applications*, vol. 117, pp. 103–111, 2019.

[33] E. Deniz, A. Şengür, Z. Kadiroğlu, Y. Guo, V. Bajaj, and Ü. Budak, "Transfer learning based histopathologic image classification for breast cancer detection," *Health Information Science and Systems*, vol. 6, no. 1, pp. 1–7, 2018.

Research Article

A Novel Method for Differential Prognosis of Brain Degenerative Diseases Using Radiomics-Based Textural Analysis and Ensemble Learning Classifiers

Manju Jain [1,2] **C. S. Rai** [1] **and Jai Jain** [3]

^{1}*University College of Information, Communication and Technology, Guru Gobind Singh Indraprastha University,
Dwarka Sector 16-C, New Delhi 110078, India*
^{2}*Meerabai Institute of Technology Maharani Bagh, New Delhi 110065, India*
^{3}*Media Agility India Ltd, New Delhi, India*

Correspondence should be addressed to Manju Jain; manju.jain69@gmail.com

We propose a novel approach to develop a computer-aided decision support system for radiologists to help them classify brain degeneration process as physiological or pathological, aiding in early prognosis of brain degenerative diseases. Our approach applies computational and mathematical formulations to extract quantitative information from biomedical images. Our study explores the longitudinal OASIS-3 dataset, which consists of 4096 brain MRI scans collected over a period of 15 years. We perform feature extraction using Pyradiomics python package that quantizes brain MRI images using different texture analysis methods. Studies indicate that Radiomics has rarely been used for analysis of brain cognition; hence, our study is also a novel effort to determine the efficiency of Radiomics features extracted from structural MRI scans for classification of brain degenerative diseases and to create awareness about Radiomics. For classification tasks, we explore various ensemble learning classification algorithms such as random forests, bagging-based ensemble classifiers, and gradient-boosted ensemble classifiers such as XGBoost and AdaBoost. Such ensemble learning classifiers have not been used for biomedical image classification. We also propose a novel texture analysis matrix, Decreasing Gray-Level Matrix or DGLM. The features extracted from this filter helped to further improve the accuracy of our decision support system. The proposed system based on XGBoost ensemble learning classifiers achieves an accuracy of 97.38%, with sensitivity 99.82% and specificity 97.01%.

1. Introduction

Medical image processing has travelled a long journey since the last two decades. The past decade has seen the bridging of medical and information technology. It led to the development of decision support systems for early identification of various brain diseases. Age and structural changes in brain cause physiological alterations, which are reflected in routine human behaviour [1, 2]. Along the years, various studies and constant attempts have been made to study dementia.

Studies [3–5] focus on specific regions of interest in brain volumes, and these are calculated from two dimensional manually traced areas. Segmentation algorithms are used to segment out gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Such volumetric studies are limited to known brain structures like hippocampus and amygdala, perirhinal, entorhinal, and parahippocampal cortex.

Voxel-based studies [6–8] provide an alternative neuro-imaging method. These studies apply a general linear model (GLM) to each voxel of an MRI and statistically compare them with standard voxel values using Jacobean matrices.

Many studies [9, 10] give detailed insights on comparisons between voxel based and volumetric studies.

Several studies [11–13] use cortical thickness measurement as a biomarker for the process of identification of various brain aging diseases.

With the advancements in the machine learning techniques for image processing and image analysis and the

availability of abundance of medical imaging data, medical informatics [14, 15] has achieved great heights. The workshop MICCAI 2014 "Challenges of Computer aided diagnosis of Dementia on Structural MRI data" addresses the challenges of applying different algorithms on the same data and the same algorithm on different data. A summary of all algorithms presented in MICCAI 2014 is listed in [16]. This paper did a standardized comparison of different studies in the domain of the computer-aided decision support system for the identification of dementia-related diseases using structured MRI data. The best performing algorithm yielded an accuracy of 63% and receiver operating characteristics area under the curve with value 78.8%.

A review of various studies used for brain disorder detection using the machine learning techniques is published in [17]. Another review of the latest image processing techniques for studying brain pathology is summarized in [18].

A set of studies have been done on how oxygen supply changes the brain functioning [19, 20].

The functional modalities of medical imaging include MRI (magnetic resonance imaging), PET (positron emission tomography), and CT (computerized tomography) giving us an insight about the pathophysiology of the organ under observation. Radiologists analyze this information with their experience and knowledge. They find this time consuming and cumbersome. In this study, we explore machine learning techniques to analyze data extracted from medical images. Machine learning is the study of algorithms that solve a problem by leaning from underlying patterns in data, as opposed to statistical heuristics or rule-based programming. Radiomics [21, 22] aids in extracting imaging-based statistical biomarkers from medical images which can be used as features for machine learning methods to get accurate predictions. Ageing leads to degeneration of the brain, which may lead to dementia, further precipitating such diseases like Alzheimer's dementia, vascular dementia, dementia with lewy body dementia, posterior cortical atrophy, and front temporal lobar degeneration. These diseases affect different regions of the brain. Clinical Dementia Rating or CDR is a five-point scale to stage dementia, ranging from 0 to 3, where 0 denotes no pathological degeneration (control patients) while any value greater than 0 indicates some pathological brain degeneration (test patients).

In this study, we propose a novel approach to develop a computational decision support system capable of differentiating control patients from test patients by analyzing features of their MRI images using Radiomics. This system can be used to assist radiologists for fast and accurate decisions.

(1) We explore the OASIS-3 dataset [23], which is a longitudinal dataset with 4096 MRI scans. This dataset also gives specific details about how the CDR value changes for a subject with respect to changes in the subject's MRI scan. These ratings can be used to label the MRI scans as healthy scans or scans showing signs of brain degeneration. Using these labels for a scan, a supervised machine learning binary classifier can be trained to support brain degeneration prognosis
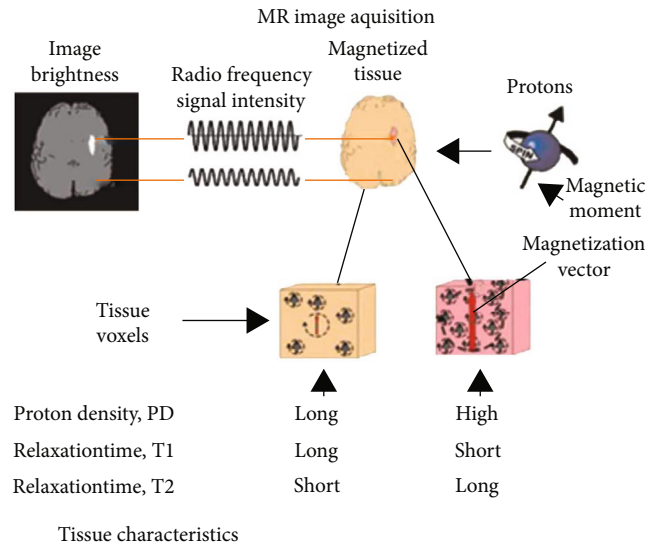
Figure 1: MRI acquisition process.

(2) We employ data preprocessing best practices such as data augmentation and feature selection which help to mitigate overfitting and underfitting of the classifier and drive it to achieve optimal accuracy on our data

(3) Feature extraction is done using Pyradiomics, which provides a python implementation of the study [24]. Pyradiomics provides a unified and standardized set of features from structured MRIs based on shape and volume as well as texture-based statistical features. Advanced Pyradiomics algorithms can handle missing data in case of low resolution MRI scans. Literature studies indicate that Radiomics has mostly been explored for oncological studies [25, 26], but not for understanding brain cognition. Our study is also an effort to determine the efficiency of Radiomics features from structural MRI scans for classification of brain degeneration diseases

(4) We explored various ensemble learning classification algorithms such as random forests, bagging-based ensemble classifiers, and gradient-boosted ensemble classifiers such as XGBoost and AdaBoost for our classification tasks. Such ensemble learning classifiers have not been used for biomedical image classification

(5) We propose a novel image texture analysis filter, Decreasing Gray-Level Matrix, which further improves the performance of our ensemble learning classifiers

We conclude the paper by comparing our novel solution with existing work in this field. Our results show that the proposed solution outperforms existing studies on various performance metrics such as accuracy, specificity, and sensitivity.

2. Materials and Methods

2.1. Data Acquisition. Magnetic resonance imaging is the process of acquiring images of anatomical structures using
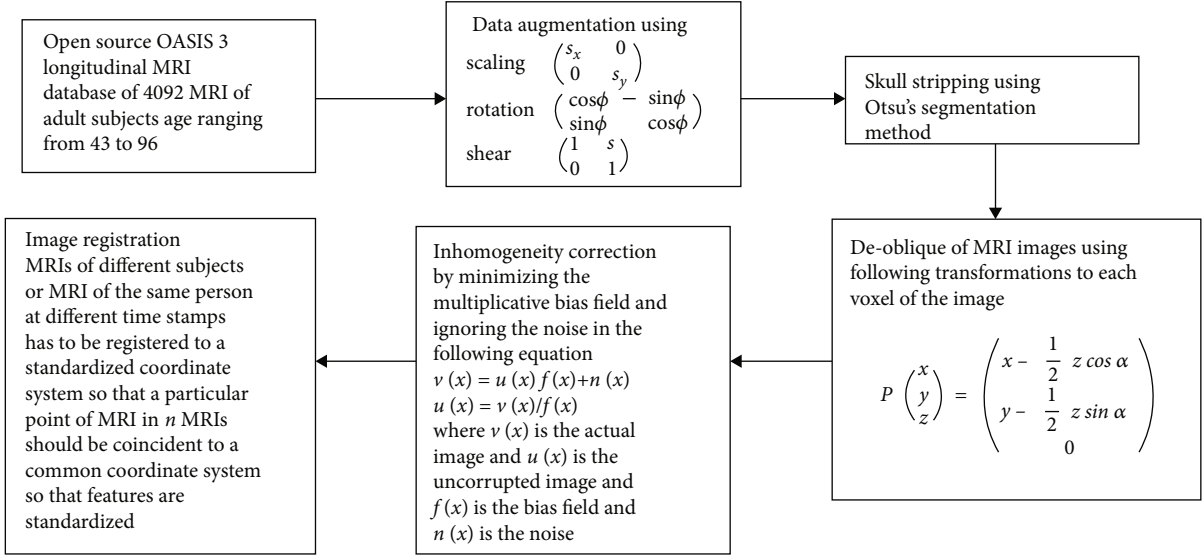
Figure 2: Data preprocessing flow.

Spatial characteristics
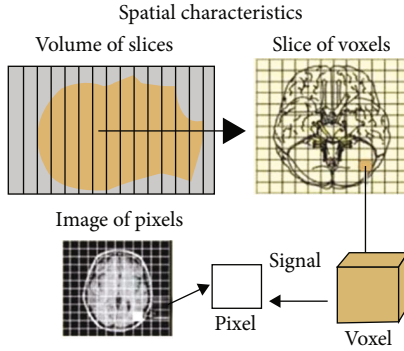
Figure 3: Spatial characteristics of MRI.
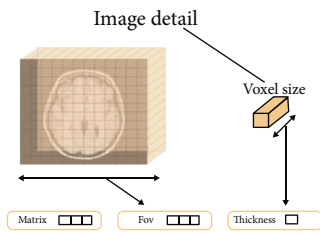
Figure 4: Voxel details.

magnetic field and radio frequency signals to detect diseases and functional problems. The "image is snapped" with different contrasts as different tissues and fluids react differently to magnetization signals. Tissue demagnetization time is different for different tissues. These times are identified as T1 and T2. Another characteristic of a tissue that affects an MRI is its proton density known as PD. Figure 1 depicts the complete MRI acquisition process.

In our study, we used the latest OASIS-3 dataset [23], which is an open source brain MRI database published in 2019. Most of the earlier studies have been done using ADNI datasets, which are cross-sectional datasets and do not include more than 500 subjects. OASIS-3 is the largest longitudinal dataset of longitudinal MRI images that consists of 1068 subjects (age group of 46 to 95), collected over a period of 15 years.

"The CDR is a 5-point scale used to characterize six domains of cognitive and functional performance applicable to Alzheimer disease and related dementias: Memory, Orientation, Judgment & Problem Solving, Community Affairs, Home & Hobbies, and Personal Care. The necessary information to make each rating is obtained through a semi-structured interview of the patient and a reliable informant or collateral source (e.g. family member)" [27].

The OASIS database also provides CDR for each subject. The CDR values of a person over a particular period of time may or may not be the same. There are multiple scans of the same subject (4-5 times) in the time period of 15 years with different CDR values. These scans can be further used as samples. Hence, the database has more than 4000 MRIs.

2.2. Data Preprocessing. We performed data preprocessing using Python and FreeSurfer [28]. Main steps of data preprocessing are listed below and more visually shown in Figure 2.

2.2.1. Data Augmentation. We augmented our data to make our classifier much more tolerant towards variance in the data (prevents overfitting) and to increase dataset size (prevents underfitting). We employed 4 augmentation techniques:

(1) *Flips*. Each image is flipped horizontally as well as vertically.

(2) *Scaling*. Each image is scaled in either "x" or "y" direction with the help of a transform matrix $\begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix}$.
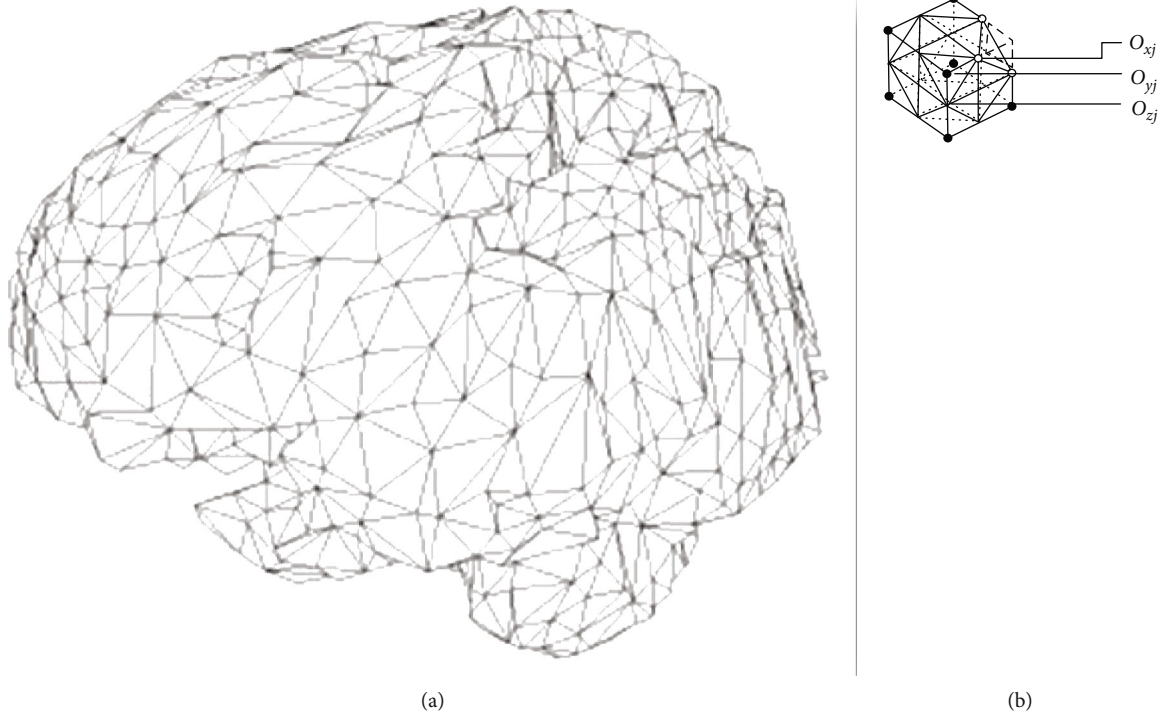
Figure 5: (a) Decimation of triangles as a mesh of the whole brain. (b) The trapezium points and edges. X_v is the number of voxels included in the masked region. V_m is the volume of the mesh mm^3. A_m the surface area of the mesh in mm^2.
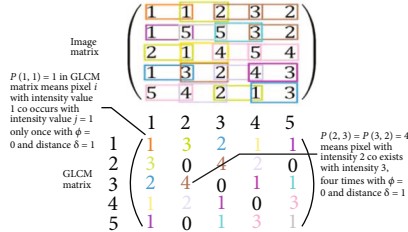
Figure 6: GLCM procedure. Different color schemes were used to track the steps.
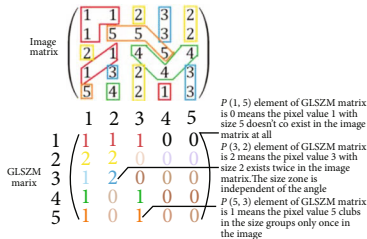
Figure 7: GLSZM procedure. Different color schemes were used to track the steps.

(3) *Rotations.* Affine transform matrix $\begin{pmatrix} \cos\varnothing & -\sin\varnothing \\ \sin\varnothing & \cos\varnothing \end{pmatrix}$ gives rotated MRI images in different directions. "$\varnothing$" was varied between 25 and 195.

Figure 8: GLRLM procedure. Different color schemes were used to track.

(4) *Shears.* Affine transform matrix $\begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$ applied to each image where Shear value changes from 0.3 to 0.7.

2.2.2. De-Oblique. During the MRI process, the subject's head may be tilted from to cover the whole brain or to avoid artefacts caused by water and air in the nose and eyes of the subject. This causes the MRI to be oblique and makes intersubject or intrasubject registration more difficult. The MRI images in our dataset were de-obliqued using the Free-Surfer software.

2.2.3. Inhomogeneity Correction. Brain consists of different types of tissues like gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), and all these tissues have different range of penetration to the magnetic field and may result into very bright or very dull artefacts in the MRI images. This may confuse a radiologist since all tissues of a

$$\begin{pmatrix} 1 & 1 & 2 & 3 & 2 \\ 1 & 5 & 5 & 3 & 2 \\ 2 & 1 & 4 & 5 & 4 \\ 1 & 3 & 2 & 4 & 3 \\ 5 & 4 & 2 & 1 & 3 \end{pmatrix}$$

Sum of absolute differences for pixel with level 1 =
|1-7/3|+|1-10/5|+|1-14/5|+|1-23/8|+|1-14/5|+|1-15/5| = 90775

(a)

$$\begin{pmatrix} 1 & 1 & 2 & 3 & 2 \\ 1 & 5 & 5 & 3 & 2 \\ 2 & 1 & 4 & 5 & 4 \\ 1 & 3 & 2 & 4 & 3 \\ 5 & 4 & 2 & 1 & 3 \end{pmatrix}$$

Sum of absolute differences for pixel with grey level 2
= |2-11/5|+|2-17 /5|+|2-24/8|+|2-17/5|+|2-14/5| = 4.8

(b)

$$\begin{pmatrix} 1 & 1 & 2 & 3 & 2 \\ 1 & 5 & 5 & 3 & 2 \\ 2 & 1 & 4 & 5 & 4 \\ 1 & 3 & 2 & 4 & 3 \\ 5 & 4 & 2 & 1 & 3 \end{pmatrix}$$

Sum of absolute differences for pixel with grey level 3
= |3-14/5|+|3-21/8|+|3-8/3|+|3-17/5|+|3-27/8| = 1.683

(c)

$$\begin{pmatrix} 1 & 1 & 2 & 3 & 2 \\ 1 & 5 & 5 & 3 & 2 \\ 2 & 1 & 4 & 5 & 4 \\ 1 & 3 & 2 & 4 & 3 \\ 5 & 4 & 2 & 1 & 3 \end{pmatrix}$$

Sum of absolute differences for pixel with grey level 4
= |4-28/8|+|4-17/5|+|4-24/8|+|4-13/3| = 3.5

(d)

$$\begin{pmatrix} 1 & 1 & 2 & 3 & 2 \\ 1 & 5 & 5 & 3 & 2 \\ 2 & 1 & 4 & 5 & 4 \\ 1 & 3 & 2 & 4 & 3 \\ 5 & 4 & 2 & 1 & 3 \end{pmatrix}$$

Sum of absolute differences for pixel with grey level 5
= |5-17/8|+|5-27/8|+|5-23/8|+|5-8/3| = 8.958

(e)

FIGURE 9: (a) NGTDM for neighbours of 1. Different color schemes were used to track NGTDM procedure to calculate the absolute sum of gray-level difference. (b) NGTDM for neighbours of 2. Different color schemes were used to track NGTDM procedure to calculate the absolute sum of gray-level difference. (c) NGTDM for neighbours of 3. Different color schemes were used to track NGTDM procedure to calculate absolute sum of gray-level difference. (d) NGTDM for neighbours of 4. Different color schemes were used to track NGTDM procedure to calculate the absolute sum of gray-level difference. (e) NGTDM for neighbours of 5. Different color schemes were used to track NGTDM procedure to calculate absolute sum of gray-level difference.

particular type should have exact intensity and brightness values. The process of correcting this is known as inhomogeneity correction.

2.2.4. Skull Stripping. The nonbrain parts (skull, neck, eyes, and nose) were removed from all MRI images to have a uniform area of study.

2.2.5. Registration. The brain consists of very fine spatial structures, due to which it is very difficult to extract and integrate the information from different images. The thickness of the cortex can be as small as 5 mm. The thickness of thalamic nuclei only extends to few millimetres. Registration is the process of aligning different MRI images in such a way that the voxels of a particular tissue from all of those images correspond to the same 3D location. We applied and adjusted the registration parameters, i.e., translations, rotations, scaling, and shear operations at voxel level to make the MRI images concurrent.

2.3. Feature Extraction

2.3.1. Features Based on Shape. To extract features based on shape, we studied spatial characteristics of MRIs as depicted in Figure 3.

Slice: an MRI is a 3D image. It consists of a set of contiguous 2D slices. These slices may either represent the axial, sagittal, or longitudinal cross section of the subject's brain.

Voxel: each slice is subdivided into rows and columns. The intersection of each row and column represents a

volume of the brain. This is known as a voxel. The field-of-view matrix of a particular size of the slice is used to determine the voxel size. The voxel details are depicted in Figure 4.

Shape features include legends of 3D size and shape. We took the whole brain area and volume as our region of interest. A triangular mesh encapsulating the whole brain area was used to extract various shape features. Figures 5(a) and 5(b) show how a brain is treated as a mesh surface [28]. The mesh has X_f number of triangles.

From this meshed surface of brain, as in Figure 5(b), we calculated the following different shape features [24].

(i) Mesh volume $V_j = O_{xi}.(O_{yj} \times O_{zj})/6$ (O_{xi}, O_{yj}, O_{zj} are the tetrahedral vertices)

$$V_m = \sum_{j=1}^{X_f} V_j. \tag{1}$$

(ii) *Voxel Volume.* $\text{Voxel}_{\text{volume}} = \sum_{j=1}^{X_v} V_m$

The whole brain volume can be obtained by multiplying the voxel volume V_j with the number of voxels in the brain.

(i) *Surface Area.* $A_J = 1/2(X_j Y_j \times X_j Z_j)$

NGTDM matrix with i grey levels and n_i pixels of the count of the ith grey level and with p_i probability and si as the sum of absolute differences for that particular grey level

$$
\begin{array}{cccc}
i & n_i & p_i & S_i \\
1 & 6 & 0.24 & 9.775 \\
2 & 6 & 0.24 & 4.8 \\
3 & 5 & 0.20 & 1.683 \\
4 & 4 & 0.16 & 3.5 \\
5 & 4 & 0.16 & 8.958
\end{array}
$$

FIGURE 10: NGTDM.

$P(1, 0)$ means grey level i with 0 dependent voxels which is none. so $P(1, 0) = 0$. $P(1, 1) = 1$ as there is only single 1 with one dependent voxel marked with red. $P(1, 2) = 2$ as there are two 1s which have 2 dependent voxels marked with green. $P(1, 3) = 1$ as there is one voxel with grey vaue 1 and three dependent voxels marked with yellow

Image matrix
$$
\begin{pmatrix}
5 & 2 & 5 & 4 & 4 \\
3 & 3 & 3 & 1 & 3 \\
2 & 1 & 1 & 1 & 3 \\
4 & 2 & 2 & 2 & 3 \\
3 & 5 & 3 & 3 & 2
\end{pmatrix}
$$

GLDM matrix P
$$
\begin{array}{c|ccccc}
 & 0 & 1 & 2 & 3 & \dots \\
1 & 0 & 1 & 2 & 1 \\
2 & 1 & 2 & 3 & 0 \\
3 & 1 & 4 & 4 & 0 \\
4 & 1 & 2 & 0 & 0 \\
5 & 3 & 0 & 0 & 0
\end{array}
$$

FIGURE 11: GLDM. Different color schemes were used to track GLDM procedure.

$$A_m = \sum_{j=1}^{X_f} A_j. \tag{2}$$

To calculate the surface area of the whole brain, it is divided into small mesh areas. We first calculate the surface area of each mesh and then sum all of them.

(i) *Ratio of the Surface Area to the Volume of the Brain.* A_m/V_m

(ii) Lower the ratio more the compactness

(iii) *Maximum 3D Diameter.* It is the largest Euclidean distance on the various mesh surfaces on the whole brain.

(iv) *The Maximum 2D Diameter of the Slice.* It is the defined as the largest Euclidean distance on the whole brain mesh surfaces where mesh vertices are in the axial plane.

(v) *Major Axis length.* $4\sqrt{\gamma_{\text{major}}}$

(vi) This feature calculates the largest axis length of the whole brain area

(vii) *Minor Axis Length.* $4\sqrt{\gamma_{\text{minor}}}$

(viii) This feature represents the minimum axis length of the whole brain area

(ix) *Elongation.* This feature gives the relationship between the largest and smallest component of the whole brain.

(x) Elongation $= \sqrt{\gamma_{\text{minor}}/\gamma_{\text{major}}}$

(xi) Flatness $= \sqrt{\gamma_{\text{least}}/\gamma_{\text{major}}}$

Image matrix
$$
\begin{pmatrix}
1 & 1 & 2 & 3 & 5 \\
1 & 5 & 5 & 3 & 4 \\
2 & 1 & 4 & 5 & 4 \\
1 & 3 & 2 & 4 & 3 \\
4 & 5 & 2 & 1 & 3
\end{pmatrix}
$$

The element $P(i, j)$ represents the count of pixels where particular pixel grey value i is coexisting with the pixel with grey value j such that $i \leq j$. The $P(1, 1) = 1$, $P(1, 2) = 1$, $P(1, 3) = 2$, $P(1, 4) = 1$ and $P(1, 5) = 1$

DGLM
$$
\begin{pmatrix}
1 & 1 & 2 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
$$
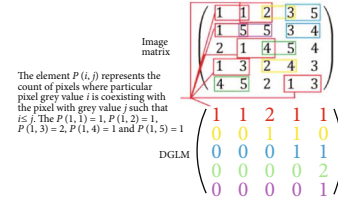
FIGURE 12: DGLM. Different color schemes were used to track DGLM procedure.

2.3.2. First-Order Features. These features are obtained by the statistical analysis of the whole brain based on values of voxel intensities [24].

Let S be a set of N_v voxels in the whole brain.

Let N_v be the discrete level of intensities in the whole brain then $X(i)$ is the first-order histogram.

The normalized first-order histogram $x(i) = X(i)/N_v$

(i) Energy $= \sum_{i=1}^{N_v} (S(i) + c)^2$

(ii) Total energy $= V_{\text{voxel}} \sum_{i=1}^{N_v} (S(i) + c)^2$

(iii) Entropy $= -\sum_{i=1}^{N_d} x(i) \log_2 (x(i) + \epsilon)$

(iv) Minimum $= \min (S)$

(v) 10th percentile of S

(vi) 90th percentile of S

(vii) Maximum $= \max (S)$

(viii) Mean $= 1/N_v \sum_{i=1}^{N_v} S(i)$ is the average gray-level intensity of the whole brain

(ix) Median $=$ the median gray level of the whole brain

(x) Range $= \max (S) - \min (S)$

(xi) Absolute mean deviation $= \sqrt{1/N_v \sum_{i=1}^{N_v} |S(i) - \bar{S}|}$

(xii) Root mean square value of the whole brain (RMS)
$$= \sqrt{1/N_v \sum_{i=1}^{N_v} (S(i) + c)^2}$$

(xiii) Standard deviation $= \sqrt{1/N_v \sum_{i=1}^{N_v} (S(i) - \bar{S})^2}$

(xiv) Skewness $= \sqrt{1/N_v \sum_{i=1}^{N_v} (S(i) - \bar{S})^3} /$
$$\left(\sqrt{1/N_v \sum_{i=1}^{N_v} (S(i) - \bar{S})^2}\right)^3 = \mu_3/\sigma^3$$

(xv) Kurtosis $= \sqrt{1/N_v \sum_{i=1}^{N_v}} (S(i) - \acute{S})^4 /$
$$\left(\sqrt{1/N_v \sum_{i=1}^{N_v} (S(i) - \bar{S})^2}\right)^2 = \mu_4/\sigma^4$$

(xvi) Variance $= 1/N_v \sum_{i=1}^{N_v} |S(i) - \bar{S}|$

2.3.3. Gray-Level Cooccurrence Matrix [24]. GLCM is a texture filter that gives the pixel distribution of a particular set of pixels i, j in a specific direction and distance. The

(a)

(b)

(c)

(d)
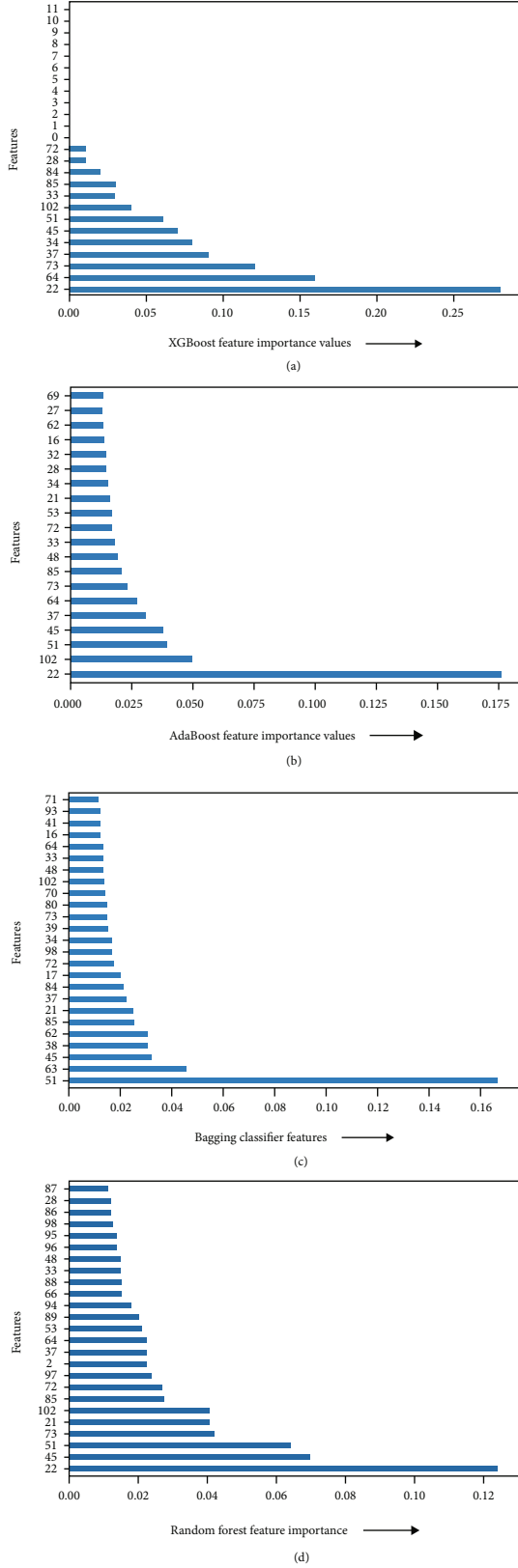
FIGURE 13: (a) Feature importance prediction by XGBoost. (b) Feature importance prediction by AdaBoost. (c) Feature importance prediction by bagging. (d) Feature importance prediction by random forest.

TABLE 1: Sensitivity, specificity, and accuracy comparison of different ensemble classifiers.

Classifiers	Sensitivity	Specificity	Accuracy
XGBoost	99.82%	97.01%	97.38%
AdaBoost	94.91%	97.76%	97.21%
Bagging classifier	74.22%	90.07%	87.56%
Random forest	94.44%	87.07%	87.72%

$p(i, j) | \varnothing, \delta|$ value of GLCM represents the number of times; the pixel with intensity i coexists with intensity j with angle $\varnothing$ and distance δ. Figure 6 shows how GLCM can be obtained from image matrix. The different color schemes indicate a particular pixel's coexistence. Generally, the following statistical features are extracted and then averaged over GLCM for each direction (angle).

   (i) Autocorrelation

   (ii) Joint average

   (iii) Entropy

   (iv) Variance

   (v) Contrast

   (vi) Energy

   (vii) Homogeneity

   (viii) Inverse of the difference movement

   (ix) Inverse variance

2.3.4. Gray-Level Size Zone Matrix [24]. GLSZM quantifies different pixel intensity values in different size zones. A size zone is defined as connected pixels/voxels with the same gray-level irrespective of direction. The $P(i, j)$ element of GLSZM represents the number of times the intensity value i of the size zone j exists in the image matrix. Figure 7 depicts how GLSZM can be obtained from the image matrix. Different colors indicate different size zones of different intensity values.

GLSZM can be used to extract the following features:

   (i) Emphasis on small areas

   (ii) Emphasis on large areas

   (iii) Gray-level nonuniformity

   (iv) Normalized gray-level nonuniformity

   (v) Size zone nonuniformity

   (vi) Normalized size zone nonuniformity

   (vii) Low gray-level emphasis on small areas

   (viii) High gray-level emphasis on small areas

   (ix) Low gray-level emphasis on large areas

   (x) High gray-level emphasis on large areas

(a)

(b)

(c)

(d)
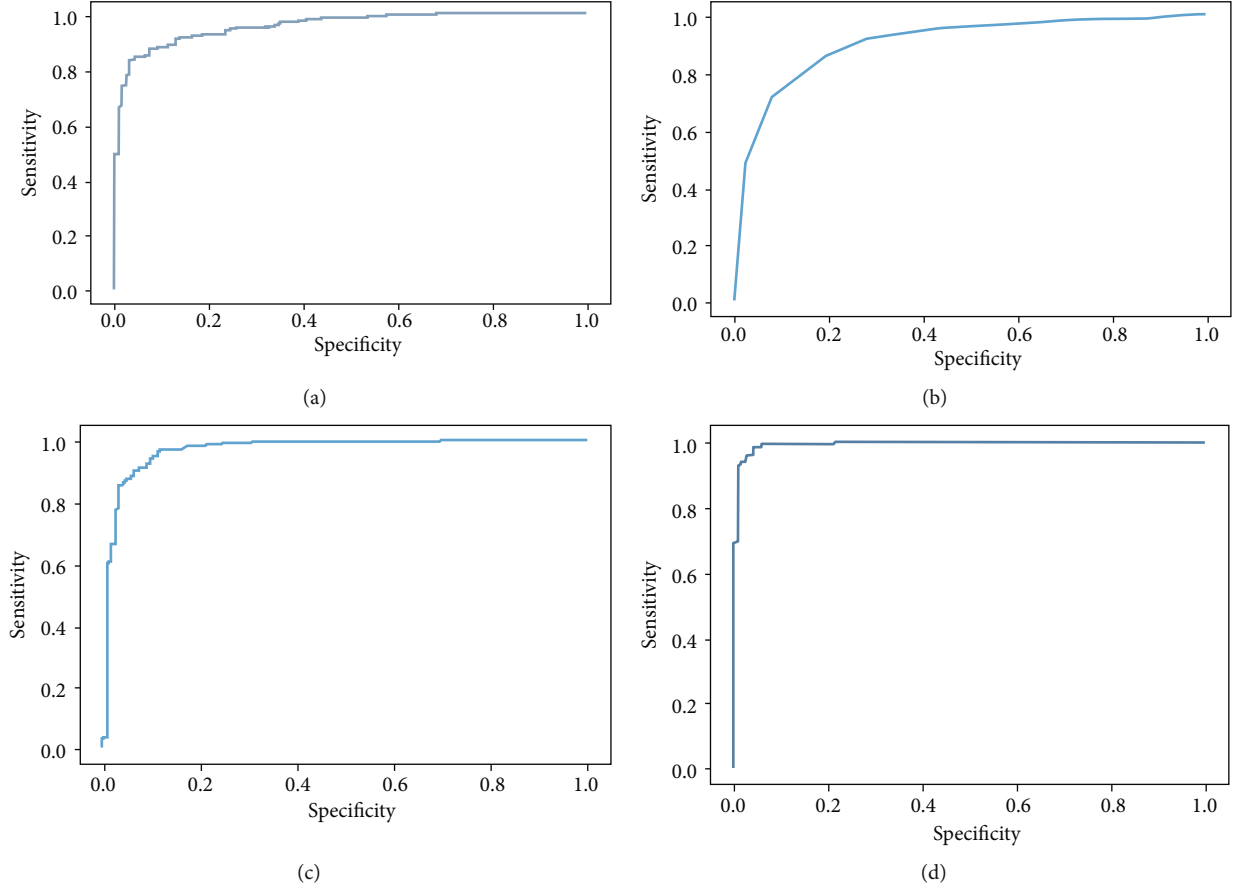
FIGURE 14: Area under the accuracy curve for different ensemble classifiers.

2.3.5. Gray-Level Run Length Matrix (GLRLM) [24]. The intensity runs in a GLRLM are defined as the length of connected pixels of equal intensity values and in a particular direction. A GLRLM element $P(i, j) \mid \varnothing$ represents the number of times a particular run length j of intensity i in direction $\varnothing$ occurs in the image matrix. Figure 8 depicts how GLRLM can be obtained from image matrix. Different colors indicate different run lengths of the particular length in a particular direction.

The GLRLM is used to extract following features:

(i) Emphasis on short run

(ii) Emphasis on long run

(iii) Nonuniform gray level

(iv) Normalized nonuniform gray level

(v) Nonuniform run length

(vi) Normalized nonuniform run length

(vii) Run percentage

(viii) Variance of gray level

(ix) Run variance

(x) Run entropy

(xi) Low gray-level run emphasis

(xii) High gray-level emphasis

(xiii) Short run low gray-level emphasis

(xiv) Short run high gray-level emphasis

(xv) Long run low gray-level emphasis

(xvi) Long run high gray-level emphasis

2.3.6. Neighbouring Gray Tone Difference Matrix [24]. Here, we consider neighbouring pixels of a particular pixel at a distance ∂ of that pixel. This matrix is the set of absolute differences of the gray levels of the voxel and its neighbouring voxels. Let P_{nv} be the set of whole brain voxels; then, $p_{nv}(i_x, i_y, i_z)$ belongs to P_{nv} where p_{nv} denotes the gray level of the voxel at position (i_x, i_y, i_z). The average gray level of the neighbourhood is given as follows: $\bar{G}_j = \bar{G}(i_x, i_y, i_z) = 1/v\sum_{k_x=-\partial}^{\partial}\sum_{k_y=-\partial}^{\partial}\sum_{k_z=-\partial}^{\partial}p_{nv}(i_x + k_x, i_y + k_y, i_z + k_z)$ where V is total number of voxels in the whole brain.

(i) Let i denote the value of gray levels in the image

(ii) Let n_i denote the number of voxels of gray level i

(iii) Let p_i denote the gray-level probability

TABLE 2: Comparison with similar studies.

Research study	Year	Dataset	Brain area	Classifier	Accuracy
Ahmad Chaddad [32]	2018	OASIS-1	Hippocampus	Random forest	84.09%
			Amygdala	Random forest	
				CNN	92.5%
Feng Feng [33]	2018	Local hospital data	Hippocampus	SVM	86.75%
Yupeng Li and Jiehui Jiang [34]	2019	Local hospital data	Hippocampus	SVM	91.5%
Kun Zhao [35]	Jan 2020	ADNI	Hippocampus	SVM	88.21%
Tao-Ran Li [36]	Dec 2020	ADNI	Right posterior and left superior cingulate Gyrus	SVM	95.9%
Current study proposed by us	2021	OASIS-3	Whole brain	Ensemble classifiers	
				XGBoost	97.38%
				AdaBoost	97.21%
				Bagging	87.56%
				Random forest	87.72%

(iv) Let s_i be the sum of absolute difference of a gray level i

Figures 9(a)–9(d) are the required NGTDM for the pixel with intensities 1-4. Figure 9(e) describes how the absolute difference of the different gray levels is calculated. Different colors are used to track down the neighbours of a particular gray level as shown in the following example where we have 5 discrete gray levels 1 to 5. Figure 10 is the final NGTDM.

Features calculated from NGTDM are as follows:

(i) The neighbourhood-based coarseness

(ii) Neighbourhood-based contrast

(iii) Rate of change of gray levels within voxels

(iv) The complexity of neighbourhood gray levels

(v) Strength of neighbourhood gray levels

2.3.7. Gray-Level Dependence Matrix [24]. GLDM represents the dependencies of one gray level on other gray levels. It is defined as a set of connected voxels within distance ∂ dependent on a central voxel. A voxel with a gray level i is dependent on another voxel of gray level j if

$$|i - j| \leq \gamma. \tag{3}$$

The $(i, j)^{\text{th}}$ element of GLDM $P(i, j)$ represents how often a voxel with the gray value i coexists with its dependent voxel having gray level j occurs in the whole brain image. Figure 11 describes how GLDM is obtained from brain MRI with $n = 5$, i.e., 5 discrete gray levels, $\gamma = 0$, and $\partial = 1$. The GLDM columns start from 0, and it can go to any finite number of dependent voxels.

The above GLDM is used to extract the following features:

(i) Small dependence significance

(ii) Large dependence significance

(iii) Gray-level heterogeneity

(iv) Dependence heterogeneity

(v) Dependence heterogeneity normalized

(vi) Gray-level deviation

(vii) Dependence deviation

(viii) Entropy of dependency

(ix) The low gray-level significance

(x) The high gray-level significance

(xi) Small dependency and low gray-level significance

(xii) Small dependency and high gray-level significance

(xiii) Large dependency and low gray-level significance

2.3.8. Decreasing Gray-Level Matrix (Novel Filter). We propose a novel filter matrix to improve feature set. The $p(i, j) | \varnothing, \delta$ pixel of DGLM represents the occurrence of the pixel with intensity i and pixel with intensity j such that $i \leq j$. Figure 12 depicts and obtains a DGLM from the image with $\varnothing = 0$ and $\delta = 1$. Colors are used to track down the location of pixels for which the condition $i < j$ holds true.

The DGLM is used to extract the following features in four directions, i.e., 0, 45, 90, and 135. Then, the average is taken to get the summary of the following features:

(i) Energy

(ii) Mean

(iii) Absolute mean deviation

(iv) Skewness

(v) Kurtosis

(vi) Entropy

(vii) Autocorrelation

2.4. Feature Selection. Feature selection is the process of evaluating and selecting the most important features from the set of all features depending on their contribution to the machine learning task at hand. This process helped us to select the features with the highest predictive relevance to our classification task. This in turn also helps to eliminate redundant features.

In our study, we focused on tree-based classification methods. These methods have intrinsic feature selection methods. Using these intrinsic methods, we found the feature relevance for that each classifier.

Figure 13(a) denotes that the novel feature "first order mean of DLGM" has the highest predictive power for XGBoost classifier; hence, it is the most important feature for this classifier. The other important features for XGBoost classifier are as follows:

(i) Decreasing Gray-Level Matrix feature first-order mean 0.26

(ii) Gray-Level Dependence Matrix feature high gray-level emphasis 0.16

(iii) Gray-Level Run Length Matrix feature gray-level run emphasis 0.14

(iv) Gray-Level Cooccurrence Matrix feature correlation 0.08

(v) Gray-Level Cooccurrence Matrix feature cluster shade 0.05

(vi) Decreasing Gray-Level Matrix feature information measure of correlation

Figure 13(b), denotes that the novel feature "first order mean of DLGM" has the highest predictive power for AdaBoost classifier; hence, it is the most important feature for this classifier. The other important top five features for AdaBoost classifier are as follows:

(i) Decreasing Gray-Level Matrix feature first-order mean 0.17

(ii) Neighbouring Gray Tone Difference Matrix feature busyness 0.05

(iii) Decreasing Gray-Level Matrix feature maximal correlation coefficient 0.04

(iv) Decreasing Gray-Level Matrix feature information measure of correlation 0.035

(v) Gray-Level Cooccurrence Matrix feature correlation 0.03

Figure 13(c) denotes that the novel feature "Maximal Correlation coefficient of DLGM" has the highest predictive power for bagging classifier; hence, it is the most important

feature for this classifier. The other important top five features for bagging classifier are as follows:

(i) Decreasing Gray-Level Matrix feature maximal correlation coefficient 0.16

(ii) Gray-Level Dependence Matrix feature large dependence emphasis 0.05

(iii) Decreasing Gray-Level Matrix feature information measure of correlation 0.03

(iv) Gray-Level Cooccurrence Matrix feature difference average 0.03

(v) Gray-Level Dependence Matrix feature high gray-level emphasis 0.03

Figure 13(d) denotes that the novel feature "first order mean of DLGM" has the highest predictive power for random forest classifier; hence, it is the most important feature for this classifier. The other important features for random forest classifier are as follows:

(i) Decreasing Gray-Level Matrix feature first-order mean 0.12

(ii) Decreasing Gray-Level Matrix feature information measure of correlation 0.07

(iii) Decreasing Gray-Level Matrix feature maximal correlation coefficient 0.06

(iv) Gray-Level Run Length Matrix feature high gray-Level run emphasis 0.05

(v) First-order mean absolute deviation 0.05

2.5. nsemble Learning Classifiers. "Ensemble learning is a machine learning paradigm where multiple learners are trained to solve the same problem. In contrast to ordinary machine learning approaches which try to learn one hypothesis from training data, ensemble methods try to construct a set of hypotheses and combine them to use" [29]. A good number of studies [30, 31] proved that the generalization capability of a set of learners is much greater than a single learner. Ensemble classifiers have been applied in diversified fields, e.g., cyber security, intrusion detection system, face recognition system, and traffic control systems. The concept of ensemble classification proceeds in two stages:

(a) Classifier generation

(b) Aggregation of results of these classifiers

There are three approaches to classifier generation and aggregation.

(i) Bagging

(ii) Boosting

(iii) Stacking

2.5.1. Bagging. In this method, different training datasets are generated by resampling the training dataset, i.e., replacing some of the samples randomly. Suppose we have the following dataset: (4,5,6,7,8,9,10) and we have 5 classification algorithms. A different dataset is created by randomly resampling our data and passed to each classifier for training:

Dataset for classifier 0: (4,5,5,7,8,10,10) by replacing 6 with 5 and 9 by 10.

Dataset for classifier 1: (4,5,7,7,9,9,10) by replacing 6 by 7 and 8 by 9.

Dataset for classifier 2: (5,5,7,7,9,9,6) by replacing 4 by 5 and 10 by 6.

The results of all these classifiers are aggregated when taking predictions and inference time.

2.5.2. Boosting. Boosting attempts to create chains of different classification algorithms. The chain with the best performance on training data is then used for inference, coming back to our previous example where we had our training dataset as (4,5,6,7,8,9,10) and 5 classification algorithms. If we are creating chains of 3 classifiers, we can create 10 such chains. A single chain of 3 classifiers is created in the following manner:

(a) A batch of training dataset is passed through classification algorithm 1, i.e., classifier 0

Dataset for classifier 0: (4,5,6,7,8,9,10)

(b) Based on the performance of classifier 0 on this training batch, the whole batch is redistributed. The incorrectly predicted samples (by classifier 0) from the training batch are chosen more often to create the training batch for classifier 1. In this manner, classifier 1 will try to improve on the mistakes done by classifier 0. This is true for each classifier in the chain

Dataset for classifier 1: (4,5,7,7,9,9,10) by replacing 6 by 7 and 8 by 9 as 7 and 8 was incorrectly predicted.

(c) The same process will be repeated for classifier 2

Dataset for classifier 2: (10,9,7,7,9,9,10) by replacing 4 by 10 and 5 by 7 as both 4 and 5 was incorrectly predicted.

In essence, boosting will create and choose the chain which is able to collectively give better results than other chains.

In this study, we have explored two boosting ensemble classifiers XGBoost and AdaBoost. As is evident from our results, the prediction accuracy with these classifiers is much higher than bagging classifiers.

2.5.3. Stacking. Stacking is usually a 2 step approach. The classifiers in step 1 are known as base learners while the classifiers in step 2 are called stacking model learners. Each step is an ensemble of few classification algorithms. Predictions from the base learners are used as dataset for stacking model learners. Note that the predictions from base-level classifiers still maintain relationships with initial dataset which the stacking level classifiers can understand. The predictions from the stacking model learners are used at inference time.

3. Results

Along with accuracy, the most important metrics to analyze a biomedical machine learning study are sensitivity and specificity.

Sensitivity is the measure of true positives, which means accurate identification of patient with the disease. The test should have more true positives and minimum false negatives. False negatives mean we may miss out the positive identification of disease. Our study is a kind of screening test hence should have more sensitivity. Table 1 shows highest sensitivity is 99.82% hence in accordance to screening test.

Specificity is the measure true negatives, which is the ability of a test to rule out the disease accurately. Target of study is to have minimum false positives. As the study is screening test, we can have false alarms and less specific. The specificity of our study is 97.01%.

The three metrics are measured with following formulae:

(i) Specificity = true negative outcomes/true negative outcomes + false positive outcomes

(ii) Sensitivity = true positive outcomes/true positive outcomes + false negative outcomes

(iii) Accuracy = true negative outcomes + true positive outcomes/true negative outcomes + false positive outcomes + true positive outcomes

3.1. Analyzing Different Ensemble Methods and Results. In our study, we observed that boosting ensemble learning classifiers such as AdaBoost and XGBoost perform better than bagging and randomized classifiers. Bagging classifiers and random forest classifiers yield almost the same accuracy of 87%. The results are listed in Table 1. The accuracy calculated from the area under the curve is depicted in Figure 14 for all four ensemble classifiers.

4. Conclusion

In this study, we have proposed to build a decision support system for radiologists in order to make fast and accurate decisions for early detection of brain degeneration by mapping CDR values to MRI images. The most important performance metrics in the field of computer-aided biomedical studies are sensitivity, specificity, and accuracy. Through this study, we have shown that better data collection and preprocessing (data augmentation and feature selection) along with gradient-boosted ensemble learning classifiers contribute to improvements in all 3 metrics.

Data is one of the most important factors for driving the accuracy of any study. In our study, we worked on the OASIS-3 dataset, which is a longitudinal dataset with 4096 MRI scans while earlier studies are performed on cross-sectional datasets with less than 500 MRI scans. This dataset also gives specific details about how the CDR value changes for a subject with respect to changes in the subject's MRI scan. Any machine learning system requires large amount of data to be optimally trained. In our study, we have also employed data augmentation techniques. Data augmentation resulted in our classifier being much more tolerant towards variance in the data; this prevents overfitting. Another major

impact of data augmentation was the increase in dataset size from 4096 to 10000 MRI scans; this prevents underfitting. Mitigating overfitting and underfitting helps to achieve optimal accuracy on any dataset, irrespective of the classifier being used.

Our domain experts (Dr. Kunal Jain and Dr. Tanu) suggested that brain degeneration is not localized and affects the brain as a whole. As such, we have utilized whole brain volumes for our study and classification.

We experimented with Radiomics features and found that, for our data, the most promising features of

(i) GLCM are correlation, cluster shade, joint average, and cluster prominence

(ii) GLRLM are gray-level run emphasis, short run high gray-Level emphasis, short run low gray-level emphasis, and gray-level variance

(iii) NGTD Matrix is busyness

(iv) GLDM are high gray-level emphasis and small dependence low gray-level emphasis

(v) GLSZM is small area low gray-level emphasis

Our study also proposes a novel texture filter DGLM. The features mean, information measure of correlation, maximal correlation coefficient, first-order entropy, and first-order skewness from novel DGLM improved the accuracy from 95.6% to 97.38%.

This study also reaffirmed the fact that ensemble learning classifiers are usually much more accurate than a single classification algorithm. The study observed that gradient-boosted classifiers do not suffer from overfitting and also help to reduce generalization error, hence improving accuracy, sensitivity, and specificity.

The study results have been compared to other different studies in this area as depicted in Table 2.

Data Availability

In this study, we used open source data. Data is available at https://www.oasis-brains.org/, the data was requested to Oasis-3 Brain team, and it provided the login and password to download the data; the same can be shared as and when needed.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

References

[1] C. L. Grady, "Functional brain imaging and age-related changes in cognition," *Biological Psychology*, vol. 54, no. 1–3, pp. 259–281, 2000.

[2] B. J. Casey, J. N. Giedd, and K. M. Thomas, "Structural and functional brain development and its relation to cognitive development," *Biological Psychology*, vol. 54, no. 1–3, pp. 241–257, 2000.

[3] C. R. Jack, R. C. Petersen, Y. C. Xu et al., "Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment," *Neurology*, vol. 52, no. 7, pp. 1397–1403, 1999.

[4] J. C. Pruessner, L. M. Li, W. Serles et al., "Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories," *Cerebral Cortex*, vol. 10, no. 4, pp. 433–442, 2000.

[5] J. C. Pruessner, S. Köhler, J. Crane et al., "Volumetry of temporopolar, perirhinal, entorhinal and parahippocampal cortex from high-resolution MR images: considering the variability of the collateral sulcus," *Cerebral Cortex*, vol. 12, no. 12, pp. 1342–1353, 2002.

[6] J. Ashburner and K. J. Friston, "Voxel-based morphometry– the methods," *Neuroimage*, vol. 11, no. 6, pp. 805–821, 2000.

[7] J. C. Baron, G. Chételat, B. Desgranges et al., "*In vivo* mapping of gray matter loss with voxel-based morphometry in mild Alzheimer's disease," *NeuroImage*, vol. 14, no. 2, pp. 298–309, 2001.

[8] J. L. Whitwell, "Voxel-based morphometry: an automated technique for assessing structural changes in the brain," *The Journal of Neuroscience*, vol. 29, no. 31, pp. 9661–9664, 2009.

[9] B. C. Emerton, M. Jerram, T. Deckersbach, D. D. Dougherty, C. Fulwiler, and D. A. Gansler, "A comparison of voxel-based morphometry and volumetry methods in the context of the neural basis of aggression," *Brain Imaging and Behavior*, vol. 3, no. 4, pp. 332–341, 2009.

[10] K. M. Kennedy, K. I. Erickson, K. M. Rodrigue et al., "Age-related differences in regional brain volumes: a comparison of optimized voxel-based morphometry to manual volumetry," *Neurobiology of Aging*, vol. 30, no. 10, pp. 1657–1676, 2009.

[11] B. Fischl and A. M. Dale, "Measuring the thickness of the human cerebral cortex from magnetic resonance images," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 20, pp. 11050–11055, 2000.

[12] A. C. Burggren, M. M. Zeineh, A. D. Ekstrom et al., "Reduced cortical thickness in hippocampal subregions among cognitively normal apolipoprotein E e4 carriers," *NeuroImage*, vol. 41, no. 4, pp. 1177–1183, 2008.

[13] A. Bakkour, J. C. Morris, and B. C. Dickerson, "The cortical signature of prodromal AD: regional thinning predicts mild AD dementia," *Neurology*, vol. 72, no. 12, pp. 1048–1055, 2009.

[14] I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," *Artificial Intelligence in Medicine*, vol. 23, no. 1, pp. 89–109, 2001.

[15] S. Tangaro, P. Inglese, R. Maglietta, and A. Tateo, "MIND-BA: fully automated method for computer-aided diagnosis of dementia based on structural MRI data," in *Proc MICCAI Workshop Challenge on Computer-Aided Diagnosis of Dementia Based on Structural MRI Data*, pp. 119–128, Boston, USA, 2014.

[16] E. E. Bron, M. Smits, W. van der Flier et al., "Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge," *NeuroImage*, vol. 111, pp. 562–579, 2015.

[17] L. Lazli, M. Boukadoum, and O. A. Mohamed, "A survey on computer-aided diagnosis of brain disorders through MRI based on machine learning and data mining methodologies with an emphasis on Alzheimer disease diagnosis and the contribution of the multimodal fusion," *Applied Sciences*, vol. 10, no. 5, p. 1894, 2020.

[18] M. Cenek, M. Hu, G. York, and S. Dahl, "Survey of image processing techniques for brain pathology diagnosis: challenges and opportunities," *Frontiers in Robotics and AI*, vol. 5, 2018.

[19] M. Georgescu, L. Haidar, A. F. Serb, D. Puscasiu, and D. Georgescu, "Mathematical modeling of brain activity under specific auditory stimulation," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 6676681, 20 pages, 2021.

[20] A. E. Kovtanyuk, A. Y. Chebotarev, N. D. Botkin, V. L. Turova, I. N. Sidorenko, and R. Lampe, "Nonstationary model of oxygen transport in brain tissue," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 4861654, 9 pages, 2020.

[21] J. J. M. van Griethuysen, A. Fedorov, C. Parmar et al., "Computational radiomics system to decode the radiographic phenotype," *Cancer Research*, vol. 77, no. 21, pp. e104–e107, 2017.

[22] J. E. van Timmeren, D. Cester, S. Tanadini-Lang, H. Alkadhi, and B. Baessler, "Radiomics in medical imaging—'how-to' guide and critical reflection," *Insights Imaging*, vol. 11, no. 1, p. 91, 2020.

[23] P. J. LaMontagne, T. L. Benzinger, J. C. Morris et al., "OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease," *Journal of Chemical Information and Modeling*, vol. 53, no. 9, pp. 1689–1699, 2013.

[24] P Organization, "Pyradiomics documentation," 2019, http://www.radiomics.io/.

[25] S. S. F. Yip and H. J. W. L. Aerts, "Applications and limitations of radiomics," *Physics in Medicine and Biology*, vol. 61, no. 13, pp. R150–R166, 2016.

[26] A. Traverso, L. Wee, A. Dekker, and R. Gillies, "Repeatability and reproducibility of radiomic features: a systematic review," *International Journal of Radiation Oncology • Biology • Physics*, vol. 102, no. 4, pp. 1143–1158, 2018.

[27] J. C. Morris, "Clinical Dementia Rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type," *International Psychogeriatrics*, vol. 9, SUPPL. 1, pp. 173–176, 1997.

[28] B. Fischl, "FreeSurfer," *NeuroImage*, vol. 62, no. 2, pp. 774–781, 2012.

[29] Z.-H. Zhou, "Ensemble learning," in *Encyclopedia of Biometrics*, Springer, 2009.

[30] R. Maclin, "Popular ensemble methods : an empirical study popular ensemble methods : an empirical study," *Journal of Artificial Intelligence Research*, vol. 11, pp. 169–198, 2016.

[31] M. P. Sesmero, A. I. Ledezma, and A. Sanchis, "Generating ensembles of heterogeneous classifiers using stacked generalization," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 5, no. 1, pp. 21–34, 2015.

[32] A. Chaddad, C. Desrosiers, and T. Niazi, "Deep radiomic analysis of MRI related to Alzheimer's disease," *IEEE Access*, vol. 6, pp. 58213–58221, 2018.

[33] F. Feng, P. Wang, K. Zhao et al., "Radiomic features of hippocampal subregions in Alzheimer's disease and amnestic mild cognitive impairment," *Frontiers in Aging Neuroscience*, vol. 10, pp. 1–11, 2018.

[34] Y. Li, J. Jiang, J. Lu, J. Jiang, H. Zhang, and C. Zuo, "Radiomics: a novel feature extraction method for brain neuron degeneration disease using18F-FDG PET imaging and its implementation for Alzheimer's disease and mild cognitive impairment," *Therapeutic Advances in Neurological Disorders*, vol. 12, 2019.

[35] K. Zhao, Y. Ding, Y. Han et al., "Independent and reproducible hippocampal radiomic biomarkers for multisite Alzheimer's disease: diagnosis, longitudinal progress and biological basis," *Scientific Bulletin*, vol. 65, no. 13, pp. 1103–1113, 2020.

[36] T. R. Li, Y. Wu, J. J. Jiang et al., "Radiomics analysis of magnetic resonance imaging facilitates the identification of preclinical Alzheimer's disease: an exploratory study," *Frontiers in Cell and Developmental Biology*, vol. 8, pp. 1–13, 2020.

Research Article

Extraction and Evaluation of Corpus Callosum from 2D Brain MRI Slice: A Study with Cuckoo Search Algorithm

K. Suresh Manic,[1] Roshima Biju,[2] Warish Patel,[3] Muhammad Attique Khan,[4] N. Sri Madhava Raja [ID],[5] and S. Uma[6]

[1]*Department of Electrical and Communication Eng., National University of Science and Tech, Muscat, Oman*
[2]*Research Scholar, Department of Computer Science Eng., Parul University, Vadodara, Gujarat 391760, India*
[3]*Department of Computer Science Eng., Parul University, Vadodara, Gujarat 391760, India*
[4]*Department of Computer Science, HITEC University, Museum Road, Taxila, Pakistan*
[5]*Department of Electronics and Instrumentation Eng., St. Joseph's College of Engineering, OMR, Chennai 119, India*
[6]*Research Scholar, Anna University, Chennai, India*

Correspondence should be addressed to N. Sri Madhava Raja; nsrimadhavaraja@stjosephs.ac.in

The work proposes a computer-based diagnosis method (CBDM) to delineate and assess the corpus callosum (CC) segment from the 2-dimensional (2D) brain magnetic resonance images (MRI). The proposed CBDM consists of two parts: (1) preprocessing and (2) postprocessing sections. The preprocessing tools have a multithreshold technique with the chaotic cuckoo search (CCS) algorithm and a preferred threshold procedure. The postprocessing employs a delineation process for extracting the CC section. The proposed CBDM finally extracts the vital CC parameters, such as total brain area (TBA) and CC area (CCA) to classify the considered 2D MRI slices into the control and autism spectrum disorder (ASD) groups. This attempt considers the benchmark brain MRI database which includes ABIDE and MIDAS for the experimental investigation. The results obtained with ABIDE dataset are further confirmed against the fuzzy C-means driven level set (FCM + LS) and multiphase level set (MLS) technique and the proposed CBDM with Shannon entropy along with active contour (SE + AC) presented improved result in comparison to the existing methodologies. Further, the performance of CBDM is confirmed on MIDAS and clinical dataset. The experimental outcomes approve that the proposed CBDM extracts the CC section from the 2D MR brain images that have higher accuracy compared to alternative techniques.

1. Introduction

Corpus callosum (CC) is one among the vital brain parts responsible for neural communication among the two brain sections. CC is the prime commissural territory in the human brain, and it is composed of nearly 200-300 axons [1]. The work by Hinkley et al. (2012) on agenesis of corpus callosum (ACC) confirms that CC plays a significant role in problem cracking schemes and swiftness in vocal processing [2]. The study of Paul et al. (2014) also presents the relation of ACC and autism [3]. Their work also confirms that the CC disorder will lead to autism. In early significant researches, many works are reported to observe autism disorder based on CC [4–7]. Some of similar research works also report the study of sexual dimorphism in CC [8–14].

Due to its clinical significance, a substantial amount of CC assessment procedures has been proposed and discussed by researchers [15, 16]. Normally, the CC region is best visible in the sagittal view of two-dimensional (2D) brain MRI. The visibility of the CC is also approximately similar to other normal brain tissues and hence, the segmentation of CC from the MRI requires some complex procedures in comparison with the separation of other brain regions. Considering literature, procedures such as manual segmentation [6], level set scheme [10], active contour method [15], and fuzzy C-means [16] are applied to

extract CC with possible accuracy. Most of these approaches consider a two-step procedure to separate the CC from the sagittal view MRIs.

In recent times, the two-step process which integrates multithresholding and segmentation is widely adopted by the investigators to obtain the region of interest (ROI) of the brain MRI documented with various modalities such as flair, T1, T1C, T2, and diffused weighting (DW) [17–22]. These approaches implement the heuristic algorithm-oriented threshold process to develop the prominence of the ROI and a preferred segmentation plan to mine the ROI. Further, the ROI is assessed in comparison with the corresponding ground truth (GT) pictures presented by a domain professional. The image similarity parameters (ISP) obtained during the ROI and GT evaluation confirm the superiority of the brain MRI assessment technique [23–25].

The earlier works confirms that the heuristic approach-based brain MRI work offers improved result. This is obtained in comparison with the existing conventional processes [25]. Hence, in this work, most successful heuristic procedure called the cuckoo search (CS) algorithm is considered during the brain MRI preprocessing. The performance of the traditional CS (TCS) is enhanced based on the chaotic operator known as Ikeda map (IM), which aided to accomplish better threshold result. The details of the IM and its application are discussed in [20, 21]. Experimental investigation of their work is then compared with the Lévy-Flight and Brownian-Walk operators, which confirmed that the chaotic cuckoo search (CCS) offers better threshold compared to the traditional CS.

During the preprocessing procedure, CCS identifies the optimal value of thresholds for brain MRI. In preprocessing, a comparative examination amongst the famous threshold approaches, such as Kapur, Tsallis, Otsu, and Shannon is performed. This helps in finding the best suited threshold scheme for CC examination using the 2D brain MRI. The role of the postprocessing plan is to demarcate the CC subjected to preprocessing. After mining the CC, an assessment in comparison with the ground truth is performed to obtain the vital ISPs.

In literature, few methods are discussed to obtain the CC present in the considered 2D brain images. Further, most of the methods are interested in computing the total brain area (TBA) and corpus callosum area (CCA) to categorize the 2D brain MRI dataset into control and autism spectrum disorder (ASD) groups. In analyzing an image which belongs to medical, it is always essential to measure the outcome of the proposed tool with a chosen image dataset. If the tool works well on the dataset, further, the developed image examination instrument can be considered to estimate the medical grade images.

The earlier works on CC examination computes only the TBA and CCA and directly implements a categorization process. To evaluate the efficacy of the developed tool, it is essential to compute the ISPs and the essential statistical measures. Further, the soft computing-based CC examination is also needed to improve the extraction accuracy. Because of these reasons, in the projected work, CCS with CBDM is proposed for examining the CC section.

This research work also presents a detailed study on (i) different threshold procedures, such as Otsu, Kapur, Tsallis, and Shannon and (ii) various segmentation approaches, such as level set (LS), Chan-Vese (CV), region growing (RG), and active contour (AC) in order to identify the appropriate pre- and postprocessing practice to mine CC.

The experimental investigation is implemented in Matlab software (Version7, Release14, Lic. No. 285705 with perpetual term) using the public autistic databases, like ABIDE (images of 60 volunteers) [19, 20] and MIDAS (images of 4×2 volunteers) [21]. The clinical implication of projected tool is confirmed with the real-time clinical MRI obtained from Proscans laboratory (images of 10×2 volunteers) [22].

2. Related Works

The MR imaging technique is extensively utilized to record the performance and malformations of internal organs of living beings. The improvement in the MRI method additionally supports the upgrading and appraisal of features recorded in 3-dimensional digital pictures. Prevailing evaluation methods which deal with 3D images are intricate. They require extraordinary swiftness in computing machines since data volume is enormous. To reduce the difficulty in assessing the MRI, reorganized 3D image is further transformed into a significant amount of 2D slices. Finally, the two-dimensional slices are assessed using a suitable image investigation system. In the proposed work, 2D brain MRI slices are considered for the study, and the stages involved in the CC segmentation and the corresponding brain abnormalities to be detected are presented in Figure 1.

Examination of CC from 2D slices of MRI with sagittal view is commonly considered by the researchers. Paul et al. (2014) proposed a practical examination to compare ACC and autism using 2D MRI of T1 and DW modalities [3]. Their work confirms that examination of CC is essential to asses ACC and autism. Wolff et al. (2015) proposed a clinical investigation to confirm that the CC region is reduced for elders and adults having autism spectrum disorder (ASD) [4]. Frazier and Hardan (2009) applied a region-based examination on CC section of patients with autism [5]. A manual segmentation scheme is considered to extract and evaluate the CC's size values and confirmed that the ASD can be predicted based on the size of the CC section. Tepest et al. (2010) inspected the size of CC and its segments associated with total brain volume (TBV) to identify the autism with respect to gender and revealed that the TBV values in males are higher than in females [6]. Lefebvre et al. (2015) proposed a work on neuroanatomical variety of CC and TBV in autism and verified the work by considering the brain MRIs of 694 volunteers [7]. The studies on the sexual dimorphism in CC and variation in size of TBV and CC also widely examined using the 2D MRI slices [8].

Previous studies authenticate the requirement of CC examination during the human brain analysis; hence, more care is essential during the segmentation of the CC region. Normally, the CC is a thin section in the brain MRI and will have the pixel intensity similar to other brain sections.
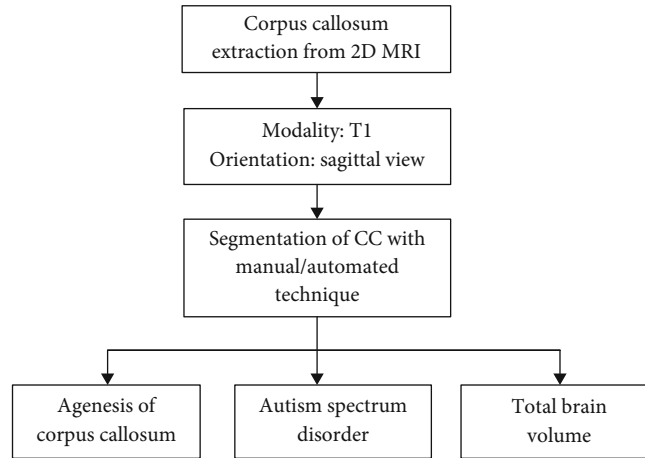
FIGURE 1: Various approaches to examine the CC to detect different brain abnormalities.

Hence, it is essential to consider an efficient image processing system to excerpt and estimate the CC from the 2D brain MRI of a chosen modality.

Fredo et al. (2014) applied a two-step process with fuzzy-C-means (FCM) clustering and multiphase level set (LS) approach to delineate the CC, cerebellum, and brain stem from the 2D MRI recorded with T1 modality and obtained a mean area of 0.87 for control (normal) cases and 0.67 for ASD cases [9]. Further, Fredo et al. (2015) implemented a similar work on the ABIDE database with 20 samples of control cases (male = 14 and female = 6) and 20 samples of ASD cases (male = 11 and female = 9) and attained mean area of 0.90 for control cases and 0.75 for ASD [10, 11]. Fredo et al. (2015) employed the reaction diffusion regularized level set (RDRLS) method to delineate CC [10]. Vachet et al. (2012) implemented the deformable active Fourier contour model [15], and İçer (2013) discussed a two-step approach based on the Gaussian mixture model and FCM to extract the CC [16]. Li et al. (2013) executed an automated two-step segmentation scheme by combining the mean shift clustering technique-based image improvement and geometric active contour (GAC) dependant segmentation of CC [26]. The work of Elsayed et al. (2010) implements a spectral segmentation with the multiscale graph decomposition process to extract CC [27]. Recent review of Cover et al. (2018) presents an elaborate evaluation of various CC evaluation schemes, MRI modalities, and performance measures existing in the literature [28]. Their work also reports that T1-weighted MRI is the widely adopted modality (44%) to examine CC. This work also discusses the merits and demerits of the existing schemes and also recommends the need for a novel evaluation tool.

The proposed procedure has a two-stage process to extract the CC section present in 2D brain MRI of T1 modality that is also implemented. For experimental investigation, the database such as ABIDE and MIDAS is utilized. Further, in this proposed method, it is implemented and validated for the clinical MR image obtained from Proscans laboratory.
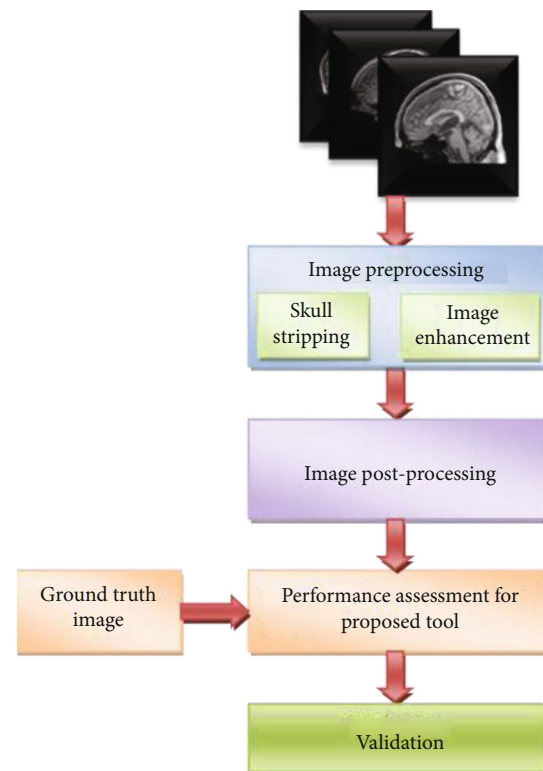
FIGURE 2: Overview of proposed tool.

3. Computer-Based Diagnosis Method

A brief summary concerning the methods adopted in this paper to provide a computer-based diagnosis method (CBDM) for the extraction of CC from 2D brain MRI is discussed in this segment. The outline of the database utilized, rudimentary tasks in examining images of brain, preprocessing, delineation, and validation are presented elaborately.

Figure 2 presents various phases used in this examination tool. Firstly, a 2D image of sagittal view MRI of chosen slice is considered with/without skull section. Early improvement of

raw MRI slice is carried out using an appropriate image pre-processing method, and a preferred postprocessing scheme is to be adopted to mine the CC section. Further, the performance of the tool can be validated with a relative analysis with ground truth (GT) image presented by a specialist. Extracted CC is then validated by a doctor by providing the decision of next step in treating the patient to normalize or provide remedy for the brain prognosis state. In all cases, the developed tool by any suitable approach can only offer a suggestion/preopinion regarding the brain abnormality, and the doctor has to provide consultation and thereby a conclusion in treating the patient further based on the condition.

3.1. MRI Database. The sagittal view of MRI adopted in this paper is collected from the public database such as ABIDE and MIDAS. Further, real clinical image obtained from Proscans is also used in this research work. These entire databases consist of 3D brain structures recorded with T1 modality. ITK-SNAP version 3.6.0 tool aids to obtain the 2D slices from the complete dataset. [29, 30], and an image normalization is implemented to obtain 2D slices of size 256×256 pixels. Similar practice is employed for the GT image of ABIDE. The ABIDE is the commonly used database in autism studies, which provides the vital details, like subject case (controlled/autistic), gender, age, area of CC, and TBV [9–11].

In this work, 60 volunteer's (age group of 13-16 years) images are considered for the examination. The MIDAS database consists of two control (normal) and two autistic volunteer's images that are recorded in the age of 2 years and follow-up in 4 years, respectively. Finally, the clinical images of a volunteer collected from Proscans are also examined using the proposed approach.

3.2. Image Preprocessing. This scheme is generally considered to improve the picture under assessment using a suitable image processing technique. This procedure will increase the ROI by uniting the similar pixel values with a set of threshold values selected. Recent related works confirm that preprocessing practice is an essential stage in two-step image processing tool.

3.2.1. Skull Stripping. Usually, the reconstructed brain MRI is associated with the outer head bone called the skull. For modalities of T2 and flair type, concentration of pixels belonging to the skull is roughly greater than soft brain tissues. Also, in T1 MRI modality, the skull intensity is similar to the intensity level of the brain tissue. Automated brain region segmentation always requires a suitable skull stripping procedure to discrete the soft region of the brain from that of the skull information [31]. Various skull removing methods discussed by the researchers can be found in [32]. If semiautomated brain section segmentation is implemented, the additional procedure of skull removal will not be required further. The methods such as LS, CV, RG, and AC fall in the semiautomated group do not take into consideration of the skull section. In this paper, the work is instigated on the 2D MRI slice, with and without skull section.

3.2.2. Cuckoo Search Algorithm. In consideration with various existing metaheuristic methods, cuckoo search (CS) presented by Yang and Deb has appeared to be one of the effective soft computing techniques [33–37]. Recently, CS is widely accepted by maximum number of researchers in solving numerous optimization tasks [38, 39]. The main advantage of CS compared with the firefly and bat algorithm is the structure of the CC that is simple and supports higher probability of getting the optimized solution. Various chaotic search procedures assisted CS can be found in [40, 41].

The mathematical expression of CS is as follows:
The CS executed with the following conventions:

(i) Each bird leaves behind an egg in randomly nominated nest of other host birds

(ii) Nest of strong surviving egg is inherited to the succeeding level. The hatching rate of this egg is faster than that of the host

(iii) The chance of categorizing the egg by host bird in CS is $p_a \in [0, 1]$ for a selected optimization task

In most of the heuristic algorithms, accomplishment in discovering a resolution for a job generally depends on its direction finding method. Typically, it is guided by Lévy Flight (LF) and chaotic strategies [42, 43].

In this paper, the Ikeda map (IM) is chosen to drive the CS, and the details on IM are available in [44, 45]. In CS optimization investigation, nascent location $(X^{(t+1)})$ naturally count on previous position $(X_i^{(t)})$. In this section, the subsequent equations are accounted to search an updated location of cuckoo:

$$X_i^{(t+1)} = X_i^{(t)} + \alpha \oplus \text{IM}, \qquad (1)$$

where $X_i^{(t+1)}$ is the updated position and signifies early position, $\oplus$ denotes the entrant multiplier, and IM shows the chaotic Ikeda map approach. Normally, the parameter "α" is assigned with a positive integer (i.e., $\alpha > 0$) and in this research, "α" is allocated as 1. Additional particulars regarding CS are cited in the works of Yang and Deb [37].

IM is one of the chaotic search operator, and its explanation and application on various heuristic algorithms can be found in [17, 18].

$$\text{IM} = R..\exp\left[Z\Phi - Z\left(\frac{\delta}{1 + \left|X_i^{(t)}\right|^2}\right)\right], \qquad (2)$$

where Z is the iteration number, ϕ is allotted as 0.1, δ is chosen as 7, and the disordered attraction constraint (R) value is given as 0.75.

Equation (2) presents the IM implemented in the recent attempt of Satapathy et al. (2018) to increase the investigation competence of the bat algorithm (BA) [46]. This work established that the IM-assisted BA offered better result in comparison with particle swarm optimization (PSO), firefly algorithm (FA), and traditional BA. Further, the work of

Abhinaya and Raja (2015) [17] and Lakshmi et al. (2016) [18] confirms the advantage of IM-based cuckoo search for the medical image processing. Hence, this work implements a chaotic IM search technique to improve the performance of the traditional cuckoo search (TCS) method. The efficacy of the proposed chaotic cuckoo search (CCS) is further confirmed with other techniques, such as particle swarm optimization (PSO) [47], bacterial foraging optimization (BFO) [48], bat algorithm (BA) [41], and TCS [41].

The subsequent initial constraints are assigned for every heuristic algorithms adopted in this paper: representative's dimension is designated as 30, exploration measure is set as 3 (a three-level), the complete iteration limit is maintained as 1500, and end criteria are given as maximized value of image measure (between class variance for Otsu and maximized entropy for Kapur, Shannon, and Tsallis thresholding schemes).

3.2.3. Image Thresholding. Thresholding is an extensively followed image enhancement process employed to process traditional and medical images [23–25]. During the threshold process, a picture frame is separated into several sections by grouping related pixels, to find and evaluate the significant information existing in the picture. Previous research works confirm the availability of a variety of threshold schemes, such as Otsu, Kapur, Shannon, and Tsallis to preprocess the gray scale and RGB pictures. This section also implements a comparative study among the above said threshold procedures.

(1) Otsu's Scheme. This scheme is one of the well-known procedures widely adopted to progress the trial picture based on the chosen threshold value. In the related works of this research, multithresholding based on Otsu's approach is widely applied by the researchers for a class of image cases based on maximizing the interclass variance.

Otsu is a nonparametric threshold scheme developed in 1979 [49], and its mathematical relation is depicted as follows;

Let $L = 256$, and the chosen threshold number is three (i.e., Th = 3), which divides the input image into three distinct groups, like $Q0$, $Q1$, and $Q2$.

Assume that the image consists the thresholds like $(t_1, t_2 \cdots t_{Th})$, which split the input picture into three groups: $Q0$; gray level values are accounted from 0 to $t - 1$, $Q1$ which has gray levels of range t_1 to $t2 - 1$, and $Q2$ contains gray levels from $t3$ to $L - 1$.

The objective function for the above case will be

$$\text{Maximize } F(T) = \varphi_0 + \varphi_1 + \varphi_2, \tag{3}$$

where $\varphi_0 = \omega_0(\mu_0 - \mu_T)^2, \varphi_1 = \omega_1(\mu_1 - \mu_T)^2, \varphi_2 = \omega_2(\mu_2 - \mu_T)^2$.

In Eq. (3), the symbols ω and μ represent the class probabilities and class means, respectively.

(2) Kapur's Technique. Kapur's entropy (KE) was originally proposed in 1985 to appraise gray scaled images in accordance with its entropy based on histogram [50]. KE aids to explore the optimal threshold of a picture on the basis of its entropy alone. Since the outcome proves to provide satisfactory results, many researches using KE are deliberated in the literature [18].

Precise model of the KE is well defined as follows:

Let $T = [t_1, t_2, \cdots, t_{L-1}]$ represent individual threshold values of the image. Further, the complete entropy of KE is represented as follows:

$$\text{Costfunction} = J_{\text{Kapur}} = F(T) = \sum_{J=1}^{L} O_j^R \text{for} R\{1, 2, 3\}. \tag{4}$$

Equation (4) designates to get the most out of value of entropy for the selected threshold.

In trilevel thresholding assignment, the objective function value is denoted as

$$
\begin{aligned}
O_1^R &= \sum_{j=1}^{t_1} \frac{Po_j^R}{\theta_0^R} \ln\left(\frac{Po_j^R}{\theta_0^R}\right), \\
O_2^R &= \sum_{j=t_{l+1}}^{t_2} \frac{Po_j^R}{\theta_1^R} \ln\left(\frac{Po_j^R}{\theta_1^R}\right), \\
O_3^R &= \sum_{j=t_{2+1}}^{L} \frac{Po_j^R}{\theta_2^R} \ln\left(\frac{Po_j^R}{\theta_2^R}\right),
\end{aligned}
\tag{5}
$$

where Po_j^R shows the likelihood distribution and $\theta_0^R, \theta_1^R, .\theta_2^R$ depicts the probability occurrence in L-levels.

(3) Shannon's Technique. Shannon's entropy (SE) procedure was established by Kannappan in 1972 [51]. Rajinikanth et al. (2017) states that the SE approach-based brain MRI examination offers better result in comparison with Kapur's and Tsallis technique [18].

In recent works, SE dependant thresholding is employed to perform preprocess medical pictures. To elucidate the SE, a picture with dimension $A \times B$ is to be under consideration. The pixel arrangement of the gray picture (h, v) is expressed as $G(h, v)$, for $h \in \{1, 2, \cdots, A\}$ and $v \in \{1, 2, \cdots, B\}$. Let L be the various levels of gray for the considered test image, and the set of all gray values $\{0, 1, 2, \cdots, L - 1\}$ can be symbolized as Z, in such a way that

$$G(h, v) \in Z \forall (h, v) \in \text{picture}. \tag{6}$$

Then, the normalized histogram will be $X = \{t_1, t_2, \cdots, t_{L-1}\}$.

For thresholding with level set to 3, Eq. (5) becomes

$$X(T) = x_0(t_1) + x_1(t_2) + x_2(t_3),$$
$$F(T) = \max_T \{X(T)\}. \tag{7}$$

Threshold value which is represented by $T = \{t_1, t_2, \cdots, t_L\}$, $X = \{x_0, x_1, \cdots, x_{L-1}\}$ denotes the normalized histogram, and $F(T)$ indicates the optimal threshold. Further information about SE can be found in [52].

(4) Tsallis Technique. Tsallis entropy (TE) is a nonextensive entropy idea derived from the SE by Tsallis [53, 54] and represented as

$$S_q = \frac{1 - \sum_{i=1}^{T}(p_i)^q}{q - 1}. \tag{8}$$

In the equation, T is the scheme prospective, q is the entropic indicator, and $p\,i$ represents the probability of each state i. Usually, the entropy value obtained with Tsallis procedure, Sq, will meet Shannon's entropy when $q \longrightarrow l$.

The entropy information is denoted using a quasiadditive instruction as

$$S_q(A + B) = S_q(A) + S_q(B) + (1 - q).S_q(A).S_q(B). \tag{9}$$

TE can be utilized to discover the finest threshold values in the image. A test picture with L gray levels which have the values $\{0, 1, \cdots, L - 1\}$ with possibility spreading $p_i = p_0, p_1, \cdots, p_{L-1}$ is considered. Thus, the Tsallis trilevel-based threshold process is achieved with the objective function:

$$F(T) = [t1, t2, t3] = \text{argmax},$$
$$F(T) = [t_1, t_2, t_3] = \text{argmax},$$
$$\left[S_q^A(T) + S_q^B(T) + S_q^C(T) + (1 - q).S_q^A(T).S_q^B(T).S_q^C(T) \right], \tag{10}$$

where

$$S_q^A(T) = \frac{1 - \sum_{i=0}^{t_1-1}\left(Pi/P^A\right)^q}{q - 1}, P^A = \sum_{i=0}^{t_1-1} Pi,$$

$$S_q^B(T) = \frac{1 - \sum_{i=t_1}^{t_2-1}\left(Pi/P^B\right)^q}{q - 1}, P^B = \sum_{i=t_1}^{t_2-1} Pi, \tag{11}$$

$$S_q^C(T) = \frac{1 - \sum_{i=t_2}^{L-1}\left(Pi/P^C\right)^q}{q - 1}, P^C = \sum_{i=t_2}^{L-1} Pi.$$

When the multilevel process is executed base on threshold, an optimal threshold value T is to be obtained such that the objective function $F(T)$ is being maximized. In this existing work, the principal part of the CCS algorithm is to discover the maximized optimal threshold "$F(T)$" in Otsu, KE, SE, and TE cases for a chosen threshold of three.

3.3. Image Postprocessing. This phase purpose is to mine the ROI (CC) from preprocessed brain MRI. The details of various automated and semiautomated separation measures prevailing in the image processing literature are presented in detail. Based on the implementation, the segmentation processes are categorized as (i) automated and (ii) semiautomated schemes. In the automated scheme, the segmentation procedure requires a minimal or nil operators' assistance. In the semiautomated method, the initiation of the segmentation task is to be done by the operator based on a trial and error approach or a by adopting a directed practice.

3.3.1. Automated Segmentation. The segmentation methods, such as watershed [55], principal component analysis [56], and clustering approaches (*k*-means, fuzzy *k*-means, etc.) [57], are some of the techniques that falls in the category of the automated segmentation approach. In these procedures, the interaction of human operator during the initiation is comparatively less.

3.3.2. Semiautomated Segmentation. Semiautomated segmentation (SAS) approaches are widely considered in medical image analysis, when a complex segmentation task is to be completed. In these methods, the operator's assistance is essential throughout the segmentation execution. The operator is responsible to begin the operation, assigning the run time/number of iteration required and assigning the terminating criterion. SAS is widely applied by the investigators to extract the ROI from a class of complex medical images [25]. Generally, SAS works based on the identification of the similar pixel values from its initial point. It will explore all the possible alike pixel values present in the preprocessed picture, until the maximum iteration value is reached. The approaches, such as level set (LS) [58], Chan-Vese (CV) [59, 60], region growing (RG) [61], and active contour (AC) [62], fall in this category. In the projected work, the AC segmentation is executed to obtain the CC, and its performance is then validated against alternative approaches, like LS, CV, and RG.

AC has an adaptable snake-like search mechanism, which modifies its direction such that it addresses all the possible comparable pixel clusters available in the image based on energy minimization theory as discuss in [63]. Because of its merit, AC is commonly adopted to inspect medical images.

AC performs operations, like (i) border recognition, (ii) preliminary curve generation with respect to the identified border, (iii) changing the snake's orientation to follow the pixel group till the energy becomes minimal, and (iv) final curve generation and extraction of the region inside the final contour.

Energy function of AC's snake is

$$\min_C \left\{ E_{\text{GAC}}(C) = \int_0^{L(C)} g(|\nabla I_0 C(s)|)ds \right\}, \tag{12}$$
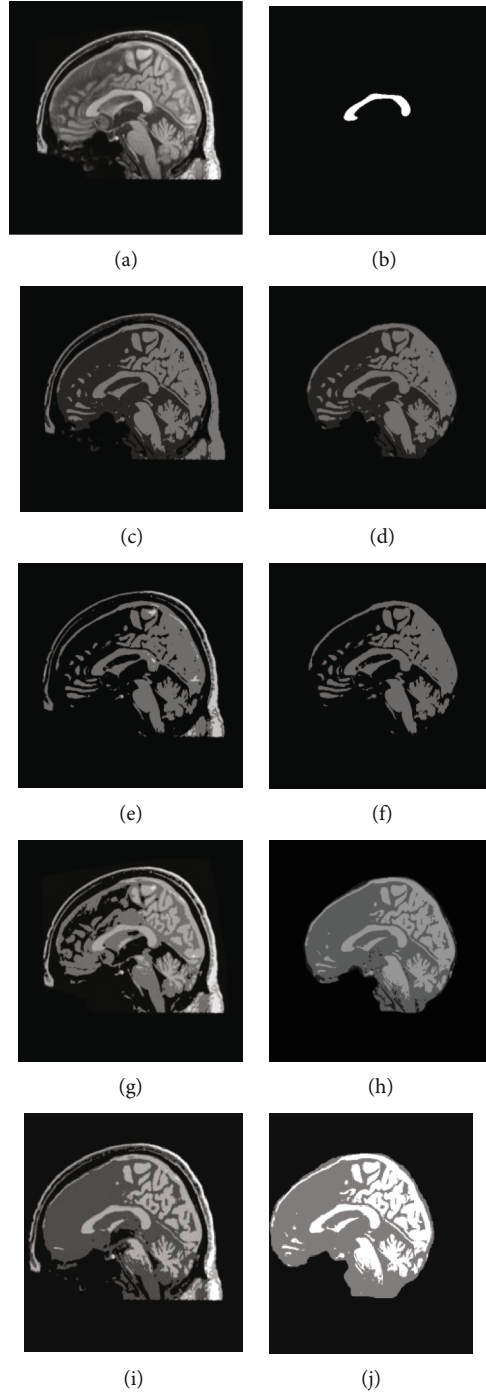
FIGURE 3: Outcome of the preprocessing approach for a chosen test image (image 1). (a, b) depicts test image and the GT, respectively, (c, d) shows Otsu's thresholding outcome for image with and without skull section, (e, f) presents the outcome of Kapur's entropy, (g, h) depicts the thresholding result of Shannon's, and (i, j) illustrates the result by Tsallis.

where ds is the Euclidean distance constituent and $L(C)$ is the length of the curvature C. It satisfies the constraints $L(C) = \int_0^{L(C)} ds$. The limitation g indicates edge, which will wane based on the objective periphery defined as

$$g(|\nabla I_0|) = \frac{1}{1 + \beta |\nabla I_0|^2}, \tag{13}$$

where I_0 signifies test image under study and β depicts a random constant. The energy value quickly declines because of the values reflected by the edges as in gradient succession quantification.

This method is scientifically characterized as

$$\partial_t C = (kg - \langle \nabla_g, M \rangle)M, \tag{14}$$
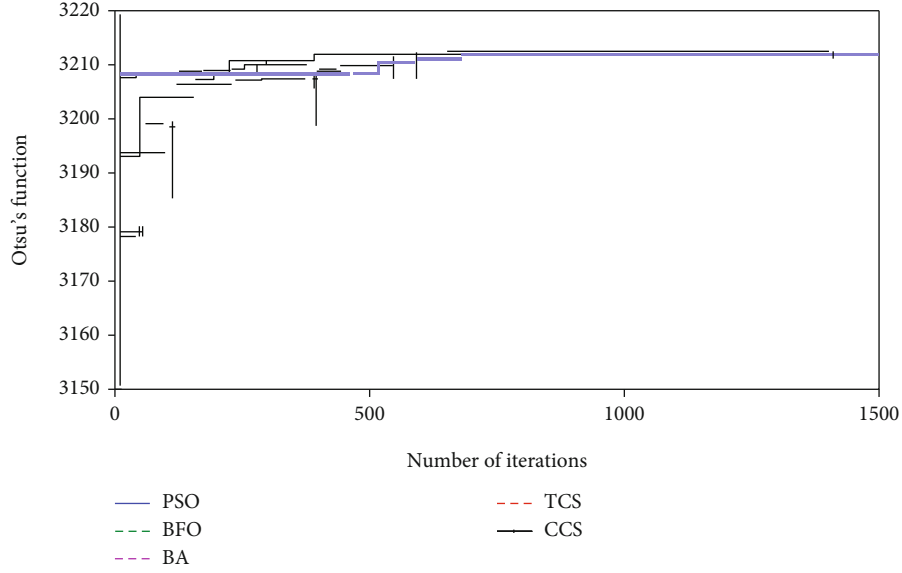
FIGURE 4: Convergence of the optimization search with Otsu's function.

where $\partial_t C = \partial C/\partial t$ indicates the changes in the snake model. t represents the repetition period. k and M are the curve and normal for the considered snake "C." In this process, the silhouette of the snake is constantly adjusted till nominal value of the energy; E_{GAC} is accomplished.

3.4. Evaluation of ROI with GT. The goal of this section focuses to examine the performance of the suggested method by employing a qualified examination amongst ROI and GT. This work deliberates standard brain MRI dataset identified as ABIDE, in which test images are associated with GT. In this study, image resemblance values, such as Jaccard, dice, false-positive rate (FPR), and false-negative rate (FNR), are computed [23–25].

The mathematical terminologies are presented in Eqs. (15)–(18):

$$\text{Jaccard}(I_G, I_C) = I_G \cap I_C I_G \cup I_C, \tag{15}$$

$$\text{Dice}(I_G, I_C) = 2(I_G \cap I_C)|I_G| \cup |I_C|, \tag{16}$$

$$\text{FPR}(I_G, I_C) = (I_G I_C)(I_G \cup I_C), \tag{17}$$

$$\text{FNR}(I_G, I_C) = (I_C I_G)(I_G \cup I_C), \tag{18}$$

where I_G signifies the GT and I_C represents the mined section.

Furthermore, the image statistical outcomes, which include sensitivity, specificity, accuracy, and precision, are also calculated [64, 65].

Expressions for these bounds are specified in Eqs. (19)–(22):

$$\text{Sensitvity} = T_P(T_P + F_N), \tag{19}$$

$$\text{Specificity} = T_N(T_N + F_P), \tag{20}$$

$$\text{Accuracy} = (T_P + T_N)/(T_P + T_N + F_P + F_N), \tag{21}$$

$$\text{Precision} = T_P(T_P + F_P), \tag{22}$$

where T_N, T_P, F_N, and F_P signify related measures.

4. Result and Discussions

The outcomes accomplished with the planned tool are elaborated. Various early works endorse the accessibility of considerable processing procedures for CC examination of the considered images. The projected work tools have a two-stage procedure to observe the well-known 2D sagittal brain MRI and the MR images obtained from the clinic. This work reflects the support of the contemporary heuristic technique known as CCS along with the well-known threshold approach. A comprehensive valuation among the prevailing segmentation processes, such as LS, CV, RG, and AC, is also presented. The developed CDT is executed with a AMD C70 Dual Core 1 GHz CPU with 4 GB of RAM PC which is equipped with Matlab software.

Firstly, the ABIDE dataset of 60 volunteers (30 control and 30 ASD class) is considered for the examination. This database contains the 2D sagittal MRI recorded with T1 modality with a pixel measurement of 256×256. This dataset is associated with relevant GT offered by a professional.

Figure 3 depicts a chosen 2D MRI and the GT of ABIDE. The preprocessing procedure is then implemented on this image by considering its original version and the skull stripped version. This figure also depicts the threshold results of various procedures reflected in this work. Figures 3(c) and 3(d) represent the enhanced image with Otsu's approach, Figures 3(e) and 3(f) depict the outcome of KE-based trilevel thresholding, and Figures 3(g) and 3(h) show that the result of SE and Figures 3(i) and 3(j) shows the results by TE. After enhancing the test picture based on a chosen threshold approach, a segmentation task is used to extract the CC section in order to find the parameters, such as TBA and CCA as discussed in [11]. During the segmentation task, every

| (a) | (b) | (c) | (d) | (e) | (f) |
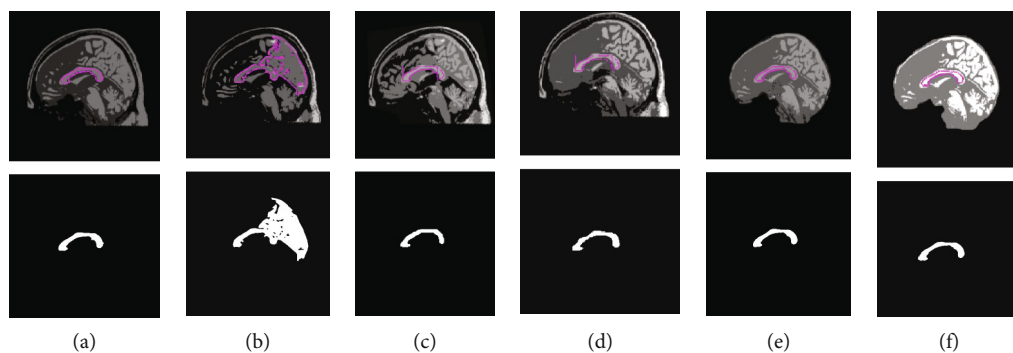
FIGURE 5: Extraction of CC using active contour segmentation: (a) Otsu's, (b) Kapur's, (c) Shannon's, (d) Tsallis, (e) Shannon's without the skull, and (f) Tsallis without the skull.

TABLE 1: Results obtained for the sample images.

Image (ASD)	Image	Test picture	GT	SE + AC	TE + AC
Male	Image 2				
	Image 3				
	Image 4				
Female	Image 5				
	Image 6				

TABLE 2: Image similarity measures achieved for the selected sample mages.

Image	Method	TPR	FNR	TNR	FPR	Jaccard	Dice
Image 1	Otsu	0.8035	0.1965	1.0000	0.0000	0.8032	0.8909
	Kapur	0.1348	0.8652	0.9999	0.0001	0.1345	0.2372
	Shannon	0.8804	0.1196	0.9992	0.0008	0.8201	0.9012
	Tsallis	0.8800	0.1200	0.9996	0.0004	0.8518	0.9200
Image 2	Shannon	0.7907	0.2093	0.9992	0.0008	0.7454	0.8541
	Tsallis	0.8339	0.1661	0.9996	0.0004	0.8078	0.8937
Image 3	Shannon	0.8405	0.1595	0.9993	0.0007	0.7841	0.8790
	Tsallis	0.4450	0.1555	0.9993	0.0007	0.7876	0.8812
Image 4	Shannon	0.8912	0.1088	0.9991	0.0009	0.8281	0.9060
	Tsallis	0.8482	0.1581	0.9988	0.0012	0.7750	0.8732
Image 5	Shannon	0.8535	0.1465	0.9998	0.0002	0.8270	0.9053
	Tsallis	0.8261	0.1739	0.9995	0.0005	0.7705	0.8704
Image 6	Shannon	0.9133	0.9986	0.0014	0.0867	0.8019	0.8901
	Tsallis	0.8396	0.1604	0.9998	0.0002	0.8263	0.9049

preprocessed test image is tested using the LS, CV, RG, and AC approaches. This test result confirms that the LS approach offered false result most of the time due to the visibility of CC. In most of the image cases, the CC pixel intensity is similar to the normal brain tissue intensity. Hence, for all the considered images, the extraction and evaluation task is implemented only with CV, RG, and AC.

Figure 4 represents the search merging of the heuristic algorithm for Otsu's trilevel threshold operation implemented on image 1. The proposed CCS is converged at 582th iteration, and the search process is terminated at 1417th iteration. This confirms that the projected CCS performs better compared to other approaches adopted in this study. Similar techniques are repeated with other threshold techniques, such as Kapur, Tsallis, and Shannon and for most of the cases, the proposed CCS offered improved outcome compared to the PSO, BFO, and BA. This confirms that the CCS works well for the chosen brain MRI thresholding problem.

Figure 5 depicts the execution of the AC-based extraction of CC from the preprocessed test images presented in Figure 3. Similar procedure is recurrent for the additional 2D sagittal images of the database, and its effects are recorded. To confirm the preeminence of the considered preprocessing approach, a relative study among the mined CC and the GT is performed, and the image match and statistical outcomes are calculated. This comparative study confirmed that, for the chosen dataset, Otsu's and KE procedures are failed to provide better result compared to the SE and TE-based procedures. Hence, the results of Otsu's and KE are ignored, and the results of SE and TE are projected in this paper. Table 1 represents the segmentation results attained for the representative images with the SE + AC and TE + AC. Similar results are attained with SE + CV, SE + RG, TE + CV, and TE + RG. Tables 2 and 3 present the similar information of the considered images and the statistical measures achieved during this experimental investigation. From Tables 2 and 3, it can also be observed that the outcome

TABLE 3: Image statistical outcomes attained for the selected sample images.

Image	Method	Sensitivity	Specificity	Accuracy	Precision
Image 1	Otsu	0.8035	1.0000	0.9976	0.9996
	Kapur	0.1348	0.9999	0.9373	0.9888
	Shannon	0.8804	0.9992	0.9980	0.9230
	Tsallis	0.8800	0.9996	0.9983	0.9637
Image 2	Shannon	0.7907	0.9992	0.9966	0.9287
	Tsallis	0.8339	0.9996	0.9976	0.9627
Image 3	Shannon	0.8405	0.9993	0.9978	0.9212
	Tsallis	0.8445	0.9993	0.9978	0.9212
Image 4	Shannon	0.8912	0.9991	0.9977	0.9212
	Tsallis	0.8482	0.9988	0.9969	0.8998
Image 5	Shannon	0.8535	0.9998	0.9987	0.9638
	Tsallis	0.8261	0.9995	0.9983	0.9197
Image 6	Shannon	0.9133	0.9986	0.9978	0.8680
	Tsallis	0.8396	0.9998	0.9979	0.9811

obtained with Otsu+AC and KE + AC is poor in comparison to the alternatives.

The performance of the projected CBDM is confirmed with a pixel level relative evaluation among the mined CC section and the GT. To demonstrate the performance, the mined CC sections SE + AC and TE + AC of image 1 are considered, and the obtained results are illustrated in Figure 6. Figure 6(a) depicts the confusion matrix of SE + AC, and Figure 6(b) presents the confusion matrix of TE + AC. From these images, it can be distinguished that the image similarity constraints (ISP) offered by the proposed CBDM are better. Similar technique is repeated with further images, and the sample consequences obtained with image 1 to image 6 are depicted in Tables 2 and 3.

Table 3 authenticates that the image measures obtained with the SE are better when compared to TE. The average
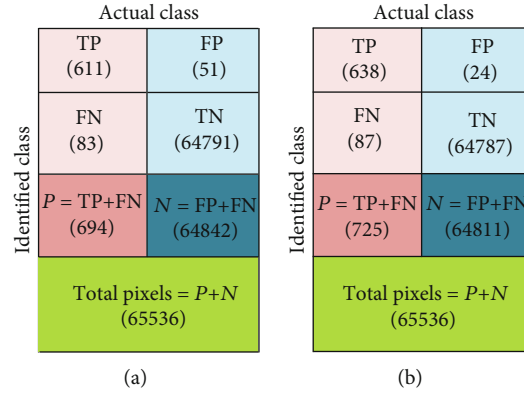
Figure 6: Confusion matrix to discuss the performance measure.

Table 4: Average values of similarity and statistical values of ABIDE dataset (60 volunteers).

Method	Jaccard	Dice	Sensitivity	Specificity	Accuracy	Precision
SE + AC	87.15	92.75	87.17	99.92	99.74	95.38
SE + CV	86.48	90.92	88.53	99.67	98.91	95.11
SE + RG	86.94	89.74	86.90	99.82	99.06	93.96
TE + AC	86.05	90.81	86.89	99.90	99.77	95.14
TE + CV	84.42	88.39	87.04	99.73	98.58	94.86
TE + RG	86.53	89.55	86.54	99.85	98.83	94.05
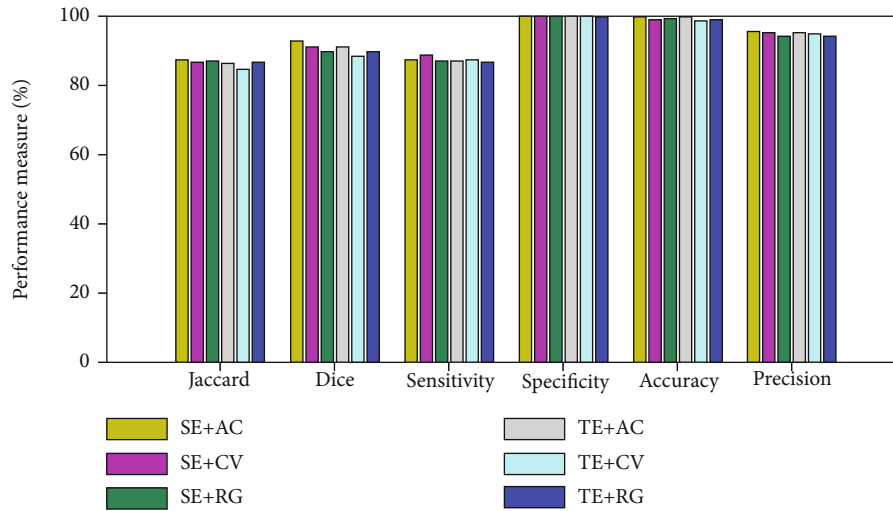
Figure 7: Performance evaluation of CBDM with chosen processing methods.

Table 5: Average values of TBA and CCA of ABIDE dataset (60 volunteers).

Parameter	FCM + LS [11]		Multiphase LS [9]		SE + AC		TE + AC	
	Controlled	ASD	Controlled	ASD	Controlled	ASD	Controlled	ASD
TBA	0.87	0.92	0.59	0.79	0.8104	0.8826	0.8685	0.9175
CCA	0.90	0.75	0.82	0.69	0.9092	0.7761	0.8917	0.7481

result computed for the ABIDE database (60 volunteers) in percentage is presented in Table 4, and its graphical representation is presented in Figure 7. This tabulation and figure confirm that the overall image similarity and the statistical outcomes obtained with SE + AC are superior compared with other approaches. This also authenticates
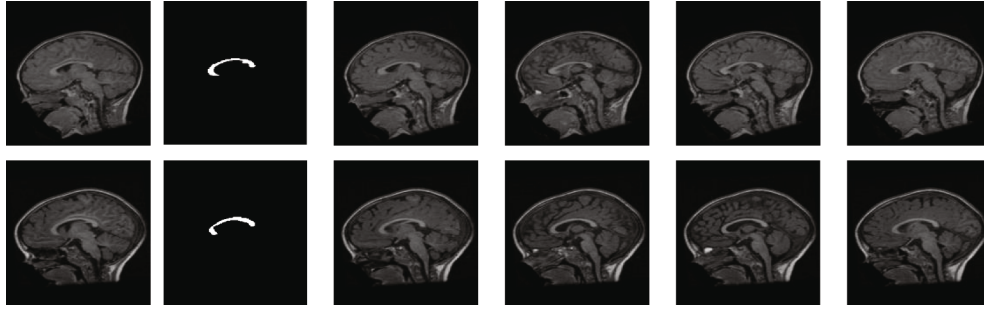
FIGURE 8: Sample test pictures and a sample result obtained with the MIDAS database.

(a)　　　　　　　　(b)

(c)　　　　　　　　(d)
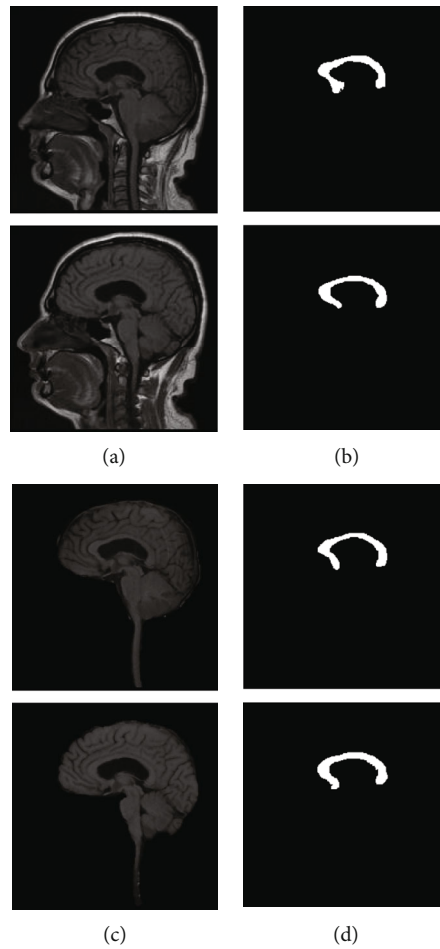
FIGURE 9: Sample 2D slices of real scan pictures and its corresponding result: (a) the sample test image with skull section, (b) segmented CC, (c) sample image without skull section, and (d) extracted CC from image (c).

that AC outperforms the CV and RG for the considered dataset. Table 5 presents the computed values of TBA and CCA, and this result also authenticates that the average results of FCM + LS, SE + AC, and TE + AC are roughly identical.

The results depicted in Figures 8 and 9 also confirm that the method based on SE + AC and TE + AC provides better result on the MIDAS and Proscans datasets. From these outcomes, it can be understood that proposed CBDM has more

efficacy in mining the CC segment from the T1 modality brain MRI slices.

This work also confirms that the average simulation period taken by SE + AC/TE + AC for ABIDE dataset is smaller (171.19 sec/168.94 sec) compared to other approaches (SE + CV = 192.16 sec, SE + RG = 174.28, TE + CV = 191.38, and TE + RG = 172.57 sec). The main limitation of the proposed technique is it implemented the semiautomated segmentation techniques, such as AC, CV, and LS procedures

to mine the CC section. In future, the segmentation methods such as super pixel [66] and local binary pattern [67, 68] can be considered to extract the CC. Further, the planned method can be considered to evaluate the medical level brain MRI collected from volunteers who are associated with autism.

5. Conclusion

This paper suggested a computerized CC extraction tool with a two-step image processing scheme. The instigated method considers the blend of CCS-assisted trilevel thresholding with Shannon's/Tsallis entropy and segmentation based on the CV/RG/AC procedure. During the investigational assessment, the benchmark datasets, such as ABIDE and MIDAS, are used for the preliminary evaluation. Further, this tool is tested on the clinical 2D sagittal MRI of T1 modality obtained from a scan centre. The experimental investigation authorizes that proposed tool extracts the CC region from the brain picture with better accuracy and helps to compute the TBA and CCA for the 2D brain MRI. A comparative study also confirms that the results are approximately similar to the result existing in the literature with fuzzy C-means+LS procedure and better than multiphase LS. The proposed CBDM also offers better segmentation result for the clinical images. Hence, for the forthcoming requirements, this method can be considered in medical clinics to estimate the sagittal view MRI recorded with T1 modality. The proposed outcome of this work could be extended further to aid the investigations in identifying the prognosis of the disease at various stages.

Data Availability

The authors of this paper thank the contributors of the ABIDE and MIDAS database. The major part of the data considered to support the findings of this study is collected from these databases.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

References

[1] S. Kollias, "Insights into the connectivity of the human brain using DTI," *Nepalese Journal of Radiology*, vol. 1, no. 1, pp. 78–91, 2011.

[2] L. B. N. Hinkley, E. J. Marco, A. M. Findlay et al., "The role of corpus callosum development in functional connectivity and cognitive processing," *PLoS One*, vol. 7, no. 8, article e39804, 2012.

[3] L. K. Paul, C. Corsello, D. P. Kennedy, and R. Adolphs, "Agenesis of the corpus callosum and autism: a comprehensive comparison," *Brain*, vol. 137, no. 6, pp. 1813–1829, 2014.

[4] J. J. Wolff, G. Gerig, J. D. Lewis et al., "Altered corpus callosum morphology associated with autism over the first 2 years of life," *Brain*, vol. 138, no. 7, pp. 2046–2058, 2015.

[5] T. W. Frazier and A. Y. Hardan, "A meta-analysis of the corpus callosum in autism," *Biological Psychiatry*, vol. 66, no. 10, pp. 935–941, 2009.

[6] R. Nandhini Abirami, P. M. Durai Raj Vincent, K. Srinivasan, U. Tariq, and C.-Y. Chang, "Deep CNN and deep GAN in computational visual perception-driven image analysis," *Complexity*, vol. 2021, Article ID 5541134, 30 pages, 2021.

[7] A. Lefebvre, A. Beggiato, T. Bourgeron, and R. Toro, "Neuroanatomical diversity of corpus callosum and brain volume in autism: meta- analysis, analysis of the autism brain imaging data exchange project, and simulation," *Biological Psychiatry*, vol. 78, no. 2, pp. 126–134, 2015.

[8] B. A. Ardekani, K. Figarsky, and J. J. Sidtis, "Sexual dimorphism in the human corpus callosum: an MRI study using the OASIS brain database," *Cerebral Cortex*, vol. 23, no. 10, pp. 2514–2520, 2013.

[9] A. R. Jac Fredo, G. Kavitha, and S. Ramakrishnan, "Segmentation and analysis of brain subcortical regions using regularized multiphase level set in autistic MR images," *International Journal of Imaging Systems and Technology*, vol. 24, no. 3, pp. 256–262, 2014.

[10] A. R. Jac Fredo, G. Kavitha, and S. Ramakrishnan, "Automated segmentation and analysis of corpus callosum in autistic MR brain images using fuzzy-c-means-based level set method," *Journal of Medical and Biological Engineering*, vol. 35, no. 3, pp. 331–337, 2015.

[11] A. R. J. Fredo, G. Kavitha, and S. Ramakrishnan, "Segmentation and analysis of corpus callosum in autistic MR brain images using reaction diffusion level sets," *Journal of Medical Imaging and Health Informatics*, vol. 5, no. 4, pp. 737–741, 2015.

[12] D. L. Floris, M. C. Lai, T. Nath, M. P. Milham, and A. D. Martino, "Network-specific sex differentiation of intrinsic brain function in males with autism," *Molecular Autism*, vol. 9, no. 1, p. 17, 2018.

[13] K. Riddle, C. J. Cascio, and N. D. Woodward, "Brain structure in autism: a voxel-based morphometry analysis of the autism brain imaging database exchange (ABIDE)," *Brain Imaging and Behavior*, vol. 11, no. 2, pp. 541–551, 2017.

[14] T. Eilam-Stock, T. Wu, A. Spagna, L. J. Egan, and J. Fan, "Neuroanatomical alterations in high- functioning adults with

autism spectrum disorder," *Frontiers in Neuroscience*, vol. 10, p. 237, 2016.

[15] C. Vachet, B. Yvernault, K. Bhatt et al., "Automatic corpus callosum segmentation using a deformable active Fourier contour model," *Proceedings of SPIE The International Society for Optical Engineering*, vol. 8317, 2012.

[16] S. Kanniappan, D. Samiayya, D. R. Vincent P M et al., "An efficient hybrid fuzzy-clustering driven 3D-modeling of magnetic resonance imagery for enhanced brain tumor diagnosis," *Electronics*, vol. 9, no. 3, p. 475, 2020.

[17] B. Abhinaya and N. S. M. Raja, "Solving multi-level image thresholding problem—an analysis with cuckoo search algorithm," in *Advances in Intelligent Systems and Computing*, vol. 339, pp. 177–186, Springer, 2015.

[18] 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 (IJCSIS)*, vol. 14, no. CIC 2016, pp. 51–56, 2016.

[19] K. R. Hiess, R. A. Alter, S. Sojoudi, B. Ardekani, R. Kuzniecky, and H. R. Pardoe, "Corpus callosum area and brain volume in autism spectrum disorder: quantitative analysis of structural MRI from the ABIDE database," *Journal of Autism and Developmental Disorders*, vol. 45, no. 10, pp. 3107–3114, 2015.

[20] "ABIDE corpus callosum and brain segmentation data," https://sites.google.com/site/hpardoe/cc_abide.

[21] "NAMIC: Brain 2-4 years old," http://insight-journal.org/midas/community/view/24.

[22] http://proscans.in/.

[23] V. Rajinikanth, S. C. Satapathy, S. L. Fernandes, and S. Nachiappan, "Entropy based segmentation of tumor from brain MR images - a study with teaching learning based optimization," *Pattern Recognition Letters*, vol. 94, pp. 87–95, 2017.

[24] V. Rajinikanth, N. S. M. Raja, and K. Kamalanand, "Firefly algorithm assisted segmentation of tumor from brain MRI using Tsallis function and Markov random field," *Control Engineering and Applied Informatics*, vol. 19, no. 3, pp. 97–106, 2017.

[25] 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.

[26] D. Polak, S. Cauley, B. Bilgic et al., "Joint multi-contrast variational network reconstruction (jVN) with application to rapid 2D and 3D imaging," *Magnetic Resonance in Medicine*, vol. 84, no. 3, pp. 1456–1469, 2021.

[27] A. Elsayed, F. Coenen, C. Jiang, M. García-Fiñana, and V. Sluming, "Corpus callosum MR image classification," *Knowledge-Based Systems*, vol. 23, no. 4, pp. 330–336, 2010.

[28] G. S. Cover, W. G. Herrera, M. P. Bento, S. Appenzeller, and L. Rittner, "Computational methods for corpus callosum segmentation on MRI: a systematic literature review," *Computer Methods and Programs in Biomedicine*, vol. 154, pp. 25–35, 2018.

[29] P. A. Yushkevich, J. Piven, H. C. Hazlett et al., "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[30] "ITK-SNAP," http://www.itksnap.org/pmwiki/pmwiki.php.

[31] P. T. Krishnan, P. Balasubramanian, and C. Krishnan, "Segmentation of brain regions by integrating meta- heuristic multilevel threshold with markov random field," *Current Medical Imaging Reviews*, vol. 12, no. 1, pp. 4–12, 2016.

[32] P. Kalavathi and V. B. S. Prasath, "Methods on skull stripping of MRI head scan images—a review," *Journal of Digital Imaging*, vol. 29, no. 3, pp. 365–379, 2016.

[33] X. S. Yang and S. Deb, "Cuckoo search via Lévy flights," in *World Congress on Nature & Biologically Inspired Computing (NaBIC)*, pp. 210–214, Coimbatore, India, 2009.

[34] X. S. Yang, S. Deb, and S. K. Mishra, "Multi-species cuckoo search algorithm for global optimization," *Cognitive Computation*, vol. 10, no. 6, pp. 1085–1095, 2018.

[35] X. S. Yang and S. Deb, "Engineering optimisation by cuckoo search," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 1, no. 4, pp. 330–343, 2010.

[36] X. S. Yang and S. Deb, "Cuckoo search: recent advances and applications," *Neural Computing and Applications*, vol. 24, no. 1, pp. 169–174, 2014.

[37] X. S. Yang and S. Deb, "Cuckoo search: state-of-the-art and opportunities," in *4th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, pp. 55–59, Mauritius, 2017.

[38] X.-S. He, F. Wang, Y. Wang, and X.-S. Yang, "Global convergence analysis of cuckoo search using Markov theory," *Nature-Inspired Algorithms and Applied Optimization*, vol. 744, pp. 53–67, 2018.

[39] A. H. Gandomi, X. S. Yang, and A. H. Alavi, "Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems," *Engineering Computations*, vol. 29, no. 1, pp. 17–35, 2013.

[40] X. S. Yang, "CUCKOO search and firefly algorithm: theory and applications," in *Studies in Computational Intelligence*, vol. 516, Springer, 2014.

[41] X. S. Yang, *Nature-Inspired Optimization Algorithms*, Elsevier Insight, London, 2014.

[42] H. Sharma, J. C. Bansal, K. V. Arya, and X. S. Yang, "Lévy flight artificial bee colony algorithm," *International Journal of Systems Science*, vol. 47, no. 11, pp. 2652–2670, 2016.

[43] A. Sharma, H. Sharma, A. Bhargava, N. Sharma, and J. C. Bansal, "Optimal power flow analysis using Lévy flight spider monkey optimisation algorithm," *International Journal of Artificial Intelligence and Soft Computing*, vol. 5, no. 4, pp. 320–352, 2016.

[44] K. Ikeda, "Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system," *Optics Communication*, vol. 30, no. 2, pp. 257–261, 1979.

[45] K. Ikeda, H. Daido, and O. Akimoto, "Optical turbulence: chaotic behavior of transmitted light from a ring cavity," *Physical Review Letters*, vol. 45, no. 9, pp. 709–712, 1980.

[46] S. C. Satapathy, N. S. M. 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.

[47] S. Lalwani, H. Sharma, S. C. Satapathy, K. Deep, and J. C. Bansal, "A survey on parallel particle swarm optimization algorithms," *Arabian Journal for Science and Engineering*, vol. 44, no. 4, pp. 2899–2923, 2019.

[48] S. P. Sahoo, S. Nayak, and M. R. Kabat, "A reference-based multiobjective bacteria foraging optimization technique for

qos multicast routing," *Arabian Journal for Science and Engineering*, vol. 43, no. 12, pp. 7457–7472, 2018.

[49] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[50] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Computer Vision, Graphics, and Image Processing*, vol. 29, no. 3, pp. 273–285, 1985.

[51] P. L. Kannappan and P. N. Rathie, "On a characterization of directed divergence," *Probability Theory and Related Fields*, vol. 22, no. 2, pp. 163–171, 1973.

[52] S. Paul and B. Bandyopadhyay, "A novel approach for image compression based on multi-level image thresholding using Shannon entropy and differential evolution," in *Students' Technology symposium (TechSym)*, pp. 56–61, Kharagpur, India, 2014.

[53] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *Journal of Statistical Physics*, vol. 52, no. 1-2, pp. 479–487, 1988.

[54] C. Tsallis, "Entropic nonextensivity: a possible measure of complexity," *Chaos, Solitons & Fractals*, vol. 13, no. 3, pp. 371–391, 2002.

[55] J. B. T. M. Roerdink and A. Meijster, "The watershed transform: definitions, algorithms and parallelization strategies," *Fundamenta Informaticae*, vol. 41, pp. 187–228, 2001.

[56] Y. Zhang and L. Wu, "An MR brain images classifier via principal component analysis and kernel support vector machine," *Progress In Electromagnetics Research*, vol. 130, pp. 369–388, 2012.

[57] K. S. Chuang, H. L. Tzeng, S. Chen, J. Wu, and T. J. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," *Computerized Medical Imaging and Graphics*, vol. 30, no. 1, pp. 9–15, 2006.

[58] 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*, vol. 19, no. 12, pp. 3243–3254, 2010.

[59] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[60] T. F. Chan and L. A. Vese, "Active contour and segmentation models using geometric pde's for medical imaging," in *Geometric Methods in Bio-Medical Image Processing*, Mathematics and Visualization, R. Malladi, Ed., pp. 63–75, Springer, Berlin, Heidelberg, 2002.

[61] F. Y. Shih and S. Cheng, "Automatic seeded region growing for color image segmentation," *Image and Vision Computing*, vol. 23, no. 10, pp. 877–886, 2005.

[62] X. Bresson, S. Esedoḡlu, P. Vandergheynst, J. P. Thiran, and S. Osher, "Fast global minimization of the active contour/snake model," *Journal of Mathematical Imaging and Vision*, vol. 28, no. 2, pp. 151–167, 2007.

[63] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "A novel content-based active contour model for brain tumor segmentation," *Magnetic Resonance Imaging*, vol. 30, no. 5, pp. 694–715, 2012.

[64] H. Lu, A. C. Kot, and Y. Q. Shi, "Distance-reciprocal distortion measure for binary document images," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 228–231, 2004.

[65] R. F. Moghaddam and M. Cheriet, "A multi-scale framework for adaptive binarization of degraded document images," *Pattern Recognition*, vol. 43, no. 6, pp. 2186–2198, 2010.

[66] Y. Zhang, X. Li, X. Gao, and C. Zhang, "A simple algorithm of superpixel segmentation with boundary constraint," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 7, pp. 1502–5014, 2016.

[67] C. Wang, D. Li, Z. Li et al., "An efficient local binary pattern based plantar pressure optical sensor image classification using convolutional neural networks," *Optik*, vol. 185, pp. 543–557, 2019.

[68] A. Tal, T. Zhao, C. Schirda, H. P. Hetherington, J. W. Pan, and O. Gonen, "Fast, regional three-dimensional hybrid (1D-Hadamard 2D-rosette) proton MR spectroscopic imaging in the human temporal lobes," *NMR in Biomedicine*, vol. 34, no. 6, article e4507, 2021.

Review Article

Hybrid Inception v3 XGBoost Model for Acute Lymphoblastic Leukemia Classification

S. Ramaneswaran [ID],[1] **Kathiravan Srinivasan** [ID],[2] **P. M. Durai Raj Vincent** [ID],[1] **and Chuan-Yu Chang** [ID][3]

[1]*School of Information Technology and Engineering, Vellore Institute of Technology (VIT), Vellore, India*
[2]*School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Vellore, India*
[3]*Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan*

Correspondence should be addressed to Chuan-Yu Chang; chuanyu@yuntech.edu.tw

Acute lymphoblastic leukemia (ALL) is the most common type of pediatric malignancy which accounts for 25% of all pediatric cancers. It is a life-threatening disease which if left untreated can cause death within a few weeks. Many computerized methods have been proposed for the detection of ALL from microscopic cell images. In this paper, we propose a hybrid Inception v3 XGBoost model for the classification of acute lymphoblastic leukemia (ALL) from microscopic white blood cell images. In the proposed model, Inception v3 acts as the image feature extractor and the XGBoost model acts as the classification head. Experiments indicate that the proposed model performs better than the other methods identified in literature. The proposed hybrid model achieves a weighted F1 score of 0.986. Through experiments, we demonstrate that using an XGBoost classification head instead of a softmax classification head improves classification performance for this dataset for several different CNN backbones (feature extractors). We also visualize the attention map of the features extracted by Inception v3 to interpret the features learnt by the proposed model.

1. Introduction

Leukemia is a malignancy that originates in cells that would otherwise develop into different types of blood cells. Most often, leukemia starts in the form of white blood cells (WBCs), but some leukemias start in other blood cell types as well. Their primary classification of leukemia is based on whether the leukemia is acute (fast-growing) or chronic (slower-growing) and whether it starts in myeloid cells or lymphoid cells. Knowing the specific type of leukemia helps doctors better predict each person's prognosis and select the best treatment.

Acute lymphocytic leukemia (ALL) is also called acute lymphoblastic leukemia. "Acute" means that if left untreated, leukemia can progress rapidly and cause fatality within months. "Lymphocytic" means it develops from early (immature) forms of lymphocytes, a type of WBC.

ALL starts in the bone marrow (the soft inner part of certain bones, where new blood cells are made). Most often, the leukemia cells invade the blood fairly quickly. They can also sometimes spread to other parts of the body, including the lymph nodes, liver, spleen, central nervous system (brain and spinal cord), and testicles (in males). Some cancers can also start in these organs and then spread to the bone marrow, but these cancers are not leukemia.

Acute lymphoblastic leukemia (ALL) is the most common type of childhood cancer and accounts for approximately 25% of pediatric cancers [1]. Approximately 74% of people under the age of twenty who are diagnosed with leukemia are diagnosed with ALL. Most cases occur between the ages of 2 and 5. ALL accounts for less than 1% of all new cancer cases worldwide and also accounts for less than 1% of all cancer-related deaths.

The 5-year survival rate gives us the percent (out of 100) of children and teenagers who live at least 5 years after being diagnosed with cancer. The 5-year survival rate for children between age 0 and 14 is 91%. The 5-year survival rate for people between ages 15 and 19 is 75%. It is rare for ALL to recur after 5 years; hence, children diagnosed with ALL who remain free from the disease after 5 years are generally considered cured.

98% of the children with ALL go into remission, and 85% of those with first-time ALL are expected to have long-term complications. However, the chance of recovery for adults is not high, as the percent of adults cured with current treatment is 20%-40%.

ALL is a life-threatening disease that can rapidly spread through children's bodies if left untreated and can cause death within a few weeks. During the diagnosis of leukemia, a necessary step is for the physician to classify the white blood cells in the bone marrow. Not only is this step difficult and complex, but it also results in increased human error and procedure time. This process can be automated by developing computerized methods to automatically classify the white blood cells. Not only does this method decreases the diagnosis time and error, but it also is economical especially with the increasing trend in digitizing microscopic images.

However, this task is not trivial; there are several challenges associated with the classification of white blood cell (WBC) images, the main challenge being the morphological similarity between the normal and the immature leukemic blast cells. Another challenging aspect in distinguishing WBCs is that they are surrounded by other blood components like red blood cells and platelets.

There are several methods and algorithms used for medical imaging; however, convolutional neural networks (CNNs) have proven to be the best choice. Pretrained neural networks such as VGGNet, ResNet, and Inception have been successfully utilized in various medical imaging applications. Moreover, these CNNs mitigate the issue of lack of sufficient training data which is a common problem in medical datasets by utilizing transfer learning, where the CNNs are trained on massive generic datasets and then trained on a specific downstream class on smaller datasets.

Our main motivation in this study is to develop a robust and efficient model for the classification of ALL from microscopic images. Medical image datasets are small; hence, it is often not feasible to train a CNN from scratch; hence, we aim to leverage the transfer learning ability of pretrained CNN architectures to learn a classifier for the C-NMC 2019 dataset. To improve the performance of these CNNs, we explore the use of different classification heads instead of a conventional softmax classification head. We aim to experiment with several data preprocessing techniques to improve the generalizability and performance of the model. We also aim to investigate and justify our choice of model design through extensive experiments presented in Ablation Study.

To this end, we introduce a hybrid Inception v3 XGBoost model which uses XGBoost as a classification head on top of an Inception v3 model fine-tuned for classification on this dataset. We perform extensive experimentation with several pretrained CNNs and different augmentation techniques.

Table 1: Comparison of the proposed approach with recent studies on leukemia detection.

Method	Accuracy	Dataset	Year
Yu et al. [29]	88.50%	DTH	2017
Mourya et al. [30]	89.62%	ISBI	2018
Kassani et al. [31]	96.17%	ISBI	2019
Bodzas et al. [10]	100%	Blood smear	2020
Kasani et al. [16]	96.58%	ISBI	2020
Shafique and Tehsin [12]	99.50%	ALL-IDB	2018
Proposed approach	98.50%	ISBI	2021

We also investigate the features learnt by the Inception v3 model visualizing the heat map of its feature maps using Grad-CAM. We have performed experiments that indicate the effectiveness of our model and justify the design; these experiments are presented in Ablation Study.

The major contributions of this proposed model are the following:

(i) The proposed model gives a high weighted F1 score of 0.98 for the C-NMC 2019 dataset

(ii) The proposed architecture involving the use of XGBoost classification head can be utilized with several CNN backbone feature extractors and results in increased performance (refer to Table 1)

(iii) The model can be interpreted using attention maps of the feature maps extracted by the Inception v3 CNN

The paper is divided into 8 sections. Recent literature pertaining to leukemia detection is reviewed in Section 2. Section 3 briefly describes the dataset used in this study. The proposed model and methodology are discussed in Section 4. The implementation details are provided in Section 5. Section 6 discusses the experimental results. Section 7 presents an ablation study for our hybrid model. Finally, we conclude the study and discuss the future directions in Section 8.

To reproduce our results, we present detailed implementation details in Implementation Details. Moreover, the full code for experiments conducted in this research is publicly available at https://github.com/ramaneswaran/lymphoblastic-leukemia-detection.

2. Literature Review

There has been a lot of research into the classification of white blood cells. Early approaches to this problem involve using traditional image processing techniques and machine learning models for classification. Jagadev and Virani [2] present an approach to classify leukemia lymphocyte images using handcrafted image features and SVM classifier. Amin et al. [3] propose yet another method involving SVM classifiers to detect acute lymphoblastic leukemia (ALL) where the geometrical and statistical features of nuclei are used to train the classifier. Rodellar et al. [4] present an approach

for morphological characterization and automatic cell image recognition using handcrafted quantitative features. Mahmood et al. [5] experiment with several models including random forest, gradient-boosted machine, and CART for the detection of pediatric ALL; from their experiments, they conclude that the best fitting model for the dataset used in the research was the CART model.

In recent literature, deep learning-based methods have been utilized for ALL classification and have met with significant success. Pretrained CNNs, as well as custom CNNs, have been successfully trained and tested on several cell classification tasks.

Macawile et al. [6] propose a method for white blood cell (WBC) classification and counting using pretrained CNNs. They use modified AlexNet, GoogleNet, and ResNet-101 in tandem to obtain classification results. Hegde et al. [7] provide a comparison between traditional image processing approaches and deep learning methods in the task of classifying WBCs. Using neural network architecture gives a significant performance increase over traditional methods. Sharma et al. [8] present a custom CNN architecture for white blood cell classification; the proposed network consists of 2D convolutions and MaxPooling layers with Relu activations. This architecture achieves high accuracy scores for both binary classification and multiclass classification settings. Habibzadeh et al. [9] present a method for utilizing the ResNet and Inception network for WBC classification. The proposed method also utilizes several augmentation techniques in the preprocessing stage. WBC classification is done using hierarchy topological feature extraction by the CNNs.

In [10], Bodzas et al. propose an approach to automatically identify ALL from peripheral blood smear images using conventional image processing techniques and ML algorithms. The approach uses an extensive preprocessing and three-phase filtration algorithm. Sixteen handcrafted features were extracted from the image and were used as input to SVM and ANN classifiers. Muntasa and Yusuf [11] present a model that detects ALL using principal object characteristics of a color image. There are four main stages in the proposed approach; these are enhancement, segmentation, feature extraction, and accuracy measurement. The proposed method archived the maximum accuracy on the ALL-IDB dataset. Shafique and Tehsin [12] compare the different methods for the early detection of ALL. The various stages in the diagnosis procedure are comparatively analyzed in their study. They also discuss the advantages and disadvantages of each method. Shafique and Tehsin [13] present an approach that uses pretrained AlexNet which is fine-tuned for the task of classification of ALL into its 4 subtypes (L1, L2, L3, L3, and normal). The last 4 layers are replaced with new linear layers, and their weights are trained from scratch. The research also employs several data augmentation techniques to generalize the model performance. The model achieves high accuracy of 99.5% for detection of ALL and 96.06% for ALL subtype classification.

Bhuiyan et al. [14] propose a framework for identifying ALL from microscopic images of WBC. A total of four different statistical models are used for classification, and their performance is compared. From the experimental results, the authors conclude that the SVM model gave the best fit for their dataset. Acharya and Kumar [15] survey various methodologies in current literature that are used to segment WBCs and provide a novel method for segmenting the nucleus and the cytoplasm of the WBC. Subsequently, models are built to extract features and perform supervised classification of the microscopic images into the four subtypes of ALL. The model achieves an accuracy of 98.6% for the dataset used. Kasani et al. [16] propose to use a pretrained CNN model in an aggregated fashion to detect ALL from microscopic WBC images. The authors use several data augmentation techniques to avoid overfitting. The proposed network consists of a VGG19 and a NASNetLarge which are used together for classification. The final ensemble produced an overall accuracy of 96.58% which is higher than any of the individual networks.

An extensive survey on the current trends and approaches to the detection of leukemia from microscopic images is presented in [17–19].

3. Dataset

The dataset used in this research is called the ISBI C-NMC 2019 dataset [20]. The dataset consists of white blood cell images collected from 60 cancer subjects and 41 healthy subjects. The dataset was prepared at Laboratory Oncology, AIIMS, New Delhi. There are a total of 10661 cell images in the dataset. The train, validation, and test splits were 75%, 15%, and 15%, respectively. Figure 1 illustrates the microscopic white blood cell images from the C-NMC 2019 Challenge dataset. Figure 2 portrays the class distribution of the C-NMC 2019 dataset.

To remove the variations in illumination, a stain normalization process has been applied to the images. The normalization procedures applied to this dataset have been described in detail in [21–25].

4. Proposed Approach

In this section, we describe our proposed model. Figure 3 shows the architecture of the proposed hybrid Inception v3 XGBoost model. Figure 4 portrays the architecture of the Inception v3 model. The proposed model consists of two components, an image feature extractor and a classification head. Generally, the classification head in a pretrained CNN for image classification tasks is a softmax classifier. In the proposed model, however, we use the XGBoost classifier as a classification head. The input features used for this XGBoost classifier are provided by the fine-tuned Inception v3 model. Through experiments, we also show that this setup works for several other pretrained CNNs too.

The proposed model is trained in two stages. In the first stage of training, we fine-tune the Inception v3 model on the training data. Through experiments, we observe that using features from fine-tuned Inception v3 leads to better classification results by the XGBoost classifier as opposed to using a pretrained Inception v3 directly as a feature extractor (refer to Figure 5).
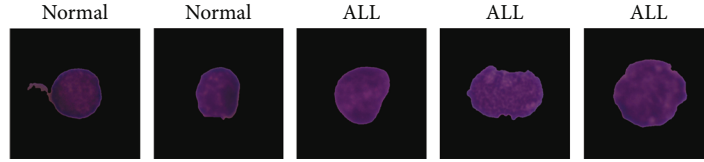
Figure 1: Microscopic white blood cell images from the C-NMC 2019 Challenge dataset.

4.1. Data Preprocessing. We have utilized the following preprocessing techniques to preprocess the dataset being used.

(1) *Center Cropping.* There is a considerable black margin in the image which is redundant to classification. Hence, the image is center cropped to size 448×448

(2) *Resizing.* The images in the dataset are of size 450×450; however, Inception v3 requires input images of size 299×299. Hence, we resize the image from 448×448 (center cropped image) to 299×299 using bicubic interpolation

(3) *Data Augmentation.* Medical image datasets are mostly limited in size owing to privacy and data acquisition issues. To prevent overfitting and improve generalization, we have applied several image augmentation techniques. Microscopic cell images are direction invariant; hence, we applied conventional image augmentation techniques such as rotation and flipping. We also used cutout [26] augmentation that acts as a regularizer by randomly masking out square regions of input during training

(4) *Normalization.* The images are normalized with ImageNet mean and standard deviation. These values are precomputed standards derived from the ImageNet database

4.2. Image Feature Extraction. Literature review on recent works of medical imaging suggests that deep convolutional networks pretrained on large datasets such as ImageNet provide the best results for medical image classification tasks. This is due to the fact that medical image datasets are difficult to collect and are usually small in size. Hence, it becomes difficult to train CNNs from scratch which often results in overfitting. However, pretrained CNNs help in avoiding this problem as we can use transfer learning to fine-tune these CNNs on medical datasets. We experiment with several popular CNN architectures such as ResNet and DenseNet to select the model which performs the best. We fine-tune these CNNs for the task of classification and choose the model with the best weighted F1 score. Refer to Experimental Results and Discussion and Table 2 for more details.

We employ an Inception v3 [27] model that is initialized with ImageNet weights and fine-tuned on the train set to extract feature maps for images. After experimenting with several pretrained CNN models for this task, Inception v3 gave the best F1 score. Inception v3 is the 3rd version of CNN from the inception family of architecture that makes several improvements. These improvements include factor-
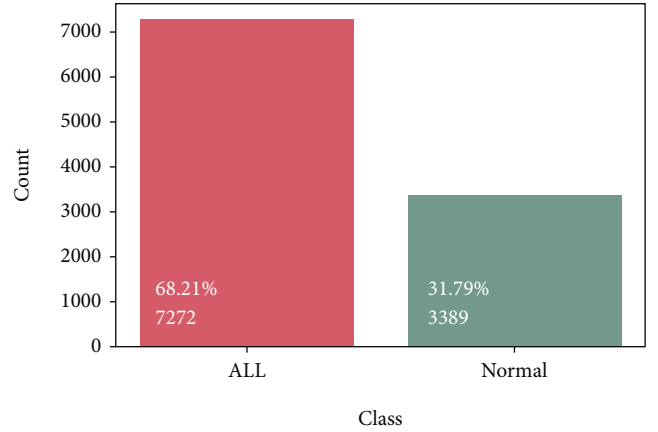
Figure 2: Class distribution of the C-NMC 2019 dataset.

ized convulsions that reduce the number of parameters without decreasing the network efficiency. It uses label smoothing to act as a regularizer. Additionally, it utilizes an auxiliary classifier to propagate label information lower down the network and further help in regularization.

4.3. Classification Head. We employ an XGBoost [28] classifier to classify the cell images as leukemic blasts or normal. XGBoost is a machine learning algorithm used for both classification and regression modelling tasks. It is an ensemble of gradient-boosted decision trees. Gradient boosting is an approach where new models are created that predict the residuals or errors of prior models and then added together to make the final prediction. It is a special case of boosting algorithms where errors are minimized by a gradient descent algorithm.

4.4. Training Details

4.4.1. Stage 1 Training. In the first stage of training, Inception v3 is trained on the training set. We employ the pretrained ImageNet weights for Inception v3. The last fully connected layer in Inception v3 is replaced with a 2-node softmax classifier. The parameters for this replaced layer were randomly initialized:

$$\text{Softmax}(x_i) = \frac{\exp\left(x_i\right)}{\Sigma_j \exp\left(x_j\right)}. \tag{1}$$

The softmax function is used to convert logits of the classifier into a probability distribution. Each element of the output lies in the interval $[0, 1]$, and the output elements sum up to 1. The input image is assigned to the class with maximum probability. Equation (1) depicts the formula for softmax function, where $\exp\left(x_i\right)$ is the exponent of the current
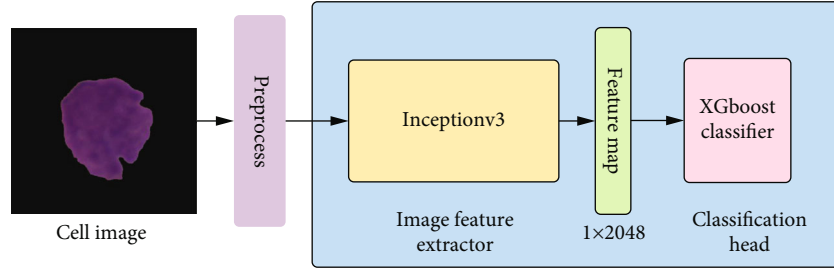
FIGURE 3: The architecture of the proposed hybrid Inception v3 XGBoost model.
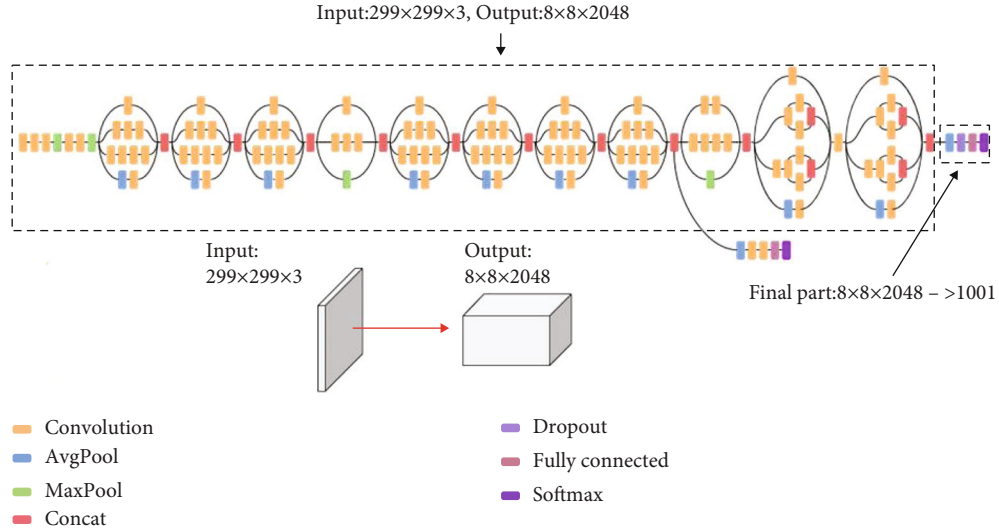
FIGURE 4: The architecture of Inception v3 model.
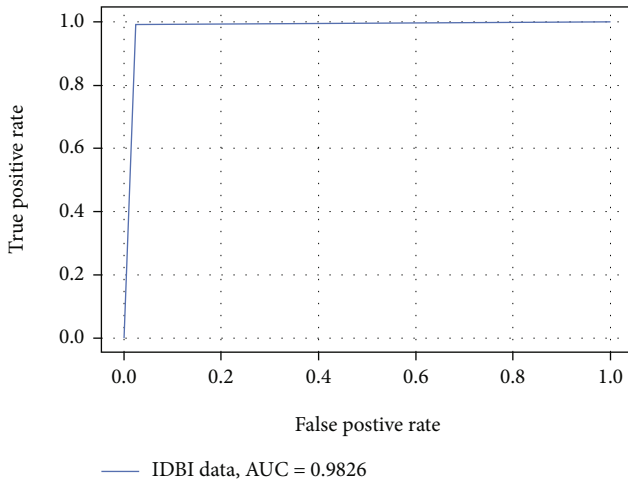
FIGURE 5: ROC curve for the hybrid model on test data.

output logit and $\Sigma_j \exp(x_j)$ is the summation of the exponent of all output logits.

From Figure 2, we can observe that the dataset has a class imbalance problem. To address this problem, we use a weighted cross-entropy loss function. This function is given by the formula

$$\mathrm{loss}(x, \mathrm{class}) = \mathrm{weight}[\mathrm{class}]\left(-x[\mathrm{class}] + \log\left(\Sigma_j \exp\left(x[j]\right)\right)\right), \quad (2)$$

where weight[class] refers to the weight assigned to each class. To minimize the effect of class imbalance, we assign larger weights for minority classes. The losses are averaged across observations for each minibatch. In this case, it is a weighted average given by

$$\mathrm{loss} = \frac{\Sigma_i^N \mathrm{loss}(i, \mathrm{class}[i])}{\Sigma_i^N \mathrm{weight}[\mathrm{class}[i]]}. \quad (3)$$

During this stage of training, we used several augmentation techniques that were mentioned in Data Preprocessing. Using the image augmentation helps the model generalize better and improve performance. Figure 6 compares the validation loss during training of two different Inception v3 models, one which uses image augmentation on the input images and the other which does not use it. Using image augmentation improves the performance of the model.

4.4.2. Stage 2 Training. In the second stage of training, an XGBoost classifier is trained to classify the cell images as normal or leukemic blasts. The XGBoost classifier is trained

Input: *A microscopic WBC image M*
Output: *Pre-processed image M_t*
Procedure $M_{cropped}$ = *CenterCropped (M)*
$M_{resized}$ = *Resizing ($M_{cropped}$)*
$M_{flipped}$ = *HorizontalFlip ($M_{resized}$)*
$M_{flipped}$ = *VerticalFlip ($M_{flipped}$)*
$M_{rotated}$ = *Rotate ($M_{flipped}$)*
M_{cutout} = *Cutout ($M_{rotated}$)*
$M_{normalized}$ = *ImageNetNormalization (M_{cutout})*
M_t = *ToTensor ($M_{normalized}$)*

ALGORITHM 1: Data preprocessing in training.

Input: *A microscopic WBC image M*
Output: *Pre-processed image M_t*
Procedure $M_{cropped}$ = *CenterCroppe (M)*
$M_{resized}$ = *Resizin ($M_{cropped}$)*
$M_{normalized}$ = *ImageNetNormalizatio (M_{cutout})*
M_t = *ToTensor ($M_{normalized}$)*

ALGORITHM 2: Data preprocessing in validation/testing.

TABLE 2: Evaluation metrics of different pretrained CNN models from stage I of training. These metrics are reported on the test set after fine-tuning on the train set.

Model	F1 score	Recall	Precision	Accuracy	AUC
AlexNet	0.889	0.894	0.901	0.894	0.832
DenseNet 121	0.871	0.869	0.876	0.869	0.861
ResNet 18	0.917	0.917	0.919	0.917	0.908
VGG 16	0.921	0.924	0.927	0.924	0.880
SqueezeNet	0.930	0.932	0.936	0.932	0.891
MobileNet v2	0.958	0.958	0.958	0.958	0.953
Inception v3	0.979	0.979	0.979	0.979	0.981

using features extracted with the Inception v3 network trained in stage 1.

To extract the features using Inception v3, we remove the softmax classifier from the network and directly obtain the feature map from the penultimate layer. The feature maps obtained are of dimension 2048×1. We use the same training, validation, and test splits that were used in stage 1 training.

5. Implementation Details

All the networks were trained on the Tesla K80 GPU provided by Kaggle's Machine learning kernels. We used the PyTorch library to develop the deep learning models. The models were optimized using Adam optimizer. For the XGBoost classifier, we used the XGBoost library. We used a grid search strategy to tune the model to optimize the loss. The detailed hyperparameter configuration for the proposed model is given in Table 3.
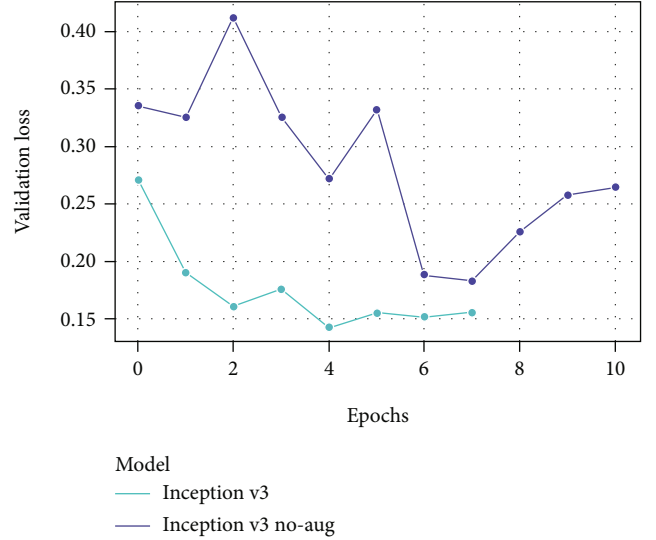
FIGURE 6: Validation loss curve for Inception v3 model during stage 1 of training.

6. Experimental Results and Discussion

In this section, we report the experimental results for our proposed model. The primary evaluation metric that we adopt is the weighted F1 score. We additionally report accuracy, precision, recall, and AUC score.

Once the model is trained, we select the best checkpoint to be used in model inference. The predicted classes are compared to the actual target classes to calculate the aforementioned metrics. We experimented with several CNN backbone feature extractors such as AlexNet and DenseNet during stage 1 of training. We experimented with these CNNs to identify which model can be used as the feature extractor for our hybrid model. Figure 7 compares the validation loss of the different CNN models during stage 1 of training. Among these, Inception, v3 was the best performing model with a weighted F1 score of 0.97. Table 2 displays the evaluation metrics of the various CNN models used during stage 1 of training.

During stage 2 of training, we extracted image features using the Inception v3 model trained in stage 1. These features were used in training an XGBoost classifier. Using an XGBoost classifier on top of this Inception v3 model gave the best result on the test set with a weighted F1 score of 0.98. Figure 8 displays the confusion matrix obtained for the proposed hybrid model. We observe that there are very few misclassified data. We observe that there is a better false positive rate when using an XGBoost classification head over a CNN; this is an essential factor when dealing with the medical diagnosis since it is better to screen a person as diseased and conduct further tests to exclude the disease than exclude a diseased person by falsely predicting a negative.

Sensitivity and specificity are two important metrics that are used to validate medical diagnosis models. Sensitivity reflects the probability that a diagnostic test will return positive for people who are diseased. Specificity on the other hand reflects the probability that a test will return negative for

TABLE 3: Hyperparameter configuration for the proposed model.

Model	Hyperparameter	Value
Adam optimizer	$[\beta_1, \beta_2]$	[0.9, 0.999]
	Learning rate	$1e-4$
XGBoost	n_estimators	1000
	Max_depth	6
	Min_child_weight	3
Loss weights	Normal class	1.5929
	ALL class	0.7330
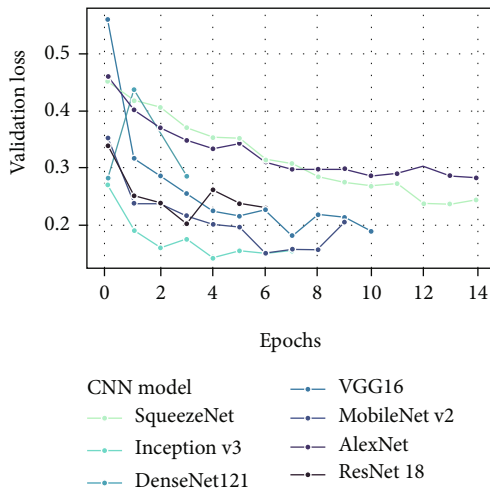
FIGURE 7: Validation loss curve for various pretrained CNNs during stage 1 training.
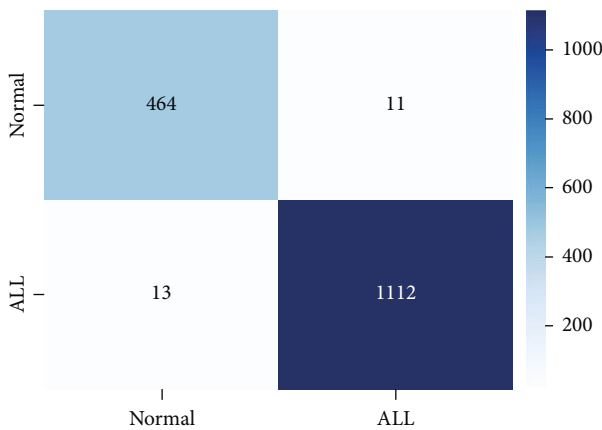
FIGURE 8: Confusion matrix for the hybrid model on test data.

persons without the disease. Clinically, these metrics are important for confirming or excluding disease. We can interpret these metrics from the confusion matrix (Figure 8). The sensitivity is 0.9884, and specificity is 0.9133.

The TPR (true positive rate) and FPR (false positive rate) are important AUC/ROC (Area Under the Curve/Receiver Operating Characteristics) metrics that help to determine the amount of information learnt by the model and how well

it is able to distinguish between the classes. In the ideal case, TPR = 1 and the FPR = 0. Refer to Figure 5 that depicts the ROC curve for the hybrid model on the test data. An AUC of near 1 indicates that a model has excellent separability. We can observe that the model achieves a high AUC of 0.9826. This shows that the proposed model has excellent separability and correctly classifies most of the samples in the test data with very few misclassifications. Also, the FPR is close to 0 and TPR close to 1 from which we can deduce that the model is performing well.

To benchmark and compare our hybrid model, we have selected the following models from recent studies on leukemia detection. These models are trained and validated on either ISBI C-NMC dataset or other similar datasets of microscopic WBC image for ALL classification. Moreover, these models use CNN for feature extraction or have some deep learning components in their model design. We have described these models in brief below.

Yu et al.: to prevent a model from fitting data noise, the authors have combined several CNNs and used their combined output to get classification results. The CNN architectures being used are ResNet50, Inception v3, VGG16, VGG19, and Xception.

Mourya et al.: this approach combines the optical density features and discrete cosine transform domain features extracted through CNN to build the classifier. They use bilinear pooling instead of average pooling after the last convolutional layer to help in fine-grained recognition.

Kassani et al.: in this approach, the image is first enhanced using several preprocessing and augmentation techniques; then, features are extracted using a hybrid VGG16 and MobileNet model. The authors have developed an integrating strategy to overcome the shortcomings of the individual models. Finally, a multilayer perceptron is trained using these features.

Bodzas et al.: in this approach, the image is segmented using a three-phase filtration; then, sixteen handcrafted features are extracted and used for classification by SVM and ANN classifiers.

Kasani et al.: the authors develop an aggregated deep learning model ALL detection. Several data augmentation techniques were applied to overcome dataset size issues, and transfer learning was utilized to accelerate learning. The authors have used the following CNN models: Inception v3, AlexNet, DenseNet201, VGGNet-16, VGGNet19, Xception, MobileNet, ShuffleNet, and two NASNet models.

Shafique and Tehsin: the authors have used a pretrained AlexNet model in their study. They have replaced the last layers with new linear layers and learnt the weights from scratch by fine-tuning on the ALL-IDB dataset. They have employed several data augmentation techniques to overcome overfitting.

We compare the models based on the accuracy obtained since this was a common metric we found in all these studies. Table 1 compares the proposed approach with its counterparts.

A common trend we noticed in these studies is that several CNN models are being aggregated and utilized to make a classification decision; we feel that this approach makes the model unnecessarily complex. Not only does this

FIGURE 9: Feature maps produced by pretrained Inception v3 and fine-tuned Inception v3, respectively. The features learnt by fine-tuned Inception v3 are more discriminative.
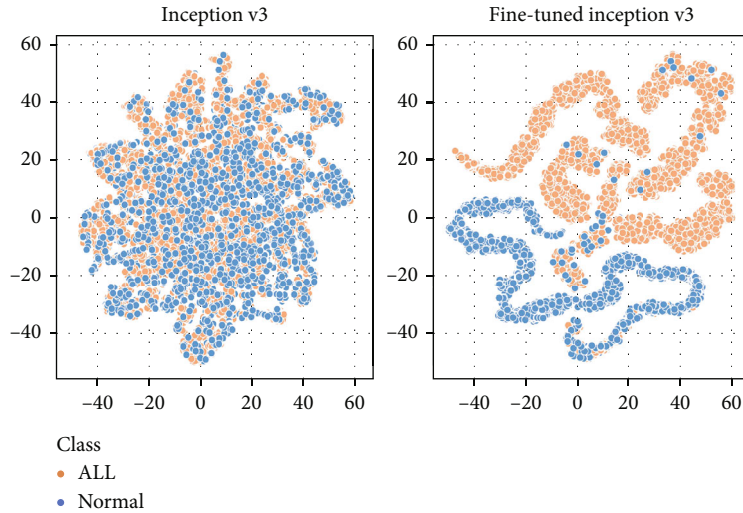
approach require a lot of computation and time resources to train and validate, but it also makes interpreting results more difficult. Our hybrid model achieves similar performance with a single CNN backbone making it simpler without losing performance.

Another limitation we noticed is that the studies do not attempt to interpret and justify classification decisions made by the models. Interpretability of models is of prime importance in building trust and towards the successful integration of these models in everyday medical use. Since we use a single CNN backbone for feature extraction, we can demystify the CNN by visualizing their activation maps of the features extracted (refer to Figure 9).

7. Ablation Study

In this section, we attempt to justify our design choices in developing the proposed hybrid model.

We investigate the effectiveness of using an XGBoost classification head with a fine-tuned CNN model. We experiment with different CNN backbones such as AlexNet and ResNet18 in our proposed model. The goal of this experiment is to demonstrate the effectiveness and generalizability of using the XGBoost classification head over the softmax classification head for this dataset. Table 4 shows the weighted F1 score of hybrid models using different CNN backbones. Table 4 shows that generally there is a significant increase in the performance of the model when used in this setting.

We check whether a pretrained CNN can be direct without fine-tuning on the train set. We conduct this experiment to check for the effectiveness and need for fine-tuning the feature extractor (stage 1 training). When we directly use the pretrained Inception v3 as a feature extractor, we notice that there is a significant drop in performance. We try to investigate the reason behind this by plotting a scatter plot of the features extracted from the Inception v3 (refer to

TABLE 4: Comparison of pretrained CNN models with softmax classifier and with XGBoost classifier. The results are the weighted F1 score on the test set.

Model name	Softmax classification head	XGBoost classification head
AlexNet	0.889	0.897
ResNet 18	0.917	0.957
VGG 16	0.921	0.924
Inception v3	0.979	0.985

Figure 9). We use t-sne to convert the high-dimensional feature maps to lower-dimensional embeddings. We observe that with fine-tuning, the Inception v3 learns better and more discriminative feature representations for the dataset which helps the XGBoost model in making better and more informed classifications, whereas the features from a pretrained off the shelf Inception v3 are not discriminative at all, which is clearly observed in Figure 9.

We also try to understand the inner workings of Inception v3 from stage 1. Being able to interpret the model can help in justifying the classification decision; this kind of interpretability will provide more confidence to medical practitioners and patients in the model prognostics. To do this, we would like to find out the parts the image Inception v3 pays attention to while making a classification decision. We visualize the feature maps to understand the active areas of the image. Figure 10 displays the heat map over the image; the highlighted areas in the image are those areas that contribute most to the classification decision. We observe that the cell nucleus is the region that contributes most to the classification decision. We also observe that the model does not pay much attention to the area surrounding the cells. This observation also justifies the choice to perform center cropping while preprocessing the data as that removes
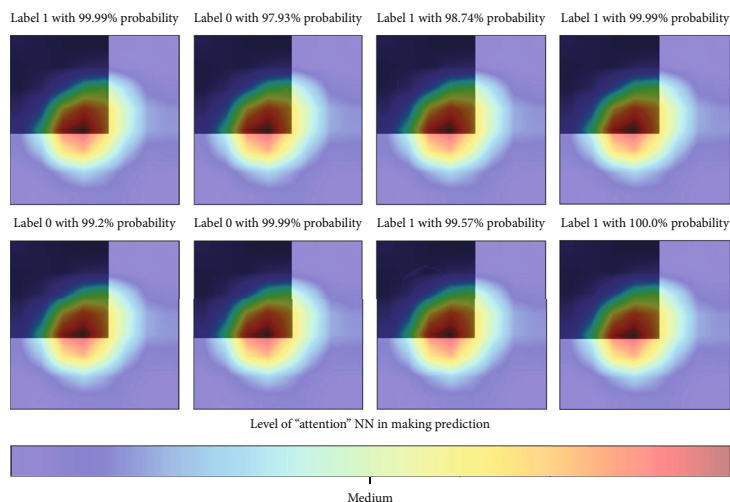
Figure 10: Attention map for feature maps extracted from fine-tuned Inception v3.

the redundant parts of the cell image. We can conclude that the feature extractor is not learning any spurious features that may be inadvertently causing data leakage.

8. Conclusion and Future Scope

In this study, we present a hybrid classification model consisting of Inception v3 (CNN backbone) and XGBoost (classification head). With the fine-tuned Inception v3 model, we extract features from microscopic white blood cell images. These learnt features are passed to an XGBoost model that acts as the classifier that makes the classification decision. Experiments indicate that the proposed hybrid model can accurately and reliably detect acute lymphoblastic leukemia cells with a F1 score of 0.986. The proposed hybrid model and training strategy work with several other pretrained CNNs too, with experimental results indicating an improvement in F1 score in the range of $[1\%, 5\%]$ over a fine-tuned CNN with a softmax classifier. We also attempt to explain the features learnt by Inception v3 by analyzing the attention map for the features extracted. This attention map demonstrates that the model pays a lot of attention to the nucleus of the cell and the center of the microscopic image where the cell is present; this is similar to how a hematologist would analyze the image.

Due to the lack of publicly available ALL datasets, we are not able to perform further analysis of the model's performance on similar datasets. In the future, we would like to collect a large dataset or synthesize an artificial dataset using GANs to improve research in this area.

A major focus on future research should be making the model more interpretable. Although we attempt to interpret the model using attention maps for feature maps, the model largely remains a black box. For future scope, we would like to make the model more explainable so that it can justify the classification decision. This kind of interpretability will provide more confidence to medical practitioners and patients in the model prognostics.

Data Availability

The dataset used in this study is available at https://wiki
.cancerimagingarchive.net/display/Public/C_NMC_2019
+Dataset%3A+ALL+Challenge+dataset+of+ISBI+2019.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

References

[1] M. Stanulla and M. Schrappe, "Treatment of childhood acute lymphoblastic leukemia," *Seminars in Hematology*, vol. 46, no. 1, pp. 52–63, 2009.

[2] P. Jagadev and H. Virani, "Detection of leukemia and its types using image processing and machine learning," in *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, pp. 522–526, Tirunelveli, India, 2017.

[3] M. M. Amin, S. Kermani, A. Talebi, and M. G. Oghli, "Recognition of acute lymphoblastic leukemia cells in microscopic images using k-means clustering and support vector machine classifier," *Journal of Medical Signals and Sensors*, vol. 5, no. 1, pp. 49–58, 2015.

[4] J. Rodellar, S. Alférez, A. Acevedo, A. Molina, and A. Merino, "Image processing and machine learning in the morphological analysis of blood cells," *International Journal of Laboratory Hematology*, vol. 40, Suppl 1, pp. 46–53, 2018.

[5] N. Mahmood, S. Shahid, T. Bakhshi, S. Riaz, H. Ghufran, and M. Yaqoob, "Identification of significant risks in pediatric acute lymphoblastic leukemia (ALL) through machine learning (ML) approach," *Medical & Biological Engineering & Computing*, vol. 58, no. 11, pp. 2631–2640, 2020.

[6] M. Macawile, V. Quiñones, A. Ballado, J. Cruz, and M. Caya, "White blood cell classification and counting using convolutional neural network," in *2018 3rd International Conference on Control and Robotics Engineering (ICCRE)*, pp. 259–263, Nagoya, Japan, 2018.

[7] R. B. Hegde, K. Prasad, H. Hebbar, and B. M. K. Singh, "Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 2, pp. 382–392, 2019.

[8] R. Sharma, "White blood cell classification using convolutional neural network," in *Soft Computing and Signal Processing*, pp. 135–143, Springer Singapore, 2019.

[9] M. Habibzadeh, M. Jannesari, Z. Rezaei, M. Totonchi, and H. Baharvand, "Automatic white blood cell classification using pre-trained deep learning models: ResNet and Inception," in *Proceedings Volume 10696, Tenth International Conference on Machine Vision (ICMV 2017)*, p. 105, Vienna, Austria, 2018.

[10] A. Bodzas, P. Kodytek, and J. Zidek, "Automated detection of acute lymphoblastic leukemia from microscopic images based on human visual perception," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 1005, 2020.

[11] A. Muntasa and M. Yusuf, "Modeling of the acute lymphoblastic leukemia detection based on the principal object characteristics of the color image," *Procedia Computer Science*, vol. 157, pp. 87–98, 2019.

[12] S. Shafique and S. Tehsin, "Computer-aided diagnosis of acute lymphoblastic leukaemia," *Computational and Mathematical Methods in Medicine*, vol. 2018, Article ID e6125289, 2018.

[13] S. Shafique and S. Tehsin, "Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks," *Technology in Cancer Research & Treatment*, vol. 17, p. 153303381880278, 2018.

[14] M. N. Bhuiyan, S. K. Rahut, R. A. Tanvir, and S. Ripon, "Automatic acute lymphoblastic leukemia detection and comparative analysis from images," in *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, pp. 1144–1149, Paris, France, 2019.

[15] V. Acharya and P. Kumar, "Detection of acute lymphoblastic leukemia using image segmentation and data mining algorithms," *Medical & Biological Engineering & Computing*, vol. 57, no. 8, pp. 1783–1811, 2019.

[16] 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.

[17] A. T. Sahlol, P. Kollmannsberger, and A. A. Ewees, "Efficient classification of white blood cell leukemia with improved swarm optimization of deep features," *Scientific Reports*, vol. 10, no. 1, p. 2536, 2020.

[18] A. Ratley, J. Minj, and P. Patre, "Leukemia disease detection and classification using machine learning approaches: a review," in *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)*, pp. 161–165, Raipur, India, 2020.

[19] H. T. Salah, I. N. Muhsen, M. E. Salama, T. Owaidah, and S. K. Hashmi, "Machine learning applications in the diagnosis of leukemia: current trends and future directions," *International Journal of Laboratory Hematology*, vol. 41, no. 6, pp. 717–725, 2019.

[20] A. Gupta and R. Gupta, "All Challenge dataset of ISBI 2019 [data set]," *The Cancer Imaging Archive*, 2019.

[21] A. Gupta, R. Duggal, S. Gehlot et al., "GCTI-SN: geometry-inspired chemical and tissue invariant stain normalization of microscopic medical images," *Medical Image Analysis*, vol. 65, p. 101788, 2020.

[22] 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," *Clinical Lymphoma Myeloma and Leukemia*, vol. 17, no. 1, p. e99, 2017.

[23] R. Duggal, A. Gupta, R. Gupta, M. Wadhwa, and C. Ahuja, "Overlapping cell nuclei segmentation in microscopic images usingdeep belief networks," in *Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)*, India, December 2016.

[24] R. Duggal, A. Gupta, and R. Gupta, "Segmentation of overlapping/touching white blood cell nuclei using artificial neural networks," in *CME Series on Hemato-Oncopathology, All India Institute of Medical Sciences (AIIMS)*, New Delhi, India, July 2016.

[25] R. Duggal, A. Gupta, R. Gupta, and P. Mallick, "SD-Layer: stain deconvolutional layer for CNNs in medical microscopic imaging," in *Medical Image Computing and Computer Assisted Intervention − MICCAI 2017*Springer, Cham.

[26] T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," 2017, http://arxiv.org/abs/1708.04552.

[27] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2015.

[28] T. Chen and C. Guestrin, "XGBoost," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California, 2016.

[29] W. Yu, J. Chang, C. Yang et al., "Automatic classification of leukocytes using deep neuralnetwork," in *2017 IEEE 12th International Conference on ASIC(ASICON)*, pp. 1041–1044, Guiyang, China, October 2017.

[30] S. Mourya, S. Kant, P. Kumar, A. Gupta, and R. Gupta, "LeukoNet:DCT-based CNN architecture for the classification of normal versusleukemic blasts in B-ALL cancer," 2018, http://arxiv.org/abs/1810.07961.

[31] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, "A hybrid deep learning architecture for leukemic B-lymphoblast classification," in *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 271–276, 2019.

Retraction

Retracted: Optimization of Correlation Filters Using Extended Particle Swarm Optimization Technique

Computational and Mathematical Methods in Medicine

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Masood, A. Zafar, M. U. Ali et al., "Optimization of Correlation Filters Using Extended Particle Swarm Optimization Technique," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 6321860, 13 pages, 2021.

Research Article

Optimization of Correlation Filters Using Extended Particle Swarm Optimization Technique

Haris Masood,[1] Amad Zafar,[2] Muhammad Umair Ali [iD],[3] Muhammad Attique Khan [iD],[4] Kashif Iqbal,[1] Usman Tariq [iD],[5] and Seifedine Kadry[6]

[1]*Wah Engineering College, University of Wah, Wah Cantt, Pakistan*
[2]*Department of Electrical Engineering, University of Lahore, Islamabad Campus, Pakistan*
[3]*Department of Unmanned Vehicle Engineering, Sejong University, Seoul 05006, Republic of Korea*
[4]*Department of Computer Science, HITEC University Taxila, Taxila 47040, Pakistan*
[5]*College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Khraj, Saudi Arabia*
[6]*Faculty of Applied Computing and Technology, Noroff University College, Kristiansand, Norway*

Correspondence should be addressed to Muhammad Attique Khan; attique@ciitwah.edu.pk

In the past few decades, the field of image processing has seen a rapid advancement in the correlation filters, which serves as a very promising tool for object detection and recognition. Mostly, complex filter equations are used for deriving the correlation filters, leading to a filter solution in a closed loop. Selection of optimal tradeoff (OT) parameters is crucial for the effectiveness of correlation filters. This paper proposes extended particle swarm optimization (EPSO) technique for the optimal selection of OT parameters. The optimal solution is proposed based on two cost functions. The best result for each target is obtained by applying the optimization technique separately. The obtained results are compared with the conventional particle swarm optimization method for various test images belonging from different state-of-the-art datasets. The obtained results depict the performance of filters improved significantly using the proposed optimization method.

1. Introduction

For the purpose of object detection and recognition in the fields of pattern recognition, computer vision, and image processing [1–5], correlation filters have been widely employed. Other fields in which correlation filters are used are object tracking [6, 7] and biometric object recognition [8–10]. The correlation filters are trained in a way to generate maximum correlation peaks pertaining to the objects desirous of being detected, while generation low peaks against illumination, clutter, and noise. Correlation filters date back to around three decades, when they were introduced primarily for object recognition [11]. Over the years, improvement in the accuracy of the correlation filters has been made, using different optimization methods [12–15].

Accurate recognition and tracking of objects can be carried out using the correlation filters. The Maximum Average

Correlation Height (MACH) and Maximum Average Correlation Energy (MACE) are used for eliminating clutter distortion and noise [16]. The MACE filter is extremely sensitive to clutter and noise, while providing distinct peaks for the detection of filter outputs [17]. The MACH filter gives maximum relative height w.r.t the expected distortions by generating the broader peaks [18].

Complex filter equations are employed for implementing the correlation filters in different software. Thus far, many correlation filters' variants have been introduced by altering values of the optimal tradeoff (OT) parameters of involved filter equations. Up till now, experimental trials have been conducted for tuning of the tradeoff parameters. The main motivation of the proposed research is to optimize the OT parameters using a technique that has not been employed before, which enables determining the best possible values. The optimization technique presented in this paper is based

on the particle swarm optimization (PSO) variant. The PSO algorithm was first introduced by Eberhart and Kennedy [19]. It is a population-oriented method that is inspired by animals and fish social behavior. The standard PSO algorithm [20] is the one used for basic optimization of parameters. EPSO proposed by Li et al. [21] is considered to be PSO's most renowned variant. This particular variant has been employed in various applications of image processing [22–28].

The paper compares the standard PSO and EPSO for OT parameter optimization. The parametric optimization relies on MACH and MACE filter cost functions. The filters produced as a result are application specific as the parameter values vary based on each target object of interest. The filter is generic in nature as given the application, it can be applied successfully on any target object of interest based on calculations of the cost functions. The proposed algorithm is novel in the sense that PSO variants have not been previously employed in conjunction with correlation filters for accurate object recognition. Previously, values suggested by Bone et al. were used for the optimization of optimal tradeoff values of a correlation filter. This is the first time that an ensemble of EPSO and correlation filters is used for the optimization of optimal tradeoff parameters for accurate object detection.

2. Proposed Methodology

2.1. Correlation Filters. The main motivation behind employing the enhanced version of the correlation filter is to exclude the peaks that make the procedure of object detection, an erroneous one. In multiplexed filters, it is usually very difficult to obtain a sharp peak using the correlation templates, which often outputs high-intensity side lobes. For easy detection of object of interest, MACE filters are employed, which are responsible for providing sharp peaks. The downside of MACE filters is that they are sensitive to distortion. In the MACE filter, the function level is evenly reduced over the entire correlation plane, with the exception of the plane center. On the contrary, MACH provides broader correlation peaks, but it comes with an added advantage of being noise and distortion tolerant. For the implementation of MACH and MACE filters, the metrics of Average Correlation Energy (ACE) and Average Similarity Matrix (ASM) are minimized, respectively. Since minimization of ASM is directly related to the reduction of dissimilarity among the correlation planes, it makes the correlation process more accurate. The amplitude of peaks of the MACH filter is higher than the MACE filter peaks [17, 18].

Equation (1) shows the energy equation [29] pertaining to the correlation filter.

$$E(f) = \alpha(\text{ONV}) + \beta(\text{ACE}) + \gamma(\text{ASM}) - \delta(\text{ACH}). \quad (1)$$

The ASM can be calculated using

$$\text{ASM} = f^+ S_x f, \quad (2)$$

$$S_x = \frac{1}{N} \sum_{i=1}^{N} (X_i - M_x)^* (X_i - M_x), \quad (3)$$

where the variable "f" depicts the chosen filter and the "$+$" sign in the superscript depicts the conjugate response in Equation (2).

The ACE of the filter can be computed using

$$\text{ACE} = f^+ D_x f,$$
$$D_x = \frac{1}{N} \sum_{i=1}^{N} X_i^* X_i. \quad (4)$$

Equation (5) is used for the calculation of output noise variance [9].

$$\text{ONV} = f^+ C f. \quad (5)$$

The variable "C" indicates a diagonal $d * d$ dimensional vector. Normally, the value of C is taken as $\delta^2 I$. Equation (6) is used for the calculation of Average Correlation Height (ACH) [29].

$$\text{ACH} = \left| \frac{1}{N} \sum_{i=1}^{N} f^T X_i \right| = \left| f^T m_x \right|, \quad (6)$$

where m_x represents the average of N vectors.

By substituting all the values, Equation (1) can be minimized into

$$E(f) = f^+ \text{If} - \delta \left| f^T m_x \right|, \quad (7)$$

where I can be described using

$$I = \alpha C + \beta D_x + \gamma S_x. \quad (8)$$

Therefore, the filter equation becomes

$$f^o = \left(\frac{\delta}{2} \right) I^{-1} m_x, \quad (9)$$

where δ represents the scaling factor and o in the superscript depicts optimal complex filter transfer function. The values of α, β, and γ are nonnegative entities. The effectiveness of the MACH filter depends mainly on the adjustment of these three parameters, i.e., α, β, and γ. By selecting $\alpha = 0$ and $\gamma = 0$, the filter transfer function is transformed into a simple MACE filter which is used for the minimization of ACE. Setting $\alpha = 0$ and $\beta = 0$ converts the filter transfer function to the MACH filter which is used for the minimization of ASM. Up till now, the optimized values as suggested by Bone et al. have been kept at $\alpha = 0.01$, $\beta = 0.1$, and $\gamma = 0.3$, since they are considered optimal for the implementation of the MACH filter. In reality, these values do not show promising results for some datasets as the conditions in different scenarios vary. In this paper, a novel method using EPSO has been

proposed for calculating the optimal OT values that can be considered optimal for every environment.

A combination of a correlation filter and an optimization technique is proposed in this paper. The combined optimized filter provides optimal values of OT parameters based on the specific target object of interest. The results of conventional PSO and proposed EPSO algorithms will be compared in the later section of the paper.

2.2. Particle Swarm Optimization. PSO, as previously described, is an optimization algorithm that is derived from fish schooling and motion of bird flocks. PSO is a very good technique used primarily for the optimization of the parameters. PSO searches the most optimal solution in a multidimensional search space with the help of several available particles who all donate towards the optimized particles. The algorithm searches the best value for each particle by the convergence method. A cost function is employed for the estimation of each optimized value. The cost function is also used to identify the most suitable value for the defined fitness function. Two primary parameters are associated with each particle: (i) the velocity of each particle $v(i)$ and (ii) the particle position of each particle $x(i)$, where the iteration index is indicated by i. Subsequently, the global best of whole swarm is obtained by extracting the best values related to all the particles and combining them. In a D-dimensional space involving a swarm of N particles, the position and velocity of each particle are updated using

$$
\begin{aligned}
v_k^d(i+1) = w.v_k^d(i) + c_1.r_{1,k}(i).\left(p_k^d - x_k^d(i)\right) \\
+ c_2.r_{2,k}(i).\left(g^d - x_k^d(i)\right),
\end{aligned}
\tag{10}
$$

$$
x_k^d(i+1) = x_k^d(i) + v_k^d(i+1),
\tag{11}
$$

where the dimensions of the particles are denoted by $d = 1, 2 \cdots, D$ and the particle index is denoted by $k = 1, 2 \cdots, N$. The constants c_1 and c_2 represent the cognitive and social coefficients, respectively. The velocity and position of the kth particle are represented by v_d^k and x_d^k, respectively, in a d-dimensional space. The particles' local best position is represented by p_d^k, while g_d represents the swarms' global best position. While searching the behavior of the swarm, the source of randomness and uniform random distribution $[0, 1]$ are the two parameters used for the derivation of $r_{1,k}$ and $r_{2,k}$.

The most popular variant of the PSO was proposed by Eberhart and Kennedy. Equation (12) shows the variant that contains a model based on inertial weight [19]. The model tends to multiply a constant factor commonly known as the weight of the inertia, with the velocity of the current iteration [19].

$$
\begin{aligned}
v_k^d(i+1) = w.v_k^d(i) + c_1.r_{1,k}(i).\left(p_k^d - x_k^d(i)\right) \\
+ c_2.r_{2,k}(i).\left(g^d - x_k^d(i)\right).
\end{aligned}
\tag{12}
$$

The momentum of the particle is controlled by the convergence of inertia weight $w \, \epsilon \, [0, 1]$. For a small value of w, negli-

gible momentum is preserved and carried forward from the previous iteration that tends to change the direction quickly. On the contrary, a larger w value means slow convergence and delayed change in particle direction. For the value of $w = 0$, the particle moves ahead without any prior knowledge of the value of velocity. The defined variant is commonly known as the standard PSO [30, 31].

The optimization algorithms have several applications associated with them. Pandey et al. [32] employed the PSO algorithm for clustering of data vectors. The algorithm was also used for the user-defined centroid of data clusters and their identification. While comparing PSO with its counterpart, i.e., the k-means clustering technique, the PSO returned the best convergence with minimal errors. The proposed PSO algorithm by Nayyar et al. is now used in conjunction with k-means for the refinement of clusters [24]. Grosan et al. [33] proposed an application pertaining to the PSO algorithm in data mining domain. They used the PSO algorithm for cloud computing such that applications were efficiently scheduled by optimizing the cloud resources. As compared to the heuristic algorithms, the proposed technique by Grosan et al. reduced the cost of data transmission and computation by one-third. As compared to the k-means algorithm, the PSO algorithm for the optimization of image clustering was utilized. Applications can be found in satellite imaging and in MRI as well.

The most commonly employed PSO variant is the extended PSO having the time-varying coefficients related to acceleration (EPSO) [22]. An optimal solution is obtained through the acceleration coefficients only by guiding the movement of particles, while coefficients related to inertia are removed. The coefficients related to acceleration move linearly with respect to time. Therefore, if at some point, velocity goes to zero, the particle is reinitialized through the use of other predefined velocities.

2.3. Extended Particle Swarm Optimization (EPSO) Algorithm. In each iteration of the conventional PSO technique, two extreme values are used for updating the state of each particle. For optimizing the algorithm's global convergence and to increase its efficiency, the global impact of many involved particles contributes towards updating the state of each particle. Such an impact caused by the multiparticle effect is commonly known as the extended particle swarm optimization (EPSO) algorithm.

In EPSO algorithms' recursive process, the optimization process includes particles that contain more information as compared to the conventional PSO. The main formula of EPSO is stated in Equations (13) and (14) [21].

$$
v_{i+1} = wv_t + \sum_{i=1}^{m} \psi_i(p_t - x_t) + \sum_{i=1}^{n} \zeta_i(\widehat{p}_t - x_t),
\tag{13}
$$

$$
x_{i+1} = x_t + v_{t+1},
\tag{14}
$$

where $\psi_i = c_{1,i} r_{1,i}(0, 1)$ and $\zeta_i = c_{2,i} r_{2,i}(0, 1)$ [21].

The number of iterations is depicted by the subscript t, v_t denotes the velocity of the particles, x_t denotes the interspaces of the involved particles, p_t signifies the extreme value

particle's position, $\widehat{p}_t$ denotes the local extreme particles, and $r_{1,I}\,(0,1)$ and $r_{2,I}\,(0,1)$ are random variables between 0 and 1, while $c_{1,I}$ and $c_{2,i}$ represent the control parameters. A comparison of Equations (14) and (12) shows that in comparison with the conventional PSO, EPSO takes into account the information associated with more particles involved in the optimization process, thus providing stronger and better global convergence. EPSO can easily be converted to the conventional PSO by setting values of $m = n = 1$.

Since the EPSO utilizes information from more particle values and it also considers more particles and more controlled parameters for the optimization process, the parameter selection has to be spot on in order for the algorithm to converge quickly. The setting strategy pertaining to the control parameters of the algorithm is obtained by testing the convergence condition of the EPSO algorithm. The iterative formula for the EPSO algorithm is calculated by substituting Equation (13) in Equation (14) and using $v_t = x_t - x_{t-1}$. Equation (15) describes the main cumulative formula [21].

$$x_{i+1} = \left(1 + w - \sum_{i=1}^{m}\psi_i + \sum_{i=1}^{n}\zeta_i\right)x_t - wx_{t-1} + \sum_{i=1}^{m}\psi_i p_i + \sum_{i=1}^{n}\zeta_i\widehat{p}_t. \tag{15}$$

Since for all the iterations, p_t and $\widehat{p}_t$ are considered constants, therefore, Equation (15) can be summarized to Equation (16) [21].

$$\begin{bmatrix} x_{t+1} \\ x_t \\ 1 \end{bmatrix} = \begin{bmatrix} 1 + w - \psi - \zeta & -w & \psi p + \zeta p \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ x_{t+1} \\ 1 \end{bmatrix}, \tag{16}$$

where $\psi = \sum_{i=1}^{m}\psi_i$, $\zeta = \sum_{i=1}^{m}\zeta_i$, $\psi p = \sum_{i=1}^{m}\psi_i p_t$, and $\zeta p = \sum_{i=1}^{n}\zeta_i\widehat{p}_t$ [21].

The velocity solution is obtained by generalizing Equation (13), and Equation (14) can be categorized into Equation (17) [22].

$$v_k^d(i+1) = c_1 . r_{1,k}(i) . \left(p_k^d - x_k^d(i)\right) + c_2 . r_{2,k}(i) . \left(g_d - x_k^d(i)\right), \tag{17}$$

where [22]

$$\begin{aligned} c_1 &= \left(c_{1f} - c_{1i}\right) * \frac{k}{\text{max ITER}} + c_{1i}, \\ c_2 &= \left(c_{2f} - c_{2i}\right) * \frac{k}{\text{max ITER}} + c_{2i}. \end{aligned} \tag{18}$$

Recursive updating of the position and velocity of the kth particle is carried out using Equations (16) and (17), respectively.

2.4. EPSO for Designing of Correlation Filter. Implementation of correlation filters is performed using complex filter trans-

Table 1: Parameter optimization of correlation filter using PSO.

Optimal tradeoff parameter estimation for correlation filter using PSO	
1.	Each particle's position and velocity parameters are randomly initialized
2.	Fitness function value estimation using Equations (15) and (17) for each particle
3.	Calculation of best value for each particle
4.	Calculation of Swarm's global best
5.	The position of particles is updated using Equation (11) [19]
6.	The velocity of particles is updated using Equation (10) [19]
7.	Fitness function value estimation using Equations (16) and (17) for each particle
8.	Calculation of local best pertaining to each particle
9.	Calculation of global best pertaining to each swarm
10.	If stopping condition is achieved, terminate the algorithm. Otherwise, go back to Step 5

Table 2: Parameter optimization of correlation filter using EPSO.

Optimal tradeoff parameter estimation for correlation filter using EPSO	
1.	Each particle's position and velocity parameters are randomly initialized
2.	Fitness function value estimation using Equations (16) and (17) for each particle
3.	Calculation of local best pertaining to each involved particle
4.	Calculation of global best pertaining to each involved swarm
5.	The position of particles is updated using Equation (15) [21]
6.	The velocity of particles is updated using Equation (17) [21]
7.	Reinitialize the velocity if the velocity of particles becomes equal to zero
8.	Fitness function value estimation using Equations (16) and (17) for each particle
9.	Calculation of local best pertaining to each particle
10.	Calculation of global best pertaining to each swarm
11.	If stopping condition is achieved, terminate the algorithm. Otherwise, go back to Step 5

Table 3: Setting of PSO parameter values.

Parameter setting	Values
Experiments	120
Iterations	320
Particles	10
Dimensions	03
$X_{\min}$	-1
$X_{\max}$	1
$V_{\min}$	-0.1
$V_{\max}$	0.1
W	0.9
C_1, C_2	2
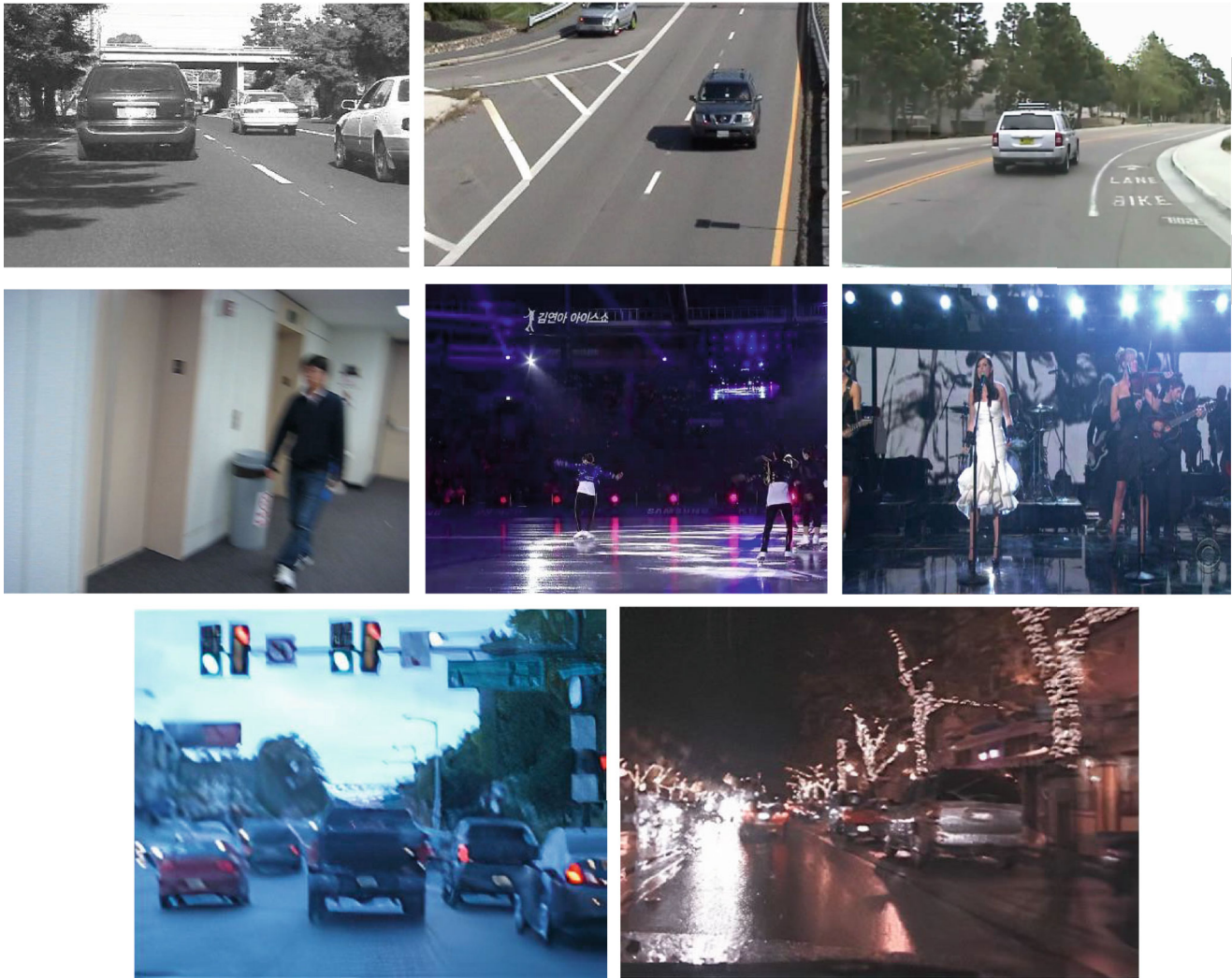
Figure 1: Datasets.

Table 4: COPI value comparison.

Dataset	Testing image (degree)	Bone et al. values (COPI)	PSO				EPSO			
			α	β	γ	COPI	α	β	γ	COPI
1	5	$4.05E-5$	0.0040	0.0402	0.0473	$2.74E-4$	$8.45E-9$	0.0954	0.1722	$2.74E-1$
2	5	$4.91E-5$	0.0040	0.0421	0.0506	$3.91E-4$	$7.39E-9$	0.0921	0.2102	$2.04E-1$
3	15	$2.05E-5$	0.0041	0.0404	0.0470	$6.05E-5$	$3.04E-8$	0.0726	0.2232	$1.24E-1$
4	15	$3.98E-5$	0.0038	0.0388	0.0525	$1.98E-3$	$2.14E-8$	0.1229	0.2212	$1.37E-0$
5	25	$4.15E-5$	0.0035	0.0389	0.0499	$3.15E-4$	$1.45E-8$	0.1021	0.1639	$1.34E-1$
6	25	$5.05E-5$	0.0039	0.0384	0.0473	$4.01E-4$	$8.45E-8$	0.1512	0.1978	$1.74E-1$
7	45	$4.95E-5$	0.0042	0.0310	0.0428	$3.15E-4$	$6.27E-9$	0.2102	0.1099	$2.44E-2$
8	45	$7.05E-5$	0.0039	0.0390	0.0478	$5.15E-4$	$7.45E-9$	0.1022	0.1877	$3.24E-2$

fer functions which are dependent on the selection of tradeoff parameters. The tradeoff parameters should be optimal in order for the filter to work in an effective manner. Better selection of tradeoff parameters will result in accurate correlation peaks and thus better object detection. Several researchers have proposed methods for the effective calcula-

tion of these values. A method proposed by Bone et al. [16] used fixed values for the optimal tradeoff parameters. The choice of selection of these values was not obvious for certain object recognition applications. A novel technique for efficient selection of these tradeoff parameters is proposed in this paper which pertains to the response of the filter. The

(a)

(b)

(c)

Figure 2: Continued.

(d)

FIGURE 2: (a) Correlation plane using Bone et al. values, (b) PSO value-based correlation plane, (c) ESPO correlation plane values, and (d) testing image.

parameter α mentioned in Equation (8) can be calculated using Equation (15) [22].

$$\alpha_k(t+1) = \alpha_k(t) + \nu_{k,\alpha}(t+1). \tag{19}$$

Similarly, using Equations (15) and (17), similar equations of β and γ can be obtained for the purpose of optimization. EPSO calculates the values of optimal tradeoff parameters via convergence of the involved fitness function. It will enable the fitness function to be calculated for specific object recognition applications by calculating the correlation output peak intensity (COPI) cost function and peak to correlation energy (PCE) cost function. The correlation pertaining to the object of interest depends on the calculation of COPI and PCE cost functions. The calculation of both of these parameters is performed for the characterization of the correlation plane [34], as mentioned in Equations (20) and (21) [9].

$$\text{COPI} = \max\left\{|C(x,y)|^2\right\}. \tag{20}$$

$C(x, y)$ depicts at location (x, y) the correlation peak output and [9]

$$\text{PCE} = \frac{\text{COPI} - \overline{|C(x,y)|^2}}{\left\{\sum\left(\left[|C(x,y)|^2 - |C(\bar{x},y)|^2\right]^2 / \left(N_x N_y - 2\right)\right)\right\}^{1/2}}, \tag{21}$$

where the average COPI is represented by $\overline{|C(x,y|^2} = \sum |C(x,y|^2 / N_x N_y$.

For maximizing the PCE cost function, the value of ACE is reduced by the MACE filter. The MACH filter is responsible for minimizing the ASM value. The height of the correlation peak is maximized due to the reduction of ASM. In the

optimization algorithms, fitness functions are defined by COPI and PCE values. The summary of the steps is mentioned in Tables 1 and 2.

3. Results and Discussion

Eight publicly available datasets [35, 36] have been used for the experiments and analysis. Five datasets are vehicle-oriented datasets in which the object of interest, i.e., the vehicle, undergoes different shift, scale, occlusion, and lightening conditions. Three remaining datasets, i.e., Singer, Blur Body, and Skating, are person-oriented datasets in which the object of interest, i.e., person, undergoes motion blur, shift, scale, and occlusion-based variations. The obtained results have been used for the comparison of results of the proposed algorithm with other similar state-of-the-art algorithms [16].

3.1. Setting of Parameters. In order to test and evaluate the optimal values of tradeoff parameters, experiments have been carried out using both the PSO and EPSO techniques. The chosen parameters are shown in Table 3.

Implementation of parameters has been ensured with a slight modification. Since there is a possibility that particles may give negative values for some particular parameters, only the magnitude is considered, while ignoring the sign. The lower limit of values has been set to -0.1 to give weightage to the lower order negative values. The results proved that the assumption was correct.

3.2. Comparison of Results of PSO and EPSO. Eight publicly available datasets shown in Figure 1 were tested for acquiring the results. The chosen datasets were based on the diversity of the conditions that the images of datasets were taken. The chosen datasets have been employed for comparison of results of the algorithms and analyzing the optimized values of each dataset. The 0-45 training images were rotated out-of-plane. Among the images, a difference of 10 was ensured. Cost function has been chosen based on the requirement in
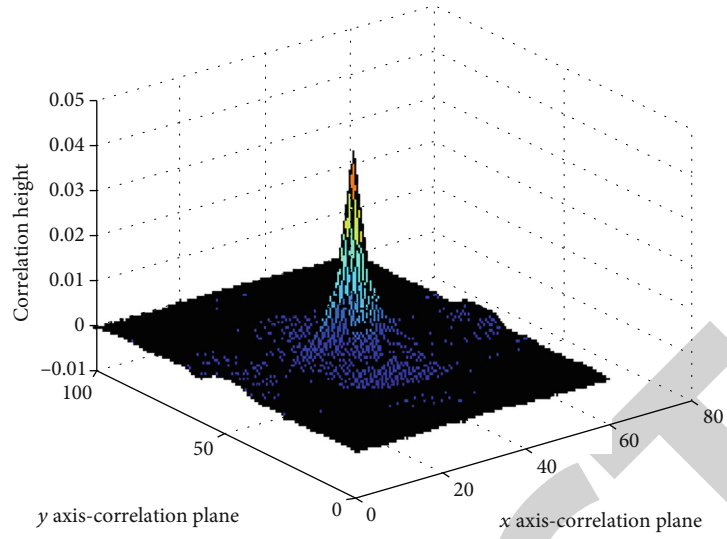
(a)
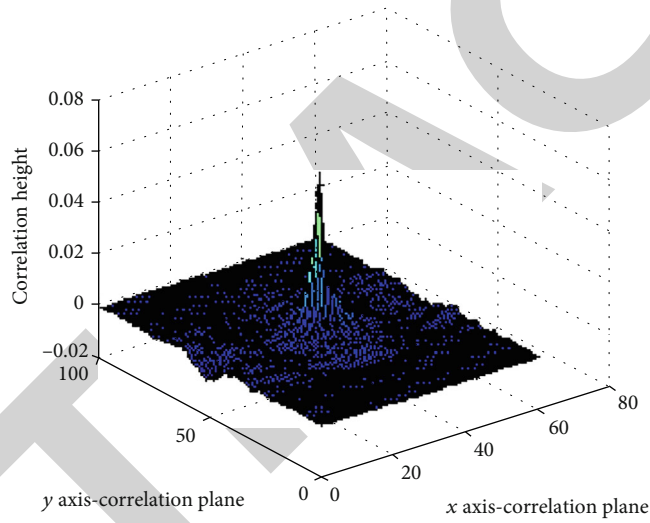
(b)
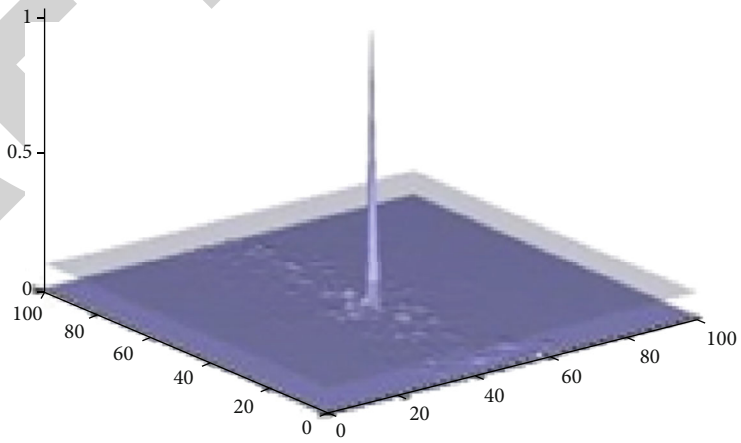
(c)

Figure 3: Continued.

(d)

Figure 3: (a) Bone et al. value-based correlation plane, (b) PSO value-based correlation plane, (c) ESPO correlation plane values, and (d) testing image.

hand. As a cost function, both the COPI and PCE values have been individually selected for comparison of the EPSO and PSO results with the Bone et al. suggested values [16].

Testing images belonging to different datasets and for different rotations are used for the analysis of the optimized values and for experimentation purposes. For the Bone et al. algorithm, the values of α, β, and γ have been set as 0.01, 0.1, and 0.3, respectively, as previously proposed. Table 4 shows the comparison based on COPI cost function between the proposed values of Bone et al. and the optimized values calculated through the proposed algorithm. The results evidently depict that the correlation peaks generated by the EPSO optimization algorithm are better than the peaks generated by PSO and Bone et al. values.

The optimization algorithms have been employed for the comparison with Bone et al.'s proposed values using one of the proposed datasets from Table 4. As a cost function, the COPI value has been used for the algorithms. Results clearly depict that in comparison with the PSO and Bone et al.'s algorithms, the optimized values from EPSO perform very well considering the COPI cost function as shown in Figure 2. Out of plane rotation of 15° is applied on the testing images. The attained COPI values in the cases of EPSO, PSO, and Bone et al. are $3.24E - 2$, $5.15E - 4$, and $7.05E - 5$, respectively. The peaks obtained from Bone et al. and PSO values are approximately the same as evident in Figures 2(a) and 2(b), respectively. However, the results of applying the PSO far outmatch the results obtained from Bone et al.'s proposed parameter values. The COPI results obtained from using the optimized parameter values of EPSO are far better than the results obtained from both the Bone et al. algorithm and the standard PSO, as shown in Figure 2(c). The EPSO result for other performance metrics is also better than the algorithms proposed by Bone et al. and standard PSO as mentioned in Figures 3–5. Table 4 also

depicts that EPSO outperforms the PSO and Bone et al. algorithm at varying degree levels which shows that the algorithm is shift tolerant. Since all of the eight datasets include images with varying scaling levels, therefore, it is evident from the results depicted in Table 4 that EPSO provides scale invariance as well.

A testing image that has been rotated out of plane by 45° is shown in Figure 3. The achieved COPI values for EPSO, PSO, and Bone et al. are $3.24E - 2$, $5.15E - 4$, and $7.05E - 5$, respectively. In the correlation plane, the presence of side lobes using optimized values of EPSO is due to the occurrence of ONV as well as due to the inclusion of full correlation process in the experimentation, i.e., full correlation of the testing image is performed with the filter. Considering the COPI cost function, the results of EPSO optimized values outmatch the results of the other values.

The maximum value of the PCE parameter can be achieved by minimizing the ACE value. This leads to a prominent and sharper peak as compared to the other methods. The pattern of optimized values is examined by experimenting on different datasets consisting of test images that have been rotated out of plane. The parameters α, β, and γ have been set as 0.01, 0.1, and 0.3, respectively. The comparison between the proposed algorithm and Bone et al.'s algorithm based on PCE cost function is shown in Table 5. The results evidently show that the PCE values generated by the EPSO optimization algorithm outmatch the values generated by PSO and Bone et al.'s values.

Different datasets have been employed for the analysis of correlation plane based on PCE function. The testing image is rotated out of plane by 45° as shown in Figure 4. The achieved values of PCE cost function in the cases of EPSO, PSO, and Bone et al. are $3.19E + 2$, $2.77E + 2$, and $2.72E + 1$, respectively. The obtained results of EPSO based on PCE cost function again outmatch the results obtained for PSO and
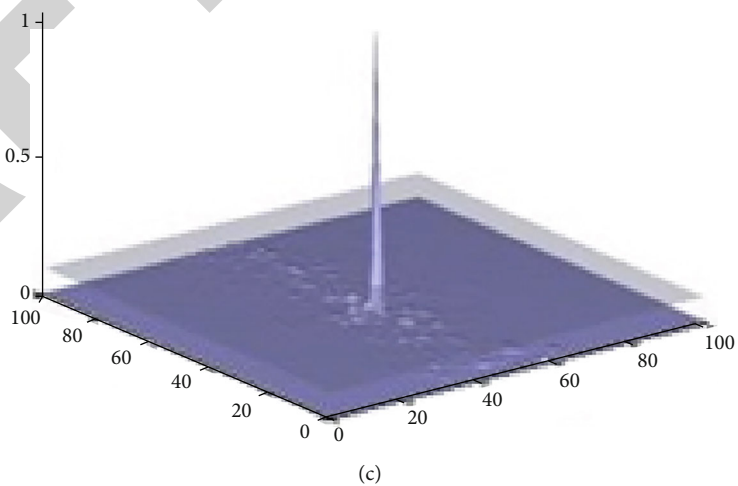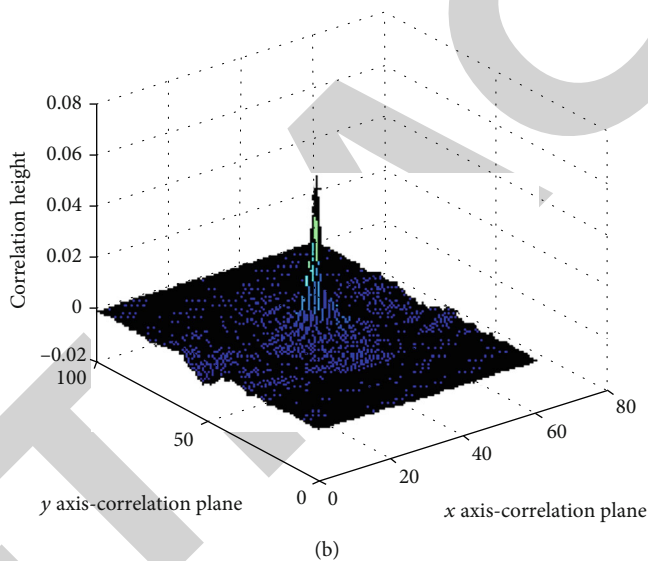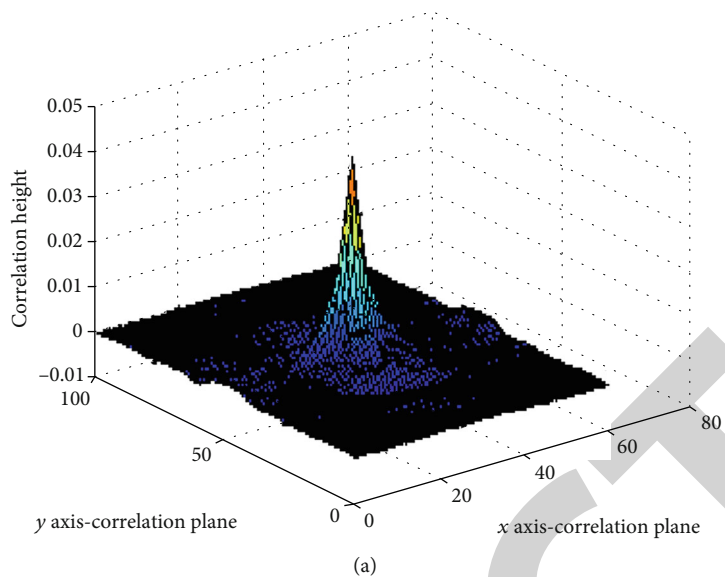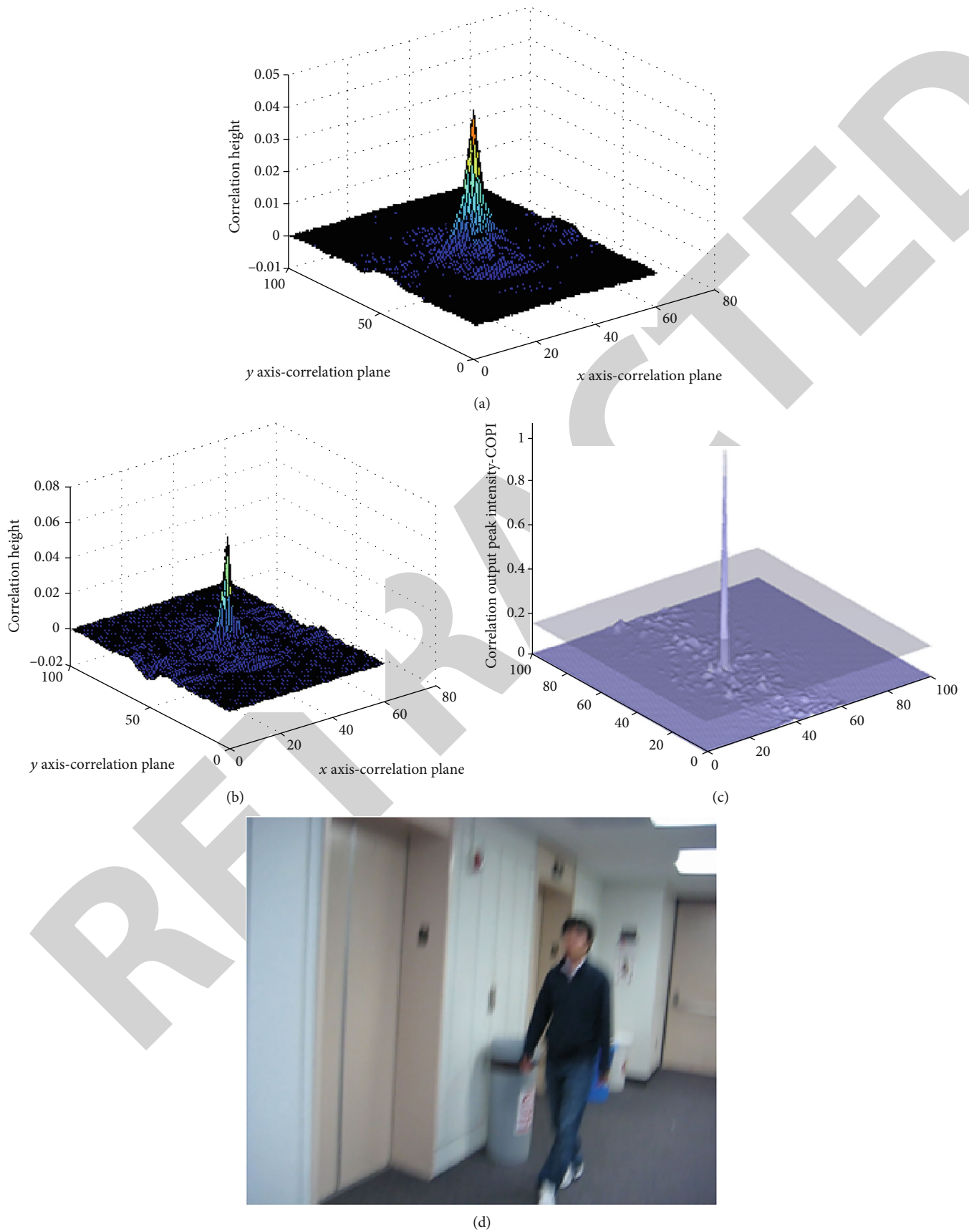
FIGURE 4: (a) Correlation plane using Bone et al. values, (b) PSO value-based correlation plane, (c) ESPO correlation plane values, and (d) testing image.
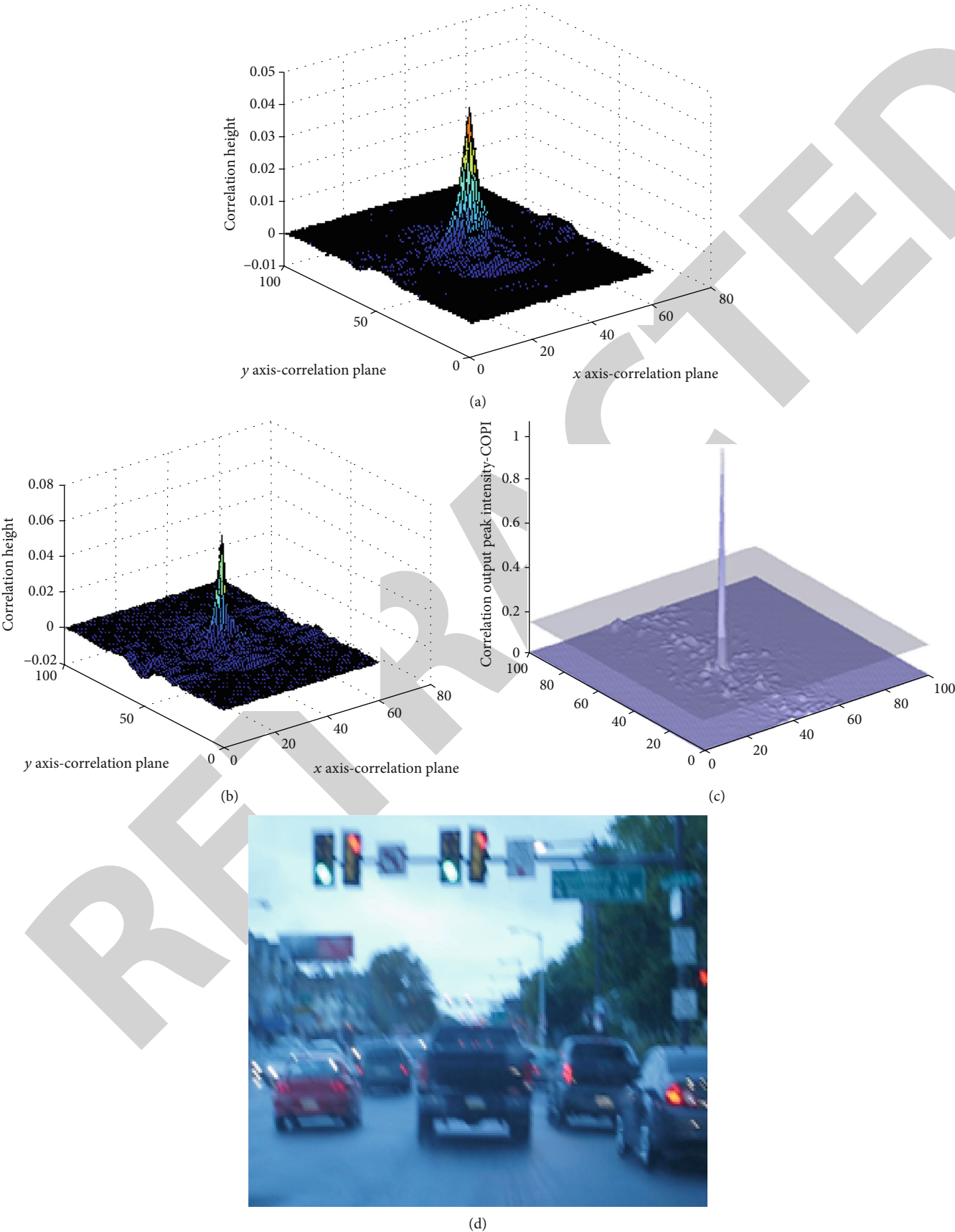
(a)

(b)

(c)

(d)

Figure 5: (a) Bone et al. value-based correlation plane, (b) PSO value-based correlation plane, (c) ESPO correlation plane values, and (d) testing image.

TABLE 5: Comparison of PCE values for EPSO, PSO, and Bone et al.'s values.

Dataset	Testing image (Deg.)	Bones values (PCE)	PSO				EPSO			
			α	β	γ	PCE	α	β	γ	PCE
1	5	$6.02E+1$	0.0034	0.7232	0.5449	$1.62E+2$	$2.19E-6$	0.7134	0.4772	$7.74E+2$
2	5	$5.14E+1$	0.0040	0.7492	0.5781	$7.52E+1$	$1.35E-6$	0.7293	0.4923	$6.34E+2$
3	15	$4.95E+1$	0.0029	0.7321	0.5322	$1.62E+2$	$1.72E-5$	0.7522	0.3982	$8.04E+2$
4	15	$3.92E+1$	0.0039	0.7002	0.5349	$5.72E+1$	$4.83E-8$	0.6578	0.4438	$5.01E+2$
5	25	$4.32E+1$	0.0031	0.7244	0.5449	$2.69E+2$	$2.44E-7$	0.7980	0.5223	$2.74E+3$
6	25	$1.48E+2$	0.0032	0.7390	0.5019	$1.42E+2$	$7.19E-7$	0.6991	0.4938	$2.14E+3$
7	45	$2.72E+1$	0.0034	0.7470	0.5709	$2.77E+2$	$4.19E-7$	0.5224	0.6220	$3.19E+2$
8	45	$3.24E+1$	0.0030	0.6295	0.5014	$1.99E+2$	$3.59E-7$	0.4900	0.5114	$2.44E+2$

Bone et al.'s proposed algorithm. The obtained optimized values clearly depict a sharper peak in the case of EPSO as compared with PSO and Bone et al.

Figure 5 shows a 15° out of plane rotated testing image. The optimized values generated from EPSO are yielding much sharper peaks as compared to the optimized values generated through PSO or the conventional values of Bone et al.'s algorithm. The PCE values for 15° out of plane rotated testing image for EPSO, PSO, and Bone et al. are $5.01E+2$, $5.72E+1$, and $3.92E+1$. The results obtained by using conventional PSO are better than the results obtained from Bone et al.'s algorithm. However, the EPSO-generated optimized values give the best results as compared to the results obtained from the optimized values generated from PSO and Bone et al., in terms of PCE and COPI cost functions. The optimized values depend on the images contained in the dataset. The cost functions and the datasets define the values of the optimized parameters.

4. Conclusion

A novel technique has been proposed which combines optimization algorithms with a correlation filter in order to improve the results of the correlation filter. The technique focuses on optimizing the tradeoff parameters pertaining to correlation filters which have not been achieved earlier. The optimization parameters achieved by using EPSO and PSO algorithms have been compared with the optimization values of the previously employed algorithms. The comparison was based on the PCE and COPI cost functions for a specific object recognition application. The values are not constant for all the object recognition applications as suggested by the previous studies. The values of optimal tradeoff parameters and the PCE and COPI cost functions are calculated for specific datasets based on their properties. The EPSO optimized values helped in the reduction of the ONV factor thus resulting in more accurate results as compared to the PSO and other previously suggested similar algorithms. The proposed work also has plenty of scope for future studies. In the future, we will try to compare EPSO and PSO with more advanced heuristic algorithms in order to achieve more accurate results. A few recent algorithms are improved GA, grasshopper, mothflame, and name a few more [37–39].

Data Availability

The imaging data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Authors' Contributions

All authors equally contributed to this study.

References

[1] M. Rashid, M. A. Khan, M. Sharif, M. Raza, M. M. Sarfraz, and F. Afza, "Object detection and classification: a joint selection and fusion strategy of deep convolutional neural network and SIFT point features," *Multimedia Tools and Applications*, vol. 78, no. 12, pp. 15751–15777, 2019.

[2] M. Rashid, M. Alhaisoni, S.-H. Wang, S. R. Naqvi, A. Rehman, and T. Saba, "A sustainable deep learning framework for object recognition using multi-layers deep features fusion and selection," *Sustainability*, vol. 12, no. 12, article 5037, 2020.

[3] N. Hussain, M. A. Khan, M. Sharif et al., "A deep neural network and classical features based scheme for objects recognition: an application for machine inspection," *Multimedia Tools and Applications*, 2020.

[4] M. Sharif, T. Akram, R. Damaševičius, and R. Maskeliūnas, "Skin lesion segmentation and multiclass classification using deep learning features and improved moth flame optimization," *Diagnostics*, vol. 11, no. 5, p. 811, 2021.

[5] I. M. Nasir, M. Raza, J. H. Shah, and A. Rehman, "Human action recognition using machine learning in uncontrolled environment," in *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, pp. 182–187, Riyadh, Saudi Arabia, 2021.

[6] X. Zhang, G. S. Xia, Q. Lu, W. Shen, and L. Zhang, "Visual object tracking by correlation filters and online learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 140, pp. 77–89, 2018.

[7] F. Changhong, L. Fuling, L. Yiming, and C. Guang, "Correlation filter-based visual tracking for UAV with online multi-feature learning," *Remote Sensing*, vol. 11, no. 5, 2019.

Research Article

An Optimized Method for Skin Cancer Diagnosis Using Modified Thermal Exchange Optimization Algorithm

Liu Wei ⓘ,[1] **Su Xiao Pan** ⓘ,[2] **Y. A. Nanehkaran** ⓘ,[3] **and V. Rajinikanth** ⓘ[4]

[1]*Gannan University of Science & Technology, Ganzhou, Jiangxi 341000, China*
[2]*Ganzhou 851, Ganzhou, Jiangxi 341000, China*
[3]*School of Informatics, Xiamen University, Xiamen, 361005 Fujian, China*
[4]*Department of Electronics and Instrumentation Engineering, St. Joseph's College of Engineering, Chennai 600119, India*

Correspondence should be addressed to V. Rajinikanth; v.rajinikanth@ieee.org

Skin cancer is the most common cancer of the body. It is estimated that more than one million people worldwide develop skin cancer each year. Early detection of this cancer has a high effect on the disease treatment. In this paper, a new optimal and automatic pipeline approach has been proposed for the diagnosis of this disease from dermoscopy images. The proposed method includes a noise reduction process before processing for eliminating the noises. Then, the Otsu method as one of the widely used thresholding method is used to characterize the region of interest. Afterward, 20 different features are extracted from the image. To reduce the method complexity, a new modified version of the Thermal Exchange Optimization Algorithm is performed to the features. This improves the method precision and consistency. To validate the proposed method's efficiency, it is implemented to the American Cancer Society database, its results are compared with some state-of-the-art methods, and the final results showed the superiority of the proposed method against the others.

1. Introduction

Cancer, as a difficult disease to treat, has long occupied the human mind [1]. Cancer occurs when cells in a part of the body grow uncontrollably, divide rapidly, invade different tissues in the body, and spread throughout the body [2]. A set of these uncontrollable cells is called a tumor [3]. One of the deadliest sorts of cancers is skin cancer. Skin cancer has grown significantly over the past decades, and the importance of its early treatment is increasing day by day [4].

Melanoma is the third most common type of skin cancer and one of the malignant cancers. Melanoma is also referred to as malignant melanoma, which changes the color of the skin due to the abnormal function of pigment-producing cells. The disease is formed by the accumulation of melanin granules and its spread to the outermost layer of the skin. Despite significant mortality, melanoma is often treatable in the early stages of diagnosis. At the same time, distinguishing between melanoma and other benign moles in the early

stages of development is a challenging task, even for dermatologists. Melanoma is known as the 19th prevalent cancer in men and women. There were about 300,000 new cases in 2018.

The data gathered by the World Health Organization (WHO) in 2018 showed that there were 17852 melanoma cases in the United Kingdom [5]. This organization predicted that the number of melanoma cases will grow by 9% to 19513 with deaths growing by 13% to 3119 by 2025.The growth of skin cancer begins when damage to skin cells (often caused by ultraviolet light) causes mutations that rapidly multiply in skin cells and form malignant tumors.

Normally, skin cells grow in a controlled and regular way. However, some newly produced cells may grow out of control and form a mass of cancer cells. Changes in the shape, size, and color of a person's mole are often the first signs of melanoma [6]. Melanoma has a black or bluish-black border; melanoma also appears as new black spots with an abnormal appearance [7]. These pigment-producing tumors are

present in the surface layer of the skin (epidermis [1]). Based on the WHO reports, melanoma with 15000 cases is ranked as the fourth prevalent cancer and with 1900 cases is the ninth deadliest cancer [8].

Diagnosis of skin cancer is difficult to distinguish due to the appearance of different types of skin lesions, especially melanoma and nevi. Even with dermoscopy, a noninvasive experimental technique, the accuracy of melanoma diagnosis by dermatologists is 84-75%. Sampling, however, provides a better diagnosis that is only possible based on surgery, which can lead to an unpleasant experience for the patient.

To prevent unnecessary sampling, researchers have reviewed several noninvasive methods for diagnosing melanoma. These methods usually involve three steps: (1) skin boundary identification, (2) feature extraction, and (3) classification [9]. The border-detection process detects the tumor in skin-related images, which is essential for the accurate classification of skin lesions. The feature extraction process uses visual properties such as color, mass shape, and texture information to classify [10]. The classification process also extracts the type of skin lesions from the image features and performs classification operations.

Navid and Ghadimi [11] proposed a method for melanoma detection in the images. Edge detection and smoothing technique were used for eliminating extra scales. Then, the segmentation method was performed. During the segmentation, mathematical morphological was used for eliminating the extra information on the melanoma boundary area. The classification of the method was performed by an optimized Artificial Neural Networks (ANN) based on World Cup Optimization (WCO) algorithm to minimize the root mean square error between the network output and the desired output. The final results indicated that the suggested technique develops the method's efficacy. Recently, several research works are introduced for the early diagnosis of skin cancers [12]. For example, Sugiarti et al. [13] introduced a method for the early diagnosis of melanoma cancer. The feature extraction method of the first order was utilized for feature extraction to achieve higher precision. The classification was performed by the Artificial Neural Network (ANN). The final results indicated that that using the proposed method provides a satisfied result for the analyzed images.

Zhi et al. [14] presented a CAD system for early detection of skin cancer. The method uses a median filter for noise reduction. Image segmentation was done based on Convolutional Neural Network (CNN) that is optimized by Satin Bowerbird Optimization (SBO). Afterward, feature extraction and feature selection were done to extract the valuable information from the segmented image. The feature selection was based on the SBO algorithm. Final features were fed to a Support Vector Machine (SVM) classifier for final recognition. The results were validated by applying them to the American Cancer Society database and comparing them with some different techniques from the literature.

Esteva et al. [15] suggested a diagnosis technique for lesion segmentation using deep learning. The analysis of the proposed method is validated by 21 clinical images to classify them into two groups of malignant and benign classes. The study analyzed two cases: the first identified the prevalent cancers, and the other one determined the deadliest skin cancer identification. The results indicated high efficacy for the suggested method.

It is clear from the literature that several applications of the deep learning in skin cancer detection still have lots of space. Therefore, in this paper, a new optimized method has been proposed for skin lesion diagnosis with higher performance based on a new modified version of the Thermal Exchange Optimization Algorithm.

The next parts of this study are structured as follows. In "Noise Reduction from the Images," the method of NLM based on the Yaroslavsky filter is used as a beneficial noise reduction tool. In "Image Segmentation," the method of image segmentation which is based on the Otsu thresholding and mathematical morphology is explained. In "Methodology," the proposed Modified Thermal Exchange Optimization Algorithm along with its application for optimal feature selection is mentioned. In "Classification," the classification method of the study which is based on the support vector machine is stated. In "Results and Discussions," the simulation results and their discussion are explained, and finally, the paper is concluded in "Conclusions."

2. Noise Reduction from the Images

Preprocessing is used to correct problems in images taken that may occur during medical imaging, such as noise or light. In medical imaging, there may be disturbances due to high-frequency reception, different brightness in the field, and problems due to distant orientation, which are corrected by artificial intelligence and image processing, and usually by default on all images before the main processing. In this paper, two modifications have been used as image preprocessing to improve the system performance [16]. Due to the stochastic physical nature of imaging systems, noise in the image is unavoidable, making it difficult to perform various image processes such as segmentation, detection, and interpretation [17]. The important point during the noise reduction is that the original image and especially its details are not damaged as much as possible and the structure of the original image is preserved. Based on this, various methods have been proposed to eliminate noise. In this study, we used the newly introduced NLM method for this purpose.

The NLM filter is an extended version of the Yaroslavsky filter [18], which uses nonlocal averaging of similar pixels (pixels with a closer brightness level) to retrieve the actual amount of pixels being processed. The main advantage of the NLM method compared to this method is that it has a more stable similarity criterion in the presence of noise, because, in addition to comparing the pixels intensity levels, a neighbor of them has also a role in determining the degree of similarity. The NLM method has a good performance in reducing most noise models, especially if the noise can be distributed collectively. The NLM method is based on the weight of all the pixels in the image, in proportion to the similarity of their neighbors; in other words, the more similar the image pixel neighbors are to the pixel neighbor being processed, the higher the weight assigned to them. The amount of pixels being processed is calculated using the total weight

found from the other pixels. The neighborhood criterion similarity in the NLM method is the weighted Euclidean principle with the Gaussian kernel, which is shown in Equation (1).

$$d = \left\| v(\eta_i) - v\left(\eta_j\right) \right\|_{2\rho}^2 = \sqrt{\sum_{k=1}^{N} w_k \left(x_{ik} - x_{jk}\right)^2}, \quad (1)$$

where $v(\eta_i)$ describes the pixel neighborhood vector under process, other pixels' neighborhood vector, and $\|.\|_{2\rho}^2$ represents the weighted Euclidean distance operator with Gaussian kernel.

In other words, in calculating the similarity of neighborhoods, the central pixel has a higher value, and by moving through the central pixel, the effect of the pixels decreases.

$$W\left(M_i, M_j\right) = \frac{1}{Z_i} \exp\left(-\frac{\left\| v(\eta_i) - v\left(\eta_j\right) \right\|_{2\rho}^2}{h^2}\right), \quad (2)$$

where

$$Z_i = \sum_j \exp\left(-\frac{\left\| v(\eta_i) - v\left(\eta_j\right) \right\|_{2\rho}^2}{h^2}\right) \quad (3)$$

where Z_i is a normalization parameter that guarantees the utilized sum of weights equals 1. h describes the main parameter of the NLM that determines the filtering intensity. If h is selected small, the value of the filtering in the image is small, and the noise effect has been not removed properly, but a large value for h makes an overfiltering for the image, and the reconstructed image is completely blurred and devoid of fine structural details. The final equation of the NLM filter with computed weighted coefficients can be formulated as follows:

$$\text{NLM}(M_i) = \sum_j W\left(M_i, M_j\right) M_j, \quad (4)$$

Although all pixels must be weighed in retrieving each pixel image, this operation is very time-consuming, so a specific area called the search window around each pixel being processed is used for the weighting operation. As explained before, NLM is a parameter filter with the following parameters: search window radius, similarity window radius, and smoothing parameter (h). Figure 1 shows a sample of noise reduction for this case.

3. Image Segmentation

3.1. Image Thresholding.
The thresholding method is used to remove unnecessary information and focus on the basic information in the image. Also, if the objects in the image and the "background" have similar gray levels, this method

(a)　　　　　　　　　(b)
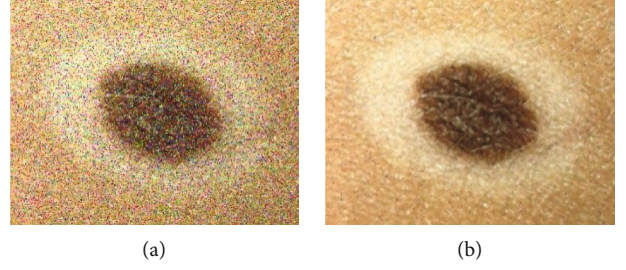
FIGURE 1: Image noise reduction: (a) before and (b) after processing.

is used to reveal hidden details in the image. Therefore, after noise reduction in the previous section, for highlighting the brain region, image thresholding has been used. One of the most popular and classic methods for finding the best threshold value is the Otsu method.

The Otsu method provides global thresholding for the input image. It uses the image histogram for maximizing the "between-class variance" of the segmented classes which consequently minimizes the "within-class variance" of the segmented classes. However, maximizing "between-class variance" needs less computational complexity than minimizing "within-class variance." During the Otsu thresholding, we look forward to a threshold level to minimize the class variance, i.e.

$$\sigma_\omega^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t), \quad (5)$$

$$\sigma_b^2(t) = \sigma^2 - \sigma_\omega^2(t) = \omega_1(t)\omega_2(t)(\mu_1(t) - \mu_2(t))^2, \quad (6)$$

where ω_i signifies the probability for two separate classes with a threshold value of t, σ_i^2 describes the variance of the classes, and $\mu_i(t)$ represents the mean value of the class and is updated alternately.

The Otsu thresholding can be briefly considered as follows:

(1) Calculate the histogram and the probabilities for each intensity level:

    (1.1) Initialize the $\omega_i(0)$ and $\mu_i(0)$ for all possible threshold levels

    (1.2) Update ω_i and μ_i

    (1.3) Calculate $\sigma_b^2(t)$

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

3.2. Morphological Operations.
After performing the thresholding stage, mathematical morphology has been used to abolish the spare parts of the region of interest in skin cancer images [19]. Mathematical morphology is based on applying a structural element (e) to the considered image. Here, a 5×5 identity matrix is used for structure element. In this study, mathematical filling, opening, and closing have been employed for this purpose. The first operation is to use mathematical filling. This operation is used to fill the empty holes

in the threshold image. This operator can be achieved by the following equation:

$$X_k = (X_{k-1} \oplus e) \cap A^c, k = 1, 2, 3 \cdots, \tag{7}$$

where A and e represent the area and the structure element, respectively.

After filling the holes, the mathematical opening operation has been performed to the image to eliminate the lighter details without deploying other gray surfaces. This is done by the following equation:

$$A \circ e = (A \ominus e) \oplus e. \tag{8}$$

The last process is to perform the mathematical closing to connect the narrow parts. The formula for this operation is given below:

$$A \bullet e = (A \oplus e) \ominus e. \tag{9}$$

Figure 2 shows a sample for skin cancer segmentation based on the explained method.

4. Methodology

In this study, a new modified metaheuristic has been proposed, and then, it has been applied for providing an optimal feature selection to get better results of diagnosing.

4.1. The Modified Thermal Exchange Optimization Algorithm.
Achieving the optimal state has been one of the most fundamental issues in the world since the creation of the universe. The scope of application of optimization-related topics is very wide. Mathematics, computer science, engineering, physics, and economics are just some of these topics. In this type of problem, the goal is to get the best decision mode from several different modes [20]. Metaheuristic algorithms can be considered one of the most important classes of optimization solutions for these types of issues. These algorithms have a lot of variety [21]. The great variety of these algorithms in solving different problems, as well as the introduction of new algorithms with different titles, has made choosing a suitable algorithm for the user who intends to use them a difficult and complex task [22]. On the other hand, each of these algorithms obtains the optimal solution with certain accuracy and speed. Therefore, it seems necessary to have a structure that can well identify the differences between these algorithms and make their comparison easier. On the other hand, the implementation of each algorithm typically requires complete knowledge of that algorithm and professional programming knowledge. Some examples of these algorithms are like the Chimp Optimization Algorithm (ChOA) [23], Black Hole (BH) [24], Crow Search Algorithm (CSA) [25], Water Strider Algorithm (WSA) [26], Ant Lion Optimizer (ALO) algorithm [27], and Thermal Exchange Optimization (MTEO) [28]. In this study, a modified version of this algorithm called the Modified Thermal Exchange Optimization (MTEO) algorithm is proposed to achieve optimal results for different parts of the diagnosis

(a) (b)
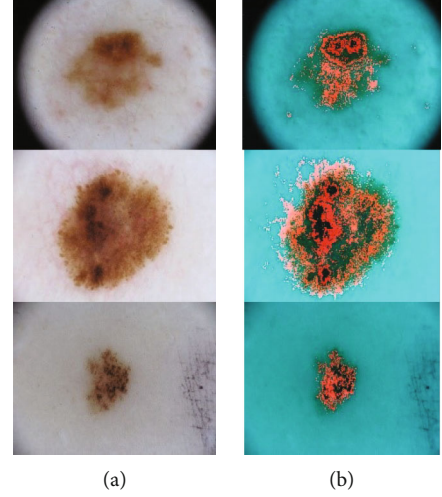
FIGURE 2: A sample for skin cancer segmentation based on the explained method: (a) input image and (b) segmented.

system. This TEO algorithm is a metaheuristic technique that is derived by the temperature behavior for the objects and their location which is exchanged between warm and cold parts and specifies the updated locations. More explanations are explained in the following.

4.1.1. The Newton Law of Cooling.
The Newton law of cooling states that the rate at which a body temperature changes is approximately proportional to the difference in temperature between the body and its surroundings. This was first discovered by Newton. When the temperature difference between the body and its surroundings is small, the average amount of heat exchanged between the body and its surroundings due to conduction, convection, and infrared radiation is approximately proportional to the difference in temperature of the body and the environment. Newton's law of cooling is the solution of a differential equation of the Fourier law which is formulated as follows:

$$\frac{dQ}{dt} = \alpha \times A \times (T_s - T_a), \tag{10}$$

where Q defines the heat, A signifies the body area surface which transmits heat, α represents the heat transfer coefficient which depends on several cases such as heat transfer mode, surface state, and object geometry, and T_b and T_a describe the body temperature and the ambient temperature.

Based on the equation, the time for losing heat is $\alpha \times A \times (T_a - T) \, dt$ which determines the change in reserved heat as the temperature falls dT, i.e.

$$V \times \rho \times c \times dT = -\alpha \times A \times (T - T_b) dt, \tag{11}$$

where c represents the specific heat (J/kg/K), ρ describes the density (kg/m^3), and V specifies the volume (m^3).

Hence

$$\frac{T - T_b}{T_M - T_b} = \exp\left(\frac{-\alpha \times A \times t}{V \times \rho \times c}\right), \tag{12}$$

where T_M represents the early high temperature. The above equation is correct when $\alpha \times A \times t / V \times \rho \times c$ is has not depended to T:

$$\zeta = \frac{\alpha \times A}{V \times \rho \times c}, \tag{13}$$

Hence, by assuming ζ as a constant

$$\frac{T - T_b}{T_M - T_b} = \exp(-\zeta t). \tag{14}$$

Accordingly

$$T = (T_M - T_b) \times \exp(-\gamma t) + T_b. \tag{15}$$

4.1.2. The Algorithm. In Thermal Exchange Optimization Algorithm, some individuals are considered cooling substances, and the other leftover individuals are considered the environment, and then, the reverse process is performed. Like any other metaheuristic algorithm, the TEO algorithm starts with initializing a definite number of randomly distributed individuals as the solution candidates. This can be presented as follows:

$$T_i^0 = T_{\min} + \delta \times (T_{\max} - T_{\min}),$$
$$i = 1, 2, \cdots, n, \tag{16}$$

where T_i^0 describes the initial population of the algorithm for the ith object, δ represents a random value limited in the range [0, 1], and $T_{\min}$ and $T_{\max}$ describe the minimum and maximum boundaries.

The cost value of all randomly generated individuals is then evaluated to indicate the cost of each algorithm. Then, the best T candidate vector positions have been stored as thermal memory (TM) to employ for developing the algorithm performance with less complexity. Some best TM candidates are then added to the individuals, and the same numbers of them that have the worst values are removed. Therefore, individuals have two equal types of environment, and the heat and cooling transfer objects can be seen in Figure 3.

To get a better conception, T_1 defines the environment object for $T_{(n/2)+1}$ cooling object, and contrariwise. If the object gives a lower value than ζ, the temperature exchanges gradually. In this situation, ζ has been achieved as follows:

$$\gamma = \frac{\text{Cos (object)}}{\text{Cos (worst object)}}. \tag{17}$$

This algorithm uses time as another significant term for the simulation. This term directly depends on iteration number. This can be mathematically formulated as follows:

$$t = \frac{\text{iteration}}{\text{Max.iteration}}. \tag{18}$$

For increasing the global searching in the algorithm, environmental temperature changing has been considered that can be considered as follows:

$$T_i^e = (1 - (m_1 + m_2 \times (1 - t) \times \text{rand})) \times T_i^{\textit{e}}, \tag{19}$$

where $T_i^{\textit{e}}$ describes the previous temperature of the object modified by T_i^e and m_1 and m_2 represent the control variables, respectively.

Considering the past models, the object new temperature can be mathematically updated by the following equation:

$$T_i^+ = T_i^e + \left(T_i^{old} - T_i^e\right) \exp(-\zeta t). \tag{20}$$

The final case which is considered in this algorithm is Pr. This term shows that a component changes in the cooling objects or not.

The Pr individuals have been compared with $R(i)$ which has a random value in the range [0, 1]. If $R(i) < \text{Pr}$, one dimension of the ith individual has been randomly selected, and the value is rewritten in the following:

$$T_{i,j} = T_j^{\min} + rnd\left(T_j^{\max} - T_j^{\min}\right) \exp(-\zeta t), \tag{21}$$

where $T_{i,j}$ describes the jth variable of the individual number i and $T_j^{\min}$ and $T_j^{\max}$ represent the lower and the upper bounds of the variable number j, respectively. Finally, the algorithm will be terminated if stopping criteria have been met.

4.1.3. Modified Thermal Exchange Optimization Algorithm. From the literature, the method is compared with DE, ECBO, CBO, PSO, GWO, GA, and lots of other optimization methods (20 other methods). The results showed that the original TEO has better convergence than most of the algorithms with a satisfied solution value. Then, the original paper concluded that TEO can be employed as a search engine in most of the optimization problems [28]. Also, it might be a source of inspiration for future algorithms or improved and hybridized with other methods. In this section, the details of the suggested modified Thermal Exchange Optimization Algorithm, named MWSA, have been presented. In a general form, metaheuristic algorithms should be efficient in two significant terms, exploitation and exploration, such that it can found an appropriate trade-off between them for better performance. The algorithm has the advantage of fast convergence and excellent local search capability, although it tends to fall into a local optima point rather than finding the global optimum [29–31]. In order to develop the algorithm efficiency by giving a proper balance between exploration and exploitation terms, a modification has been applied to it in this study. Opposition-based learning and chaos map are two modification mechanisms that are used here for improving algorithm efficiency.

The first mechanism, the opposition-based learning (OBL) mechanism, was first presented by Tizhoosh [32]. This mechanism contains a strong mathematical concept for improving the global searching of the algorithm. As
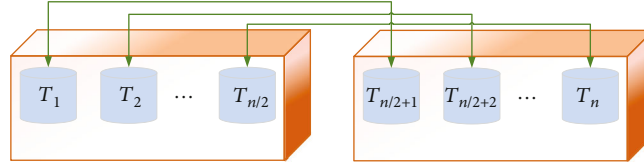
FIGURE 3: The pairs of environment and the heat and cooling transfer objects.

TABLE 1: The information about the utilized test functions.

No.	Test function	Minimum value	Boundary
1	$F_1 = \sum\limits_{n=1}^{N} x_n^2$	0	$-\infty \leq x \leq \infty$
2	$F_2 = \sum\limits_{n=1}^{N-1} \left(100 \times \left[x_{n+1} - x_n^2 \right]^2 + \left[1 - x_n \right]^2 \right)$	0	$-\infty \leq x_n \leq \infty$
3	$F_3 = \sum\limits_{n=1}^{N} |x_n| - 10 \cos \left(\sqrt{|10 x_n|} \right)$	0	$-\infty \leq x_n \leq \infty$
4	$F_4 = x \sin (4x) + 1.1 y \sin (2y)$	-18.5547	$0 \leq x, y \leq 10$
5	$F_5 = \left[\sum\limits_{n=1}^{N} n x_n^4 \right] + N_n(0,1)$	Varies	$-\infty \leq x \leq \infty$
6	$F_6 = 10N + \sum\limits_{n=1}^{N} \left[x_n^2 - 10 \cos (2\pi x_n) \right]$	0	$-\infty \leq x_n \leq \infty$
7	$F_7 = 1 + \sum\limits_{n=1}^{N} \frac{x_n^2}{4000} - \prod\limits_{n=1}^{N} \cos (x_n)$	0	$-\infty \leq x_n \leq \infty$
8	$F_8 = \frac{1}{2} + \frac{\sin^2 \sqrt{x^2 + y^2} - 0.5}{1 + 0.1(x^2 + y^2)}$	-0.5231	$-\infty \leq x, y \leq \infty$

aforementioned, the initializing step in TEO is completely random, and the aim is to find the best points in the solution space. Here, if the generated variables have a proper value close to the solution space, the proper solution will be achieved. But, if the algorithm starts with values too distant from the optimal solution, the time for finding the global value will be extended or even makes a premature convergence in some cases. The OBL is a mechanism to modify this issue by generating opposite values from the originally generated population. So, for every single solution, its original cost value and its opposite cost value have been compared, the best one will have remained, and the other will be removed. This can be mathematically formulated as follows:

$$\widehat{T}_i^+ = T_{\max} + T_{\min} - T_i^+, \tag{22}$$

where $\widehat{T}_i^+$ describes the opposite position of T_i^+ and $T_{\min}$ and $T_{\max}$ describe the variables upper and the lower bounds in the problem, respectively.

The new position provides a higher opportunity to get the best solution. The second mechanism is the chaos map. This mechanism utilizes chaotic conception to generate unpredictable variables instead of random variables. This mechanism accomplishes simple searches at a higher convergence rate than probability-based random searches [33]. By

employing chaotic variables instead of random ones in metaheuristics, better exploration has been generated for the solution space because of the dynamic behavior of the sequence [34]. Several functions have been introduced as chaos functions [35]. This study employed a sinusoidal chaotic map function to modify the convergence speed of the TEO and make a balance between its exploitation and exploration terms. By considering the sinusoidal map in the TEO algorithm, environmental temperature changing is considered:

$$T_i^e = \left(1 - \left(m_1 + m_2 \times (1-t) \times k_{i+1} \right) \right) \times T_i^{\varepsilon}, \tag{23}$$

$$k_{i+1} = \alpha \times k_i^2 \sin (\pi . k_i), \tag{24}$$

where k_{i+1} describes a chaotic random number made by current iteration and k_i describes the chaotic random number made by the previous iteration. $P = 2.3$ defines the control parameter, and the k_0 is considered a random value in the range [0, 1].

4.1.4. Algorithm Authentication. In this paper, in order to demonstrate the effectiveness of the suggested MTEO, eight standard benchmark functions have been selected which are listed in Table 1. To provide a comprehensive analysis on the optimization performance, the results of the proposed

TABLE 2: The parameters setting utilized for the comparative algorithms utilized in this study.

Algorithm	Parameter	Value	Algorithm	Parameter	Value
BBO [36]	P_{habit}	1	EPO [38]	$\vec{A}$	[-1.5, 1.5]
	P_{imig}	[0,1]		T'	[1, 1000]
	Step size	1		M	2
	E	1		f	[2, 3]
	I	1		S	[0, 1.5]
	P_{mutation}	0.005		l	[1.5, 2]
	F	0.6		$\vec{M}$	[0.5, 1]
LS [37]	L	1	SHO [39]	$\vec{h}$	[5, 0]
	g	20			

MTEO have been compared with some different new state-of-the-art metaheuristics, including the Biogeography-Based Optimizer (BBO) [36], Locust Swarm Optimization (LS) [37], Emperor Penguin Optimizer (EPO) [38], Spotted Hyena Optimize (SHO) [39], and original Thermal Exchange Optimization Algorithm [40]. Table 1 indicates the information about the utilized test functions.

The experiment environments are MATLAB 2019b, the Core™ i7-4720HQ with 1.60 GHz CPU, 16 GB RAM with Windows 10. Table 2 indicates the parameters setting utilized for the comparative algorithms utilized in this study.

This study, considers two important measures including mean value and standard deviation value results from the applying optimization algorithms on the benchmark functions after 35 independent runs. To achieve a fair comparison between the proposed MTEO and the comparative algorithms, the population size for all of them and the iteration number are considered 100 and 200, respectively [41]. Table 3 illustrates the performance analysis of the comparison.

As can be observed from Table 3, the proposed MTEO algorithm provides the smallest value for the mean value of the benchmark functions. This shows that the proposed MTEO algorithm has the highest accuracy compared with the other algorithms. Also, the standard deviation value achieved by the algorithms shows the minimum value based on the MTEO algorithm that shows consequently the higher reliability of the proposed method against the other compared methods.

4.2. Feature Extraction and Selection. After segmentation of the region of interest from the input images, the main information (features) has been extracted from the images to reduce the complexity of the diagnosis process by considering only vital characteristics. In other words, feature extraction provides an easy way for demonstrating and analyzing the images. Recently, several algorithms have been performed for proper feature extraction of the images. During the feature extraction with different methods, all of the patterns in the features should be searched and determined. In this study, 20 different features are employed to extract the beneficial features from the segmented skin cancer for the

diagnosis. In this study, three groups of features, i.e., geometric features, statistical features, and texture features, are utilized. In the following, the formulation of the utilized features is explained:

$$\text{Mean} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} p(i, j), \tag{25}$$

$$\text{Variance} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (p(i, j) - \mu), \tag{26}$$

$$\text{Std} = \sqrt{\text{variance}}, \tag{27}$$

$$\text{Contrast} = \sum_{i=1}^{M} \sum_{j=1}^{N} p^2(i, j), \tag{28}$$

$$\text{Area} = \sum_{i=1}^{M} \sum_{j=1}^{N} p(i, j), \tag{29}$$

$$\text{Rectangularity} = \frac{\text{Area}}{a \times b}, \tag{30}$$

$$\text{Elongation} = \frac{2\sqrt{\text{Area}}}{a\sqrt{\pi}}, \tag{31}$$

$$\text{Irregularity index} = \frac{4\pi \times \text{Area}}{\text{Perimeter}^2}, \tag{32}$$

$$\text{Form Factor} = \frac{\text{Area}}{a^2}, \tag{33}$$

$$\text{Eccentricity} = 2a^{-1} \left(a^2 - b^2\right)^{0.5}, \tag{34}$$

$$\text{Entropy} = \sum_{i=1}^{M} \sum_{j=1}^{N} p(i, j) \log p(i, j), \tag{35}$$

$$\text{Perimeter} = \sum_{i=1}^{M} \sum_{j=1}^{N} b_p(i, j), \tag{36}$$

$$\text{Homogeneity} = \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{p(i, j)}{1 + |i - j|}, \tag{37}$$

$$\text{Energy} = \sum_{i=1}^{M} \sum_{j=1}^{N} p^2(i, j), \tag{38}$$

$$\text{Correlation} = \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{p(i, j) - \mu_r \mu_c}{\sigma_r \sigma_c}, \tag{39}$$

$$\begin{aligned}
\varphi_1 &= \eta_{20} + \eta_{02}, \\
\varphi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \\
\varphi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2, \\
\varphi_4 &= (\eta_{30} + 3\eta_{12})^2 + (3\eta_{21} + \mu_{03})^2,
\end{aligned} \tag{40}$$

where b_p signifies the external side length of the boundary pixel, $p(i, j)$ represents the pixel intensity value at point (i, j),

TABLE 3: The performance analysis of the comparative algorithms applied to studied standard benchmarks.

Algorithm Function		BBO [36]	LS [37]	EPO [38]	SHO [39]	TEO	MTEO
f_1	Min	2.615e-25	1.1100e-29	-3.2688e-26	2.3086e-27	2.4400e-30	9.2082e-32
	Std	1.448e-20	3.3826e-28	4.0754e-27	1.8827e-28	1.0062e-32	3.2681e-33
f_2	Min	6.0652e-4	8.3420e-3	5.6024e-3	1.4527e-4	2.4352e-5	7.6700e-5
	Std	4.1073e-5	3.0718e-4	1.0056e-4	2.4807e-5	3.0537e-6	1.0142e-5
f_3	Min	-6.1442	-9.0464	-9.86	-8.0826	-9.86	-9.86
	Std	0.31	0.42	0.23	0.11	0.11	0.06
f_4	Min	-6.1735	-17.020	-16.0035	-15.2816	-17.0095	-17.0572
	Std	3.015	1.183	2.280	4.089	1.520	0.980
f_5	Min	12.35e-10	1.486e-15	3.0765e-8	4.0802e-8	1.7085e-22	2.6827e-23
	Std	7.831e-11	3.0862e-16	1.1832e-9	5.4403e-9	3.7786e-24	6.0826e-25
f_6	Min	5.165e-10	3.1842e-11	1.0856e-20	1.0846e-9	3.0008e-20	4.5013e-22
	Std	8.186e-11	2.4253e-13	5.1738e-22	4.7080e-11	1.2058e-21	2.5387e-23
f_7	Min	3.512e-14	2.2621e-9	4.0305e-8	2.6517e-10	1.5670e-9	7.2837e-16
	Std	1.056e-15	3.0856e-11	3.8253e-9	2.1825e-12	2.0834e-10	3.1175e-18
f_8	Min	0.0056	-0.1361	-0.2381	-0.4735	-0.4680	-0.4162
	Std	0.542	0.356	0.274	0.704	0.141	0.089

MN describes best the image size, a and b represent the major and the minor axis, respectively, and μ and σ represent the mean value and the standard deviation value, respectively.

Because of the higher volume of feature information and the presence of some useless features, some of these features should be then eliminated before the classification stage. This is done by using a method, called feature selection. To achieve an optimal diagnosis system, the suggested modified Thermal Exchange Optimization Algorithm has been utilized that is explained in the following.

Feature selection is the process of reducing the data dimension by choosing the best features and eliminating the others. Furthermore, however, some features are useless, but once they blend with other features, they have been beneficial. This study uses a definite cost function where by minimizing it, the optimal features can be selected. The cost function is formulated in the following:

$$ CF = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TN + FP) \times (TP + FP) \times (TP + FN) \times (TN + FN))}}, \tag{41} $$

where TP, FP, TN, and FN represent the true positive, false positive, true negative, and false negative, respectively.

The main idea is to minimize the above function. This is performed by the proposed modified Thermal Exchange Optimization Algorithm.

5. Classification

The classification in this study is based on the Support Vector Machine (SVM). The SVM consists of a set of points in the n-
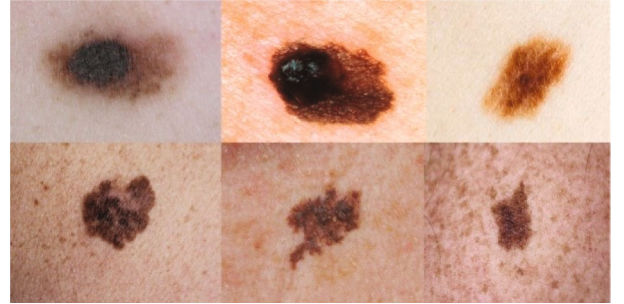
FIGURE 4: Some examples of the American Cancer Society (ACS) database [43].

dimensional space of data which indicates the class boundaries and organizes them and can be altered with the rearrangement of one of these two cases. The SVM provides the best results for separating the data with the criterion for placement of the support vectors. This classifier organizes the best separation surface by the following equation:

$$ y = \text{sgn}\left(\sum_{i=1}^{N} y_i \alpha_i K(x, x_i) + b \right), \tag{42} $$

where $K(x, x_i)$ describes a kernel function, x signifies a test set vector with d dimensions, x_i describes the ith training set vector, y represents the output class by labeling -1 or 1, N is the number of the training set, and b and $\alpha = [\alpha_1 \cdots \alpha_N]$ represent the model parameters, respectively.

The present study uses SVM for the classification of the extracted features achieved by the feature selection in the previous stage into two parts of healthy and cancerous groups.
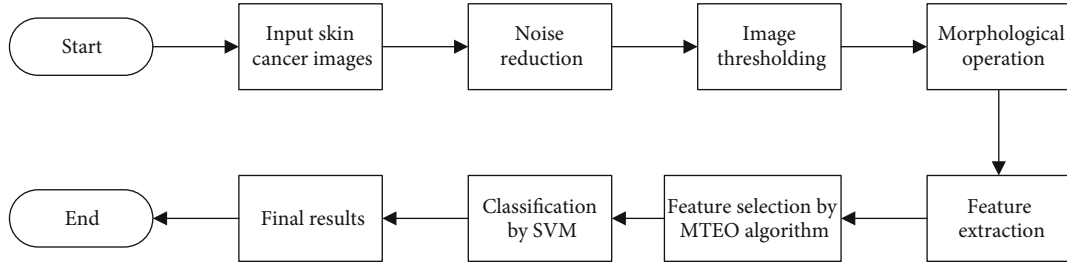
Figure 5: The pipeline of the proposed methodology.

Table 4: The validation results of the compared method for skin cancer diagnosis.

| Method | Performance metric | | | | |
	NPV	PPV	Specificity	Accuracy	Sensitivity
PSO	93.69	89.19	88.29	88.29	90.99
m-Skin Doctor	83.87	65.76	61.26	81.98	83.78
GFAC	88.28	77.48	82.88	86.48	89.19
ANN	82.88	58.56	56.76	67.57	81.98
GA	85.58	74.77	79.28	81.08	79.28
Proposed method	93.69	85.58	89.19	92.79	90.99

6. Results and Discussions

The main purpose of this study is to present a computer-aided automatic method for optimal diagnosis of skin cancer from the dermoscopy images. The idea is to utilize a metaheuristic-based method to achieve the best feature selection, and consequently the best diagnosis.

6.1. Database. To validate the proposed skin cancer diagnosis system, the so-called American Cancer Society (ACS) database has been employed. This database contains 68 pairs of XLM and TLM images that are collected from the Nevoscope system. 51 XLM images and 60 TLM images have been manually classified by a dermatologist since other images do not show pigmentation [42]. Therefore, the validation has been based on comparing our results with these manually segmented results. For giving less complexity to the analysis, all of the images are resized to 256 to 256 pixels. Some examples of this database are given in Figure 4 [43].

6.2. Simulation Results. The present study in this subsection has been verified on the ACS database, and the results have been validated based on some different state-of-the-art techniques. Simulations have been validated by the MATLAB 2019b environment with the following hardware configuration: Core™ i7-4720HQ 1.60 GHz with 16 GB RAM. The overall procedure of the suggested method is illustrated in Figure 5.

The present study uses 85% of the data for training and 15% for testing the data. The training stage is based on applying 750 iterations and is iterated 20 times independently to achieve a guaranteed result. Five measurement indicators

including PPV, NPV, specificity, accuracy, and sensitivity are used for validation that is formulated as follows:

$$PPV = \frac{\text{correctly detected skin cancer cases}}{\text{detected skin cancer cases}}, \quad (43)$$

$$NPV = \frac{\text{correctly detected healthy skin cases}}{\text{detected healthy skin cases}}, \quad (44)$$

$$Specificity = \frac{\text{correctly detected healthy skin cases}}{\text{total healthy skin cases}}, \quad (45)$$

$$Accuracy = \frac{\text{correctly detected cases}}{\text{total cases}}, \quad (46)$$

$$Sensitivity = \frac{\text{correctly detected skin cancer cases}}{\text{Total skin cancer cases}}. \quad (47)$$

To give a fair analysis on the proposed method, its results have been compared with some different state-of-the-art methods including the Particle Swarm Optimization- (PSO-) based method [44], m-Skin Doctor [45], GFAC [46], ANN [47], and Genetic Algorithm (GA) [42]. The results of the validation are tabulated in Table 4.

It can be observed from Table 4 that the proposed optimized methodology with 92.79% accuracy has the highest precision against the other comparative methods. Similarly, with 90.99% sensitivity, it has proper reliability compared with the other methods. This is also proved for the specificity, NPV, and PPV compared with the others. The higher value of NPV and PPV including 93.69% and 85.58%, respectively, which are the highest among the comparative methods, provides the higher prevalence of the condition to diagnose the likelihood of a test cancer diagnosis system. Finally, the better results of the sensitivity and specificity for the proposed method indicate the suggested method's higher prevalence-independent results.

7. Conclusions

Skin cancer is one of the most dangerous diseases among different cancers in the world. However, early detection of this disease can be so beneficial for cancer treatment. In the present study, a new hierarchical methodology was proposed for the optimal diagnosis of skin cancer from dermoscopy images. According to the suggested method, after performing noise reduction of the input dermoscopy images, the considered area has been segmented based on a simple Otsu. Then,

feature extraction has been performed to the processed image to extract valuable features from the images. To provide an optimized result, the best features have been selected by a modified metaheuristic method, called the Modified Thermal Exchange Optimization Algorithm to modify the network performance in terms of precision and consistency. Final results were obtained by applying support vector machine as the final classifier. To give a proper validation, the results of the proposed method were applied to the American Cancer Society (ACS) database, and its results were compared with some different methods including Particle Swarm Optimization- (PSO-) based method, m-Skin Doctor, GFAC, ANN, and Genetic Algorithm (GA). The final results indicated that according to different measurement indicators, the proposed methodology has the best results for the other compared methods. As can be observed from the explanations, the proposed method has good results for skin cancer detection. However, this can be an inspiration to our future work to use different hybrid and developed versions of different new computational intelligence algorithms like the Monarch Butterfly Optimization (MBO) [48], Earthworm Optimization Algorithm (EWA) [49], Elephant Herding Optimization (EHO) [50], Moth Search (MS) algorithm [51], Slime Mold Algorithm (SMA) [52], and Harris hawks optimization (HHO) [53] to improve the system efficiency.

Data Availability

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] 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 (IJSIR)*, vol. 11, no. 3, pp. 1–18, 2020.

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

[3] 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.

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

[5] C. Mattiuzzi and G. Lippi, "Cancer statistics: a comparison between world health organization (WHO) and global burden of disease (GBD)," *European Journal of Public Health*, vol. 30, no. 5, pp. 1026-1027, 2020.

[6] R. Pugalenthi, M. 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.

[7] 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.

[8] 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. e10, 2018.

[9] 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.

[10] N. S. M. Raja, S. 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.

[11] F. R. S. Navid and R. N. Ghadimi, "A hybrid neural network–world cup optimization algorithm for melanoma detection," *Open Medicine*, vol. 13, pp. 9–16, 2018.

[12] 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.

[13] Y. Sugiarti, 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.

[14] Y. Zhi, W. Weiqing, W. Haiyun, and R. Navid, "New approaches for regulation of solid oxide fuel cell using dynamic condition approximation and STATCOM," *International Transactions on Electrical Energy Systems*, vol. 31, article e12756, 2021.

[15] 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.

[16] S. C. Satapathy, N. S. M. 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.

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

[18] J. Salmon, R. Willett, and E. Arias-Castro, "A two-stage denoising filter: the preprocessed Yaroslavsky filter," in *2012 IEEE Statistical Signal Processing Workshop (SSP)*, pp. 464–467, Ann Arbor, MI, USA, 2012.

[19] R. Navid, B. S. Mousavi, B. Sadeghi, and M. Khalilpour, "Image thresholding optimization based on imperialist competitive algorithm," in *3rd Iranian Conference on Electrical and Electronics Engineering (ICEEE2011)*, Qom, Iran, 2011.

[20] R. Navid, M. Ashourian, and Z. Foroozandeh, *Metaheuristics and Optimization in Computer and Electrical Engineering*, Springer International Publishing, 2021.

[21] R. Navid, V. V. Estrela, H. J. Loschi, and W. Fanfan, "A Comprehensive Survey of New Meta-Heuristic Algorithms," in

Recent Advances in Hybrid Metaheuristics for Data Clustering, Wiley Publishing, 2019.

[22] M. Ramezani, D. Bahmanyar, and R. Navid, "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.

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

[24] A. Hatamlou, "Black hole: a new heuristic optimization approach for data clustering," *Information Sciences*, vol. 222, pp. 175–184, 2013.

[25] A. Askarzadeh, "A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm," *Computers & Structures*, vol. 169, pp. 1–12, 2016.

[26] A. Kaveh, A. D. Eslamlou, and N. Khodadadi, "Dynamic water strider algorithm for optimal design of skeletal structures," *Periodica Polytechnica Civil Engineering*, vol. 64, no. 3, pp. 904–916, 2020.

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

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

[29] A. Kaveh and A. Dadras, "Structural damage identification using an enhanced thermal exchange optimization algorithm," *Engineering Optimization*, vol. 50, no. 3, pp. 430–451, 2018.

[30] A. Kaveh, A. Dadras, and T. Bakhshpoori, "Improved thermal exchange optimization algorithm for optimal design of skeletal structures," *Smart Structures and Systems*, vol. 21, no. 3, pp. 263–278, 2018.

[31] Z. Xing and H. Jia, "Modified thermal exchange optimization based multilevel thresholding for color image segmentation," *Multimedia Tools and Applications*, vol. 79, no. 1-2, pp. 1137–1168, 2020.

[32] H. R. Tizhoosh, "Opposition-based learning: a new scheme for machine intelligence," in *International conference on computational intelligence for modelling, control and automation and international conference on intelligent agents, web technologies and internet commerce (CIMCA-IAWTIC'06)*, vol. 1, pp. 695–701, Vienna, Austria, 2005.

[33] G. Zhang, C. Xiao, and R. Navid, "Optimal parameter extraction of PEM fuel cells by meta-heuristics," *International Journal of Ambient Energy*, pp. 1–10, 2020.

[34] Z. Yuan, W. Wang, H. Wang, and R. Navid, "A new technique for optimal estimation of the circuit-based PEMFCs using developed Sunflower Optimization Algorithm," *Energy Reports*, vol. 6, pp. 662–671, 2020.

[35] M.-W. Tian, S.-R. Yan, S.-Z. Han, S. Nojavan, K. Jermsittiparsert, and R. Navid, "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 119414, 2020.

[36] D. Simon, "Biogeography-based optimization," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 6, pp. 702–713, 2008.

[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, 2020.

[38] G. Dhiman and V. Kumar, "Emperor penguin optimizer: a bio-inspired algorithm for engineering problems," *Knowledge-Based Systems*, vol. 159, pp. 20–50, 2018.

[39] 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.

[40] M. Jamil and X.-S. Yang, "A literature survey of benchmark functions for global optimisation problems," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 4, no. 2, pp. 150–194, 2013.

[41] M. S. Doshi, *Automated Segmentation of Skin Cancer Images*, University of Houston, 2004.

[42] 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, Cancun, Mexico, 2020.

[43] ACS, *Skin Cancer Database*, 2019, https://www.cancer.org/.

[44] T. Y. Tan, L. Zhang, S. C. Neoh, and C. P. Lim, "Intelligent skin cancer detection using enhanced particle swarm optimization," *Knowledge-Based Systems*, vol. 158, pp. 118–135, 2018.

[45] M. A. Taufiq, N. Hameed, A. Anjum, and F. Hameed, "m-Skin Doctor: a mobile enabled system for early melanoma skin cancer detection using support vector machine," in *eHealth 360°*, pp. 468–475, Springer, 2017.

[46] T. Sreelatha, M. Subramanyam, and M. G. Prasad, "Early detection of skin cancer using melanoma segmentation technique," *Journal of medical systems*, vol. 43, no. 7, p. 190, 2019.

[47] T. Kanimozhi and A. Murthi, "Computer aided melanoma skin cancer detection using artificial neural network classifier," *Singaporean Journal of Scientific Research (SJSR), Journal of Selected Areas in Microelectronics (JSAM)*, vol. 8, no. 2, pp. 35–42, 2016.

[48] G.-G. Wang, S. Deb, and Z. Cui, "Monarch butterfly optimization," *Neural Computing and Applications*, vol. 31, no. 7, pp. 1995–2014, 2019.

[49] G.-G. Wang, S. Deb, and L. D. S. Coelho, "Earthworm optimisation algorithm: a bio-inspired metaheuristic algorithm for global optimisation problems," *International Journal of Bio-Inspired Computation*, vol. 12, no. 1, pp. 1–22, 2018.

[50] G.-G. Wang, S. Deb, and L. D. S. Coelho, "Elephant herding optimization," in *2015 3rd International Symposium on Computational and Business Intelligence (ISCBI)*, pp. 1–5, Bali, Indonesia, 2015.

[51] D. Pelusi, R. Mascella, L. Tallini, J. Nayak, B. Naik, and Y. Deng, "An Improved Moth-Flame Optimization algorithm with hybrid search phase," *Knowledge-Based Systems*, vol. 191, p. 105277, 2020.

[52] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, "Slime mould algorithm: a new method for stochastic optimization," *Future Generation Computer Systems*, vol. 111, pp. 300–323, 2020.

[53] 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.

Research Article

An Effective Method for Detecting and Classifying Diabetic Retinopathy Lesions Based on Deep Learning

Abdüssamed Erciyas [ID] **and Necaattin Barışçı** [ID]

Department of Computer Engineering, Gazi University, Ankara, Turkey

Correspondence should be addressed to Abdüssamed Erciyas; abdussamed.erciyas@gazi.edu.tr

Diabetic retinopathy occurs as a result of the harmful effects of diabetes on the eyes. Diabetic retinopathy is also a disease that should be diagnosed early. If not treated early, vision loss may occur. It is estimated that one third of more than half a million diabetic patients will have diabetic retinopathy by the 22nd century. Many effective methods have been proposed for disease detection with deep learning. In this study, unlike other studies, a deep learning-based method has been proposed in which diabetic retinopathy lesions are detected automatically and independently of datasets, and the detected lesions are classified. In the first stage of the proposed method, a data pool is created by collecting diabetic retinopathy data from different datasets. With Faster RCNN, lesions are detected, and the region of interests are marked. The images obtained in the second stage are classified using the transfer learning and attention mechanism. The method tested in Kaggle and MESSIDOR datasets reached 99.1% and 100% ACC and 99.9% and 100% AUC, respectively. When the obtained results are compared with other results in the literature, it is seen that more successful results are obtained.

1. Introduction

Diabetes occurs as a result of insufficient production of insulin or insufficient use of produced insulin [1]. There are many organs damaged by diabetes. For example, diabetic nephropathy damaging kidney nephrons, diabetic neuropathy damaging brain neurons, and diabetic retinopathy damaging eye retina can be given [2]. Diabetic retinopathy (DR) is a type of type II diabetes in which the retina of the eye is damaged and if left untreated, the disease can progress to vision loss [3]. DR's effect on the eye is often blurred or complete loss of vision [4]. The risk of blindness in diabetic patients is many times higher than in a healthy person. Therefore, DR is one of the leading causes of blindness in the world between the ages of 20 and 65 [5]. The World Health Organization (WHO) stated that up to half a million people are at risk of DR [6]. The economies of low- and middle-income countries suffer seriously from diabetes. By 2040, it is estimated that 33% of 600 million diabetic patients worldwide will have diabetic retinopathy [7].

Deep learning (DL) started with the work of LeCun et al. [8]. DL's popularity began in 1998 with the success of the convolutional neural network (CNN), a DL method used by his student Krizhevsky [9] at the 2012 ImageNet [10] competition. In the years after AlexNet on ImageNet, GoogleNet [11], InceptionV3 [12], VGGNet [13], ResNet [14], and DenseNet [15], networks were developed, and more successful results were achieved. Improvements in GPU hardware have a great impact on the success here. Because as the depth increases in the developed networks, the number of trained parameters increases in direct proportion. While the number of parameters in GoogleNet is 6.8 M, there are 144 M parameters in the deeper VGG19. While CNN image classification was done, the CNN structure was modified for segmentation and object detection in the image. Region-based CNN (RCNN) [16], Fast RCNN [17, 18], Faster RCNN [19], Single Shot multiBox Detector (SSD) [20] and You Only Look Once (YOLO) [21, 22] appeared with this change. Experts believe that deep learning will facilitate medical studies in the coming years of medicine. The successes obtained in the works [23–30] on the subject support this idea; it is about the
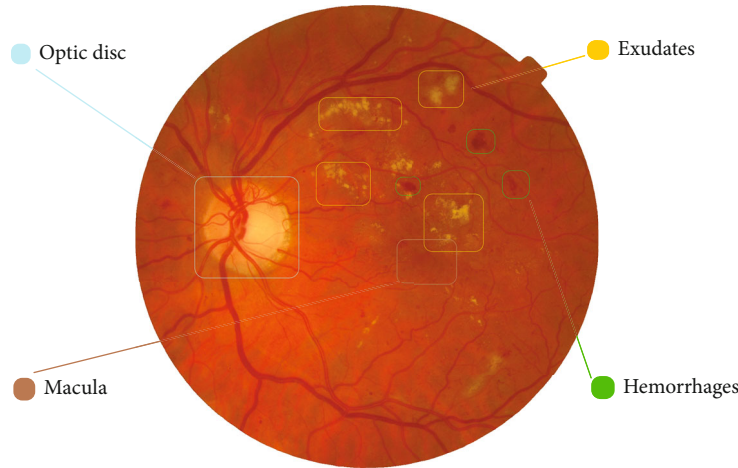
FIGURE 1: EX, HM, optic disc (OD), and macula in the DR retina.

improvement, classification, segmentation, and detection of medical images and related to the images and taking vital precautions. Moreover, Limwattanayingyong et al. showed that DL was more successful when they compared sight-threatening DR (STDR) screening with educated human grading and DL grading [31].

When the studies about DR classification in the literature were examined in detail, each study performed a preprocessing stage before training the network with CNN. The reason for this is that the lesions do not have a certain shape or form and are scattered in the image. This causes classification errors by reducing the clarity of the lesions in the image. These preprocessing phases were generally traditional image processing methods. Also, each study focused on operations for a particular dataset, and different methods were used for each dataset. This is because the grading system of each dataset is different. In this study, we proposed the 2-stage method that detecting independent from the dataset and classifying diabetic retinopathy lesions, completely based on deep learning. In the first stage, we created a pool of selected DR datasets and trained with Faster RCNN. We automatically determined the lesion region of interests in the images without any special process for the images in different DR datasets and prepared a pretrained model for the classification process, which is the second stage of the work. We completed the classification process by training images with the attention mechanism we added to pretrained ImageNet models.

In the second part of the work, literature research was made, and DR features, related studies, and results were mentioned. In the third chapter, features of the proposed method used datasets, and DL methods used were mentioned. In the fourth chapter, the results obtained with the proposed method and the comparison of the results in the literature were mentioned. In the fifth and last section, information was given about the success, effects, and future works of the method.

2. Literature Review

2.1. Diabetic Retinopathy Datasets. There are many datasets belonging to DR in open access. Some of these are MESSIDOR [32], DIARETDB [33], IDRiD [34], and Kaggle 2015

FIGURE 2: Confusion matrix.

DR Competition Dataset [35]. These datasets has been reviewed and graded by ophthalmologists. Each dataset can be used in a different grading system. For example, DR levels were graded from 0 to 4 in Kaggle, while in MESSIDOR, they were graded from 0 to 3. The MESSIDOR dataset contains 1200 images classified into 4 levels [36]. MESSIDOR was published in 2008 by Criann [37].

DIARETDB consists of 219 retinal images containing 25 healthy and 194 with DR symptoms. Images were classified as exudate (soft and hard), spots (red), and bleeding. The detected lesions were expressed in 5 different degrees with 0.25 intervals between 0 and 1. Kaggle dataset images were shared with an award-winning DR determination contest. Approximately, 90,000 right and left eye retinal images were reserved for the test of approximately 40% and 60% of the training set. Images were graded in five different classes according to the ETDRS [38] grading method. IDRiD is a dataset with DR lesions created in India. The dataset presented for ME detection classified DR in five levels according to the ETDRS grading method. The dataset contains 516 images (413 training sets, 103 test sets) [39].

2.2. Diabetic Retinopathy Symptoms. Microaneurysms (MA): these are deformations of the blood vessel walls of 1-3 pixels in images [40, 41].

Bleeding/hemorrhages (HM): bleeding/hemorrhages is a blood leaking from damaged capillaries [40, 42].

TABLE 1: Studies using MESSIDOR dataset and results.

Authors	Training type	Method	Process type	ACC	AUC	SEN
Zhang et al. [49]	ZFNet	TL	Optic disc localization	99.9		—
Alghamdi et al. [50]	CNN	EE	Optic disc localization	99.2	—	—
Xu et al. [51]	CNN	TL	Optic disc localization	99.4	—	—
Abràmoff et al. [52]	CNN	EE	Lesion detection	—	—	100
Grinsven et al. [53]	CNN	EE	Lesion detection	—	97.9	93.1
Gulshan et al. [54]	CNN	TL	Fundus classification	—	99.0	87.0
Costa and Campilho [55]	SURF + CNN	EE	Fundus classification	—	90.0	—
Gargeya and Leng [56]	CNN	EE	Fundus classification	—	94.0	—
Wang et al. [57]	Zoom	EE	Fundus classification	91.1	95.7	—
Chen et al. [58]	SI2DRNet	EE	Fundus classification	91.2	96.5	—

TABLE 2: Studies using Kaggle dataset and results.

Authors	Training type	Method	Process type	ACC	AUC	SEN
Grinsven et al. [53]	CNN	EE	Lesion detection	-	91.7	84.8
Mansour [59]	AlexNet + SVM	TL	Fundus classification	97.9	-	100
Quellec et al. [60]	CNN	EE	Fundus classification	-	95.5	-
Colas et al. [61]	CNN	EE	Fundus classification	-	94.6	96.2
Pratt et al. [62]	CNN	EE	Fundus classification	75.0	-	95.0
Jinfeng et al. [63]	CNN	TL	Fundus classification	80.3	-	-
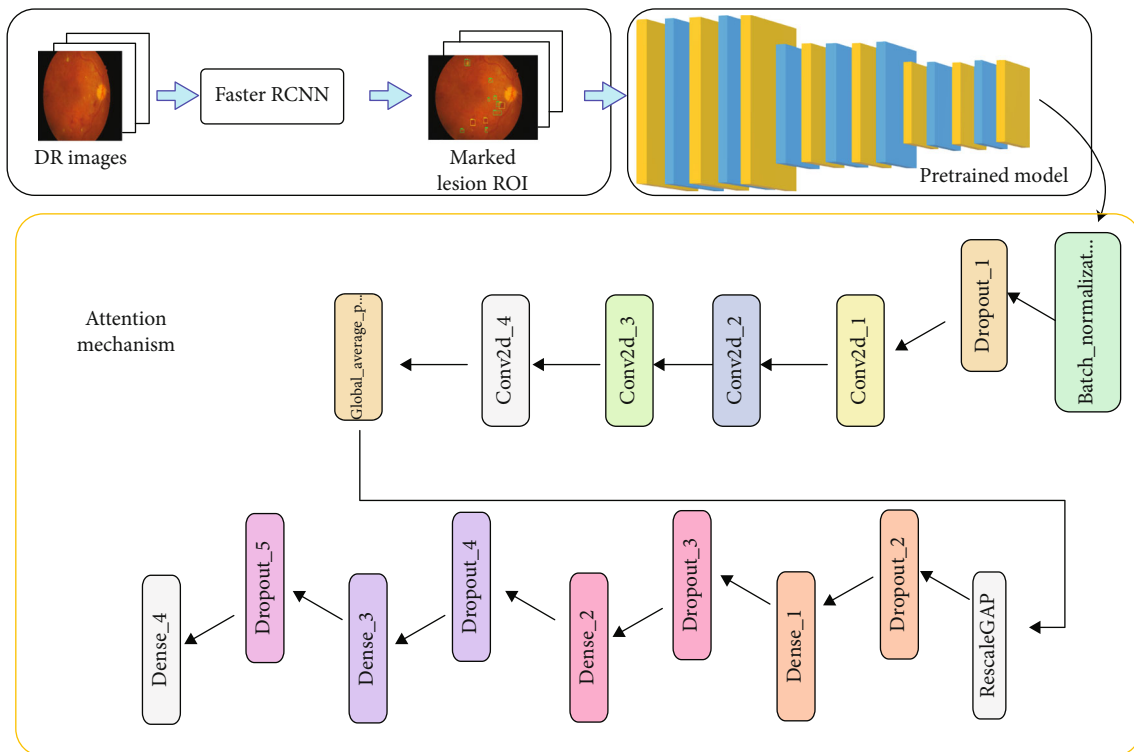
FIGURE 3: Developed model structure.

(a)

(b)

(c)

(d)
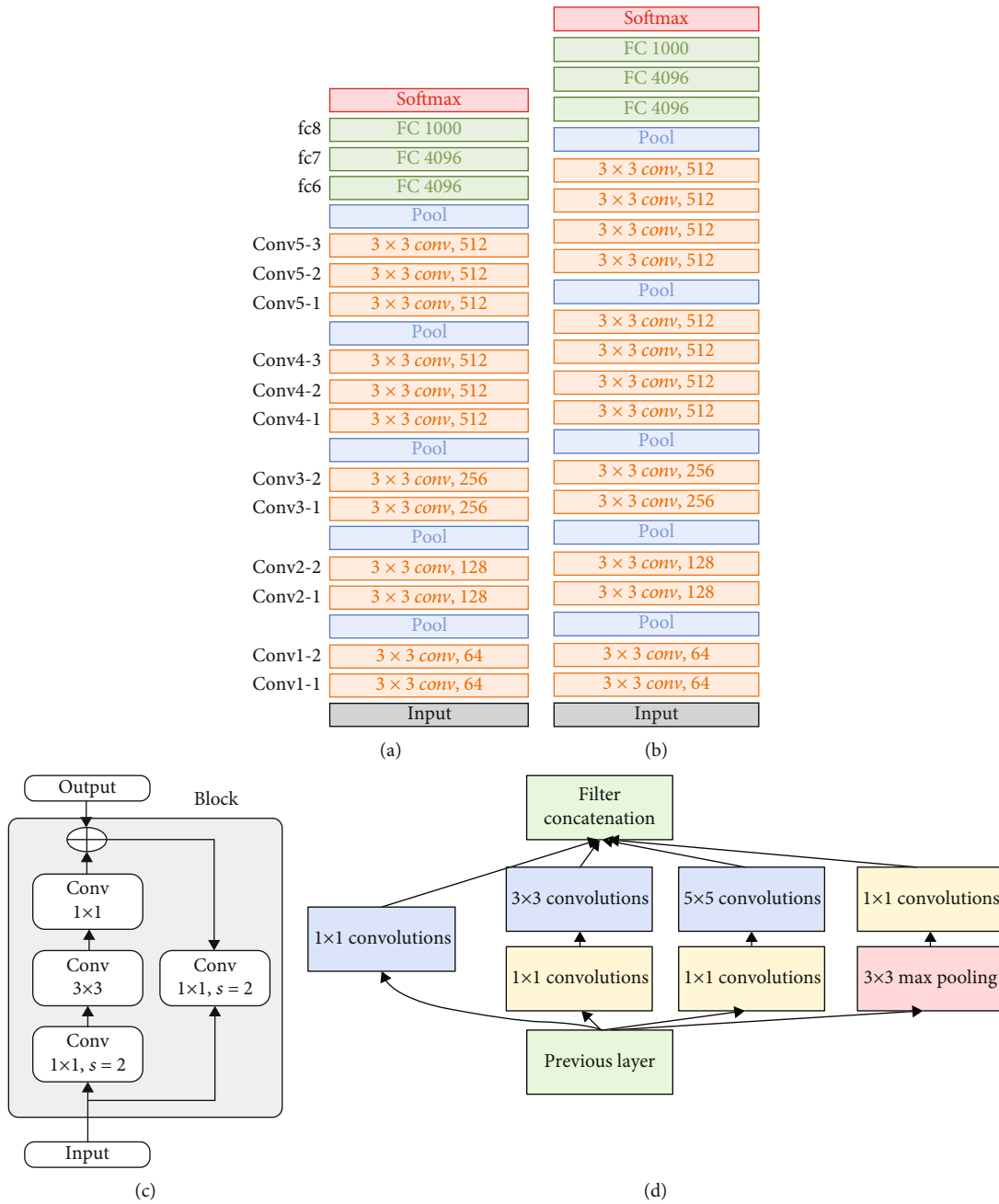
FIGURE 4: Continued.
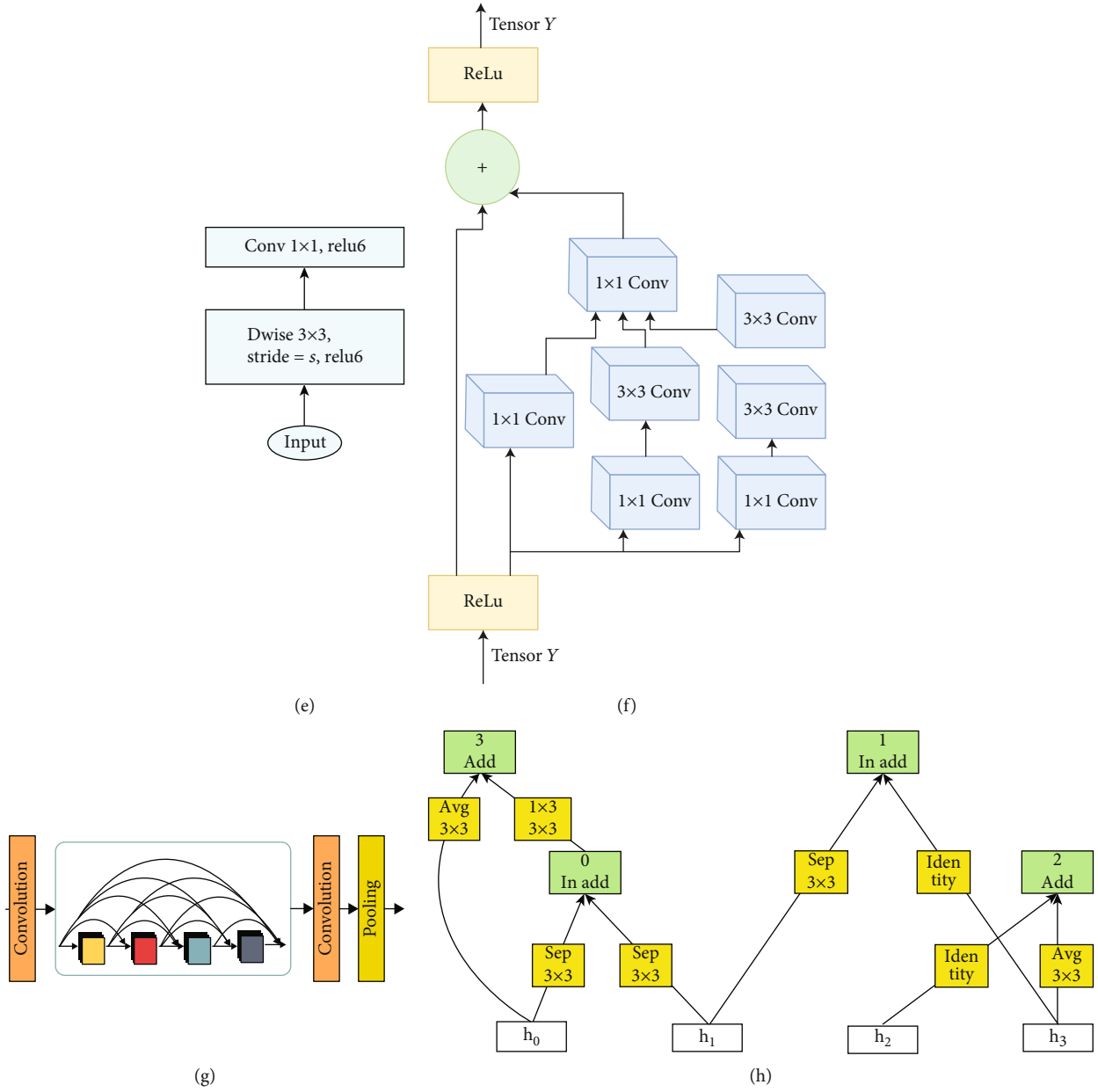
FIGURE 4: Pretrained models: (a) VGG16, (b) VGG19, (c) ResNet, (d) Inception, (e) MobileNet, (f) InceptionResNet, (g) DenseNet, and (h) NasNet.

Exudates/exudates (EX): when blood leaks more through capillaries, it causes exudates that are usually yellow in the retina [43].

*Macular edema (*ME): it occurs when there is leakage from the vessels around the macula [44].

Neovascularization (NV): it occurs when veins grow into the vitreous [45].

Figure 1 shows the EX, HM, optic disc (OD), and macula in the DR retina. The OD is the reference point for DR detection [45–47].

2.3. Performance Metrics. The confusion matrix in Figure 2 shows the predicted number of outcomes for 2 classes (0 and 1). Accordingly, when the classification value is 1 and the obtained value is 1 then true positive (TP); else then false

negative (FN) is obtained. When the classification value is 0 and the obtained value is 0 then true negative (TN); else then false positive (FP) is obtained.

Accordingly, performance metrics can be calculated with the following equations:

$$\text{Sensitivity} = \text{TP Rate (TPR)} = \frac{\text{TP}}{\text{FN} + \text{TP}}, \tag{1}$$

$$\text{Specificity (SPE)} = \frac{\text{TN}}{\text{FP} + \text{TN}}, \tag{2}$$

$$\text{Accuracy (ACC)} = \frac{\text{TN} + \text{TP}}{\text{FP} + \text{FN} + \text{TP} + \text{TN}}, \tag{3}$$

$$\text{FP Rate (FPR)} = 1 - \text{SPE}, \tag{4}$$
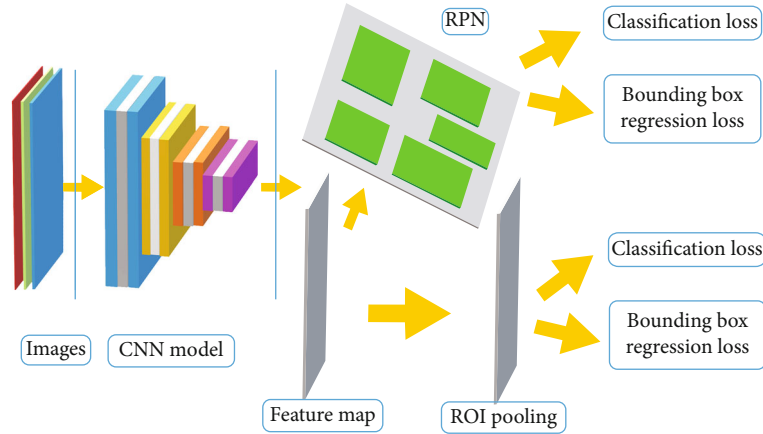
FIGURE 5: Faster RCNN architecture.

TABLE 3: The number of images in datasets and the number of training and test images used for detection and classification.

Dataset	Total images	Number of images for detection stage training	Number of images for detection stage testing	Number of images for classification stage training	Number of images for classification stage testing
Kaggle	80,000	100	4900	3920	980
MESSIDOR	1200	100	1100	880	220
IDRiD	516	100	-	-	-
DIARETDB	219	100	-	-	-

AUC (area under curve) is the area under the receiver operator characteristics (ROC) curve obtained with the change rates of FPR and TPR.

3. Related Works

There have been 747 studies on about DR in the literature [48]. In this section, studies on DR detection with deep learning are examined. Some of the studies created their own CNN models and used end-to-end learning (EE), while others used transfer learning (TL) using pretrained models available on ImageNet. In the studies, optic disc localization, lesion detection, and fundus classification procedures were performed on the DR images. Most of the studies used the MESSIDOR dataset. In end-to-end training, there are studies that create their own special models such as Zoom, ZFNet, and SI2DRNet.

The authors in [49] developed the ZFNet based on the Faster R-CNN in their work on the localization of the optical disc using a Hessian matrix. This study was conducted using the MESSIDOR dataset. Alghamdi et al. [50] first classified the images as OD or non-OD with the CNN they developed. Detected OD locations were classified by the second CNN module as normal, suspect, or abnormal. The MESSIDOR dataset was used in this study. In [51], the authors made changes before the last FC layer of the VGG model to find the OD, thresholding the probability map and obtaining the center of gravity of the pixels. This study was conducted using the MESSI-DOR dataset. The authors in [52] developed a controlled CNN model to classify the ME lesion type. This study
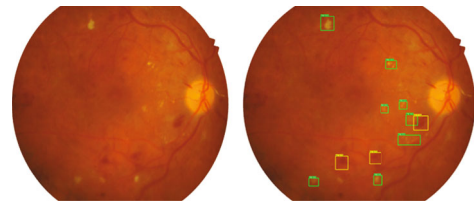
FIGURE 6: DR image, whose lesions are detected automatically with the trained model.

was conducted using the MESSIDOR dataset. In [53], HM is detected, and a 41-pixel square image containing HM was extracted from the original image. The resulting image was classified and labeled according to the number of HM removed. It was then given to the CNN network for training. The method was tested on a Kaggle and MESSIDOR datasets using a 10-layer CNN model. The authors in [54] used TL to determine DR in 1748 samples from the MESSIDOR dataset and DR in 9963 samples from the EyePACS dataset. Each image was graded 3 to 7 times by ophthalmologists. In [55], they created a CNN model by extracting rare local features with the structure they call Bag of Visual Words (BoVW) and Speed-Up Robust Properties (SURF). This study was conducted using the MESSIDOR dataset. Gargeya and Leng [56] proposed a CNN for DR detection by modifying ResNet. They evaluated the method with MESSIDOR. The authors of [57] proposed a pretrained CNN model that includes the attention network and crop network to detect suspicious patch sites called Zoom for DR detection. The management was developed using the MESSIDOR
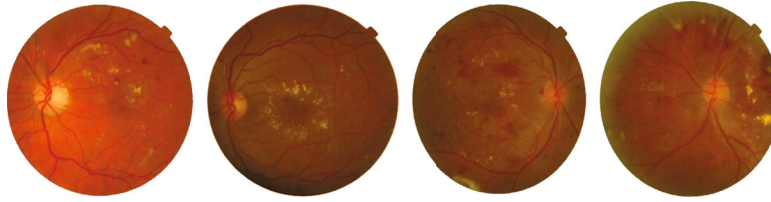
FIGURE 7: Proliterative DR images.

dataset. The authors in [58] created SI2DRNet-v1 by scaling the kernel size from 3×3 to 5×5 after each pooling layer in CNN. MESSIDOR was used in the model. The author in [59] developed a method for localizing blood vessels and a pretreatment for bound component analysis. Linear separation analysis was then used to reduce dimensionality. SVM was used for classification in this method. Kaggle dataset was used in this study. Quellec et al. [60] developed a CNN model to detect DR lesions. Heat maps created by this method were not optimized for diagnosis. In this study, Kaggle dataset was used. The authors of [61], proposed a method for EX detection using the LeNet model. They dismissed the EX zones and gave them input to the LeNet network for training. They made data replication before the training. The work was developed using the Kaggle dataset. In [62], the authors dealt with overfitting and skewed datasets in DR detection. They used data amplification to train the CNN model, which consists of 13 layers. Kaggle dataset was used in this study. In the work of Jinfeng et al.'s [63], an ensemble technique and two deep CNN models were proposed to detect all stages of DR using balanced and unbalanced datasets. First, they created 3 sub-datasets by dividing the Kaggle dataset into 3 parts. In the first model, they trained 3 datasets separately with DenseNet-121 and ensembled their results. In the second model, they trained 3 dataset separately with DenseNet-121, ResNet50 and Inception-V3, and ensembled their results. Then, the models were compared with each other.

When examined Table 1, the highest SEN value among the studies was 100, and Abramoff et al. have achieved. With the highest AUC of 99.0, Gulshan et al. have achieved. The highest ACC value of 99.4 was obtained by Xu et al. that have achieved.

When Table 2 was examined, the highest SEN and ACC values were 100 and 97.9, respectively, Mansour; with the AUC value of 95.5, Quellec et al. have achieved.

4. Materials and Methods

Based on the abovementioned shortcomings, a 2-stage method was proposed where all types of DR datasets could be trained using DL completely without preprocessing in traditional ways. If it is explained in more detail, since the use of CNN directly to classify DR is insufficient, the lesions should be clarified by preprocessing. In order to clarify the lesions, the region of interests(ROIs) of the lesion must be determined first. These regions can be made clear by using regional CNN with DL. As the regional CNN only detects objects, a CNN structure is needed for classification. For
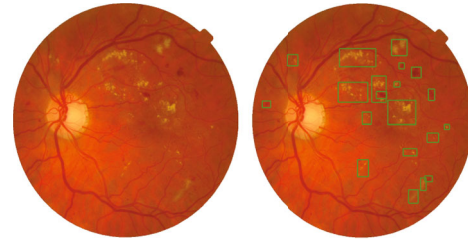
FIGURE 8: DR image with marked lesion region of interest.

these reasons, Faster RCNN and CNN were used together, and a 2-stage method was developed. The first stage of the 2-stage method is the automatic detection of lesions and marking of the lesion ROIs, and the second stage is the classification of marked images with a model created by transfer learning and attention mechanism [64] (Figure 3).

4.1. Used DL Methods. CNN has a structure that learns these properties by determining the image properties. CNN consists of certain layers. The convolution layer (conv), as evident from its name, performs a filter operation by convolution of the input image with the kernel matrix. This layer reveals the details in the image. Pooling layer pools the input image with one of the maximum (max pool) or global average pooling (global avg pool-GAP) methods, resulting in an image smaller than the image size. The aim is to delete unnecessary details and make learning easier. The fully connected (FC/Dense) layer helps the classification process by image features at the end of the network. In this study, VGG [65], DenseNet [66], ResNet [67], Inception [68], NasNet [69], MobileNet [70], and InceptionResNet [71], which are pretrainig models in ImageNet, were used in order to make faster training (Figure 4).

Regional training in CNN is needed to focus on specific objects in the image and to identify and segment them. RCNN structures have been developed to perform these operations. In simple terms, RCNN returns the box corridors of the regions detected in the image and the classification results. The first developed RCNN [72] creates weak candidate regions, while Fast R-CNN [73] feeds an input image directly to the CNN and reshapes it to be passed to the FC layer by ROI pooling. Faster R-CNN [74] uses region proposal network (RPN) instead of the selective search algorithm, unlike Fast R-CNN (Figure 5).

5. Results and Discussion

5.1. Used Datasets. In the proposed 2-stage method, a total of 6400 image data were used, including 1200 from MESSIDOR,

ROC curve for Non-DR

NDR (AUC = 100.00)

(a)

ROC curve for proliterative DR

PDR (AUC = 100.00)

(b)

ROC curve for Non-DR

PDR (AUC = 99.98)

(c)

ROC curve for proliterative DR

PDR (AUC = 99.97)

(d)
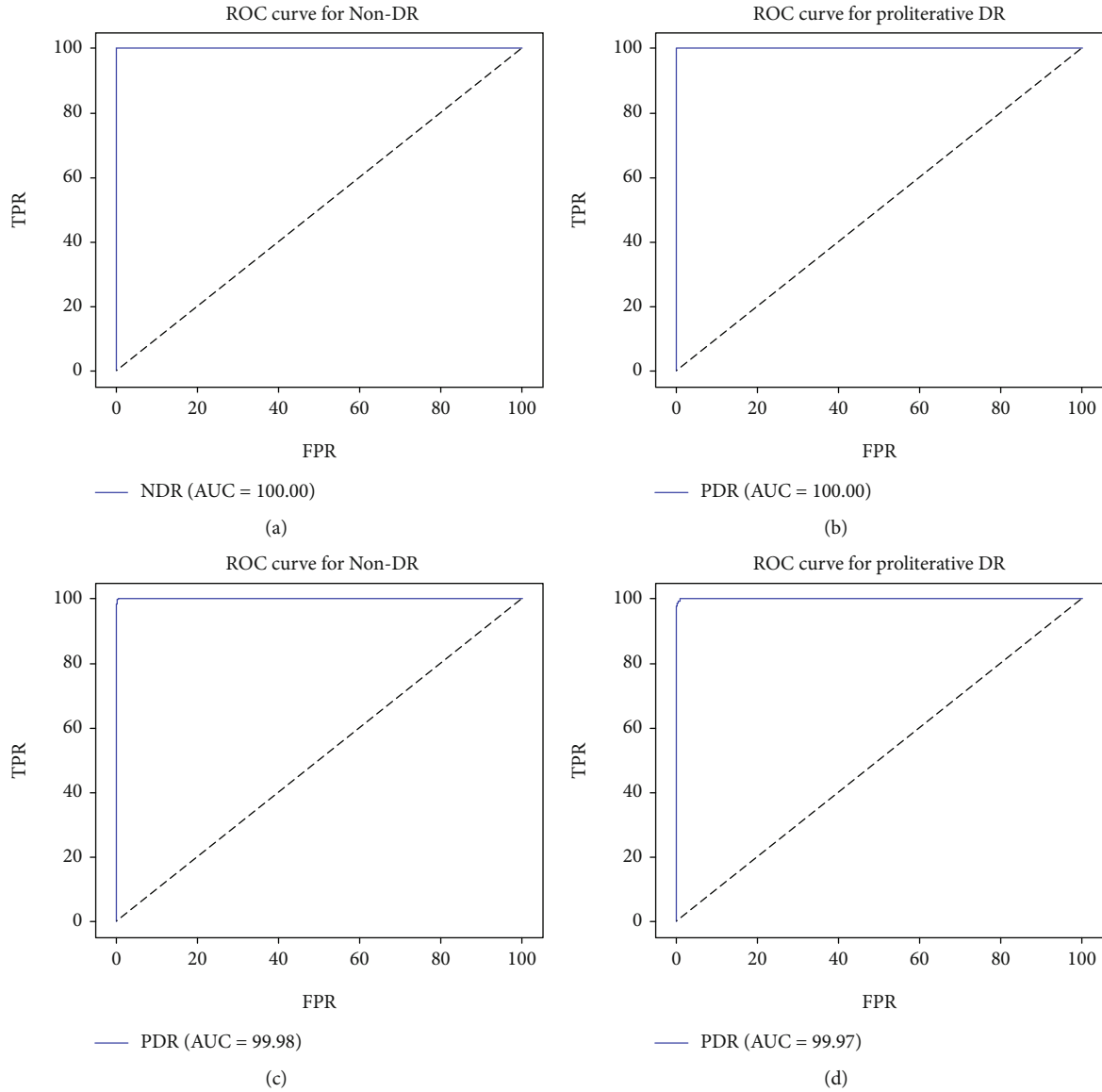
Figure 9: AUC prediction values in the ROC curve for the model trained with VGG16 in the second stage for (a) MESSIDOR non-DR, (b) MESSIDOR proliterative DR, (c) Kaggle non-DR, and (d) Kaggle proliterative DR.

5000 from Kaggle, and 100 from DIARETDB and IDRiD datasets. In the first stage, the dataset was divided into 400 training and 6000 tests to determine DR lesion ROIs. In the second stage, the marked 6000 data used for testing in the first stage were used. In the first stage, MESSIDOR, Kaggle, DIARETDB, and IDRiD datasets were used together to automatically detect lesions in different datasets. Since MESSIDOR and Kaggle datasets were used in the second phase, the test data of the first phase were used from these datasets. The training, test, and validation set of the data used in the two DL methods were given in detail in the relevant sections. Table 3 shows the number of images in the datasets used in the proposed method and the number of training and test images used for each stage.

5.2. Detection of Lesions with Region-Based CNN. In this stage, EX and HM lesion ROIs on DR datasets were determined by training with Faster RCNN. For Faster RCNN

Table 4: Results obtained by using MESSIDOR dataset and different pretrained models on the proposed method.

Model	TP	FN	TN	FP	ACC	AUC	SEN
VGG16	220	0	660	0	100	100	100
VGG19	220	0	660	0	100	100	100
DenseNet201	217	3	650	10	98.5	100	98.5
DenseNet121	195	25	658	2	96.9	97.6	88.6
DenseNet169	192	28	647	13	95.3	91.3	87.2
MOBILENET	192	28	574	86	87.0	94.5	87.2
NASNet	190	30	570	90	86.3	96.5	86.4
InceptionV3	200	20	593	67	90.1	94.2	90.1
InceptionResNetV2	192	28	574	86	87.0	87.0	87.2
Resnet50	186	34	560	100	84.7	89.8	84.5

training, a total of 400 data including EX and HM lesions from MESSIDOR, Kaggle, DIARETDB, and IDRiD datasets were selected randomly and labeled as EX and HM. 1100 remaining data from MESSIDOR and 4900 remaining data from Kaggle were used for the test of 6000 data in total. 80 of the 400 data used for training were used for validation. The purpose of using all datasets together in training is to diversify training and to automatically detect lesions for any dataset related to DR. With the trained model in the first step, the lesion ROIs were predicted in 6000 data as EX or HM and marked on the images as in Figure 6.

The marked images obtained in the first stage will be classified in the second stage by adding the attention layer to the pretrained ImageNet models. In the proposed model, the lesion ROIs were made clear so that the attention mechanism can work more efficiently.

When Figure 7 is analyzed, some images of proliferative DR are EX-weighted, and some are HM-weighted; some have only EXs while some have only HMs. With this information, it is seen that when grading DR, the density of the lesions is taken into account, not the type. Therefore, the ROIs in the lesion were displayed in one color, and the training phase was started as shown in Figure 8.

5.3. Classification of Detected Lesions. In this stage, the lesion ROIs detected in the DR images were classified by adding the mechanism of attention to the pretrained ImageNet CNN models. In this section, MESSIDOR and Kaggle datasets, which were used for testing at the first stage and marked on the image of the ROIs of the lesion, were used for DR classification. By ophthalmologists, the MESSIDOR dataset was divided into 4 classes (0-3) and the Kaggle dataset into 5 classes (0-4). The grading was not based on EX or HM lesions detected in the retina, but according to the intensity of any of the lesions in the retina, as seen in Figure 7. Therefore, lesion ROIs detected in the first stage are marked with the same color. During the training phase, the model was aimed to learn the lesion density by focusing on the marked lesion ROIs on the image and to give more accurate results. For this reason, the last layer of ImageNet models was changed with the mechanism of attention. The reason for the addition of the mechanism of attention is that the GAP added after pretrained models is simple because the prominent lesion ROIs are more important than others. Therefore, 4 convolution layers were added to unlock pixels in space before pooling. Then, the global weighted average pooling (GWAP) layer is created in which attention was multiplied by features and then divided by the sum of attention. Let $\{x_1, x_2, x_3, \cdots, x_n\}$ be a finite nonempty array and the weights of the x in this array be $\{w_1, w_2, w_3, \cdots, w_n\}$. In this case, the weighted average ($\bar{x}$) of the array is calculated as follows [75]:

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}. \tag{5}$$

Let the dimensions of a 3D image be expressed by x, y, and z, respectively. Let IF (x, y, z) expresses image features,

TABLE 5: Results obtained by using Kaggle dataset and different pretrained models on the proposed method.

Model	TP	FN	TN	FP	ACC	AUC	SEN
VGG16	971	9	3887	33	99.1	99.9	99.1
VGG19	969	11	3889	31	99.1	99.7	98.9
DenseNet201	919	61	3865	55	97.6	98.6	93.8
DenseNet121	952	28	3894	26	98.9	99.6	97.1
DenseNet169	836	144	3811	109	94.8	98.5	85.3
MOBILENET	684	296	3630	290	88.0	92.6	69.8
NASNet	508	472	3496	424	81.7	83.8	51.8
InceptionV3	674	306	3683	237	88.9	92.3	68.8
InceptionResNetV2	540	440	3726	194	87.0	84.3	55.1
Resnet50	202	778	3142	778	68.2	82.4	20.6

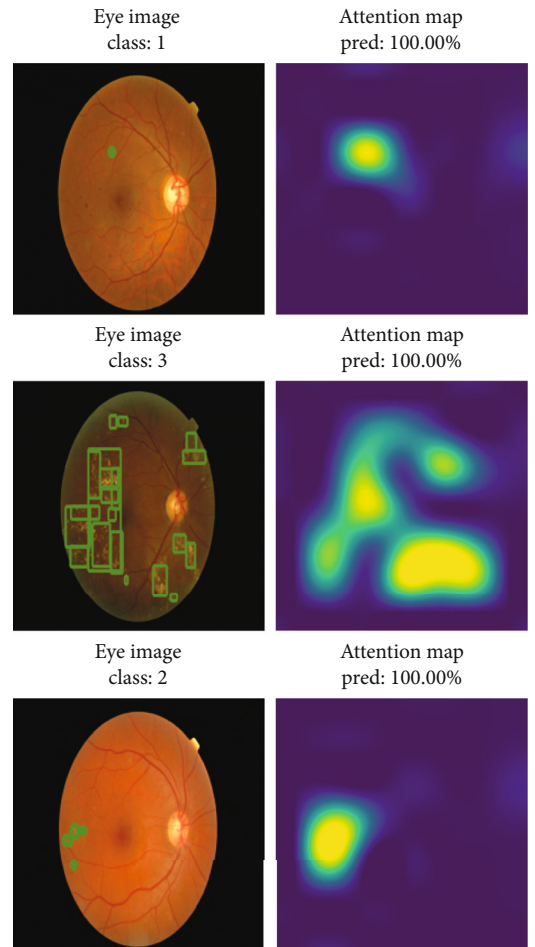FIGURE 10: The predicted results of the training using the VGG16 model of the proposed method and the MESSIDOR dataset and the attention map obtained in the attention layer.

and AF (x, y, z) expresses attention features. GWAP in image pixels is calculated according to Equation (5) as follows:

$$GWAP(x, y, z) = \frac{\sum_x \left(\sum_y AF(x, y, z) IF(x, y, z) \right)}{\sum_x \left(\sum_y AF(x, y, z) \right)}. \tag{6}$$

TABLE 6: Comparison of studies conducted with the MESSIDOR dataset and the proposed study.

Authors	Training type	Method	Process type	ACC	AUC	SEN
Gulshan et al. [54]	CNN	TL	Fundus classification	—	99.0	87.0
Costa and Campilho [55]	SURF + CNN	EE	Fundus classification	—	90.0	—
Gargeya and Leng [56]	CNN	EE	Fundus classification	—	94.0	—
Wang et al. [57]	Zoom	EE	Fundus classification	91.1	95.7	—
Chen et al. [58]	SI2DRNet	EE	Fundus classification	91.2	96.5	—
Ours	Faster RCNN + CNN	TL	Fundus classification	100	100	100

TABLE 7: Comparison of studies conducted with the Kaggle dataset and the proposed study.

Authors	Training type	Method	Process type	ACC	AUC	SEN
Mansour [59]	AlexNet + SVM	TL	Fundus classification	97.9	-	100
Quellec et al. [60]	CNN	EE	Fundus classification	-	95.5	-
Colas et al. [61]	CNN	EE	Fundus classification	-	94.6	96.2
Pratt et al. [62]	CNN	EE	Fundus classification	75.0	-	95.0
Jinfeng et al. [63]	CNN	TL	Fundus classification	80.3	-	-
Ours	Faster RCNN + CNN	TL	Fundus classification	99.1	99.9	99.1

The Lambda layer was then added to the rescaling results by pixel count to include the missing values in the attention model. Finally, the model was obtained by adding 4 dense layers. The resulting model's hyper parameters were finely tuned for each ImageNet model individually to achieve the best results.

For classification, a total of 6000 data were used 1100 from MESSIDOR and 4900 from Kaggle whose lesion ROIs were marked on the image as a result of the test in the first stage. Since DR classes for MESSIDOR and Kaggle are not the same, they were evaluated by training and testing separately for the two datasets. In MESSIDOR, 880 data were used for training, and 220 data were used for testing. 176 of the 880 data used for training were used for validation. In Kaggle, we used 3920 data for training and 980 data for testing. 784 of 3920 data used for training were used for validation.

Figure 9 shows the ROC curve and AUC values drawn with the classification prediction results for the non-DR (DR level 0) and proliferative DR (MESSIDOR DR level 3, Kaggle DR level 4) classes in the MESSIDOR and Kaggle datasets in the second stage. While calculating the ROC curve, the average of each FPR and TPR prediction result formed with 980 test data in Kaggle and 220 test data in MESSIDOR reserved for the classification test was taken. Detailed performance criteria obtained as a result of the prediction in the second stage were explained in Tables 4 and 5.

Table 4 shows the results obtained by using the method with different pretrained models in the MESSIDOR dataset. According to the results, VGG16 and VGG19 achieved 100% value in all metrics. DenseNet201 achieved 100% in AUC.

Table 5 shows the results obtained by using the method with different pretrained models in the Kaggle dataset. According to the results, the best result in the SEN value was obtained with VGG16 with 99.1%, and the best results in the AUC value with 99.9% in VGG16 and the best results in the ACC value with 99.1% were obtained in VGG16 and VGG19.

Figure 10 shows the prediction results of marked DR images selected randomly and in different classes, obtained with the test data of the trained model using VGG16 and MESSIDOR dataset in the proposed method. The figure also shows the attention map obtained in the attention layer.

In Table 6, the results obtained in the studies that made the MESSIDOR dataset fundus classification were compared with our proposed study. Accordingly, our method achieved a better result than other methods in all metrics.

In Table 7, the results obtained in studies developed with the Kaggle dataset were compared with our proposed study. Accordingly, our method achieved a better result than other methods with 99.1% in ACC and 99.9% AUC values. With a sensitivity value of 100%, Mansour achieved better results than our method.

6. Conclusions

Deep learning gives successful results in disease detection. In this work, a deep learning-based method has been proposed in which diabetic retinopathy lesions were detected automatically and independently of datasets, and the detected lesions were classified. In the first stage, lesions were detected with the regional CNN, and the images obtained in the second stage were classified using the transfer learning and attention mechanism for diabetic retinopathy grading. When the method tested in Kaggle and Messidor datasets was evaluated, 99.1% and 100% ACC, and 99.9% and 100% AUC were obtained, respectively. When the obtained results were compared with other results in the literature, it was seen that more successful results were obtained.

In future studies, the algorithms using the method will be developed to use minimum system resources.

Data Availability

Previously reported diabetic retinopathy datasets were used to support this study and are available at https://www.adcis.net/en/third-party/messidor/, https://www.kaggle.com/c/diabetic-retinopathy-detection/data, https://www.it.lut.fi/project/imageret/diaretdb0/, https://www.it.lut.fi/project/imageret/diaretdb1/, and https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid. These datasets are cited at relevant places within the text as references [32–35].

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

References

[1] R. S. Biyani and B. M. Patre, "Modified Alexnet architecture for classification of diabetic retinopathy images," *Biomedicine and Pharmacotherapy*, vol. 107, pp. 681–688, 2018.

[2] T. Shanthi and R. S. Sabeenian, "Modified Alexnet architecture for classification of diabetic retinopathy images," *Computers and Electrical Engineering*, vol. 76, pp. 56–64, 2019.

[3] T. J. Jebaseeli, C. A. D. Durai, and J. D. Peter, "Segmentation of retinal blood vessels from ophthalmologic diabetic retinopathy images," *Computers and Electrical Engineering*, vol. 73, pp. 245–258, 2019.

[4] A. Melville, R. Richardson, A. McIntosh et al., "Complications of diabetes: Screening for retinopathy and management of foot ulcers," *Quality in Health Care*, vol. 9, no. 2, pp. 137–141, 2000.

[5] R. Klein, B. E. Klein, S. E. Moss, M. D. Davis, and D. DeMets, "The Wisconsin epidemiologic study of diabetic retinopathy: II. Prevalence and rRisk of diabetic retinopathy when age at diagnosis is less than 30 years," *Archives of Ophthalmology*, vol. 102, no. 4, pp. 520–526, 1984.

[6] P. Porwal, S. Pachade, M. Kokare et al., "IDRiD: Diabetic Retinopathy - Segmentation and Grading Challenge," *Medical Image Analysis*, vol. 59, p. 101561, 2020.

[7] H. V. Doctor, S. Nkhana-Salimu, and M. Abdulsalam-Anibilowo, "Health facility delivery in sub-Saharan Africa: successes, challenges, and implications for the 2030 development agenda," *BMC Public Health*, vol. 18, no. 1, p. 765, 2018.

[8] 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.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, *Imagenet Classification with Deep Convolutional Neural Networks*, NIPS'12 Curran Associates Inc., USA, 2012.

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

[11] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, Boston, MA, USA, June 2015.

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

[13] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition, The 3rd International Conference on Learning Representations (ICLR2015)," 2015, https://arxiv.org/abs/1409.1556.

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

[15] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2017.

[16] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, June 2014.

[17] R. B. Girshick, "Fast R-CNN," CoRR, 2015, https://arxiv.org/abs/1504.08083.

[18] G. Gkioxari, R. Girshick, and J. Malik, "Contextual action recognition with r∗ cnn," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 1080–1088, 2013.

[19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks, CoRR," 2015, https://arxiv.org/abs/1506.01497.

[20] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," CoRR, 2017, https://arxiv.org/abs/1703.06870.

[21] W. Liu, D. Anguelov, D. Erhan et al., "SSD: single shot multibox detector," CoRR, 2015, https://arxiv.org/abs/1512.02325.

[22] O. Russakovsky, J. Deng, H. Su et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[23] E. Topol, *Deep medicine: how artificial intelligence can make healthcare human again*, Hachette, UK, 2019.

[24] L. M. Prevedello, S. S. Halabi, G. Shih et al., "Challenges related to artificial intelligence research in medical imaging and the importance of image analysis competitions," *Radiology: Artificial Intelligence*, vol. 1, no. 1, article e180031, 2019.

[25] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1289–1296, 2018.

[26] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015*, N. Navab, J. Hornegger, W. Wells, and A. Frangi, Eds., vol. 9351 of Lecture Notes in Computer Science, pp. 234–241, Springer, Cham, 2015.

[27] G. Haskins, U. Kruger, and P. Yan, "Deep learning in medical image registration: asurvey," 2019, https://arxiv.org/abs/1903.02026.

[28] Y. Xu, A. Hosny, R. Zeleznik et al., "Deep learning predicts lung cancer treatment response from serial medical imaging," *Clinical Cancer Research*, vol. 25, no. 11, pp. 3266–3275, 2019.

[29] P. Mobadersany, S. Yousefi, M. Amgad et al., "Predicting cancer outcomes from histology and genomics using convolutional networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, no. 13, pp. E2970–E2979, 2018.

[30] Y. Wang, S. Nazir, and M. Shafiq, "An overview on analyzing deep learning and transfer learning approaches for health monitoring," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 5552743, 10 pages, 2021.

[31] J. Limwattanayingyong, V. Nganthavee, K. Seresirikachorn et al., "Longitudinal screening for diabetic retinopathy in a nationwide screening program: comparing deep learning and Human graders," *Journal of Diabetes Research*, vol. 2020, Article ID 8839376, 8 pages, 2020.

[32] E. Decencière, X. Zhang, G. Cazuguel et al., "Feedback on a publicly distributed image database: the Messidor database," *Image Analysis & Stereology*, vol. 33, no. 3, pp. 231–234, 2014.

[33] V. V. Kamble and R. D. Kokate, "Automated diabetic retinopathy detection using radial basis function," *Procedia Computer Science*, vol. 167, pp. 799–808.

[34] P. Porwal, S. Pachade, R. Kamble et al., "Indian diyabetic retinopathy image dataset (IDRid)," in *IEEE Dataport*, 2018.

[35] Kaggle Diabetic Retinopathy Detection Competitionhttps://www.kaggle.com/c/diabetic-retinopathy-detection.

[36] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, and H. Kang, "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening," *Data*, vol. 3, no. 3, article 25, 2018.

[37] MESSIDOR datasethttps://www.adcis.net/en/third-party/messidor/.

[38] C. P. Wilkinson, Ferris FL 3rd, R. E. Klein et al., "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, 2003.

[39] P. Porwal, S. Pachade, R. Kamble et al., "Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research," *Data*, vol. 3, no. 3, p. 25, 2018.

[40] N. Salamat, M. M. S. Missen, and A. Rashid, "Diabetic retinopathy techniques in retinal images: a review," *Artificial Intelligence in Medicine*, vol. 97, pp. 168–188, 2019.

[41] 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.

[42] M. Zhang, *Blood Vessel Detection in Retinal Images and Its Application in Diabetic Retinopathy Screening, [Ph.D. thesis]*, Texas University, College Station, TX, USA, 2008.

[43] N. Asiri, M. Hussain, F. Al Adel, and N. Alzaidi, "Deep learning based computer-aided diagnosis systems for diabetic retinopathy: a survey," *Artificial Intelligence in Medicine*, vol. 99, p. 101701, 2019.

[44] M. Usman Akram, S. Khalid, A. Tariq, S. A. Khan, and F. Azam, "Detection and classification of retinal lesions for grading of diabetic retinopathy," *Computers in Biology and Medicine*, vol. 45, pp. 161–171, 2014.

[45] A. Sopharak, B. Uyyanonvara, S. Barman, and T. H. Williamson, "Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods," *Computerized Medical Imaging and Graphics*, vol. 32, no. 8, pp. 720–727, 2008.

[46] J. B. Jonas, G. C. Gusek, and G. O. H. Naumann, "Optic disk morphometry in high myopia," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 226, no. 6, pp. 587–590, 1988.

[47] G. D. Joshi, J. Sivaswamy, and S. R. Krishnadas, "Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment," *IEEE Transactions on Medical Imaging*, vol. 30, no. 6, pp. 1192–1205, 2011.

[48] A. Chu, D. Squirrell, A. M. Phillips, and E. Vaghefi, "Essentials of a robust deep learning system for diabetic retinopathy screening: a systematic literature review," *Journal of Ophthalmology*, vol. 2020, Article ID 8841927, 11 pages, 2020.

[49] D. Zhang, W. Zhu, H. Zhao, F. Shi, and X. Chen, "Automatic localization and segmentation of optical disk based on faster r-cnn and level set in fundus image," in *Medical Imaging 2018: Image Processing*, vol. 10574, p. 105741U, International Society for Optics and Photonics, 2018.

[50] H. S. Alghamdi, H. L. Tang, S. A. Waheeb, and T. Peto, "Automatic optic disc abnormality detection in Fundus images: a deep learning approach," in *Proceedings of the Ophthalmic Medical Image Analysis Third International Workshop*, Athens, Greece, October 2016.

[51] P. Xu, C. Wan, J. Cheng, D. Niu, and J. Liu, "Optic disc detection via deep learning in fundus images," in *Fetal, Infant and Ophthalmic Medical Image Analysis*, pp. 134–141, Springer, 2017.

[52] M. D. Abràmoff, Y. Lou, A. Erginay et al., "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning," *Investigative Opthalmology & Visual Science*, vol. 57, no. 13, pp. 5200–5206, 2016.

[53] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sanchez, "Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1273–1284, 2016.

[54] V. Gulshan, L. Peng, M. Coram et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, pp. 2402–2410, 2016.

[55] P. Costa and A. Campilho, "Convolutional bag of words for diabetic retinopathy detection from eye fundus images," *IPSJ Transactions on Computer Vision and Applications*, vol. 9, no. 1, p. 10, 2017.

[56] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017.

[57] Z. Wang, Y. Yin, J. Shi, W. Fang, H. Li, and X. Wang, *Zoom-in-Net: Deep Mining Lesions for Diabetic Retinopathy Detection*, Springer, International Conference on Medical Image Computing and Computer-Assisted Intervention, 2017.

[58] Y.-W. Chen, T.-Y. Wu, W.-H. Wong, and C.-Y. Lee, "Diabetic retinopathy detection based on deep convolutional neural networks," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1030–1034, Calgary, AB, Canada, April 2018.

[59] R. F. Mansour, "Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy," *Biomedical Engineering Letters*, vol. 8, no. 1, pp. 41–57, 2018.

[60] G. Quellec, K. Charrière, Y. Boudi, B. Cochener, and M. Lamard, "Deep image mining for diabetic retinopathy screening," *Medical Image Analysis*, vol. 39, pp. 178–193, 2017.

[61] E. Colas, A. Besse, A. Orgogozo, B. Schmauch, N. Meric, and E. Besse, "Deep learning approach for diabetic retinopathy screening," *Acta Ophthalmologica*, vol. 94, no. S256, 2016.

[62] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Procedia Computer Science*, vol. 90, pp. 200–205, 2016.

[63] G. Jinfeng, S. Qummar, Z. Junming, Y. Ruxian, 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.

[64] "InceptionV3 for Retinopathy GPU-HR," https://www.kaggle.com/kmader/inceptionv3-for-retinopathy-gpu-hr.

[65] K. Simonyan and A. Zisserman, "Very deep convolutional networksfor large-scale image recognition," pp. 1-2, 2014, https://arxiv.org/abs/1409.1556.

[66] Y. Sun, D. Liang, X. Wang, and X. Tang, "DeepID3: Face Recognition with Very Deep Neural Networks," 2015, https://arxiv.org/abs/1502.00873.

[67] M. Guillaumin and V. Ferrari, "Large-scale knowledge transfer for object localization in ImageNet," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3202–3209, June 2012.

[68] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," *AAAI*, vol. 31, no. 1, 2017.

[69] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," 2017, https://arxiv.org/abs/1707.07012.

[70] A. G. Howard, M. Zhu, B. Chen et al., "MobileNet: efficient convolutional neural networks for mobile applications," 2017, https://arxiv.org/abs/1704.04861.

[71] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI'17). AAAI Press*, pp. 4278–4284, 2017.

[72] G. Gkioxari, B. Hariharan, R. B. Girshick, and J. Malik, "R-CNNs for Pose Estimation and Action Detection," 2014, https://arxiv.org/abs/1406.5212.

[73] X. Zhang, J. Zou, K. He, and J. Sun, "Accelerating Very Deep Convolutional Networks for Classification and Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 10, pp. 1943–1955, 2016.

[74] A. Salvador, X. Giro-i-Nieto, F. Marques, and S. Satoh, "Faster R-CNN Features for Instance Search," in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 394–401, Las Vegas, NV, USA, 2015.

[75] D. F. Gatz and L. Smith, "The standard error of a weighted mean concentration–I. Bootstrapping vs other methods," *Atmospheric Environment*, vol. 29, no. 11, pp. 1185–1193, 1995.

Research Article

Automatic Segmentation of Left Ventricle in Echocardiography Based on YOLOv3 Model to Achieve Constraint and Positioning

Zhemin Zhuang ⓘ,[1,2] **Pengcheng Jin** ⓘ,[1] **Alex Noel Joseph Raj** ⓘ,[2] **Ye Yuan** ⓘ,[1] **and Shuxin Zhuang** ⓘ[1]

[1]*Department of Electronic Engineering, Shantou University, Shantou 515063, China*
[2]*Guangdong Provincial Key Laboratory of Digital Signal and Image Processing, Shantou University, Shantou 515063, China*

Correspondence should be addressed to Shuxin Zhuang; sxzhuang@stu.edu.cn

Cardiovascular disease (CVD) is the most common type of disease and has a high fatality rate in humans. Early diagnosis is critical for the prognosis of CVD. Before using myocardial tissue strain, strain rate, and other indicators to evaluate and analyze cardiac function, accurate segmentation of the left ventricle (LV) endocardium is vital for ensuring the accuracy of subsequent diagnosis. For accurate segmentation of the LV endocardium, this paper proposes the extraction of the LV region features based on the YOLOv3 model to locate the positions of the apex and bottom of the LV, as well as that of the LV region; thereafter, the subimages of the LV can be obtained, and based on the Markov random field (MRF) model, preliminary identification and binarization of the myocardium of the LV subimages can be realized. Finally, under the constraints of the three aforementioned positions of the LV, precise segmentation and extraction of the LV endocardium can be achieved using nonlinear least-squares curve fitting and edge approximation. The experiments show that the proposed segmentation evaluation indices of the method, including computation speed (fps), Dice, mean absolute distance (MAD), and Hausdorff distance (HD), can reach 2.1–2.25 fps, $93.57 \pm 1.97\%$, 2.57 ± 0.89 mm, and 6.68 ± 1.78 mm, respectively. This indicates that the suggested method has better segmentation accuracy and robustness than existing techniques.

1. Introduction

Cardiovascular diseases (CVDs) are one of the most common diseases affecting humans. "Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019," published in [1], shows that the incidence and mortality of CVD worldwide have been increasing since 1990 and that the mortality of CVD ranks first and is far higher than that of other diseases. Therefore, early detection and diagnosis of cardiac disease through various means is crucial for reducing the prevalence and mortality of CVD and improving the quality of life of patients [2].

Compared with X-ray coronary angiography, myocardial contrast echocardiography, computed tomography, and magnetic resonance imaging, the use of ultrasound for the screening and diagnosis of heart function and disease has great advantages. Using an ultrasound instrument, the heart and blood vessels, the movement of the ventricular wall, and the opening and closing of the valve can be observed dynamically in real time through flexible operation from multiple directions and angles. In addition, ultrasound has many advantages, such as safety and noninvasiveness, high diagnostic accuracy, and rapid inspection, and has become one of the most used and important examination methods for heart disease.

At present, the diagnosis of heart diseases based on ultrasound technology usually focuses on the analysis of the left ventricle (LV). The LV is responsible for blood supply to the body. Based on the changes in the LV, indicators such as LV end-diastolic volume, LV end-systolic volume, LV ejection fraction (EF), and LV stroke volume can be obtained. To obtain the indicators above, accurate positioning and segmentation of the LV on echocardiography are very important.

Clinically, the segmentation methods for LV ultrasound images can be classified into manual and automatic methods. The manual segmentation method requires the user to outline

the region of interest manually. Marking the position or contour of the LV manually is tedious and time-consuming, and there are subjective differences among different observers. The automatic segmentation method is superior to the manual segmentation method [3, 4]. Usually, the automatic segmentation method of LV ultrasound images includes two steps.

First, it is necessary to determine the position of the LV in the ultrasound images. Methods such as scale-invariant feature transformation [5] and histogram of oriented gradient [6] can be used to determine the position of the LV. However, the shape and appearance of the LV corresponding to different individuals are usually different, so these methods cannot accurately identify the position of the LV, and the segmentation accuracy of LV is also affected. Recently, the application of deep learning models for target detection and localization has attracted increasing attention [7, 8]. Compared with the faster R-CNN model [9] and the single-shot multibox detector model [10], the YOLOv3 model [11] has a higher detection speed and accuracy. Therefore, a method based on the YOLOv3 model is proposed herein for accurate positioning and segmentation of the LV.

Second, after the LV in the ultrasound image is accurately located, the LV can be segmented. Methods such as structured random forest based on machine learning [12] have been proposed for LV segmentation; however, such methods require manual selection of space features. Dong et al. [13] developed a deep fusion network and deformable model to achieve LV segmentation in 3-D echocardiography. Smistad et al. [14] successfully segmented the LV in two-dimensional ultrasound images based on the U-Net method. Oktay et al. [15] further extended the U-Net model to improve the accuracy of LV segmentation. However, these methods usually require significant morphological features or prior knowledge and have the disadvantages of poor real-time performance and high computing power requirements. Traditional image processing methods, such as a motion-based method (Kalman filter) [16], deformable models (BEAS, level-set) [17, 18], graph-based approach (graphcut) [19], active appearance model [20], and atlas-based method [21], have been proven to have high segmentation speed and robustness in heart image segmentation. Therefore, the YOLOv3 model and the traditional statistical shape model are combined in this study to achieve fast and accurate LV segmentation in ultrasound images.

Herein, an automatic segmentation method based on the YOLOv3 model to satisfy the relevant constraints and achieve appropriate positioning is proposed for accurate segmentation of the LV endocardium. The results of experiments conducted using the proposed method show that the segmentation evaluation indices, including the computation speed (fps), Dice, mean absolute distance (MAD), and Hausdorff distance (HD), can reach 2.1–2.25 fps, $93.57 \pm 1.97\%$, 2.57 ± 0.89 mm, and 6.68 ± 1.78 mm, respectively.

2. Method

To obtain clinical indicators such as EF, strain, and strain rate of the LV on echocardiography, accurate segmentation of the LV is crucial. In this study, the YOLOv3 model is first used to determine the three positions of the apex and bottom of the

LV, as well as the location of the LV region. Then, based on the Markov random field (MRF) model with the iterated conditional mode (ICM), preliminary identification and binarization of the myocardium of the LV subimages are performed, and under the three constraint points of the LV, the left and right parts of the myocardium in the LV subimages are located. Finally, when approaching the edge of the myocardium, the B-spline method is used to smooth the edge of the endocardium, and then, accurate segmentation and extraction of the LV endocardium are achieved. Speckle noise and artifacts in ultrasound images can lead to the loss of borders and edges during image segmentation; therefore, when approaching the LV endocardium, a morphological mask is applied to eliminate the interference from speckle noise and edge artifacts inside the LV cavity. Figure 1 presents the block diagram of the proposed technique.

3. Segmentation of LV Endocardium Based on YOLOv3 for Positioning and Restraint

3.1. LV Localization and Collection of Restraint Points Based on the YOLOv3 Model. There are large differences in the shape of the LV in different echocardiogram frames. In addition, due to the interference of the mitral valve, as well as the influence of noise, artifacts, and frame-to-frame drift, traditional methods cannot locate the LV position well or extract the endocardium accurately. Therefore, this study proposes to use the target detection model YOLOv3 to realize the positioning of the LV region and the three ventricular constraint points in echocardiography. From Figure 2, the YOLOv3 model consists of the following: a general feature extraction network based on the Darknet-53 network, a multibranch deep feature extraction network, and a multiscale target area bounding box detection network.

In Figure 2, for the general feature extraction network, the convolutional network (Conv), batch normalization (BN) layer, and linear activation function (Leaky ReLU) constitute Darknetconv2d BN Leaky (DBL), which extracts the general features of cardiac ultrasound images. The DBL is also the basic block of deep feature extraction networks. Concurrently, to solve problems such as the disappearance of gradients due to the deep network structure, DarkNet53 uses the jump structure to form Res_unit, Resblock_body, and Res_Module in multiple DBLs.

For the deep-level feature extraction network, YOLOv3 forms a multibranch network and a Concat layer through the route structure. Simultaneously, YOLOv3 uses a bilinear upsampling layer to expand the feature map to form three branch networks for locating target areas of three different scales; through these three branch networks, the feature matrix of the LV ultrasound image can be obtained. In practice, it is difficult to obtain enough labeled LV images, and to avoid overfitting, transfer learning is applied in this study to train the entire feature extraction network: first, load the weight parameters obtained based on the VOC dataset [22] and then fine-tune the weight parameters of the feature extraction network using the labeled heart dataset.

After the feature matrices of the LV ultrasound images of the heart are obtained, they are input into the detection
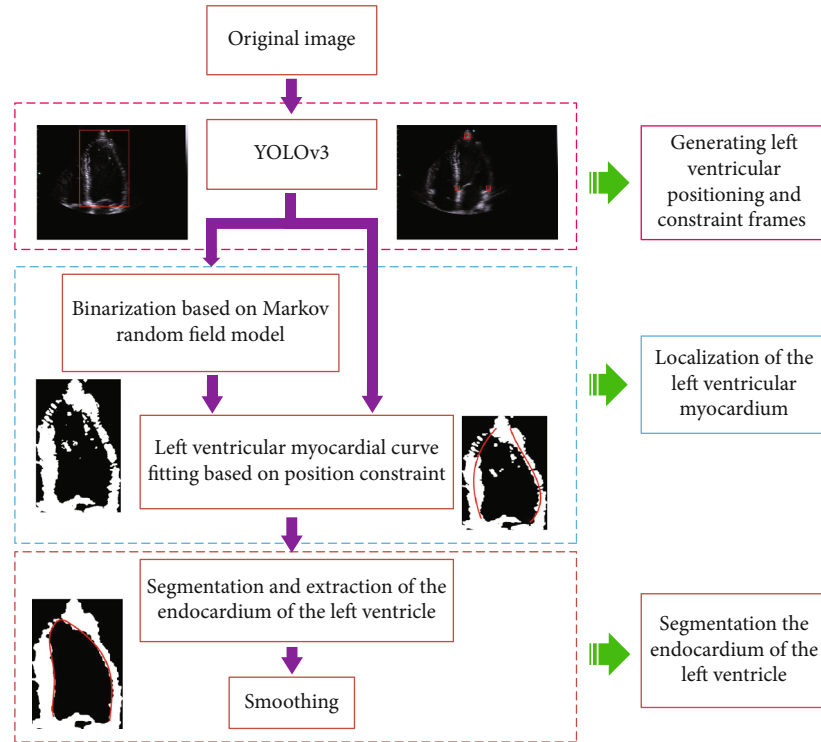
Figure 1: Block diagram of the proposed method.
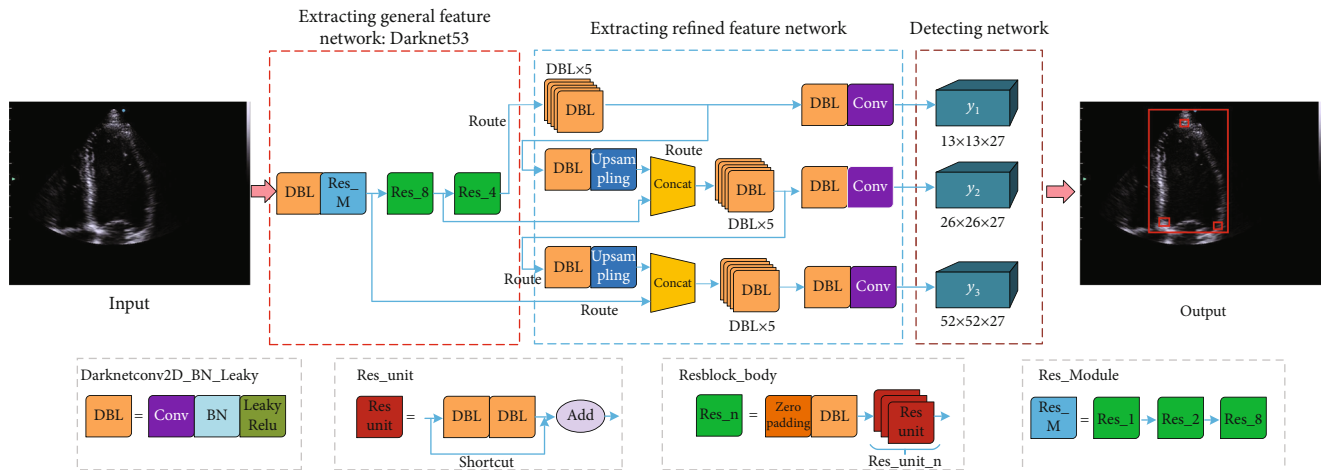
Figure 2: Detection of LV based on YOLOv3.

network to obtain the positioning matrices. The YOLOv3 model divides the original input images into three types of $S \times S$ grids (i.e., 13×13, 26×26, and 52×52) for positioning the target area; hence, three types of positioning matrices with different dimensions are obtained. As shown in Figure 2, the y_1 matrix corresponding to a 13×13 grid is used to detect a large target area and is used to locate the LV area in this study; the y_2 and y_3 matrices correspond to the 26×26 and 52×52 grids, respectively, which are used to locate three ventricular restraint points in this study.

Each grid corresponds to a $(B + O) \times$ anchors-dimensional positioning vector, where B is the bounding box of the target area, composed of $(b_x, b_y, b_w, b_h, b_c)$, corresponding to the center abscissa, ordinate, width, height, and confidence from the center of the target area, respectively, and O is the number of types of the target area. In this study, there are four types of targets: the LV region and three ventricular constraint points. anchors are the number of anchor frames in the positioning matrix; the number of anchor frames with three scales in this study is three.

The anchor box is used to describe the length and width of the target area in this study, and the relationship between the anchor box and bounding box is shown in Equation (1).

$$
\begin{aligned}
b_x &= \sigma(t_x) + C_x, \\
b_y &= \sigma(t_y) + C_y, \\
b_w &= P_w e^{t_w}, \\
b_h &= P_h e^{t_h},
\end{aligned}
\tag{1}
$$

where C_x, C_y, P_w, and P_h are the abscissa, ordinate, and the width and height of the upper left corner of the grid where the center point of the anchor frame is located, respectively; $\sigma(\bullet)$ is the sigmoid activation function; t_x and t_y are the abscissa and ordinate offsets of the center of the anchor frame; and t_w and t_h are the changes in the length and width of the anchor frame.

In this study, the target area in the training set is divided into nine anchor boxes using the K-means [23] clustering algorithm, and each anchor box is represented as (w, h). For these anchor boxes, three small anchor boxes $((0 \times 0), (11 \times 13),$ and $(11 \times 15))$ (i.e., the y_1 matrix in Figure 2) are used to locate the LV area: three medium anchor boxes $((13 \times 15), (14 \times 20),$ and $(15 \times 17))$, and three large anchor boxes $((16 \times 22), (110 \times 218),$ and $(146 \times 323))$ (i.e., the y_2 and y_3 matrices in Figure 2) are used for the positioning of three constraint points.

3.2. Extraction of Endocardium Based on Constraint Points. The three positions of the apex and bottom of the LV, as well as the positioning of the LV area, can be found by the YOLOv3 model mentioned above. Then, based on the MRF model, the binarization and preliminary identification of the LV myocardial region in the subimages can be performed. Under the constraints of the three position points of the apex and bottom of the LV, curve fitting was performed on the left and right myocardial parts in the LV subimages, and the edge of the endocardium was approximated to realize accurate segmentation of the LV endocardium, and the B-spline method was also employed to smooth the edge of the LV endocardium.

3.2.1. Binarization of LV Myocardium Based on MRF Model. Before the LV myocardial images are binarized, to reduce the influence of speckle and noise in echocardiograms, the echocardiograms are denoised on the premise of preserving the characteristics of the LV myocardium. First, the LV subimages are smoothed via 2-D adaptive Wiener noise-removal filtering [24], the local neighborhood size is set to (5×5), and then, the pixel-wise Wiener filter can be constructed using Equation (2).

$$
b(n_1, n_2) = \mu + \frac{\sigma^2 - \nu^2}{\sigma^2} (a(n_1, n_2) - \mu),
\tag{2}
$$

where μ and σ^2 are the local mean and variance around each pixel, respectively, and ν^2 is the variance of the noise. The Wiener filter adjusts itself to the local image variance, i.e., when the variance is large, a minor smoothing operation is performed by the Wiener filter whereas when variance is small, the Wiener filter performs a major smoothing.

The MRF model utilizes the correlation between the upper and lower adjacent pixels in the image; thus, the spatial connectivity and edge smoothness of the binarized region can be improved. Therefore, an MRF model based on the ICM algorithm was used in this study to binarize and initially identify the myocardial region.

Assume that X and Y are random fields on a two-dimensional plane, where $X = \{x_i, i = 1, 2, 3, \cdots, M \times N\}$ represents the input image and $Y = \{y_i, i = 1, 2, 3, \cdots, M \times N\}$ represents the labeling field, where M and N represent the rows and columns of the image, respectively. In this study, the K-means clustering method was used to obtain the initial marker field, and the category was set to 2.

Considering the input images as an MRF model, the image segmentation problem can be transformed into an optimization problem using the ICM algorithm. According to the Bayesian principle, the posterior probability distribution of MRF is as follows:

$$
P(X = x \mid Y = y) = \frac{P(Y = y \mid X = x)P(X = x)}{P(Y = y)},
\tag{3}
$$

$= x)$ is the prior probability of the label domain, and $P(Y = y \mid X = x)$ is the likelihood function.

When binarizing the LV images, the optimal labels can be obtained by maximizing the posterior probability of Equation (3).

$$
P(X = x \mid Y = y)_{MAP} = \mathrm{argmax}\{P(Y = y \mid X = x)P(X = x)\}.
\tag{4}
$$

The prior probability $P(X = x)$ in the MRF neighborhood system can be expressed using the Gibbs distribution function [25]. Then, based on the Gibbs distribution, the prior probability $P(X = x)$ of the marker field can be characterized as follows:

$$
P(X = x) = \frac{1}{Z} \exp\left[-\frac{E(x)}{T}\right],
\tag{5}
$$

where $Z = \sum_{x \in \Omega} \exp[-E(x)/T]$ is a normalized constant, $E(x) = \sum_{c \in S} V_c(x)$ is the energy function, $V_c(x)$ is the potential function, and T is the temperature parameter, which is usually set to 1 [26].

Similarly, the posterior probability $P(X = x \mid Y = y)$ can also be expressed by an energy function, as shown in Equation (6).

$$
P(X = x \mid Y = y) = \frac{1}{Z} \exp\left[-\frac{E(x \mid y)}{T}\right].
\tag{6}
$$

Substituting Equations (5) and (6) into Equation (4), and taking the logarithms on both sides of the equation simultaneously, the product form is transformed into a summation form, and the result is as follows:
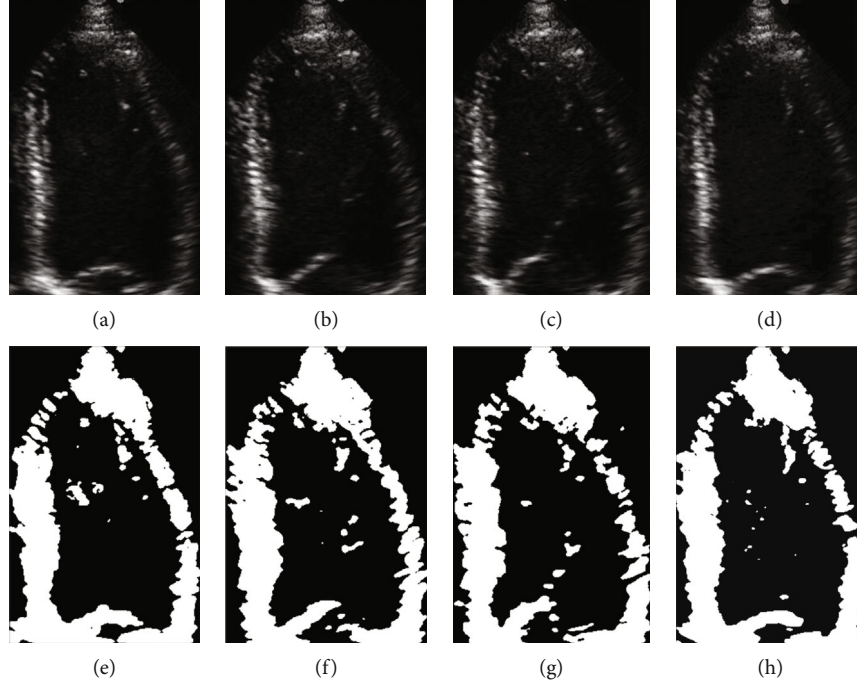
FIGURE 3: Binarization results of LV endocardium images based on MRF: (a–d) the original frames extracted at the equal interval from the same echocardiogram and (e–h) the corresponding binarization results.

$$E(x \mid y) = \mathrm{argmax}\{E(y \mid x) + E(x)\}, \qquad (7)$$

where $E(x \mid y)$ represents the minimized energy function, $E(y \mid x)$ is the likelihood function energy of pixel x, and $E(x)$ is the prior probability energy corresponding to pixel x. Therefore, the final energy relationship can be expressed as Equation (8).

$$E(x_F \mid y) = E(y \mid x) + E(x), \qquad (8)$$

where x_F is the final segmentation mark.

The ICM algorithm is used to optimize Equation (8), i.e., to minimize the energy function $E(x_F \mid y)$. Finally, the binarization results of LV myocardium images can be obtained and are shown in Figure 3. As shown in Figure 3, the LV myocardium can be clearly observed after the original LV images are binarized using the MRF model.

3.2.2. Segmentation and Extraction of LV Myocardium Based on Position Constraints.
After binarizing the original LV images based on the MRF model, the positioning curve of the LV myocardium will be fitted based on the position constraints. Firstly, divide the LV into the left and right regions and then use the nonlinear least squares (NLS) method to perform curve fitting on the two regions. Because only the LV endocardium is approximated in this study, three constraint points are used to limit and constrain the fitted curve.

In this study, a polynomial model based on the NLS method is employed to fit the left and right segments, respectively, as shown in Equation (9).

$$F(x) = a_1 x^m + a_2 x^{m-1} + \cdots + a_m x + a_{m+1}, \qquad (9)$$

where $a_1, a_2, \cdots, a_{m+1}$ represents the fitting coefficient of the polynomial, and m is the polynomial degree; in this study, the polynomial degree m is set to 3.

For a given set of coordinate points $\{(x_i, y_i): i = 1, 2, \cdots, n\}$, the polynomial fitting error equation can be written as Equation (10).

$$V = BX - L, \qquad (10)$$

where

$$
B = \begin{pmatrix} x_1^m & x_1^{m-1} & \cdots & x_1 \\ x_2^m & x_2^{m-1} & \cdots & x_2 \\ \vdots & \vdots & \vdots & \vdots \\ x_n^m & x_n^{m-1} & \cdots & x_n \end{pmatrix},
$$

$$
X = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{m+1} \end{pmatrix}, \qquad (11)
$$

$$
L = \begin{pmatrix} F(x_1) \\ F(x_2) \\ \vdots \\ F(x_n) \end{pmatrix}.
$$

(a)　　　　　(b)　　　　　(c)
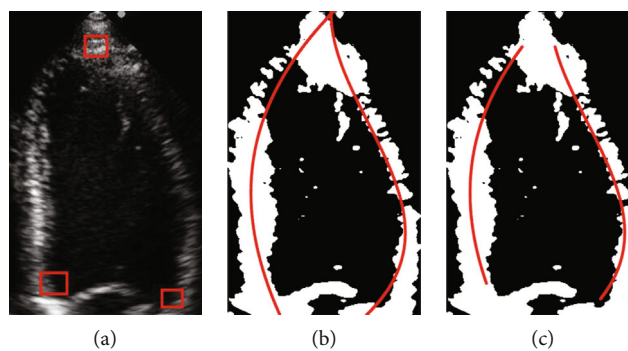
FIGURE 4: Fitting curve results of the LV subgraph in different sequences without and with constraints.

(a)　　　　　(b)

(c)　　　　　(d)

(e)
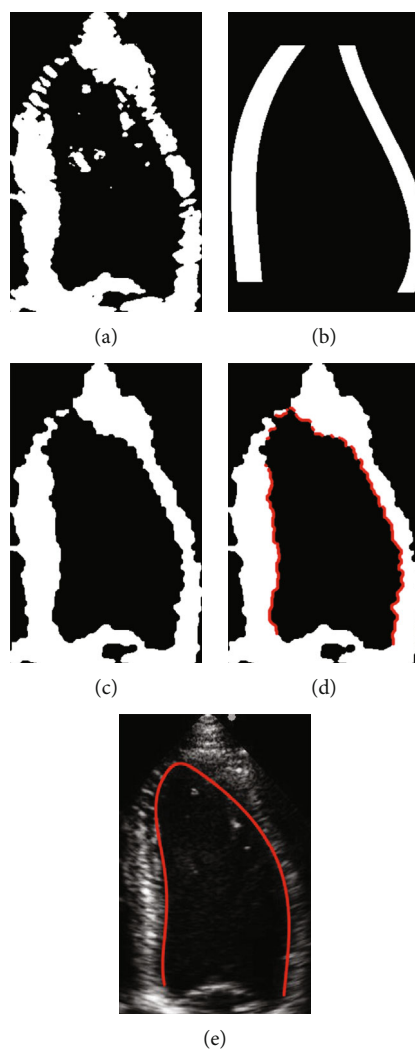
FIGURE 5: Extraction and segmentation of LV endocardium. (a) Result of binarization of the LV image using MRF model. (b) Morphological mask generated according to the fitted positioning curve. (c) Binary myocardial image obtained after mask processing. (d) Approximation result of the endocardium based on the three constraint points. (e) Result of smoothing the myocardium based on the B-spline method.

Based on the NLS method, the estimated value of X can be obtained as follows:

$$X = \left(B^T B\right)^{-1} B^T L. \tag{12}$$

Substituting the result obtained from Equation (12) into Equation (9), the LV myocardial positioning fitting curve can be obtained, as shown in Figure 4, where the red boxes in Figure 4(a) are the constraint points obtained by the YOLOv3 model. The obtained three constraint points are used for the constraint of the myocardial fitting curve. Figure 4(b) is the positioning fitting curve without restraint, and Figure 4(c) is the constrained positioning fitting curve.

As shown in Figure 4, under the three constraint points obtained based on the YOLOv3 model, the positioning curve of the myocardium can be accurately determined in the LV.

To mitigate the influence of speckle noise around the LV myocardium, the binary LV images obtained based on the MRF model are processed using the morphological masking method. After the initial positioning of the LV myocardium is achieved, the endocardium is approached based on the three constraint points, the edge of the endocardium is smoothed by the B-spline method [27], and the segmentation and extraction of the LV endocardium can be realized as shown in Figure 5.

4. Results

The cardiac ultrasound imaging data used in this study were provided by the Ultrasound Imaging Department of the First Affiliated Hospital of Medical College of Shantou University.

4.1. Evaluation Criteria. For target detection tasks, the average precision (AP) indicator [28] is commonly used to evaluate whether a model can detect a target class accurately. The AP is computed as the intersection of union (IOU) between the detection bounding box and the label bounding box. When the IOU of the detection bounding box and the label bounding box is greater than the set IOU threshold, it is considered that the model detects the target correctly. Subsequently, the AP value of the target class is calculated. In practice, the IOU threshold is usually set to 0.5, and the corresponding AP indicator is called AP50. For a model used to detect multiple target classes, the mean average precision (mAP) can comprehensively evaluate the performance of the model, i.e., compute the average value of the AP values of all target classes.

A precision-recall (*P-R*) curve [29] is shown with precision and recall as the vertical and horizontal axis, respectively. Also the size of the area under the *P-R* curve can comprehensively reflect the performance of a model for detecting the target.

AP can be expressed as

$$AP = \int_0^1 P(R) dR, \tag{13}$$

where P and R represents the precision and recall rates, respectively. The precision and recall in the *P-R* curve are

TABLE 1: Evaluation results of the YOLOv3-based LV and three bounding box positioning model using AP50.

	LV	Left_down	Right_down	Top
AP	100.00%	92.33%	95.44%	94.50%
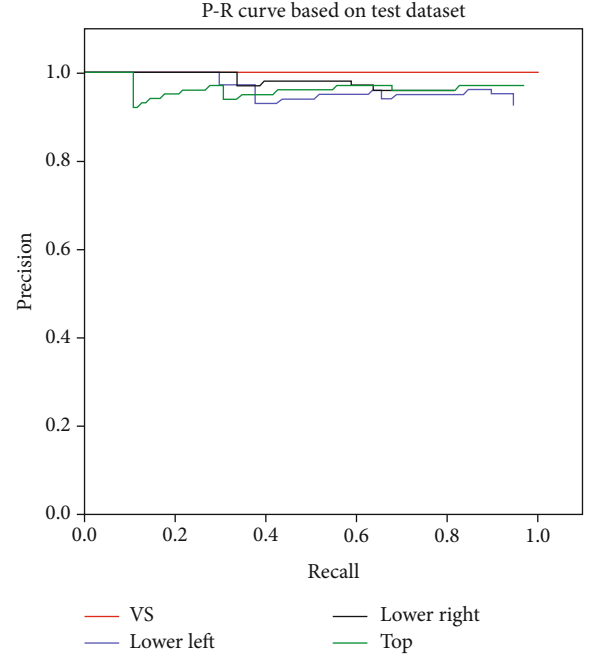
FIGURE 6: *P-R* curve of the LV identification and constraint box based on the proposed method in this paper. (VS: ventriculus sinister; Lower left: the constraint point in the lower left corner; Lower right: the constraint point in the lower right corner; and Top: the constraint point on the top of the myocardial wall).

calculated using Equations (14) and (15), respectively.

$$\text{precision} = \frac{TP}{TP + FP}, \tag{14}$$

$$\text{recall} = \frac{TP}{TP + FN}, \tag{15}$$

where TP, FP, and FN represent the true positive, the false positive and the false negative, respectively.

The Dice coefficient [30], MAD [31], and HD [32] parameters are used to evaluate the segmentation results of the LV endocardium:

$$\text{Dice}(S, G) = \frac{2\text{Area}(S \cap G)}{\text{Area}(S) + \text{Area}(G)},$$

$$\text{MAD}(A, B) = \frac{1}{2} \left\{ \frac{1}{m} \sum_{i=1}^{m} d(a_i, B) + \frac{1}{n} \sum_{j=1}^{n} d(b_i, A) \right\},$$

$$\text{HD}(A, B) = \max \left\{ \max_i \{ d(a_i, B) \}, \max_j \{ d(b_j, A) \} \right\}, \tag{16}$$

| | | | |
| (a) | (b) | (c) | (d) |

| | | | |
| (e) | (f) | (g) | (h) |

FIGURE 7: Binarization results of ultrasonic LV images using the Otsu method, K-means clustering method, and MRF model. The area enclosed by the blue line in (a) and (e) is the gold standard for the LV myocardium. (b, f) The binarization results obtained by using the Otsu method. (c, g) The binarization results obtained by using the K-means clustering algorithm. (d, h) The binarization results obtained by using the method proposed in this paper.

where S represents the myocardial area data obtained by different binarization methods, G is the gold standard data of the myocardial area, $A = \{a_1, a_2, \cdots, a_m\}$ is the endomyocardial edge data obtained by the method proposed in this paper, and $B = \{b_1, b_2, \cdots, b_n\}$ is the gold standard endomyocardial edge data.

4.2. LV and Restraint Point Positioning Model Based on YOLOv3.
Table 1 illustrates the performance of the YOLOv3-based LV and bounding box positioning model on the test dataset using AP50. From Table 1, all the AP50 values of the four target regions formed by the LV and the three bounding boxes are above 92%, and the mAP value reaches 95.57%, which ind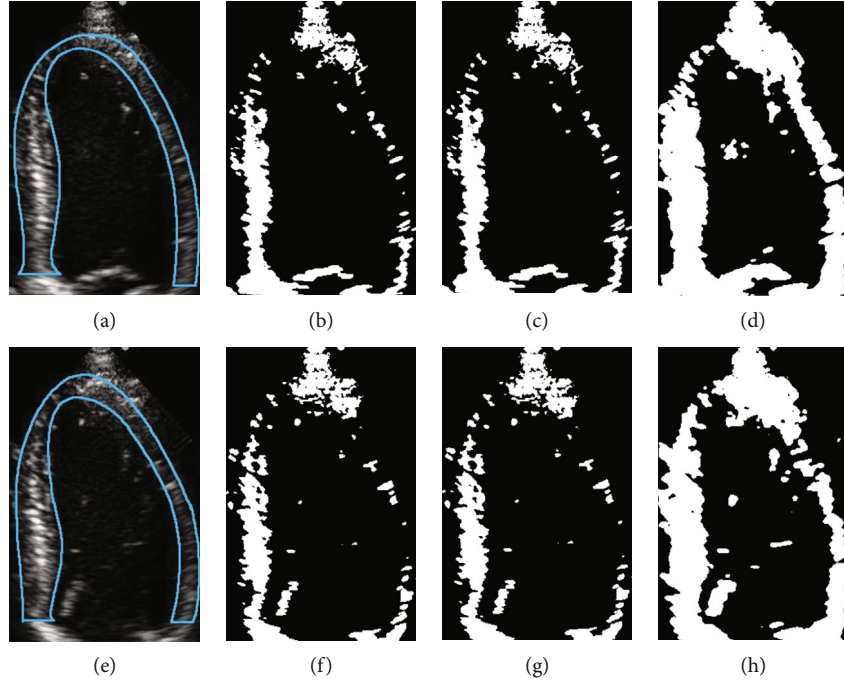icates that the model designed in this study can detect the LV and the three bounding box areas well and meet the requirements of LV myocardium segmentation.

The P-R curve, which can intuitively evaluate whether the model can detect a target class well, is drawn based on the precision-recall value pairs calculated from different confidence values when the model detects a target class. The value of the area enclosed by the P-R curve is the AP value.

The P-R curve of the model on the test dataset is shown in Figure 6. It can also be seen from Figure 6 that the area under the four P-R curves is sufficiently large, which indicates that the performance of the model is satisfactory.

4.3. LV Binarization.
To analyze the effect of the MRF model on the binarization of the ultrasound LV images, the proposed method, traditional Otsu method, and K-means clustering algorithms were used to binarize the same LV image for

TABLE 2: Comparison of binarization results obtained by different methods and gold standards.

	Otsu	K-means	MRF
Dice	0.58 ± 0.07	0.59 ± 0.06	0.88 ± 0.03

comparison; the binarization results obtained by different methods were also compared with the gold standard, and the results are shown in Figure 7.

From Figure 7, it can be verified that the myocardial area obtained using the proposed model is closest to the gold standard.

For quantitative analysis, the Dice index is used for evaluation. The LV myocardial regions obtained by the Otsu method, K-means clustering algorithm, and the method based on MRF proposed in this paper are compared with the gold standard myocardial region obtained by manual segmentation by senior clinicians, and the corresponding Dice indices are obtained, and the results are shown in Table 2.

It can be seen from Table 2 that the Dice value corresponding to the proposed binarization method based on the MRF model is 0.88 ± 0.03, which is far greater than the Dice values corresponding to the Otsu method and K-means clustering algorithm, namely, the performance of the binarization method based on the MRF model proposed in this paper is much better than the other two methods. Therefore, the binarization method proposed in this study can fully meet the requirements for extraction of the LV myocardial region.

TABLE 3: Comparison of endocardial segmentation results by different methods.

Methods	Training set size (frame)	Computation speed (fps)	Dice (%)	MAD (mm)	HD (mm)
Hansson et al. [33]	0	0.3	—	2.58 ± 0.85	—
Qin et al. [34]	450	0.01	90.8 ± 1.7	2.0 ± 0.42	6.86 ± 1.71
Carneiro and Nascimento [35]	496	0.2	—	1.94 ± 0.51	—
The method without constraints	252	1.5–1.8	80.103 ± 2.13	7.06 ± 0.85	10.34 ± 3.51
Proposed method	252	2.1–2.25	93.578 ± 1.97	2.57 ± 0.89	6.68 ± 1.78

4.4. LV Endocardium Segmentation. In order to evaluate the performance of the method proposed in this paper, the same LV ultrasound images were segmented using different methods (listed in Table 3) along with the proposed method, and the segmentation results by different methods were compared with the gold standard obtained by manual segmentation by cardiologists, and five evaluation indicators including training set size, computation speed, Dice coefficient, MAD, and HD were used to evaluate the segmentation results. The results are shown in Table 3.

It can be seen from Table 3 that the proposed segmentation technique is superior to other methods in terms of various evaluation indicators. In particular, for the computation speed index, the method proposed in this paper has a great advantage, and owing to the use of transfer learning, the method uses less training data to obtain a better segmentation effect.

5. Discussion

In this paper, an automatic LV segmentation method based on the YOLOv3 model is proposed to determine the constraints and positioning. Through the YOLOv3 model, the three positions of the apex and bottom of the LV and LV area are positioned, and based on the MRF model, the LV myocardium subimages are binarized; under the limitation of the three constraint points of the LV, combined with NLS curve fitting and B-spline smoothing, the accurate segmentation and extraction of the LV can be realized. Experiments show that the suggested method can accurately and automatically identify and segment the LV in cardiac ultrasound images.

In the experimental section, a comparison is presented with other segmentation models. Hansson et al. [33] proposed an unsupervised segmentation method based on a Bayesian probability map. Although MADs corresponding to the aforementioned method and the method proposed herein are similar (which means that the two methods are similar in terms of segmentation accuracy), the computation speed of the latter is much higher than that of the former (see the computation speed indicator). The level set segmentation method proposed by Qin et al. [34] is unsupervised, does not require a training dataset, and can yield accurate segmentation results. However, owing to the need for sparse matrix transformation to identify the right ventricle, this method requires many training sets and a large processing time; in addition, it is necessary to readjust the parameters according to the movement of the heart, which will lead to unstable results. Compared with that of the aforementioned method,

the MAD of the method proposed herein this paper is slightly lower, but the Dice value is better. In fact, the method proposed by Qin et al. is similar to our method in terms of segmentation accuracy. However, the method proposed herein is far superior in terms of the computation speed indicator. The method proposed by Carneiro and Nascimento [35] uses a deep neural network method to segment the systolic and end-diastolic contours and achieves high segmentation accuracy; however, a large number of datasets is required, and thus, a set of 496 images had to be established. Compared with this method, the method proposed herein only requires a small amount of data (252 frames) to obtain a suitable positioning effect; in terms of calculation speed, the method proposed herein this paper is significantly better than that proposed by Carneiro and Nascimento (see the corresponding computation speed index in Table 3). Finally, according to the computation speed, Dice, MAD, and HD, the automatic LV segmentation method based on constraints and positioning are better in the proposed technique than unconstrained positioning segmentation methods in terms of segmentation accuracy and computation speed.

In summary, if the segmentation accuracy indices (i.e., Dice, MAD, and HD) are considered, the method proposed is not the best, but it can be said that the method proposed in this paper is one of the best methods in terms of segmentation accuracy; however, if the computation speed, data volume, and segmentation accuracy are considered comprehensively, it can be said that the method proposed in this paper is the best. Compared with other methods, the proposed segmentation technique has significant advantages in terms of computation speed and the amount of data required. The method proposed in this study uses fewer data to obtain a good segmentation effect. It is well known that it is very difficult to obtain medical data in practice, thus obtaining a good segmentation effect based on a small amount of data is conducive to the clinical application of the algorithm. The computation speed is another important factor that affects the application of algorithms in clinical practice, and the algorithm proposed has significant advantages in terms of computation speed over the other methods.

6. Conclusions

Here, an automatic LV segmentation method based on the YOLOv3 model for constraint and positioning determination is proposed. Through the YOLOv3 model, the three positions of the apex and bottom of the LV and LV area are positioned, and based on the MRF model, the LV myocardium subimages are binarized; under the limitation of the

three constraint points of the LV, combined with NLS curve fitting and B-spline smoothing, the accurate segmentation and extraction of the LV can be realized. Experiments show that the method can accurately and automatically identify and segment the LV in cardiac ultrasound images, and related indicators such as fps, Dice, MAD, and HD can reach 2.1–2.25 fps, 93.57 ± 1.97%, 2.57 ± 0.89 mm, and 6.68 ± 1.78 mm, respectively. Compared with other methods, the proposed method has a better segmentation accuracy and robustness. In particular, our method has a high computational speed, which is very important for real-time evaluation of cardiac function based on echocardiography. In addition, our method uses less training data to achieve better segmentation results. In short, our method can accurately segment LV ultrasound images, which is important for the accurate acquisition of clinical indicators for cardiac function evaluation, such as the EF, strain, and strain rate of the LV on echocardiography and will play a vital role in assisting doctors in clinical diagnosis.

Data Availability

The cardiac ultrasound imaging data used in this study were provided by the Ultrasound Imaging Department of the First Affiliated Hospital of Medical College of Shantou University, China, which is not open to the public because it would breach the privacy of the research.

Conflicts of Interest

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

Acknowledgments

References

[1] G. A. Mensah, G. A. Roth, and V. Fuster, "The global burden of cardiovascular diseases and risk factors," *Journal of the American College of Cardiology*, vol. 74, no. 20, pp. 2529–2532, 2019.

[2] N. Benyounes, C. Van Der Vynckt, S. Tibi et al., "Echocardiography in confirmed and highly suspected symptomatic COVID-19 patients and its impact on treatment change," *Cardiology research and practice*, vol. 2020, Article ID 4348598, 9 pages, 2020.

[3] V. Tirronen, F. Neri, T. Krkkinen, K. Majava, and T. Rossi, "An enhanced memetic differential evolution in filter design for defect detection in paper production," *Evolutionary Computation*, vol. 16, no. 4, pp. 529–555, 2008.

[4] A. Caponio, F. Neri, and V. Tirronen, "Super-fit control adaptation in memetic differential evolution frameworks," *Soft Computing*, vol. 13, no. 8, pp. 811–831, 2009.

[5] P. C. Ng and S. Henikoff, "SIFT: predicting amino acid changes that affect protein function," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3812–3814, 2003.

[6] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, San Diego, CA, USA, 2005.

[7] L. Teng, Z. L. Fu, Q. Ma et al., "Interactive echocardiography translation using few-shot GAN transfer learning," *Computational and mathematical methods in medicine*, vol. 2020, Article ID 1487035, 9 pages, 2020.

[8] Y. Lu, B. Li, N. Liu et al., "CT-TEE image registration for surgical navigation of congenital heart disease based on a cycle adversarial network," *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 4942121, 8 pages, 2020.

[9] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, vol. 39, pp. 91–99, 2015.

[10] W. Liu, D. Anguelov, D. Erhan et al., "SSD: single shot multibox detector," in *European conference on computer vision*, pp. 21–37, Cham, 2016.

[11] J. Redmon and A. Farhadi, "Yolov3: an incremental improvement," 2018, https://arxiv.org/abs/1804.02767.

[12] S. Leclerc, T. Grenier, F. Espinosa, and O. Bernard, "A fully automatic and multi-structural segmentation of the left ventricle and the myocardium on highly heterogeneous 2D echocardiographic data," in *2017 IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, Washington, DC, USA, 2017.

[13] S. Dong, G. Luo, K. Wang, S. Cao, Q. Li, and H. Zhang, "A combined fully convolutional networks and deformable model for automatic left ventricle segmentation based on 3D echocardiography," *BioMed research international*, vol. 2018, Article ID 5682365, 16 pages, 2018.

[14] E. Smistad, A. Ostvik, B. O. Haugen, and L. Lovstakken, "2D left ventricle segmentation using deep learning," in *2017 IEEE international ultrasonics symposium (IUS)*, pp. 1–4, Washington, DC, USA, 2017.

[15] O. Oktay, E. Ferrante, K. Kamnitsas et al., "Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation," *IEEE transactions on medical imaging*, vol. 37, no. 2, pp. 384–395, 2018.

[16] E. Smistad and F. Lindseth, "Real-time tracking of the left ventricle in 3D ultrasound using Kalman filter and mean value coordinates," *MICCAI Challenge Echocardiogr. Three-Dimensional Ultrasound Segmentation (CETUS)*, pp. 65–72, 2014.

[17] D. Barbosa, D. Friboulet, J. D'hooge, and O. Bernard, "Fast tracking of the left ventricle using global anatomical affine optical flow and local recursive block matching," *MICCAI Challenge Echocardiogr. Three-Dimensional Ultrasound Segmentation (CETUS)*, pp. 17–24, 2014.

[18] C. Wang and Ö. Smedby, "Model-based left ventricle segmentation in 3D ultrasound using phase image," *MICCAI Challenge*

Echocardiogr. Three-Dimensional Ultrasound Segmentation (CETUS), pp. 81–88, 2014.

[19] M. Bernier, P. Jodoin, and A. Lalande, "Automatized evaluation of the left ventricular ejection fraction from echocardiographic images using graph cut," *MICCAI Challenge Echocardiogr. Three-Dimensional Ultrasound Segmentation (CETUS)*, pp. 25–32, 2014.

[20] M. Van Stralen, A. Haak, K. E. Leung, G. van Burken, and J. G. Bosch, "Segmentation of multi-center 3D left ventricular echocardiograms by active appearance models," *MICCAI Challenge Echocardiographic Three-Dimensional Ultrasound Segmentation (CETUS)*, pp. 73–80, 2014.

[21] O. Oktay, W. Shi, K. Keraudren, J. Caballero, and D. Rueckert, "Learning shape representations for multi-atlas endocardium segmentation in 3D echo images," *MICCAI Challenge Echocardiogr. Three-Dimensional Ultrasound Segmentation (CETUS)*, pp. 57–64, 2014.

[22] T. Ahmad, Y. Ma, M. Yahya, B. Ahmad, and S. Nazir, "Object detection through modified YOLO neural network," *Scientific Programming*, vol. 2020, Article ID 8403262, 10 pages, 2020.

[23] A. Likas, N. Vlassis, and J. Verbeek, "The global K-means clustering algorithm," *Pattern Recognition*, vol. 36, no. 2, pp. 451–461, 2003.

[24] S. Aja-Fernndez, C. Alberola-Lpez, and C.-F. Westin, "Noise and signal estimation in magnitude MRI and Rician distributed images: a LMMSE approach," *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1383–1398, 2008.

[25] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Advances in artificial intelligence*, vol. 2009, 19 pages, 2009.

[26] S. Z. Li, *Markov Random Field Modeling in Image Analysis*, Springer, London, U. K, 2009.

[27] P. H. C. Eilers and B. D. Marx, "Splines knots and penalties," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 6, pp. 637–653, 2010.

[28] S. Robertson, "A new interpretation of average precision," in *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 689-690, New York, NY, USA, 2008.

[29] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in *Proceedings of the 23rd international conference on Machine learning*, pp. 233–240, New York, NY, USA, 2006.

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

[31] X. Huang, "Contour tracking in echocardiographic sequences via sparse representation and dictionary learning," *Medical Image Analysis*, vol. 18, no. 2, pp. 253–271, 2014.

[32] A. A. Taha and A. Hanbury, "An efficient algorithm for calculating the exact Hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 11, pp. 2153–2163, 2015.

[33] M. Hansson, S. S. Brandt, J. Lindström et al., "Segmentation of B-mode cardiac ultrasound data by Bayesian probability maps," *Medical Image Analysis*, vol. 18, no. 7, pp. 1184–1199, 2014.

[34] X. Qin, Z. Cong, and B. Fei, "Automatic segmentation of right ventricular ultrasound images using sparse matrix transform and a level set," *Physics in Medicine & Biology*, vol. 58, no. 21, pp. 7609–7624, 2013.

[35] G. Carneiro and J. Nascimento, "Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2592–2607, 2013.

Research Article

Diagnosis of Alzheimer's Disease Severity with fMRI Images Using Robust Multitask Feature Extraction Method and Convolutional Neural Network (CNN)

Morteza Amini [1], **Mir Mohsen Pedram** [2,3] **AliReza Moradi** [4,5] **and Mahshad Ouchani** [6]

[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]*Institute for Cognitive and Brain Science, Shahid Beheshti University, Tehran, Iran*

Correspondence should be addressed to Mir Mohsen Pedram; pedram@khu.ac.ir

The automatic diagnosis of Alzheimer's disease plays an important role in human health, especially in its early stage. Because it is a neurodegenerative condition, Alzheimer's disease seems to have a long incubation period. Therefore, it is essential to analyze Alzheimer's symptoms at different stages. In this paper, the classification is done with several methods of machine learning consisting of K-nearest neighbor (KNN), support vector machine (SVM), decision tree (DT), linear discrimination analysis (LDA), and random forest (RF). Moreover, novel convolutional neural network (CNN) architecture is presented to diagnose Alzheimer's severity. The relationship between Alzheimer's patients' functional magnetic resonance imaging (fMRI) images and their scores on the MMSE is investigated to achieve the aim. The feature extraction is performed based on the robust multitask feature learning algorithm. The severity is also calculated based on the Mini-Mental State Examination score, including low, mild, moderate, and severe categories. Results show that the accuracy of the KNN, SVM, DT, LDA, RF, and presented CNN method is 77.5%, 85.8%, 91.7%, 79.5%, 85.1%, and 96.7%, respectively. Moreover, for the presented CNN architecture, the sensitivity of low, mild, moderate, and severe status of Alzheimer patients is 98.1%, 95.2%,89.0%, and 87.5%, respectively. Based on the findings, the presented CNN architecture classifier outperforms other methods and can diagnose the severity and stages of Alzheimer's disease with maximum accuracy.

1. Introduction

In fluorodeoxyglucose-positron emission tomography research, cognitive impairment in AD has been correlated with localized brain metabolic damage in systematic and functional imaging experiments [1–3]. Blood-oxygen-level-dependent imaging was seen to reflect healthy functional networks, including default mode (DMN), visual (VIS), and executive networks (EN) [4], within a given resting state. Unlike task-related functional MRI (fMRI), patients' capability to recognize and memorize the instructions for executing a given task is not confounded by resting-state fMRI, which makes it useful for the survey of individuals with cognitive decline [5]. Besides, convincing literature-wide data confirms the application of resting-state connectivity as an AD biomarker [6]. Machine learning (ML) is an artificial intelligence

field that typically utilizes factual methods to allow computers to "learn" through data from stored datasets. A subset of ML [7] is fundamental deep learning (DL). The DL is a neural network that uses several variables and layers to define. There are a variety of simple network architectures [8], including CNNs, mainly a standard spatial mutual weight neural network [9].

The CNN is designed to identify images that see the edges of a known target on the image by making convolutions inside [10]. (ii) Recurrent neural networks are names of artificial neural networks where a graph is generated by specific associations between nodes in the temporal chain. RNNs can use their internal condition to handle the sequences of inputs, unlike feedforward neural networks. RNN is meant to identify sequences such as a voice signal or a text [9], for example. (iii) In recursive neural networks, the input sequence does not include a time dimension, and the input must be hierarchically evaluated in a tree form [8, 10]. Various external inputs usually contribute to distinct brain functions, and various functional brain representations are displayed by different brain activities [11]. For that function, the classification of images plays an essential role in detecting various brain functions. Several deep learning approaches have recently been suggested to carry out image recognition for various brain activities [12, 13]. A deep neural network feedforward has been employed by Koyamada et al. [12] to identify different brain functions, including preferences; motor, social, emotional, and language activities; and work memory, using functional magnetic resonance imaging (fMRI) images. A SoftMax layer and various secret layers were used in the feedforward deep neural network. Similarly, to get high-level latent properties, these hidden layers were used. In contrast, the SoftMax layer has been applied to calculate a subject's ability in a class. To boost the final classification efficiency, dropout, minibatch stochastic decrease [14], and main sensitivity analyses [15] were also integrated into the deep feedforward neural network. Jang et al. newly exploited deep neural networks and hidden layers completely connected to feedforward to distinguish different sensor roles, including visual attention and stimuli and right-hand and left-hand clenching, are included. The DL classification of MRI images included other classifications above and below the classifications, such as diagnosis of stroke [16], age predictions [17], classification of attention-deficit hyperactivity disorder (ADHD) [18], prejudice against cerebellar ataxia [19], and predictive emotional response [20]. Due to science, computer-aided diagnosis systems (CADs) were developed to play an important role in enhancing the understanding of medical imagery among researchers and physicians. The application of the machine learning technique, in particular DL strategies in CAD models to diagnose and classify stable control patients with average (CN), AD, and mild cognitive impairment (MCI), has exponentially grown [21, 22]. The automatic diagnosis of AD performs an essential role in human health, especially in the early stages. AD has a considerable incubation period because it is a neurodegenerative disorder.

Thus, the AD symptoms need to be analyzed at various levels. Currently, several scholars have discussed using image

classification to carry out AD diagnosis. Several DL approaches have been suggested to use MRI images to introduce multiple AD patients' severity [22, 23]. The higher the image quality, the better the outcomes achieved, known in image analysis. However, the quality of image relies on image processing, and when the picture is acquired higher, the image quality is higher. MRI retains noninvasive and good contrasting properties of soft tissue but does not expose to people ionizing with high radiation. As MRI can produce a great deal of priceless knowledge of tissue frameworks such as position, size, and type, more attention is paid to computerized diagnostics and clinical routine [24, 25]. Functional and structural imaging can be classified into MRI. T1-weighted MRI (T1w), diffusion tensor imaging (DTI), and T2-weighted MRI (T2w) [26] are used in structural imaging. Functional imagery includes functional MRI task status (ts-fMRI) and functional MRI resting state (rs-fMRI). Medical diagnostic data systems are employed for medical centers and doctors to treat diseases, and analytical tools to improve management and diagnosis are critical. Given the crucial function of medical data in humans' lives, computer scientists have been involved in this area. Healthcare professionals may make their decisions, including medical diagnoses and the effects of severe conditions, by contributing to the medical details' classification. In addition to the number of these conditions, a data collection of diseases comprises patient symptoms as characteristics. The extensive patient evidence available can be used for health treatment. Data mining may be used in medical center studies to provide appropriate origins of disease for prohibiting and prompt diagnosis and avoiding the significant costs of diagnostic tests [27].

In this paper, machine learning methods are utilized for Alzheimer's disease classification. Moreover, robust multi-task methods are utilized for feature extraction of fMRI images from the ADNI dataset. In the output layer, the main aim is to find the severity of Alzheimer's diseases. Therefore, the results of MMSE are used. For classification and diagnosis of Alzheimer's disease severity, the machine learning methods are trained. Input and output features are applied for six classifiers including, KNN, SVM, DT, LDA, RF, and CNN. Finally, performance analysis consists of the confusion matrix and the ROC curve illustrates the classification results.

2. Research Background

AD recognition has been extended to many different methods focused on deep learning. Nevertheless, several controversial findings encouraged us to participate in the literature review to determine the current operating condition and what could be the potential innovations. In this section, the primary study concern is if DL techniques have been able to classify AD using neuroimaging data. The training dataset scale is considered to significantly impact the classifier's output over an undefined test range [28]. In each dataset, the amount of AD and MCI topics can be minimal, inadequate for deep models to be evaluated. For multimodality experiments, the condition is worse. Any experiments, however, have mixed datasets. While it can result in more heterogeneity by integrating multiple datasets, this may advance a broad

TABLE 1: Summary research on Alzheimer's disease diagnosis methods.

Author	Year	Database	Modality	Method	Accuracy
Suk and Shen [32]	2013	Alzheimer's Disease Neuroimaging Initiative (ADNI)	PET, MRI, CSF	Stacked autoencoder, SVM	95.9
Suk al.[33]	2014	ADNI	PET, MRI	Deep Boltzmann machine	95.4
Liu et al. [34]	2016	ADNI	MRI	Influence of subclass number, multiview feature extraction, subclass clustering-based feature selection, SVM	93.8
Zu et al. [35]	2016	ADNI	PET, MRI	Label-aligned multi-task feature selection, support vector machine	96.0
Sarraf and Tofighi [36]	2016	ADNI	fMRI	LeNet-5	96.85
Sarraf and Tofighi [37]	2016	ADNI	MRI, fMRI	LeNet, GoogleNet	98.84
Li et al. [38]	2017	ADNI	MRI	CNN	88.31
Amoroso et al. [39]	2018	ADNI	MRI	Random Forest, deep neural network, fuzzy logic	38.8
Liu et al. [40]	2018	ADNI	MRI, PET	2D and 3D CNN,	93.26
Yang et al. [41]	2018	ADNI	MRI	The convolutional neural network, 3DVGGNET, 3DRESNET	76.6
Wang et al. [42]	2018	Open Access Series of Imaging Studies	MRI	CNN	97.65
Khvostikov et al. [43]	2018	ADNI	MRI, DTI	CNN	96.7
Shi et al. [44]	2018	ADNI	MRI, PET	Multimodal stacked deep polynomial network, SVM	97.13
Ramzan et al. [45]	2019	ADNI	fMRI	Off-the-shelf and fine-tuned	97.88
Parmar et al. [46]	2020	ADNI	fMRI	3D CNN	96.55
Duc et al. [47]	2020	ADNI	fMRI	3D CNN and SVM-RFE	85.27
Li et al. [48]	2020	ADNI	4D fMRI	3D CNN and C3d-LSTM	89.47
Al-Khuzaie et al. [49]	2021	Alzheimer Network (AlzNet)	2D fMRI	CNN	99.30
Bhaskaran and Anandan [50]	2021	Research Anthology on Diagnosing and Treating Neurocognitive Disorders	rsfMRI	Graph metrics and lateralization	97.54
Luo et al. [51]	2021	Population-specific Chinese brain atlas	rsfMRI	Graph metrics and false discovery rate (FDR)	95.67
Ahmadi et al. [52]	2021	Harvard Medical School	MRI	Robust PCA and CNN method	96

and stable classification and prediction model. Using data augmentation is another means of addressing the small number of topics in a dataset. Data increase is a technique that increments the data range of training model applications without additional data being obtained. In approximately 20 percent of research aimed at enhancing classification performance, data enhancement strategies like random translation, rotation, reflection, adding noise, gamma filter, blurring, cutting, and scaling were used where appropriate [29].

Moreover, at various time points, longitudinal datasets include multiple brain scans per subject; it may also be employed for data increase in time, while their main objective was to analyze disease development [30]. While implementing a DNN from scratch is completed in some experiments, it is always impossible to do so: the training phase can use much time, or the sample may be tiny [31]. Even though there are millions of images in datasets of object detection and etiquette, neuroimaging datasets contain hundreds of

images that help overfit the planning. It is generally beneficial to start tested, previously trained CNN with one dataset and retrain them with just the fine-tuning of CNN on another dataset (transfer learning). It is feasible since more general characteristics in the lower CNN layers can profit certain classification activities that can be moved from one program domain to another. CNN classifier is one of the effective methods for classification for all brain diseases. Besides, finding the best way for classification impacts diagnosis accuracy and process time. Therefore, our presented method is justified computationally.

Transfer learning is also more comfortable with small projects and produces higher performance than planning from the beginning [53]. Payan and Montana proposed classifying AD stages, namely, MCI, AD, and standard control [54] (NC). The algorithms were designed to implement a 3D CNN to separate brain scans employing autoencoding systems and 2D CNN. For 3D CNN and 2D CNN versions, an accuracy of 89.47 percent and 85.53 percent was reached.

Liu et al. have also achieved a classification accuracy of about 85.53 percent with the identical network structure for 2D CNNs [34]. A study for the classification of AD was done by Sarraf and Tofighi [36]. The research was focused on classifying AD patients using MRI and fMRI scans from normal control subjects. For binary classification, two network architectures have been implemented. LeNet-5 and GoogleNet were the foundations for these CNN-based architectures. It obtained an approximate accuracy of 99 percent with LeNet and 100 percent with GoogleNet utilizing fMRI data. An analysis of research that focuses on AD classification using deep learning techniques is given in Table 1. Structural MRI or PET scans have been used in many experiments that concentrate on characterizing a few stages of the disorder, i.e., AD, MCI, and CN. In multiclass AD diagnosis and grouping, a restricted number of researches have employed fMRI findings.

3. Methods and Materials

3.1. Quantum Matched-Filter Technique (QMFT). Initially, a preprocessing step with a noise reduction would take place. In conjunction with the local threshold and the active contour, each image is displayed employing a two-dimensional pixel array, the value of which is an integer in the [0, 255] scale. In two stages, local thresholds initialize images. Then, the input noise picture is named the main image to which image noise reduction is implemented. This procedure is used explicitly by the quantum matched-filter technique (QMFT) as a local search operator to improve the initial images. In this article, the utilization of local thresholds and active contours was considered since it is faster computationally than other approaches in the literature. Thus, there will be a decomposed picture at the end of the first stage. Thresholding is performed on the thorough coefficients in the second step, and each of the decomposed pieces is randomly picked and submitted to a reconstruction process. It is possible to describe the restoration portion [55]:

(i) Gaussian Blur: a Gaussian filter is used to filter an image. The filter size is chosen unintentionally, between 3×3 pixels and 5×5 pixels

(ii) Mean filter (averaging filter): the picture is filtered utilizing an average filter

(iii) Intensity change: a randomly selected associated criterion in [0.7, 1.3] range is used to multiply all the image pixels

(iv) Integrate light-intensive parts that conduct the QMFT in quantum and reverse processing

Then, it executes the following procedures:

(i) One-point row: random selection of a pixel row

(ii) One-point column: it is similar to the preceding method, except that it is regarded instead of a row

(iii) Point-to-point random: every pixel is incorrectly chosen until a new image is produced from decomposition

(iv) Mark points in rows and columns of the picture as QMFT to diminish the bulk of the noise

If the range value [0.1] chosen in the QMFT is lower than the rate of local search, the current image will be passed to the local search operator after a review. Its pixel value sorts the entire picture until the decomposition is complete. The best aspect ratio of the picture is then known in the sequel as a quantum value. The signal can be split into multiple displaced or revamped characteristic displays located at the feature's extraction point in fMRI photos. For the study of an image in its elements, local thresholds and active contours may be used. After implementing QMFT alongside local and active contouring thresholds, it is feasible to execute image classification operations. In this case, it is possible to destroy the local threshold coefficients and the QMFT-based active contour to delete certain information. Local thresholds and active contours based on QMFT have a significant advantage when details are separated into an image. It is possible to employ active contour to isolate excellent image information. Simultaneously, extensive details can be identified by local thresholds, integrating fine and extensive details and linearly and diagonally reading all rows and columns. Quantum reaches QMFT, so noise in the fMRI image can be minimized. A light display can be used to create a QMFT display with local thresholds and active contours. The local and active QMFT contouring mechanism has two key features: the oscillation or wave presence function, as in the following equation [55]:

$$\int_{-\infty}^{0} \left| \Psi(t) \right|^2 dt < \infty. \tag{1}$$

The energy in $\Psi(t)$ is confined to a short period as

$$\int_{-\infty}^{0} \Psi(t) dt = 0. \tag{2}$$

Generally, the suggested approach is estimated to decrease the noise in

$$\Omega(I) = \left(\sum_{\Omega} \sqrt{1 + \beta^2 |\nabla I|^2} \right) + \frac{\lambda}{2} (I - I_0)^2. \tag{3}$$

Within Equation (3), the term $(I - I_0)^2$ guarantees the rated image and a certain degree of authenticity and consistency in the original image, where I denotes the rated picture and I_0 corresponds to the noisy picture. The parameter ∇I is described as the number of times of variable change, β and λ are balancing variables, and Ω is the sum of the image's pixels. The purpose of reducing Equation (3) is to diminish the broad variety of images while retaining accuracy and validation. For both β and λ, balancing values are modified from 1 to the image size to decrease Equation (3) [55].

3.2. Robust Multitask Feature. This paper is aimed at simultaneously catching common characteristics among several similar tasks and detecting outer work using the robust multitask

learning function algorithm (rMTFL). The rMTFL will estimate the correct assessment and the true underlying weights. Also, if the true underlying weights are over noise thresholds, rMTFL will achieve exact sparsity patterns. Also, rMTFL optimization can be easily solved, and rMTFL scales can be used to solve significant problems [56]. Presume that there are m learning tasks relevant to the $\{(X_1, y_1), \cdots, (X_m, y_m)\}$, training results, where $X_i \in R^{d \times n_i}$ is the ith task data matrix with column as a sample; $y_i \in R^{n_i}$ is the ith task response (y_i has continuous regression values and discrete classification values); d is the dimensionality of the data; and n_i is the number of ith task samples. The data were normalized to satisfy X_i's (j, k)th input, which is referred to as $x_{jk}^{(i)}$ [56]:

$$\sum_{k=1}^{n_i} \left(x_{jk}^{(i)}\right)^2 = 1, \quad j \in \mathbb{N}_d. \tag{4}$$

The linear function of learning is

$$y_{ii} \approx f_i\left(x_j^{(i)}\right) = \left(x_j^{(i)}\right)^T w_i, \quad i \in \mathbb{N}_m, j \in \mathbb{N}_{n_i}. \tag{5}$$

The sum of two elements, P and Q, for each task and for decomposing of the weight matrix $W = [w_1, \cdots, w_m] \in R^{d \times m}$. To manipulate relationships between tasks, various regularization conditions on P and Q are used. The rMTFL model, theoretically, is developed as

$$\min_{W, P, Q} \quad \sum_{i=1}^{m} \frac{1}{mn_i} \left|\left|X_i^T w_i - y_i\right|\right|^2 + \lambda_1 ||p||_{1,2} + \lambda_2 \left|\left|Q^T\right|\right|_{1,2}$$

$$\text{s.t.} \quad W = P + Q. \tag{6}$$

When P reports the mutual functions between tasks and Q learns the second term's outer tasks, λ_1 and λ_2 are nonnegative parameters to handle these two terms [56].

3.3. Convolutional Neural Network.
CNNs have been widely employed for DL and the most prominent classes of neural networks, mostly in extensive data such as images and videos. It is a multilayer neural network architecture caused by cortex neurobiology. It consists of convolutional layers and fully connected layers. Between these two layers, subsampling layers can exist. The best of DNNs is achieved, which are challenging to scale along with multidimensional input data associated locally well. Therefore, CNN can be automatically applied in databases where comparatively large numbers of nodes and parameters are trained (e.g., image processing) [57].

3.3.1. Convolutional Layer.
This is the essential building block of a CNN that determines the output of associated inputs in the field of reception. These kernels' findings translate into data height and width, calculate the point product between inputs and filter values, and then create a 2D filter map enabled. It helps the CNN quickly find the filters that enable when an input temporarily detects a specific type of function [57].
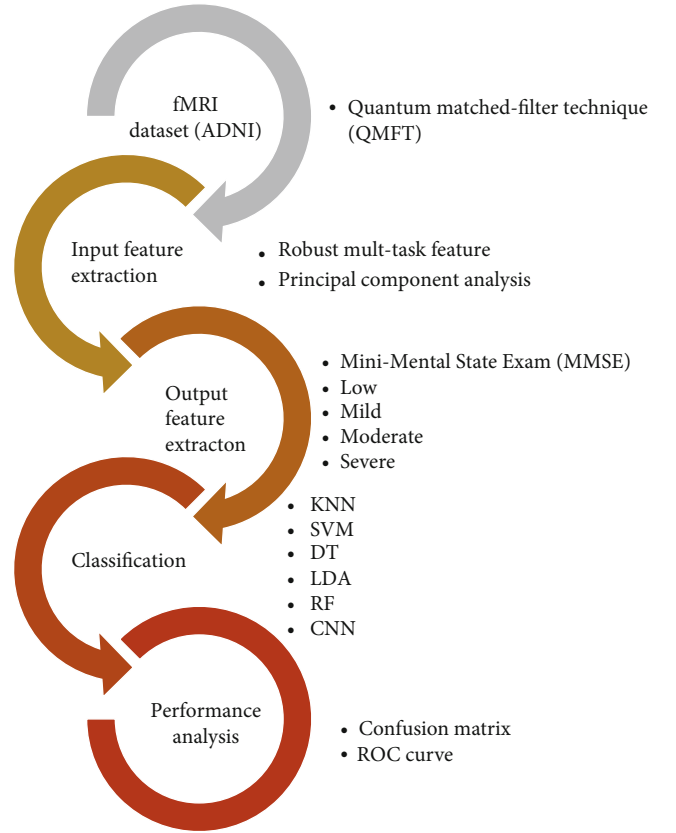

FIGURE 1: The conceptual flowchart of the presented process.

3.3.2. Nonlinearity Layer.
Nonlinear characteristics have a high degree of importance and curvature. This layer's primary purpose is to convert the input signal into the output signal, which is used as an input in the next layer. Sigmoid or logistical forms, Tanh, ReLU, PReLU, ELU, and more, are not linear.

3.3.3. Pooling Layer.
The CNN may be locally or globally sampled to link the neuron outputs to an established neuron on a single layer in the following layer. The critical task is to limit the number of parameters and equations within the model to reduce spatial depiction volume [57]. It not only speeds up calculations but also takes the issue of overfitting into account. The most popular method of pooling is max pooling.

3.3.4. Fully Connected Layer.
FC layers are deep NNs typical for the regression or classification of the activation to construct the predictions. A description of the multilayer perceptron (MLP) neural system is equivalent to the typical neural system. The entire relationship with each activation is formed in the antecedent layer. Activation can be determined by the matrix multiplication and a bias offset [57].

3.3.5. Loss/Classification Layer.
The loss layer defines how the training eliminates the disparity between the actual and projected marks, ensuring that the training phase of NN is directly guided by it. Various loss functions for different
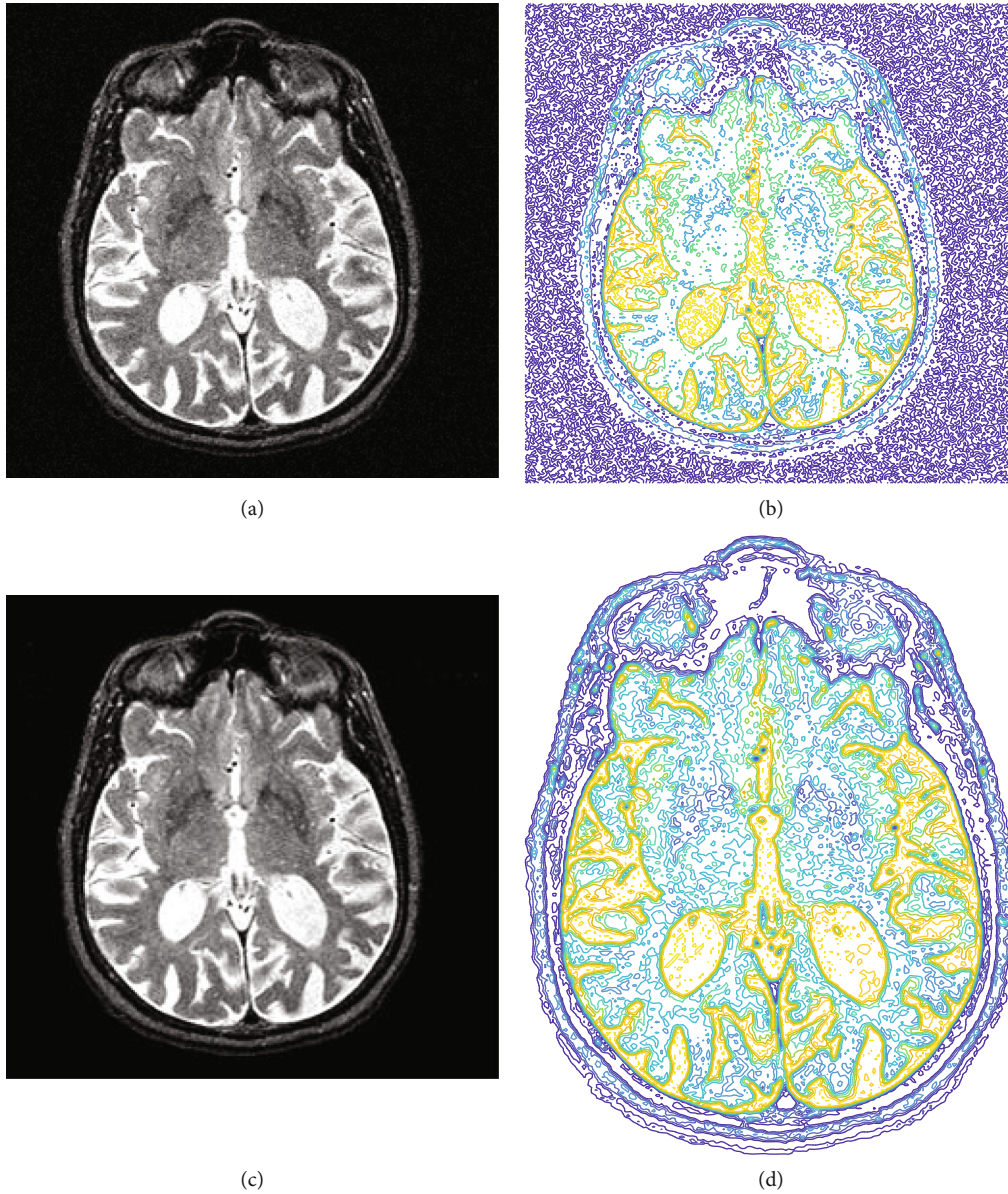
(a)

(b)

(c)

(d)

Figure 2: Results of noise reduction using QMFT: (a) input image; (b) input image contour form; (c) noise-reduced image; (d) contour form of noise-reduced image.

commands such as SoftMax and crossentropy may be used in DCNN. SoftMax losses are used to measure a solo class of K mutually exclusive classes. The SoftMax layer is used to calculate the likelihood, i.e., the total output values for 1. Furthermore, this layer is a responsive max-output layer type, such that irregularities are distinguishable and often scalable. Sigmoid crossentropy loss is used to foresee K -free probability values [58]. The sigmoid capability yields negligible probabilities, and lines can be used for grouping various groups alongside these probabilities. A problem with sigmoid is that the gradient disappeared after the saturation had been achieved. Euclidean failure is used to regress to fully appreciated names. The following is an overview of the neural network model's programs, database, results, and implementations.

4. Results and Discussion

In this paper, machine learning methods are utilized for Alzheimer's disease classification. First of all, the input image is filtered with the QMFT method to reduced input fMRI images. To imply the classifier in fMRI images, feature extraction should be done for both the input and output layers. Therefore, robust multitask methods are used for feature extraction of input layers. Then, for reducing the number of features, the PCA method is chosen. In the output layer, the main aim is to find the severity of Alzheimer's disease. Therefore, the results of MMSE are the best choice. It consists of four categories: the low, mild, moderate, and severe patients' severity. The next step is to train the machine learning methods. Input and output features are applied for
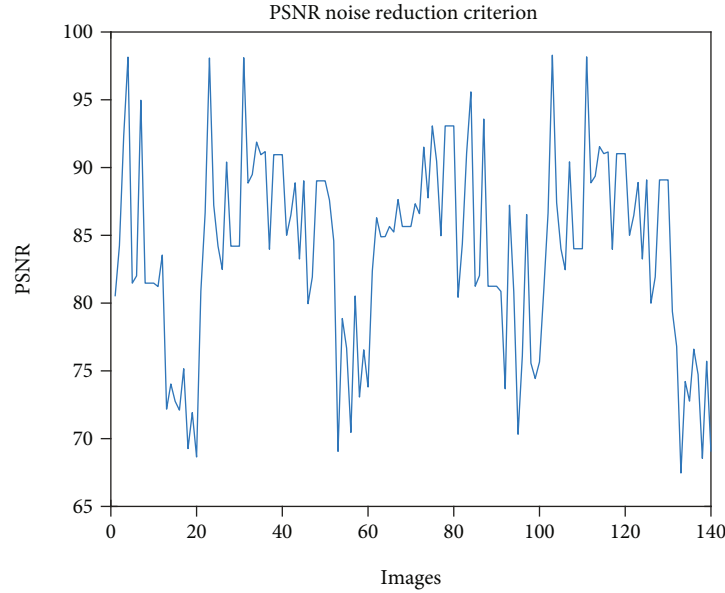
FIGURE 3: The PSNR value of noise reduction from fMRI images.

TABLE 2: Scoring system of MMSE and the severity of Alzheimer's disease.

Score	Severity	Psychometric analysis	Day-to-day functioning
25-30	Low	If there are clinical symptoms of cognitive disability, a formal cognition test can be useful	Clinically significant, however mild, deficits may be available. Only the most stressful everyday life tasks are expected to be affected
20-25	Mild	To further assess the trend and nature of deficits, a systematic examination can be useful	Meaningful effects. Any monitoring, assistance, and aid may be needed
10-20	Moderate	The formal assessment of whether there are clear health indications may be helpful	Obvious deficiency. 24-hour surveillance could be required
0-10	Severe	The patient will not be testable	Impairment labelled. 24-hour surveillance and support with ADL are likely to be required

six classifiers including, KNN, SVM, DT, LDA, RF, and CNN. Finally, performance analysis consists of the confusion matrix and the ROC curve illustrates the classification results. The conceptual diagram of the method is presented in Figure 1.

4.1. Preprocessing of Dataset. Data used in this paper's preparation was obtained from the ADNI database. Each subject's standard format was a series of 140 $64 \times 64 \times 48$ 3D NIFTI files and a single T1-weighted structural MRI file. Each 3D NIFTI file represented the patient's brain's rs-fMRI data from a 3-Tesla MRI scanner. Multiple subjects had nonstandard fMRI size (e.g., $96 \times 96 \times 48$, $80 \times 80 \times 48$) and were filtered out as well.

First, subjects were arbitrarily categorized into groups for training and testing. Around 80 percent of the details were required for training, and the remaining 20 percent was used for testing. For the training and testing datasets, similar preprocessing was implemented. First, the skull and neck voxels, which are the MRI scans' nonbrain regions, were removed from the T1-weighted image that corresponded to each subject. The resting-state fMRI contained 140 time steps per subject and was corrected for motion artifacts. Then, regular slice timing correction was applied to each time series

because later steps assume all slices were acquired halfway through the relevant acquisition time. Slice timing correction shifts each time series by the appropriate fraction. Spatial smoothing was carried out next using a Gaussian kernel (5 mm full width at half maximum). Then, low-level noise was removed from the data using quantum matched-filter technique (QMFT). The noise reduction results can be shown by the 2D section of images in Figure 2.

Based on the results of QMFT in Figure 2, the prominent image noise was removed from 3D fMRI images. For better illustration of noised and reduced images, the contour form of image matrixes is shown in Figures 2(b) and 2(d). The peak signal-to-noise ratio (PSNR) is shown in Figure 3. Results of reduction for 140 images are depicted in Figure 3. The average value of PSNR for the tested images is 83.9731. The reduction of noise gives an exciting outcome that enables a proper extraction of features.

4.2. Feature Extraction and Input Features. The ADNI database is adopted for feature extraction of fMRI images. The fMRI of 675 patients is included in the results. fMRI data include 285 features classified into five types: average cortical thickness, the standard deviation of cortical thickness, the
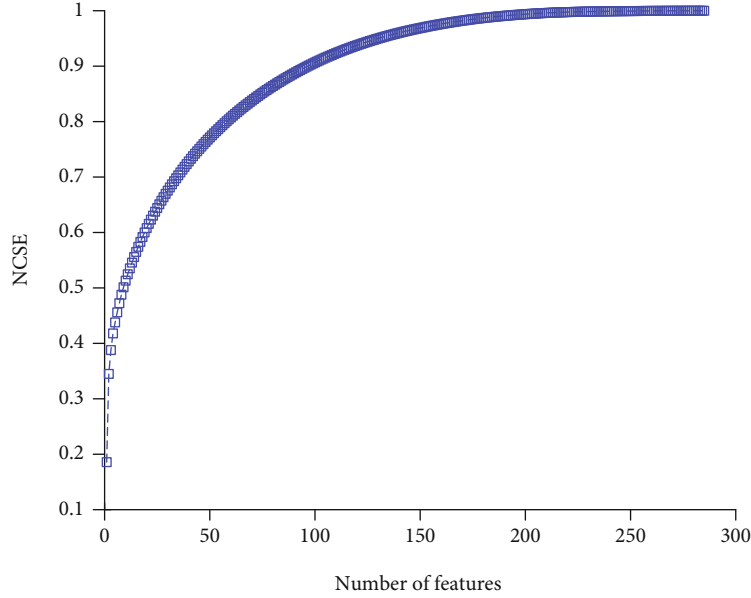
FIGURE 4: The cumulative summation of sorted eigenvalues.

volume of cortical parceling, white matter, and surface area. The result is the score from 6 separate time points of the Mini-Mental State Examination: M06, M12, M18, M24, M36, and M48. The samples that fail to track the consistency of fMRI and missing results are removed.

4.3. Mini-Mental State Exam (MMSE). According to certain risk factors, the cognitive function can decrease (e.g., hypertension, elevated cholesterol, cardiac arrhythmias). The physical and life quality of older people may be adversely affected. Dementia is a significant disorder and a cause of elderly disabilities. The second leading source in the dementation of AD is brain vascular disease or multi-infarct dementia. The Mini-Mental State Exam (MMSE) is an elderly cognitive function test commonly used; it requires orientation, attention, memory, language, and visual-spatial ability. The MMS is broken into two parts; the first only includes vocal responses and encompasses orientation, memory, and attention; 21 is the highest score. The second section checks the ability to name, obey verbal and written orders, automatically write a phrase, and copy a complex Bender-Gestalt figure-like polygon; the highest score is 9. Patients with seriously affected vision can have some added difficulties due to the reading and writing involved in part II, which can typically be eased by broad writing and allowed for in the scoring. There is a full cumulative score of 30 [59] (see Table 2).

For this paper, the relationship between Alzheimer's patients' functional magnetic resonance imaging (fMRI) images and their MMSE scores is assessed. Furthermore, a machine learning model's training is done on sample data consisting of 285 features (extracted from an fMRI image) and the patients' respective MMSE scores. The training data contained information for 800 patients with normalized features. The test sample consists of 200 datasets of features and a corresponding MMSE score as well.

4.4. Dimensionality Reduction. For function collection and reduction, the well-known PCA approach is used. PCA is a commonly utilized strategy for reducing dimensionality, extraction of features, and visualization of results. PCA can be described as the information's orthogonal projection into a low-dimensional, linear space known as the principal spaces. The predicted data variance rises. PCA diminishes the mean projection cost, defined as the mean square distance between the data points and their projections [60]. The value of characteristics is sorted in a descending order to find a sufficient number of characteristics. The total standard value summation (NCSE(i)) is then calculated as the corresponding sorted value:

$$\text{NCSE}(i) = \frac{\sum_{n=1}^{i} \text{eigenvalue}(n)}{\sum_{n=1}^{N_f} \text{eigenvalue}(n)}, \quad 1 \leq i \leq N_f, \qquad (7)$$

where the nth function's value is eigenvalue (n) and the dimensionality of the function vector obtained by the PCA method is N_f. The result of feature reduction is depicted in Figure 4. Based on the chart, the minimum value of features with maximum variance should be chosen. Based on results, 167 features contain 98% variance of all 285 features. Therefore, classification should continue with these 167 features, regarding this reduction number of features decremented by 41.4%.

The results of classification with several methods of machine learning consisting of KNN, SVM, decision tree (DT), linear discrimination analysis (LDA), and random forest (RF) are illustrated in Figure 5. Regarding the confusion matrix of Figure 5, the green arrays show the true values, and red elements indicate false ones. The classification is performed based on four classes, including low, mild, moderate, and, severe based on the MMSE scoring system. The
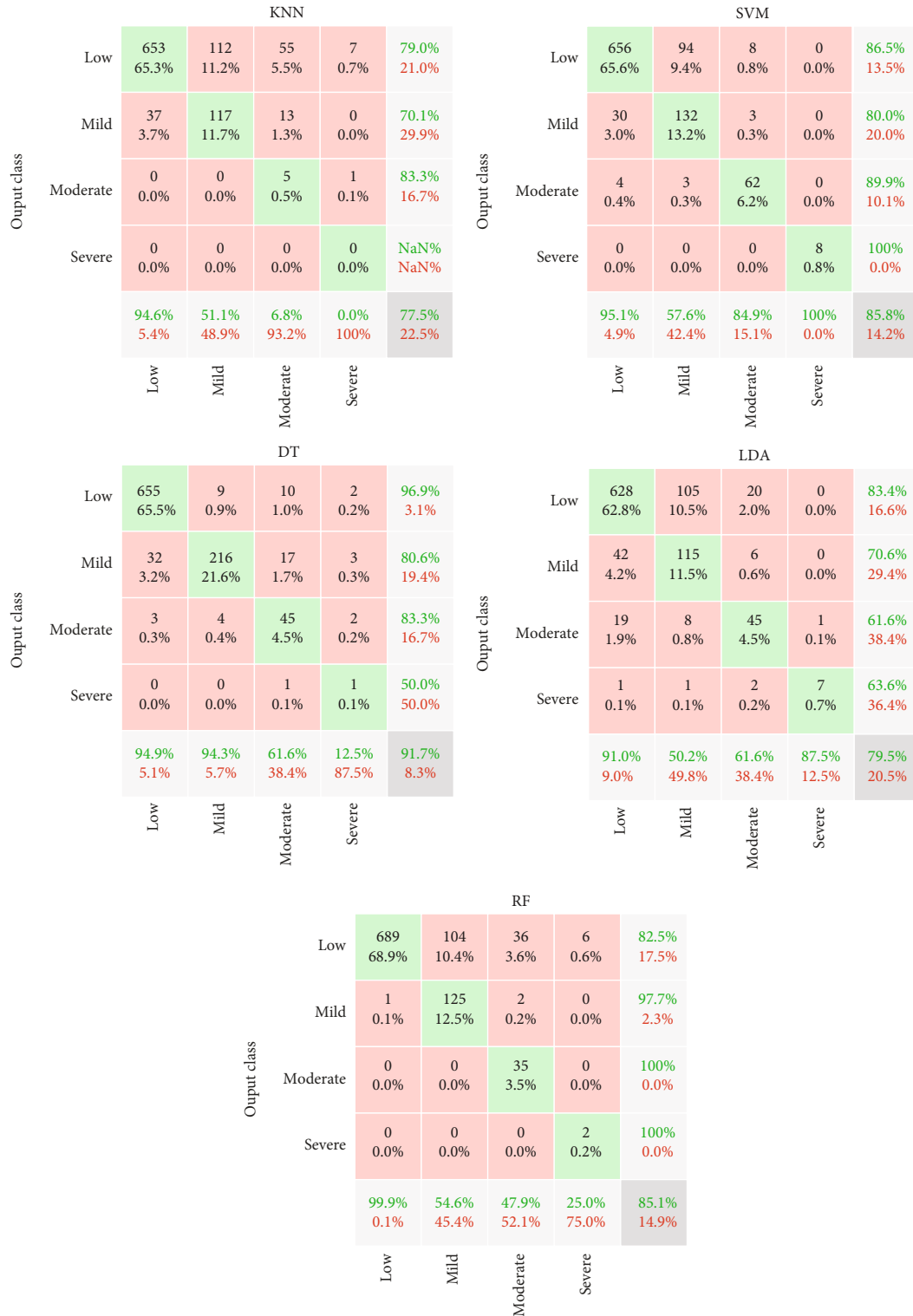
FIGURE 5: Confusion matrix of machine learning methods.

horizontal gray cells indicate sensitivity, and vertical cells illustrate precision values for each class. For instance, in the SVM method, from 690 patients with low severity, 656 (94.1%) are diagnosed correctly. However, 30 of them are misdiagnosed with mild, and four are detected with moderate severity. In other words, the sensitivity of low, mild, moder-

ate, and severe is 95.1%, 57.6%, 84.9%, and 100%, respectively. Moreover, in the RF classifier, from all detected patients in the mild class, 97.7% (precision) are true. On the other hand, the precision of low, mild, moderate, and severe classes for RF classifier is 82.5%, 97.7%, 100%, and 100%, respectively. The value of the lower-
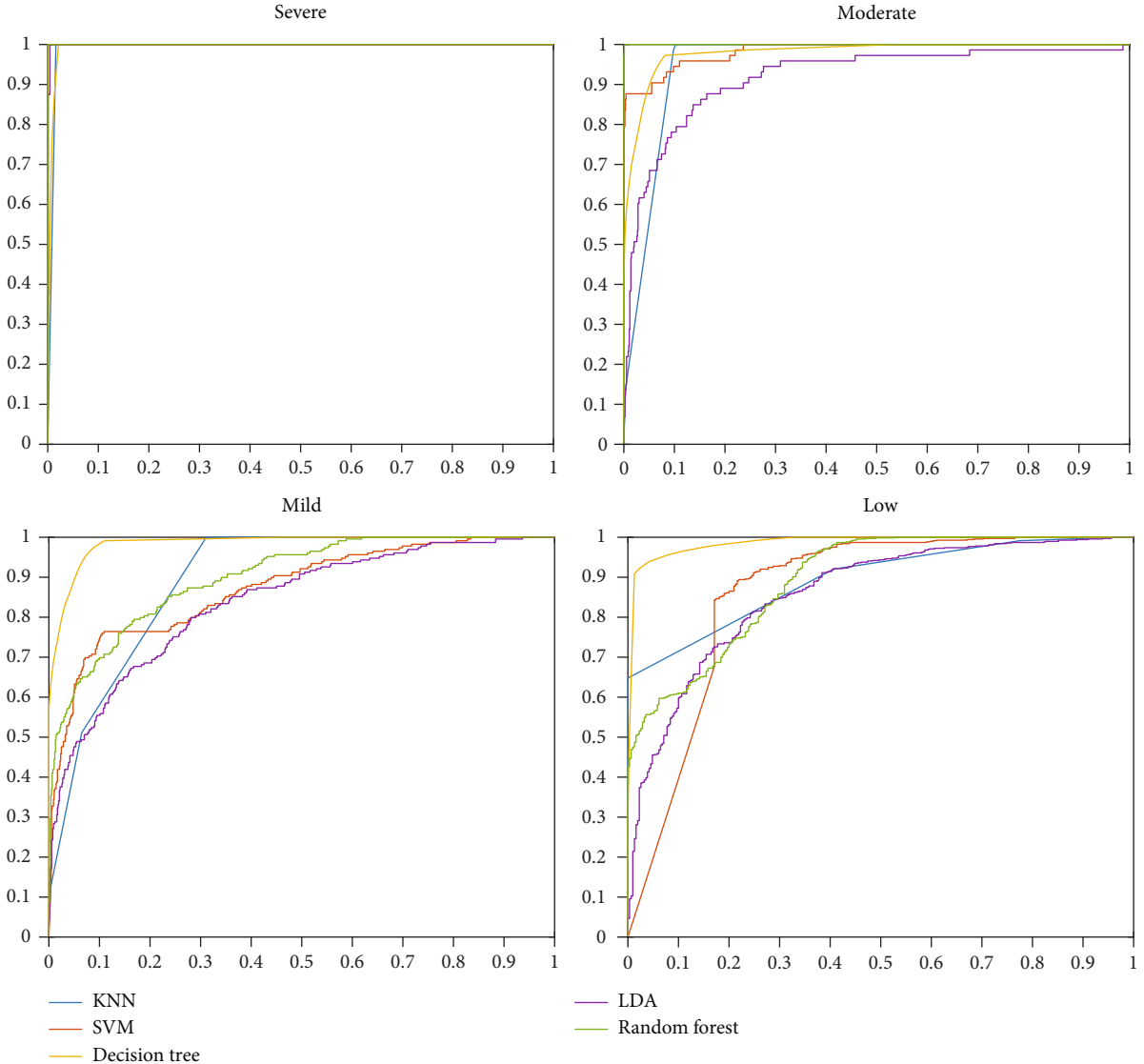
Figure 6: ROC curves of machine learning methods.

Table 3: The architecture of the presented CNN method.

Layer	Type	Properties
1	Feature input	$167 \times 1 \times 1$ images
2	Convolution	16 (5×5) convolutions with stride [1]
3	ReLU	$F(x) = \max(0, x)$
4	Fully connected	384 fully connected layer
5	Fully connected	384 fully connected layer
6	Fully connected	Four fully connected layer
7	SoftMax	$\sigma(x)_i = \dfrac{e^{x_i}}{\sum_{j=1}^{K} e^{x_j}}, \, i = 1, \cdots, K \, x = (x_1, \cdots, x_K)$
8	Classification output	For multiclass grouping problems with mutually exclusive groups, the crossentropy loss

right corner cell in the confusion matrix is the total accuracy value. To conclude, the results show that the accuracy of KNN, SVM, DT, LDA, and RF methods is 77.5%, 85.8%, 91.7%, 79.5%, and 85.1%, respectively. Moreover, the total error value of the classifier is illustrated in the lower-right corner with red text. Results indicated that from all traditional classifiers, DT results with high accuracy than other methods.

FIGURE 7: The accuracy and the loss value for the presented CNN architecture.

FIGURE 8: The confusion matrix of the presented CNN method.

Table 4: Comparison of the diagnosis methods used in this paper.

	Class	KNN	SVM	DT	LDA	RF	Presented CNN
Sensitivity	Low	94.6%	95.1%	94.9%	91.0%	99.9%	98.1%
	Mild	51.1%	57.6%	94.3%	50.2%	54.6%	95.2%
	Moderate	6.8%	84.9%	61.6%	61.6%	47.9%	89.0%
	Severe	0.0%	100%	12.5%	87.5%	25.0%	87.5%
Precision	Low	79.0%	86.5%	96.9%	83.4%	82.5%	98.1%
	Mild	70.1%	80.0%	80.6%	70.6%	97.7%	92.4%
	Moderate	83.3%	89.9%	83.3%	61.6%	100%	97.0%
	Severe	0%	100%	50.0%	63.6%	100%	100%
Accuracy		77.5%	85.8%	91.7%	79.5%	85.1%	96.7%

For a better analysis of the machine learning classifiers, the ROC curve is represented in Figure 6. For each of the classes, the ROC curve is different because it is plotted based on binary classification. The horizontal axis displays the ROC curve's false-positive trend, and its vertical axis shows the true-positive rate. In other words, the ROC curve is depicted, with consideration of each class as the positive state. Based on the ROC curve, if the values are observed with a low, false-positive rate and high true-positive rate, it is considered desirable. One of the essential criteria for the classifier's performance analysis is the area under the curve of ROC curve called AUC. It can be seen that the DT classifier resulted in high AUC than other methods. Furthermore, the AUC value for the severe class is almost identical, almost 100%.

Based on robust multitask features and MMSE score results, a CNN architecture for assessing or diagnosing Alzheimer's patient severity in this article is presented. The input layer consists of 167 features for every 1000 patients. Therefore, input matrix size is 167×1. For the convolutional layer, 16 filters with 5×5 size are used with stride [1] and zero padding. Moreover, for activating the layers, the ReLU function is used to vanish the negative values. Then, four fully connected layers are used with 384, 384, 384, and 4, respectively. Finally, the SoftMax layer is used to find probability and to activate the final layers. Then, the classification layer is used based on the crossentropy considering mutually exclusive classes. The architecture of the CNN layer is shown in Table 3.

The results of the classification process are indicated in Figure 7. The process is performed with core i7, Intel processor with 3 GHz CPU and 12GB RAM. The training process is done for 420 iterations. The accuracy and loss value of the training process is depicted in Figure 7. Furthermore, the confusion matrix of the presented CNN method is illustrated in Figure 8. Based on the low, mild, moderate, and severe status of Alzheimer patients, the sensitivity is 98.1%, 95.2%, 89.0%, and 87.5%, respectively. Moreover, the precision value for low, mild, moderate, and severe is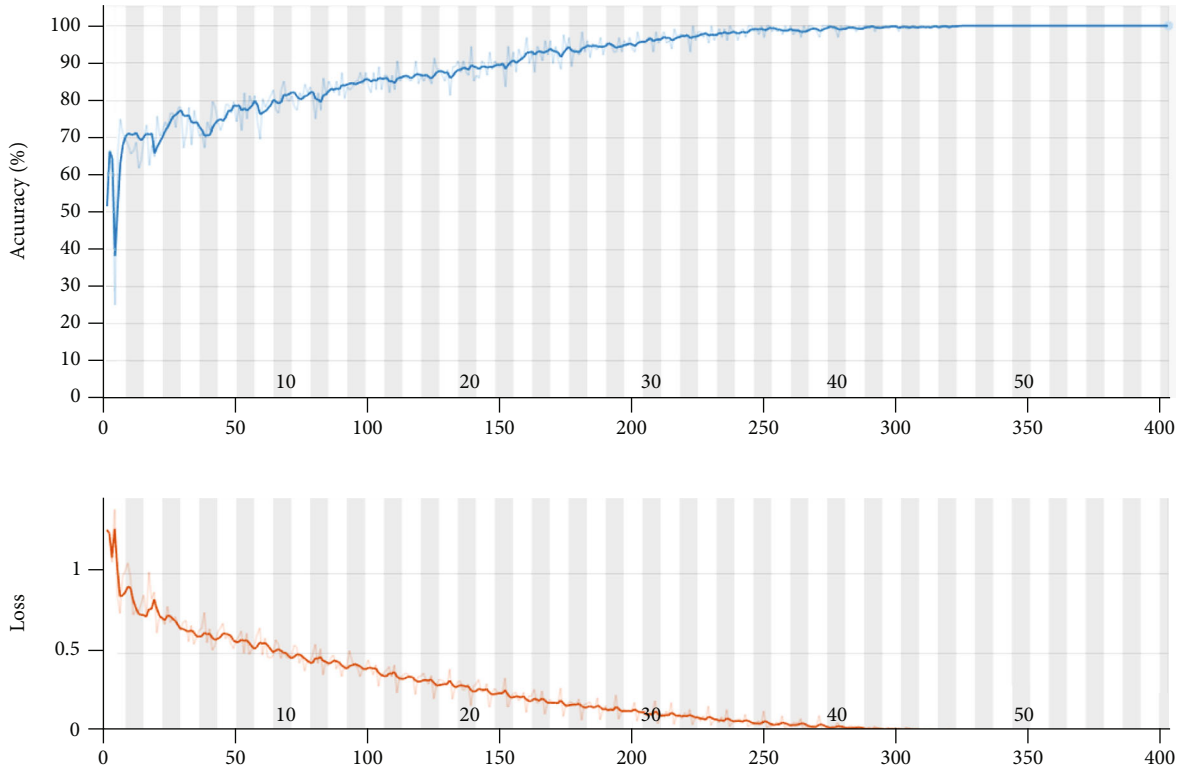 98.1%, 92.4%, 97.0%, and 100%, respectively. The absolute accuracy is also 96.7%. The summary of the results and comparison of the different classifiers are indicated in Table 4.

The results of the comparison between the presented architecture and traditional machine learning methods are shown in Table 4. Based on results, the sensitivity of the pre-sented method outperforms other approaches. The sensitivity indicates the power of the method to diagnose disease severity based on the inputs. Therefore, the magnitude of it represented the potential of the classifiers. In other words, the sensitivity of the proposed CNN architecture is higher than that of other methods. The precision also shows the potential of results or reliability of the method. For instance, the precision of the CNN method is 98.1% for the low class. It means that, from all patients that the CNN recognized as low-severity patients, 98.1% are correct. To conclude the results, the presented CNN method's accuracy is 96.7% and higher than other methods. In the next priority, DT, SVM, RF, LDA, and KNN indicate high accuracy, respectively.

5. Conclusion

AD is an incurable brain illness affecting a large percentage of the planet. To enhance patients' lives and establish effective care and targeted drugs, early detection of AD is critical. The machine learning approaches are used to diagnose the seriousness of AD focused on fMRI images. To start the training process, matched-filter technique is applied to increase the contrast of the 3D images and decrease the noise or outlier of images. The ADNI containing fMRI data of 675 patients is used. The fMRI data include 285 features base on the robust multitask feature learning algorithm. The response (target) is the Mini-Mental State Examination score that shows the severity of AD including low, mild, moderate, and severe categories.

Furthermore, the machine learning model's training task is implemented using sample data consisting of 285 features (extracted from an fMRI image) and the patients' respective MMSE scores. The training data contained information for 800 patients with normalized features. The test sample consists of 200 datasets of features and a corresponding MMSE score as well. Then, the PCA approach is used for feature selection and reduction. Based on results, 167 features contain 98% variance of all 285 features. The classification is performed with several machine learning methods consisting of KNN, SVM, DT, LDA, random forest (RF), and CNN. The results show that the accuracy of the KNN, SVM, DT, LDA RF, and presented CNN method is 77.5%, 85.8%, 91.7%, 79.5%, 85.1%, and 96.7%, respectively. For the presented CNN architecture, for the low, mild, moderate, and severe

status of Alzheimer patients, the sensitivity is 98.1%, 95.2%,89.0%, and 87.5%, respectively. Moreover, the precision value for low, mild, moderate, and severe is 98.1%, 92.4%, 97.0%, and 100%, respectively. In the next priority, DT, SVM, RF, LDA, and KNN indicate high accuracy, respectively. The detection of the severity of AD could help discover medications by having improved pathogenesis for evaluating the efficacy of target therapies that can delay the development of the disease. It can help recognize patterns of brain structural changes associated with the progression of Alzheimer's by combining clinical imaging with DL methods that can help identify risk factors and prognostic markers.

Data Availability

Data used in this paper's preparation was obtained from the ADNI database (http://adni.loni.usc.edu/).

Conflicts of Interest

The authors declare that there are no financial or other conflicts of interest in this research and its publication.

References

[1] J. Ottoy, E. Niemantsverdriet, J. Verhaeghe et al., "Association of short-term cognitive decline and MCI-to-AD dementia conversion with CSF, MRI, amyloid-and 18F-FDG-PET imaging," *NeuroImage: Clinical*, vol. 22, p. 101771, 2019.

[2] R. Balachandar, S. Bharath, J. P. John et al., "Resting-state functional connectivity changes associated with visuospatial cognitive deficits in patients with mild Alzheimer disease," *Dementia and geriatric cognitive disorders*, vol. 43, no. 5-6, pp. 229–236, 2017.

[3] S. Gardini, A. Venneri, F. Sambataro et al., "Increased functional connectivity in the default mode network in mild cognitive impairment: a maladaptive compensatory mechanism associated with poor semantic memory performance," *Journal of Alzheimer's Disease*, vol. 45, no. 2, pp. 457–470, 2015.

[4] C. Rosazza and L. Minati, "Resting-state brain networks: literature review and clinical applications," *Neurological sciences*, vol. 32, no. 5, pp. 773–785, 2011.

[5] D. M. Cole, S. M. Smith, and C. F. Beckmann, "Advances and pitfalls in the analysis and interpretation of resting-state FMRI data," *Frontiers in systems neuroscience*, vol. 4, p. 8, 2010.

[6] A. Badhwar, A. Tam, C. Dansereau, P. Orban, F. Hoffstaedter, and P. Bellec, "Resting-state network dysfunction in Alzheimer's disease: a systematic review and meta-analysis," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 8, no. 1, pp. 73–85, 2017.

[7] T. H. H. Aldhyani, M. Alrasheedi, A. A. Alqarni, M. Y. Alzahrani, and A. M. Bamhdi, "Intelligent hybrid model to enhance time series models for predicting network traffic," *IEEE Access*, vol. 8, pp. 130431–130451, 2020.

[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] M. Z. Alom, T. M. Taha, C. Yakopcic et al., "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, no. 3, p. 292, 2019.

[10] D. Ravi, C. Wong, F. Deligianni et al., "Deep learning for health informatics," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 4–21, 2017.

[11] T. Horikawa, M. Tamaki, Y. Miyawaki, and Y. Kamitani, "Neural decoding of visual imagery during sleep," *Science*, vol. 340, no. 6132, pp. 639–642, 2013.

[12] S. Koyamada, Y. Shikauchi, K. Nakae, M. Koyama, and S. Ishii, "Deep learning of fMRI big data: a novel approach to subject-transfer decoding," 2015, https://arxiv.org/abs/1502.00093.

[13] H. Jang, S. M. Plis, V. D. Calhoun, and J. H. Lee, "Task-specific feature extraction and classification of fMRI volumes using a deep neural network initialized with a deep belief network: evaluation using sensorimotor tasks," *Neuro Image*, vol. 145, no. Part B, pp. 314–328, 2017.

[14] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," 2012, https://arxiv.org/abs/1207.0580.

[15] S. Koyamada, M. Koyama, K. Nakae, and S. Ishii, "Principal sensitivity analysis," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 621–632, Ho Chi Minh City, Vietnam, 2015.

[16] T. Schmah, G. E. Hinton, R. S. Zemel, S. L. Small, and S. Strother, "Generative versus discriminative training of RBMS for classification of fMRI images," in *21st International Conference Neural Information Processing Systems*, pp. 1409–1416, Vancouver, Canada, 2008.

[17] J. H. Cole, R. P. K. Poudel, D. Tsagkrasoulis et al., "Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker," *NeuroImage*, vol. 163, pp. 115–124, 2017.

[18] D. P. Kuang and L. H. He, "Classification on ADHD with deep learning," in *2014 International Conference on Cloud Computing and Big Data*, pp. 27–32, Wuhan, China, 2014.

[19] Z. Yang, S. H. Zhong, A. Carass, S. H. Ying, and J. L. Prince, "Deep learning for cerebellar ataxia classification and functional score regression," in *International Workshop on Machine Learning in Medical Imaging*, pp. 68–76, Boston, MA, USA, 2014.

[20] H. C. Kim and J. H. Lee, "Deep neural network predicts emotional responses of the human brain from functional magnetic resonance imaging," *NeuroImage*, vol. 186, pp. 607–627, 2019.

[21] H. T. Gorji and J. Haddadnia, "A novel method for early diagnosis of Alzheimer's disease based on pseudo Zernike moment from structural MRI," *Neuroscience*, vol. 305, pp. 361–371, 2015.

[22] J. S. Benoit, W. Chan, L. Piller, and R. Doody, "Longitudinal sensitivity of Alzheimer's disease severity staging," *American Journal of Alzheimer's Disease & Other Dementias®*, vol. 35, Article ID 1533317520918719, 2020.

[23] F. Li, L. Tran, K. H. Thung, S. W. Ji, D. G. Shen, and J. Li, "A robust deep model for improved classification of AD/MCI patients," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1610–1616, 2015.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1097–1105, Lake Tahoe, CA, USA, 2012.

[25] J. Liu, M. Li, Y. Pan et al., "Complex brain network analysis and its applications to brain disorders: a survey," *Complexity*, vol. 2017, Article ID 8362741, 27 pages, 2017.

[26] J. Liu, Y. Pan, M. Li et al., "Applications of deep learning to MRI images: a survey," *Big Data Mining and Analytics*, vol. 1, no. 1, pp. 1–18, 2018.

[27] O. S. Soliman and E. A. El Hamd, "Classification of breast cancer using differential evolution and least squares support vector machine," *International Journal of Emerging Trends and Technology in Computer Science*, vol. 3, no. 2, pp. 155–161, 2014.

[28] S. Wang, H. Wang, A. C. Cheung, Y. Shen, and M. Gan, "Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease," in *Deep learning applications*, pp. 53–73, Springer, Singapore, 2020.

[29] S. Esmaeilzadeh, D. I. Belivanis, K. M. Pohl, and E. Adeli, "End-to-end Alzheimer's disease diagnosis and biomarker identification," in *International Workshop on Machine Learning in Medical Imaging*, pp. 337–345, Springer, Cham, 2018.

[30] A. Ortiz, J. Munilla, F. J. Martínez-Murcia, J. M. Górriz, and J. Ramírez, "Alzheimer's disease neuroimaging initiative. Learning longitudinal MRI patterns by SICE and deep learning: assessing the Alzheimer's disease progression," in *Annual conference on medical image understanding and analysis*, pp. 413–424, Cham, 2017.

[31] C. D. Billones, O. J. Demetria, D. E. 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 (TENCON)*, pp. 3724–3727, Singapore, 2016.

[32] H. I. Suk and D. Shen, "Deep learning-based feature representation for AD/MCI classification," in *International conference on medical image computing and computer-assisted intervention*, pp. 583–590, Berlin, Heidelberg, 2013.

[33] H. I. Suk and S. W. Lee, "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis," *NeuroImage*, vol. 101, pp. 569–582, 2014.

[34] S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis, and D. Feng, "Early diagnosis of Alzheimer's disease with deep learning," in *2014 IEEE 11th international symposium on biomedical imaging (ISBI)*, pp. 1015–1018, Beijing, China, 2014.

[35] C. Zu, B. Jie, M. Liu, S. Chen, D. Shen, and D. Zhang, "Label-aligned multitask feature learning for multimodal classification of Alzheimer's disease and mild cognitive impairment," *Brain imaging and behavior*, vol. 10, no. 4, pp. 1148–1159, 2016.

[36] S. Sarraf and G. Tofighi, "Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data," in *2016 future technologies conference (FTC)*, pp. 816–820, San Francisco, CA, USA, 2016.

[37] S. Sarraf and G. Tofighi, "Alzheimer's disease neuroimaging initiative," in *DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI*, no. - article 070441, 2016BioRxiv, 2016.

[38] F. Li, D. Cheng, and M. Liu, "Alzheimer's disease classification based on combination of multi-model convolutional networks," in *2017 IEEE international conference on imaging systems and techniques (IST)*, pp. 1–5, Beijing, China, 2017.

[39] N. Amoroso, D. Diacono, A. Fanizzi et al., "Deep learning reveals Alzheimer's disease onset in MCI subjects: results from an international challenge," *Journal of neuroscience methods*, vol. 302, pp. 3–9, 2018.

[40] M. Liu, D. Cheng, K. Wang, and Y. Wang, "Multi-modality cascaded convolutional neural networks for Alzheimer's dis-

ease diagnosis," *Neuroinformatics*, vol. 16, no. 3-4, pp. 295–308, 2018.

[41] C. Yang, A. Rangarajan, and S. Ranka, "Visual explanations from deep 3D convolutional neural networks for Alzheimer's disease classification," *AMIA Annual Symposium Proceedings*, , , American Medical Informatics Association, p. 1571, 2018.

[42] S. H. Wang, P. Phillips, Y. Sui, B. Liu, M. Yang, and H. Cheng, "Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling," *Journal of medical systems*, vol. 42, no. 5, p. 1, 2018.

[43] A. Khvostikov, K. Aderghal, J. Benois-Pineau, A. Krylov, and G. Catheline, "3D CNN-based classification using sMRI and MD-DTI images for Alzheimer's disease studies," 2018, https://arxiv.org/abs/1801.05968.

[44] 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, 2018.

[45] F. Ramzan, M. U. Khan, A. Rehmat et al., "A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks," *Journal of medical systems*, vol. 44, no. 2, pp. 1–6, 2020.

[46] H. S. Parmar, B. Nutter, R. Long, S. Antani, and S. Mitra, "Deep learning of volumetric 3D CNN for fMRI in Alzheimer's disease classification," in *Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 11317, article 113170C, International Society for Optics and Photonics, 2020.

[47] N. T. Duc, S. Ryu, M. N. 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.

[48] W. Li, X. Lin, and X. Chen, "Detecting Alzheimer's disease based on 4D fMRI: an exploration under deep learning framework," *Neurocomputing*, vol. 388, pp. 280–287, 2020.

[49] F. E. Al-Khuzaie, O. Bayat, and A. D. Duru, "Diagnosis of Alzheimer disease using 2D MRI slices by convolutional neural network," *Applied Bionics and Biomechanics*, vol. 2021, 9 pages, 2021.

[50] B. Bhaskaran and K. Anandan, "Assessment of graph metrics and lateralization of brain connectivity in progression of Alzheimer's disease using fMRI," in *Research anthology on diagnosing and treating neurocognitive disorders*, pp. 589–610, IGI Global, 2021.

[51] Y. Luo, T. Sun, C. Ma et al., "Alterations of brain networks in Alzheimer's disease and mild cognitive impairment: a resting state fMRI study based on a population-specific brain template," *Neuroscience*, vol. 452, pp. 192–207, 2021.

[52] 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. 30, pp. 1–3, 2021.

[53] A. Valliani and A. Soni, "Deep residual nets for improved Alzheimer's diagnosis," in *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, p. 615, Boston Massachusetts, USA, 2017.

[54] A. Payan and G. Montana, "Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks," 2015, https://arxiv.org/abs/1502.02506.

[55] 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, 16 pages, 2021.

[56] P. Gong, J. Ye, and C. Zhang, "Robust multitask feature learning," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 895–903, Beijing, China, 2012.

[57] A. Dhillon and G. K. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection," *Progress in Artificial Intelligence*, vol. 9, no. 2, pp. 85–112, 2020.

[58] 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 110170, 2020.

[59] M. F. Folstein, S. E. Folstein, and P. R. McHugh, ""Mini-mental state": a practical method for grading the cognitive state of patients for the clinician," *Journal of psychiatric research*, vol. 12, no. 3, pp. 189–198, 1975.

[60] S. Theodoridis and K. Koutroumbas, "Pattern recognition and neural networks," in *Advanced Course on Artificial Intelligence*, pp. 169–195, Springer, Berlin, Heidelberg, 1999.

Research Article

Diagnosis of Alzheimer's Disease by Time-Dependent Power Spectrum Descriptors and Convolutional Neural Network Using EEG Signal

Morteza Amini [ID],[1] **Mir Mohsen Pedram** [ID],[2,3] **AliReza Moradi** [ID],[4,5] **and Mahshad Ouchani** [ID][6]

[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]*Institute for Cognitive and Brain Science, Shahid Beheshti University, Tehran, Iran*

Correspondence should be addressed to Mir Mohsen Pedram; pedram@khu.ac.ir

Using strategies that obtain biomarkers where early symptoms coincide, the early detection of Alzheimer's disease and its complications is essential. Electroencephalogram is a technology that allows thousands of neurons with equal spatial orientation of the duration of cerebral cortex electrical activity to be registered by postsynaptic potential. Therefore, in this paper, the time-dependent power spectrum descriptors are used to diagnose the electroencephalogram signal function from three groups: mild cognitive impairment, Alzheimer's disease, and healthy control test samples. The final feature used in three modes of traditional classification methods is recorded: k-nearest neighbors, support vector machine, linear discriminant analysis approaches, and documented results. Finally, for Alzheimer's disease patient classification, the convolutional neural network architecture is presented. The results are indicated using output assessment. For the convolutional neural network approach, the accurate meaning of accuracy is 82.3%. 85% of mild cognitive impairment cases are accurately detected in-depth, but 89.1% of the Alzheimer's disease and 75% of the healthy population are correctly diagnosed. The presented convolutional neural network outperforms other approaches because performance and the k-nearest neighbors' approach is the next target. The linear discriminant analysis and support vector machine were at the low area under the curve values.

1. Introduction

The term "dementia" refers to many neurodegenerative illnesses caused by neuronal failure and death that interrupt cognitive and behavioral activities. The most prevalent of the several types of dementia is Alzheimer's disease (AD), with about 70% of worldwide dementia cases. It affects the individual over 65 years, and the rate of occurrence increases exponentially at the age of 65 years [1–3]. To date, AD has not been resolved by palliative therapies, which have been temporarily slow to deteriorate in patients and caregiver living [4]. Today, only postmortem diagnosis of definitive AD is possible after examining the structural brain injury that is typical of the condition. Accuracies of up to 90% have usually been recorded for modern testing procedures, such as neurological assessments and medical history. The National Institute on Aging and Alzheimer's Association has established the existing standards of clinical diagnosis of AD, and the Alzheimer's Association has established them [5]. These standards are an advancement in the previous guidelines,

which had been developed in 1984 by the National Institute of Neurological And Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) [6]. It is part of the NINCDS-ADRDA guideline. These revised suggestions require neuroimagery and the use of biomarkers and cerebrospinal fluid to diagnose AD for symptomatic people [5].

A guideline for diagnosing and monitoring AD [7] was established by the European Federation of Neurological Societies. The Mini-Mental State Assessment [8, 9] is the most used AD diagnosis method to test cognitive ability. The revised Montreal Cognitive Assessment [10] is commonly used in therapeutic functional applications and the revised Addenbrooke Cognitive Evaluation [11]. Another example of neurological testing is the Extreme Cognitive Disorder, Alzheimer's Cognitive Disease Evaluation Scale, Neuropsychological Assessment Battery, and Serious Impairment Battery. The Trail Making Test [12] and the clock drawing test [13], by contrast, focus not only on testing thinking abilities but also on concentrating and administrative work. In comparison, the visual learning test and the Rey Auditory Fluency Assessment assess all patient support practice skills [14]. In some instances, AD is also associated with other diseases that cause dementia as brain vascular injury, Lewy body disease, and Parkinson's disease [15]. The early diagnosis of AD and these problems is improved using methods that gain biomarkers as early signs overlap [16–19]. Electroencephalogram (EEG) is a technology that enables the recording by postsynaptic potentials of a thousand neurons of identical spatial orientation of the time of cerebral cortex electrical activity. Scalp-positioned electrodes measure the electrical potentials. EEG's spatial resolution refers to the number and location of electrodes on the scalp. The most used configuration is the international 10-20 system, which consists of 21 electrodes; the 10-20 system is often used for higher density versions, for instance, 10-10 and 10-5, usually 64 and 128 electrodes, and the Maudsley [20] and geodesic positions [21] alternating layouts. Reliable therapeutic methods have been shown in recent years for the diagnosis and analysis of disorders and cortical conditions like the Huntington syndrome, the autism spectrum disorder [22], epilepsy and seizure [23], brain ischemia [24], frontotemporal dementia [25], and Parkinson's dementia [26]. Furthermore, EEG evaluations were carried out on the comparative diagnosis of AD and other dementia-contributing diseases such as brain vascular injuries [27, 28] and Lewy disease [29, 30]. Theta (θ) 4–8 Hz, delta (δ) 0.1–4 Hz, beta (β) 12–30 Hz, alpha (α) 8–12 Hz, and gamma (γ) > 30 Hz are typically divided between 5 major frequency bands in the analytics. Also, more divisions into these bands (high alpha, low alpha, low beta) are considered, but the subband frequency limits are not uniform in all studies. The different data on brain function and synchronization are given in each frequency band [31–33]. There has been a comprehensive study of the possible use of electroencephalography to diagnose dementia and AD [34]. EEG is a high time resolution noninvasive, comparatively inexpensive, and potentially mobile technology (about milliseconds). It was studied primarily as an AD diagnostic tool when comparing EEG records in AD patients with control subjects (healthy individuals) [35, 36]. AD is generally known to decrease the complexity of EEG signals and synchronous change in EEG.

These improvements have been used as discriminatory features for AD diagnosis in EEG recordings. Several methods of assessing the complexity of EEG signals have been established. The connection factor and the first positive exponent of Lyapunov have been used frequently [37–42]. EEG signals from AD patients have been shown to show lower (lower complexity) values of certain tests than signals from age-matched control subjects. Other information-theoretical methods, in specific entropy-based approaches, have appeared as theoretically useful EEG indicators for AD: epoch-based entropy [43, 44], sample entropy [45], Tsallis entropy [46], approximate entropy [47, 48], multiscale entropy [49], and complexity of Lempel-Ziv [50]. These approaches relate the strength of a signal to unpredictability: irregular signals are more complicated than regular ones because they are erratic. Different detection algorithms have been suggested in previous studies for epileptiform EEG data [51]. Current seizure detection methods use hand-built feature extraction techniques from EEG signals [52], including frequency domain, time domain, time-frequency domain, and nonlinear evaluation of signals [53, 54]. The features selected must be listed after the feature extraction to identify various EEG signals using all forms of classifiers [55]. Hamad et al. employed a differential wavelet transformation procedure to obtain the feature collection, then trained the radial reference algorithm with the support vector machine (SVM), demonstrating an epilepsy diagnosis with the suggested SVM gray wolf optimizer [56]. For the refinement of the SVM parameters based on genetic algorithms and particle swarm optimization, Subasi et al. developed a hybrid model. The proposed SVM hybrid model shows that neuroscientists use EEG as an essential method for diagnosing epileptic seizures [57]. However, the manual function selection criteria are not eradicated by these strategies [58]. The feature extraction is an important stage in classification determination, as it determines its specificity in large part. A system for classifying without the removal of complicated properties was suggested. Furthermore, recent advancements in deep learning have shown a new way of coping with this problem. Deep learning has in recent years reached the recognized form of computer vision and machine learning and has demonstrated that almost all human and superhuman functions such as pattern recognition and sequence learning perform numerous functions [59], among other things. Feature extraction before classification is more advantageous than entering raw EEG samples directly into the classifier. Nevertheless, several recent research types have not performed feature extraction, but instead, raw EEG signals were used for the deep learning model [60, 61].

In this paper, the time-dependent power spectrum descriptor (TD-PSD) method is utilized for feature extraction of the EEG signal from three categories of MCI, AD, and HC sample test. The final feature with labeling is used in three types of traditional classification methods, including k-nearest neighboring (KNN), SVM, and linear discriminant analysis (LDA) approaches, and the results are recorded.

TABLE 1: Summary of machine learning method for brain disease diagnosis with EEG signal.

Author	Year	Disease	Feature extraction	Classification	Results
Xin et al. [80]	2021	Epilepsy	Dimensionality reduction principal component analysis (PCA)	Convolution SVM	The method's accuracy, sensitivity, and specificity reach up to 99.56%, 99.72%, and 99.52%, respectively
Aliyu and Lim [81]	2021	Epilepsy	Discrete wavelet transforms (DWT)	LSTM network	Reduction of the number of LSTM trainable parameters needed to achieve extreme accuracy
Tuncer [82]	2021	Epileptic seizure	Nonlinear textural feature extraction (Hamsi hash)	k-nearest neighborhood	This model has an accuracy in the EEG dataset of 99.20% for five classes and has 100.0% accuracy in other conditions
Cicalese et al. [83]	2020	AD	Pearson correlation coefficient-based feature selection (PCCFS)	LDA	The EEG-fNIRS feature set combination was expected to obtain greater precision (79.31%) by combining its supplementary properties as compared with the EEG (65.52%) or fNIRS alone (58.62%). Moreover, AD development is associated with the right and left parietal lobe
Ferri et al. [84]	2020	AD	rsEEG + sMRI	Low-resolution brain electromagnetic tomography (LORETA)	Classification accuracy of 80%, 85%, and 89% using rsEEG, sMRI, and rsEEG + sMRI features, respectively, discriminates against them
Trambaiolli et al. [85]	2017	AD	Feature selection (FS)	SVM classifier	Since eliminating 88.76 ± 1.12% of the initial elements, the filtered subset evaluator technique obtained the highest efficiency gain, both on a per-patient basis (91.18% accuracy) and on a per-epoch basis (85.29 ± 21.62%)
Nobukawa et al. [86]	2020	AD	Functional connectivity	SVM	A novel interpretation of neural network functions in healthy brains and unhealthy disorders can be provided by applying a mixture of both machine learning approaches to neurophysiological evidence
Kulkarni and Bairagi [87]	2017	AD	Extracting salient features that are spectral-, wavelet-, and complexity-based	SVM	The increased performance in AD diagnosis
Vecchio et al. [88]	2020	AD	—	SVM	A low-cost and noninvasive process uses readily available tools that, when integrated, achieve high sensitivity/specificity and optimum individual classification accuracy (0.97 of AUC)

Finally, an architecture of convolutional neural network (CNN) for AD patients' classification is provided. The results are indicated using performance analysis.

2. Literature Review

EEG signals' complex and nonlinear nature implies creating new methods to study machine and signal processing [62, 63]. Recent progress has been made to enhance high-level abstractive methods for the automated removal of complex data features in the field of deep learning methodologies [64–66]. In the last years, these deep learning methods are usually used in image processing, natural language processing [67–70], speech processing, and video games [71]. The biomedical area has also been identified with these methods [72–74]. Acharya et al. [60] suggested a deep, 13-layer neural CNN that distinguishes normal, preictal, and EEG signals of seizure. In the study, 300 EEG signals were used for registering a classification rate of 88.67%. The same group proposed

a deep neural network approach for an innovative EEG-based depression screening method [60]. This investigation's outcomes are reported in 15 regular and 15 depressive patients, 93.5% (left hemisphere) and 96.0% (right hemisphere). Oh et al. [75] suggested that further studies used EEG signals to diagnose Parkinson's disease. A 13-layer CNN model of 20 healthy and 20 Parkinson patients reached an accuracy of 88.25%. Hong et al. [76] propose a mathematical model employing Long Short-Term Memory (LSTM), a recurrent neural network (RNN) that predicts the mild cognitive impairment of AD. The data is taken in image form in this research, and the preprocessing is done by skull strip, normalization, registration, smoothing, and segmentation. The training is carried out by feeding sequential data with time steps to the model after preprocessing, and the model projects the state of the next six months. During model testing, when the feature data for the 18th and 24th months is presented, it forecasts the state of the subject for the 30th month. Similarly, Aghili et al. [77] suggest an RNN approach
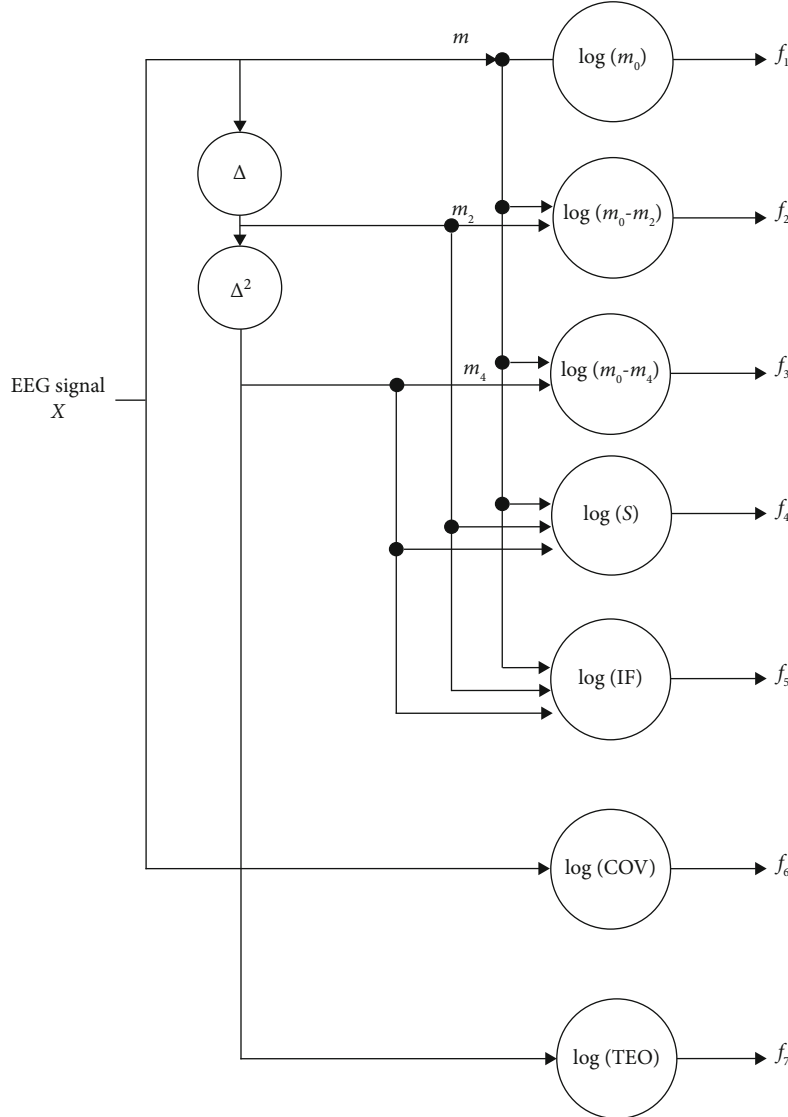
Figure 1: The block diagram of the TD-PSD feature extraction method for EEG signals.

to evaluate longitudinal data to differentiate between stable people against AD individuals. The input is preprocessed, and the function is normalized. It is fed into LSTM and gated recurring units after preprocessing the data. In the model of LSTM and gated regular units, each subject's time point data is provided to the corresponding cell along with its final diagnosis mark to learn the pattern of data transition. The effect models are contrasted with the effects of nonrecurrent networks, i.e., multilayer perceptron (MLP), for all data arrangements. For each patient, the data is fed into the MLP once. There are many trainable parameters in the LSTM models that need to be substantially trained for sequential data and are vulnerable to overfitting the training data. The purpose of unsupervised feature learning is to define AD using the principle of unsupervised feature learning. An approach used sparse filtering to learn the expressive characteristics of brain images [78]. The SoftMax regression is trained to classify the circumstances. In the first step, there are three phases: sparse filtering is trained, and its W weight matrix is calculated. The

sparse filtering learned is used to extract from each sample the local characteristics. In terms of negative matrix factorization (NMF) and SVM with limitations of certainty, Padilla et al. [79] offer a novel conclusion technique for the early determination of Alzheimer's disease. Through implementing the Fisher discriminant proportion and non-NMF for highlighting preference and extraction of the most important highlights, the single-photon emission computed tomography and positron emission tomography datasets are studied (see Table 1).

3. Research Methodology

The proposed study uses the EEG signal to describe the phases of the disorder. It is suggested that a deep CNN network architecture is learned to distinguish multichannel human EEG signal data into different stages and that increases the efficiency of classification. This work includes the modules below:

(i) Preprocessing

(ii) Feature extraction

(iii) Classification

3.1. Feature Extraction. The EEG trace is expected to be explicated in a function of frequency $X[k]$, using the discrete Fourier transform, as a product of the sampled representation of the EEG signal as $x[j]$ with $j = 1, 2, \cdots, N$, length N, and sampling frequency fs (Hz). Parseval's theorem explains that the function's total square is the complete square of its transformation; the process starts with the extraction of features.

$$\sum_{j=0}^{N-1} |x[j]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]X^*[k]| = \sum_{k=0}^{N-1} P[k]. \quad (1)$$

According to the above equation, $P[k]$ is the phase-excluded power spectrum. It means that the frequency index is obtained by multiplying $X[k]$ by the $X^*[k]$ conjugate divided by N, where the phase-excluded power spectrograph is $P[k]$, i.e., $X[k]$ has its conjugate $X^*[k]$, separated by N, which is compounded by k, and frequency index. The full definition of the frequency as obtained by the Fourier transform is usually well-known to be symmetrical concerning zero frequency; i.e., it has similar sections extending to the frequencies, which are both positive and negative [89]. The whole spectrum, including positive and negative frequencies, is free from this symmetry. Access to spectral power from the time domain has not been completed. According to the concept of a one-minute m of the order $nP[k]$ of the power spectral density, all irregular moments are also zero by the frequency distribution model's statistical approach.

$$m_n = \sum_{k=0}^{N-1} k^n P[k]. \quad (2)$$

In the latter equation, the Parseval theorem of Equation (1) may be used where $n = 0$ is used, and the Fourier transform time-differentiation feature for nonzero quantities of n is used. Such a feature explicitly indicates the n' th equal to multiply the k augmented by the spectrum to the n' th power, the derivative of a time-domain function referred to as Δ^n for various time signals.

$$F[\Delta^n x[j]] = k^n X[k]. \quad (3)$$

To this end, as seen in Figure 1, the description of the characteristics used is as follows:

Root squared zero-order moment ($\overline{m_0}$): this is a function that shows the overall power of the frequency domain and is as follows:

$$\overline{m_0} = \sqrt{\sum_{j=0}^{N-1} x[j]^2 s}. \quad (4)$$

All channels may also standardize their corresponding zero-order moments by dividing all channels into zero-order moments.

Root squared second and fourth-order moments: according to Hjorth [89], the second time is used as power, but then, a spectrum shifted $k^2 P[k]$, referring to the frequency function:

$$\overline{m_2} = \sqrt{\sum_{k=0}^{N-1} k^2 P[k]} = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} (kX[k])^2} = \sqrt{\sum_{j=0}^{N-1} (\Delta x[j])^2}. \quad (5)$$

A repetition of this procedure gives the moment:

$$\overline{m_4} = \sqrt{\sum_{k=0}^{N-1} k^4 P[k]} = \sqrt{\sum_{j=0}^{N-1} (\Delta^2 x[j])^2}. \quad (6)$$

The total energy of the signal is decreased by consideration of the second and fourth signals; thus, a power transformer is implemented to normalize the domain of m_0, m_2, and m_4 to minimize the noise effect on all moment-based features as follows:

$$m_0 = \frac{\overline{m_0}^\lambda}{\lambda},$$

$$m_2 = \frac{\overline{m_2}^\lambda}{\lambda}, \quad (7)$$

$$m_4 = \frac{\overline{m_4}^\lambda}{\lambda}.$$

The experimental setting of λ is 0.1. From these parameters, consequently, the first three features extracted are described as follows:

$$f_1 = \log (m_0),$$

$$f_2 = \log (m_0 - m_2), \quad (8)$$

$$f_3 = \log (m_0 - m_4).$$

Sparseness: this feature measures the amount of vector energy in just several more elements. It is followed as

$$f_4 = \log \left(\frac{m_0}{\sqrt{m_0 - m_2} \sqrt{m_0 - m_4}} \right). \quad (9)$$

A feature represents a vector with all elements equivalent to a zero-sparseness index, i.e., m_2 and $m_4 = 0$, due to differentiation and $\log (m_0/m_0) = 0$, while it should require a value greater than zero for all other sparseness levels [90].

Irregularity factor (IF): a measure that expresses the ratio of peak numbers divided by zero crossings. According to [91], only in terms of their spectral instances can the number of upward zero crossings (ZC) and the number of peaks (NP) in a random signal be specified. It is necessary to write the

corresponding feature as

$$f_5 = \frac{\text{ZC}}{\text{NP}} = \frac{\sqrt{m_2/m_0}}{\sqrt{m_4/m_2}} = \sqrt{\frac{m_2^2}{m_0 m_4}} = \frac{m_2}{\sqrt{m_0 m_4}}. \tag{10}$$

Covariance (COV): COV function is the ratio of the standard deviation on arithmetic averages as follows:

$$f_6 = \log \left(\frac{\sqrt{\left(\sum_{j=0}^{N-1} (x - \bar{x})^2 \right)/n}}{\bar{x}} \right). \tag{11}$$

Teager energy operator (TEO): it mainly displays the magnitude of the signal amplitude and instantaneous changes that are very susceptible to minor changes. While TEO was proposed for nonlinear speech signal modeling, it was later used to process audio signals. It is formed as follows:

$$f_7 = \log \left(\Psi(\text{x}[j]) \right) = \log \left(\sum_{j=0}^{N-1} x^2[j] - x[j-1]x[j+1] \right). \tag{12}$$

In conjunction with the schematics in Figure 1, from each EEG record x, first, the seven features are extracted. In the classification method, the features described by the corresponding vector f are used. These characteristics can be assumed to reflect the EEG behavior in the form of cepstrum. Contrary to the well-known voice cepstral features [92], our EEG features have been obtained as the orientation between characteristics derived from a nonlinear EEG record and an initial EEG record following the equation. In the case of EEG classifications at differing levels of force, orientation-based feature extraction processes have recently been demonstrated to be of considerable significance for research on intact-limbed subjects as force generation relies on multiple hand muscle coordination [93].

There have not been any prior attempts to test the efficacy of specific features in amputees. In the coming subsection, the suggested orientation-based feature is adequate for amputees to classify EEG signals with different classes. In the remainder of the essay, the last feature f is defined, together with the time-dependent descriptor spectrum, from all channels, given as TD-PSD.

3.2. Convolutional Neural Network.
CNN is one of the learning networks inspired by the MLP in this type of neural network. This deep network comprises an input layer, an output layer, and a deeply hidden layer. Firstly, the problem's signal or data are identified and trained into the algorithm. The hidden weights of the output layer appear in many forms [94]. If the algorithm output includes numerous numerical components, such as a binary number or index (e.g., signal classification, normal = 1, abnormal = 2), then the algorithm presented is a classification or detection algorithm. That is, the outcomes are weighted after the training of several signals. When a new signal is added to the algorithm other than the training signals, the signal form is identified, for example,
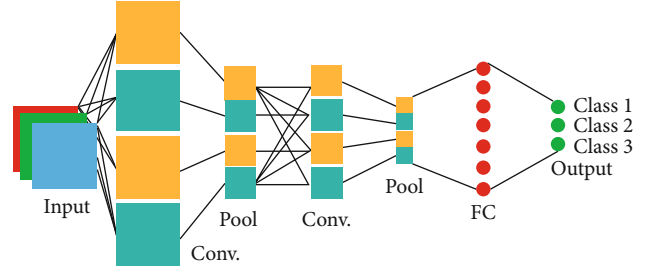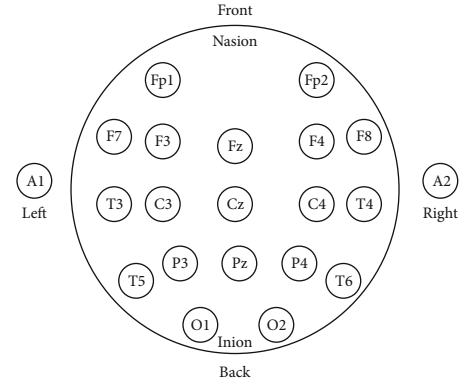

Figure 2: A conventional sublayer in CNN.

Figure 3: The place of electrodes of EGG signals.

whether a matrix of various kinds of signals is sent to the algorithm and trains the machine, signals of benign or malignant types of cancer, Alzheimer's, sarcoma, or brain tumor, for example. The type of disease can be identified by the algorithm with the weights obtained. CNN consists of numerous hidden sublayer forms that are explained as follows.

3.3. Convolutional Sublayer.
The basic of the CNN is the convolutional sublayer, and it is possible to view its output matrix as a three-dimensional neuron matrix. For a deeper explanation of this, imagine traditional neural networks. Each layer was a little more than a list (one-dimensional as a rectangle) of neurons in regular neural networks in which each neuron generated its output, and gradually, a collection of outputs referring to each neuron was produced. However, instead of a single list, it is presented with a three-dimensional list (one cube) where the neurons are organized in three dimensions. Therefore, a three-dimensional matrix would also be the production of this cube. This principle and the distinction between the two are illustrated in the images below [95].

Let the size of the input matrix be $20 \times 16 \times 16$. Thus, utilizing a receptive field of 3×3, each neuron would have $3 \times 3 \times 20 = 180$ connections to the input matrix in the conventional layer. Notice that space's connection is local (for example, 3×3 here) but covers the maximum depth (Figure 2). An input displays the left image (for example, a $3 \times 32 \times 32$ image). The neuron matrix is observed in the conventional blue sheet. In terms of spatial coordinates (length and width) in the input matrix, each neuron in the conventional layer is related to one local region only, but this connection extends
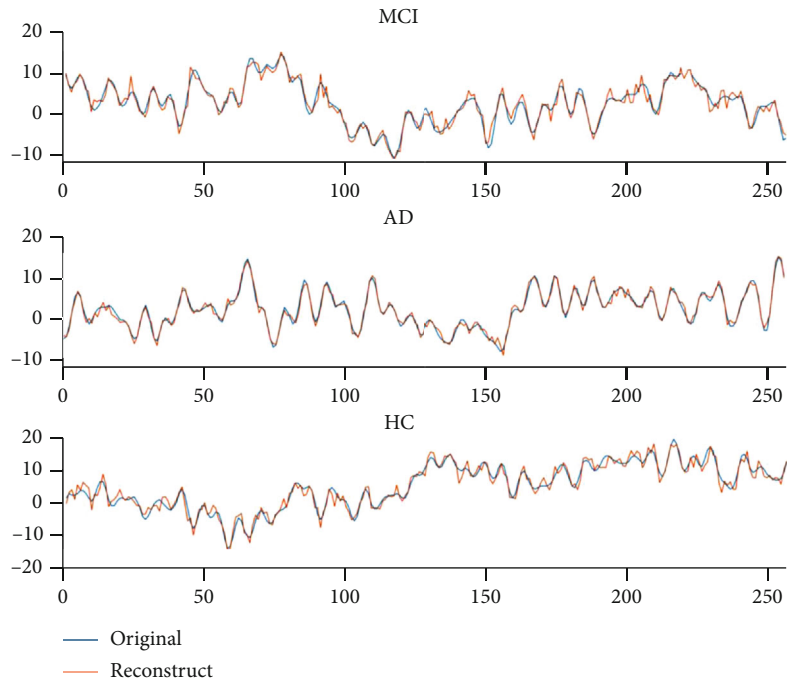
FIGURE 4: Plots of sampling from EEG signals for three categories of MCI, AD, and HC.
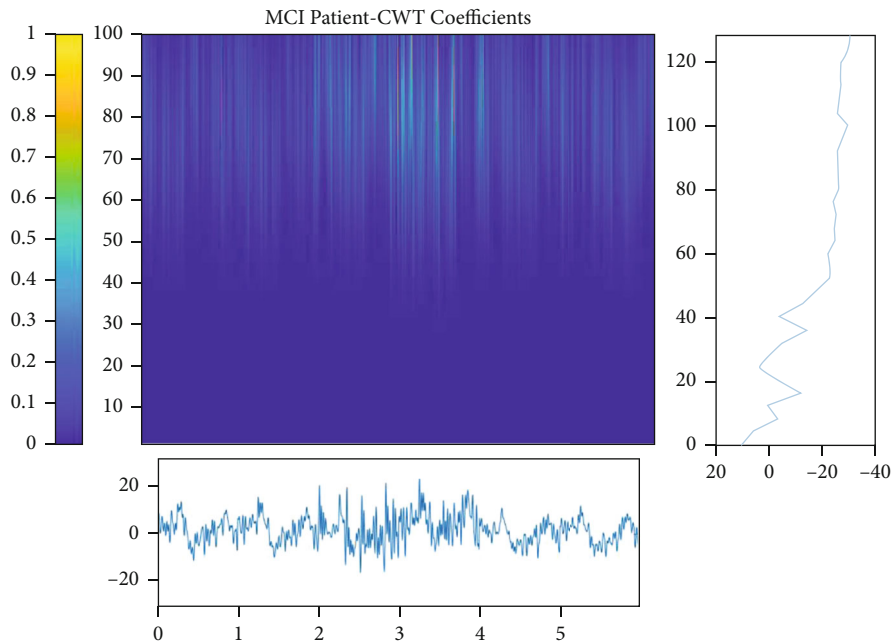

FIGURE 5: Time-frequency analysis of EEG signals for MCI patients.

in-depth (i.e., covers all color channels). There are depths (in this case, neurons) that all look at one place at the entrance.

The hyperparameter regulates the output matrix's dimension. The depth, stride, and zero-padding layer are these three parameters. The parameter can be used in the depth of the output matrix. This parameter regulates the number of neurons that bind to a region in the conventional layer's input matrix. This variable is analogous to multiple neurons in a hidden layer all attached to one input in classical neural networks. All these neurons learn to function on various feedback features in which the deep columns around the spatial dimensions must be defined (width and height). When the stride is equal to 0, the spacing coordinates of just five spaced points are allocated to a new depth column of neurons. Also, in large output matrixes, this leads to receptive areas of overlap between columns. Alternatively, the
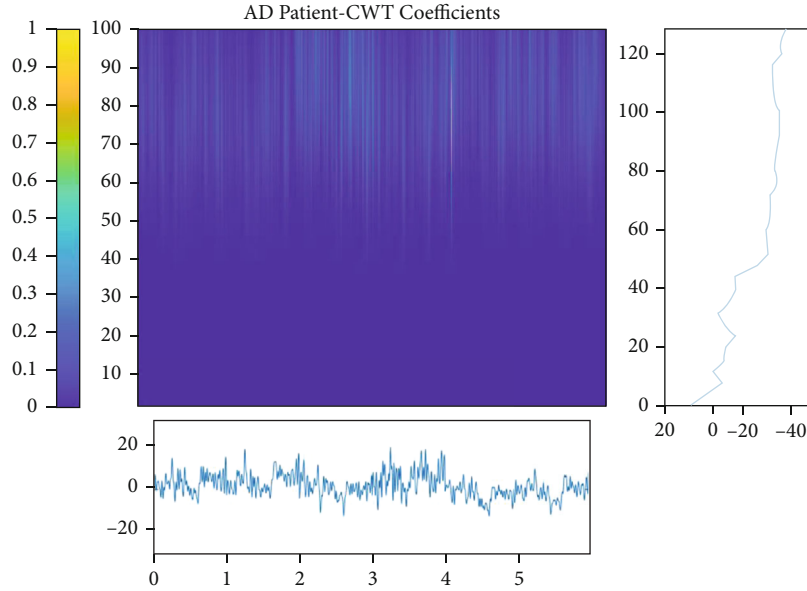
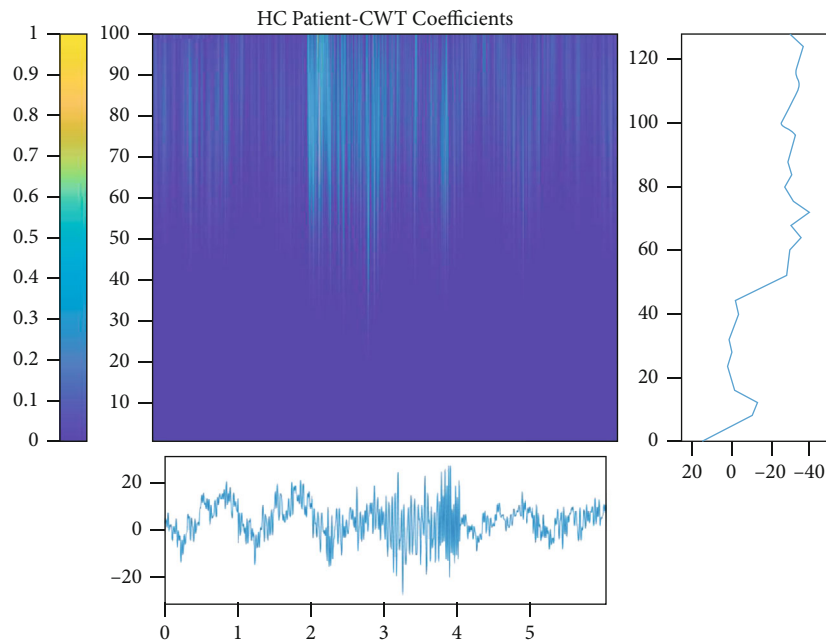FIGURE 6: Time-frequency analysis of EEG signals for AD patients.

FIGURE 7: Time-frequency analysis of EEG signals for HC samples.

receptive areas are less frequent if the measures taken are bigger, and the output mass is smaller in the spatial dimension [96]. Overriding the inbound matrix with a zero pad is also more precise. In other terms, fill zero with the input picture circle. Our signal is put within a zero signal like inserting row 1 and column 2 at the beginning and end of a signal.

3.4. Max Pool Sublayer. One standard technique in conventional architecture is the positioning of a pooling layer between many successive layers. This layer's purpose is to minimize the matrix (input) size (width and height) by reducing the number of variables and calculations within

the grid and thereby overfitting the monitor. The pooling layer functions and uses it on each depth cut of the input matrix independently. The MAX function resizes the spot. The most typical way to utilize this layer is to use this layer with filters of 2×2 sizes with phase $S = 2$ that eliminates any depth cuts at the input by deleting two elements from the width and two elements from the height and deleting 2% of the values [97].

3.5. Activation Function. Artificial neural networks' activation function determines the node's output node or "neuron" according to the input or group of inputs. In the next node,
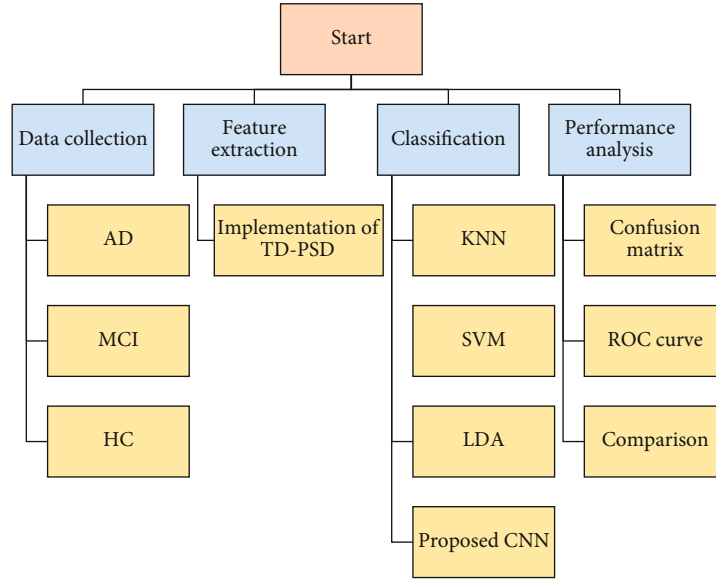
FIGURE 8: Conceptual diagram of the process used in this paper.

this output is known as the input. It follows before a solution to the issue is sought. The outcome values are translated to the target set, such as 0 to 1 or -1 to 1 (depending on the activation function selected). Using the logistic activation function, for example, transforms all inputs into true real ranges between 0 and 1. Another essential characteristic of an activation function is that it must be derivative to execute the optimization technique for backpropagation error and measure the weight gradient error in the network and use gradient descent or another optimal approach. Another optimization is optimizing weight to reduce weight. The rectified linear unit (ReLU) is used in this paper for the use of functions as follows:

$$f(x) = \begin{cases} 0, & x < 0, \\ x, & x \geq 0. \end{cases} \tag{13}$$

Some activation functions are also not unique to a single variable and refer to the vector or different variables used in this article, such as SoftMax [98]:

$$f_i\left(\vec{x}\right) = \frac{e^{x_i}}{\sum_{j=1}^{J} e^{x_j}}. \tag{14}$$

Also, to normalize input results, a batch normalization layer is applied to the network to speed up the training process and reduce network sensitivity between convolutional layers and nonlinearities. Also, to create an abnormal signal augmentation, a dropout layer on the fully connected layers is applied. Fully connected layers are used at the end of the hidden layer, which has been known to distinguish signals. The deep learning layer's outcome leads to a fully connected layer that drives the final classification judgment.

3.6. Receiver Operating Characteristic (ROC) Curve. In 2004, the ROC curves were developed, which were used to detect a radio noise signal [99]. These curves have recently been discovered to have important uses in medical decision-making. Presume having two kinds of individuals, one is normal and the other is a patient. It is a screening test on both our patients and healthy people, and the spectrum of values from the test ranges from 0 to the large number scale. In this case, the greater the test outcome, the greater the risk of the disease. (For certain things, the action can be the opposite.)

The ROC curve is established by the true-positive rate (TPR) projection in different threshold settings against the false-positive rate (FPR). The TPR is often identified in ML as sensitivity, recall or detection probability. Beginning from the ROC's left side, both the FPR and the TPR are zero at this point. (This argument indicates that the threshold line, which is the most significant number of test outcomes, is very large.) The TPR and FPR values in this case are measured and the next curve is drawn. The definition of TPR equals TP/Y, and the definition of FPR equals FP/N.

For a reduced number of previous values, let us reduce the threshold line. For lower values, the trend is replicated and eventually reaches the rightmost point of the ROC curve, which is equal to the baseline of the lowest value of the test outcome in this case. In this function, there is one TPR and one FPR. Accuracy relies entirely on allocating random errors and does not correspond with the real value or the value specified. In terms of bias, correctness is conveyed. A complete structural error can consist of one or more components of a systematic error. In the comparison value, a strong bias implies a large disparity. The two variables' sensitivity and specificity in statistics were used to determine the binary classification outcome (duality). When the data can be broken into positive and negative classes, using sensitivity and attribute indicators, the consistency of a test's
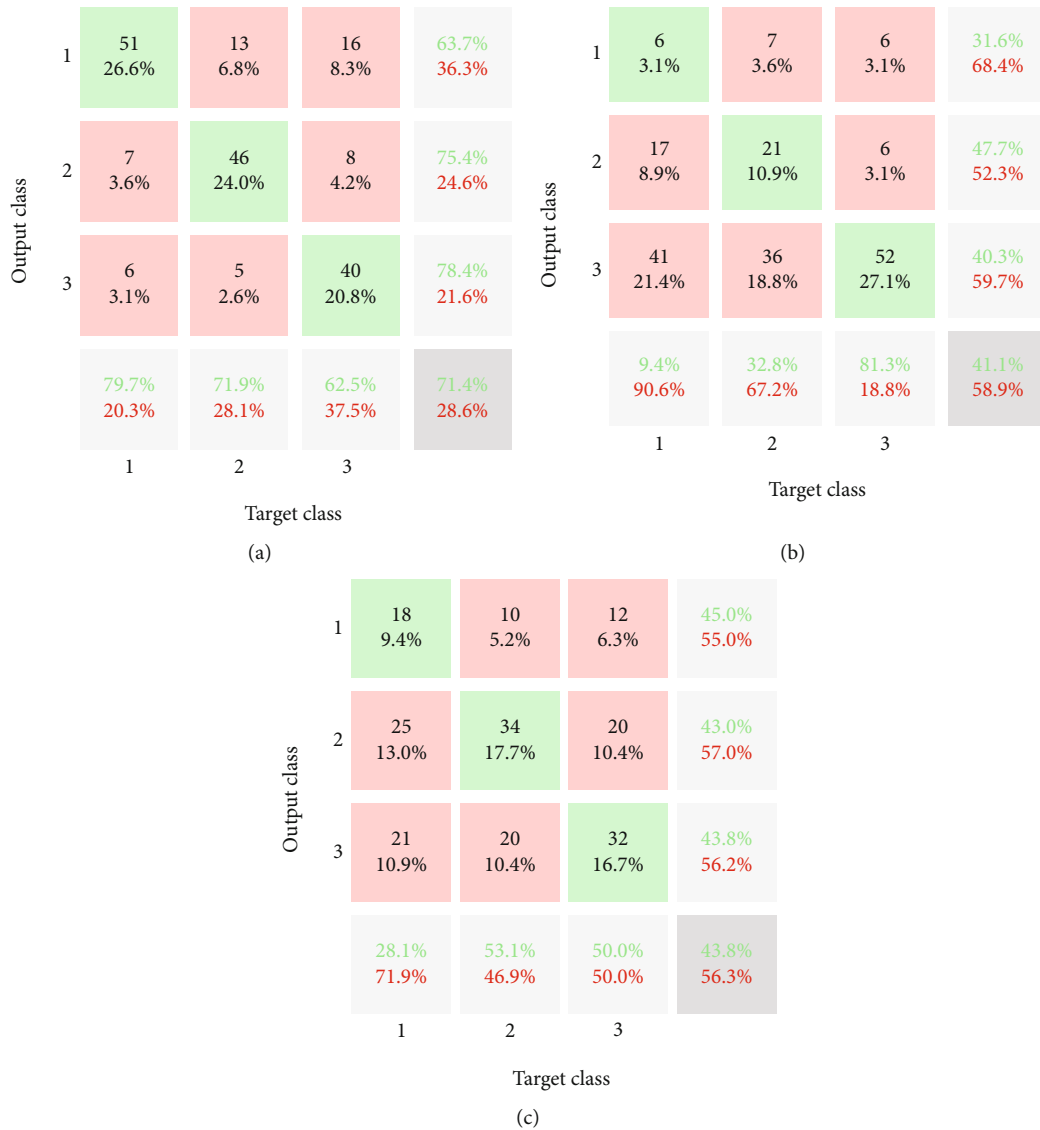
Figure 9: Confusion matrix of classification methods: (a) KNN, (b) SVM, and (c) LDA.
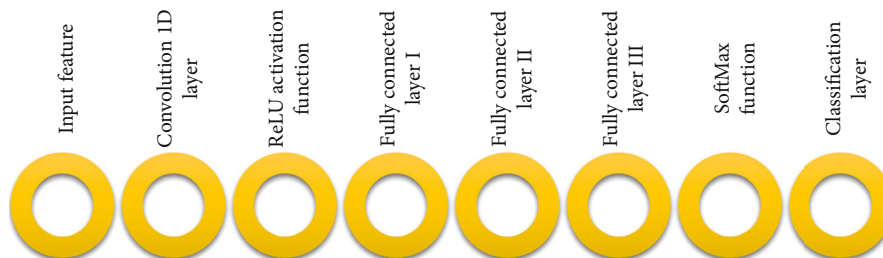
Figure 10: The architecture of the CNN layer for classification Alzheimer's disease.

outcomes that separate the information into these two divisions is observable and descriptive. Sensitivity means the number of positive cases that would be accurately checked as positive. Specificity means the number of negative cases that accurately label them as negative (positive = special illness, negative = other cases).

True positive (TP): the positive signal is accurately detected.

False positive (FP): the negative signal is detected with mistakes.

True negative (TN): the negative signal is detected accurately.

FIGURE 11: Accuracy and loss values for classification process of the presented CNN.

False negative (FN): the positive signal is detected with mistakes.

The sensitivity divides TP cases into the sum of true-positive and false-negative cases in statistical terms.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (15)$$

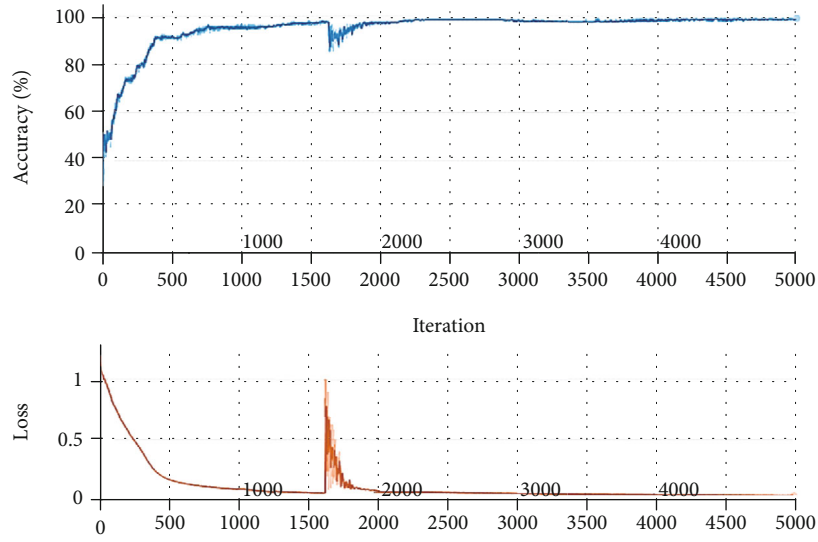The confusion matrix is the role of the algorithms described in the field of artificial intelligence. Usually, for supervised learning algorithms, such a demonstration is being used, but it is often utilized in unsupervised learning. An instance of the predicted value is seen in each column of the matrix. Suppose it includes an actual (true) instance in each row. This matrix's name is also gained, making it possible to mistake and mess with the results. This matrix is commonly called a contingency matrix or an error matrix outside of artificial intelligence.

4. Results and Discussion

4.1. Data Collection. Multichannel EEG signals were captured using earlobe-electrode hallmark monopolar connections [100]. The location of the electrodes over the scalp was obtained according to the 10-20 International Electrode Positioning Method (i.e., Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2). In Figure 3, an example of EEG electrode placement is shown. The electrodes calculate the weak electrical potential in the microvolt range produced by brain activities. Recordings were conducted with closed eyes in a sleeping state. In this way, multiple brain regions can be believed to be governed by the exact hierarchical mechanism.

Data with a signal length of 300 seconds and a sampling frequency of 1024 and 256 samples per second were obtained. Just 180 seconds is extracted for each signal (i.e., from 60 to 240 seconds) to minimize the EEG context artifacts and convert each one to 256 samples per second. The

FIGURE 12: Confusion matrix of the presented CNN method for classification of patients suffering from Alzheimer's.

sampling frequency, or sample rate, is the number of equal-spaced samples per unit of time. For instance, if there are 1024 equally spaced observations per second, the sampling rate is 1024/second or 1024 Hz. In this paper, for each second, 256 samples are used. Therefore, the length of the time series is $300 * 256 = 76800$. Moreover, the time interval from 60 to 240 seconds (of 300-second signals) is used with length $240 - 60 = 180$. Finally, the length of the signal is $180 * 256 = 46080$. To summarize, the EEG capacity of human samples belonging to three categories has been recorded:

(1) Patients who have Alzheimer's disease (AD)

(2) Patients who suffer from mild cognitive impairment (MCI)

(3) Healthy control (HC)

Figure 4 gives an example of extracted EEG recordings of 256 samples for each group.
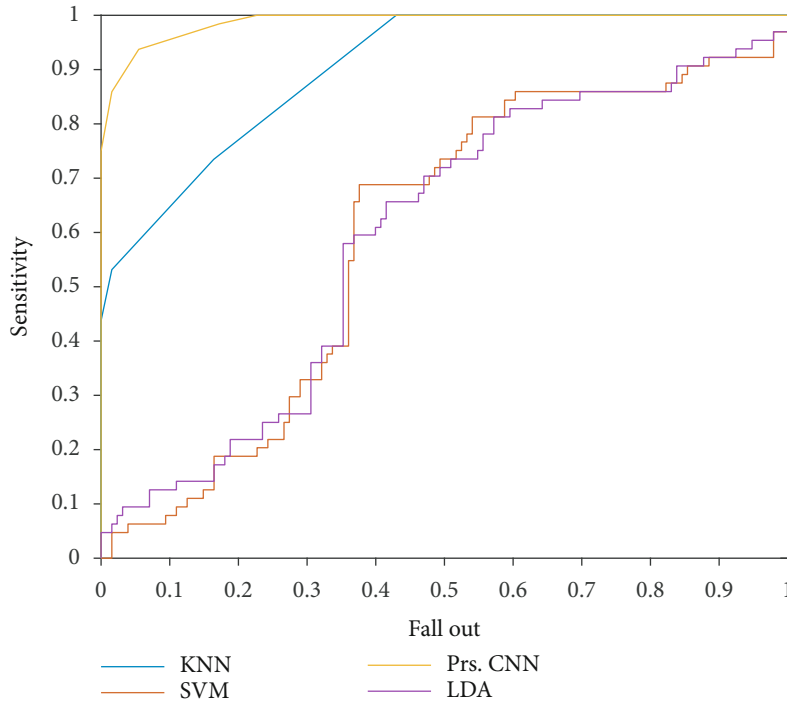
Figure 13: The ROC curve for the presented methods.

Figures 5–7 show each category's time-frequency analysis based on continuous threshold wavelet transform. The EEG signal processing strategies are addressed in this section to extract the necessary quality information from clinical limitations and simultaneously determine patients' status. The dispersion coefficient is a strong AD marker that can separate AD EEG from MCI and HC.

4.2. Implementation of the Proposed Pattern Detection Method. In this research, the automatic classification of normal and abnormal EEG signals was built in a deep learning model. EEG signals are used in the EEG database. Documents are separated into two sections of the EEG database: training and testing samples. During the learning stage, the training data were used, and during the evaluation stage, assessment data were used for training the algorithms with 80% used as validation data, while the remaining with 20%.

These data distributions have been randomly chosen. For validation purposes, many datasets have been used since the model parameter is set in several steps. However, experimental results with a specific random seed value are obtained to ensure that the model is reproducible and uniform. Figure 6 provides a thorough description of the data considered for this work. The block diagram of the process is described in Figure 8. Based on the process, firstly, TD-PSD is used for feature extraction of input EEG signals. The results of feature extraction are the produced seven features of values from 256 EEG samples.

Regarding this fact, the number of input variables for each person reduced from 256 to 7 features. It not only reduced the number of inputs but also increased the classification process speed. Signals with the time-dependent value

are needed to be transformed into a meaningful value. Therefore, these values are extracted features. Altogether, it is 64 persons for each category of MCI, AD, and HC that each one has seven features. Therefore, the input matrix for each category has 7×64 elements.

In the next step, the output features are used for classification methods. For achieving this purpose, three powerful traditional methods of classification of EEG signal and diagnosis of AD and MCI patients from HC or healthy people are used. Used methods include KNN, SVM, and LDA. All the methods are chosen from powerful machine learning techniques. The results of classification for mentioned methods are as follows (Figure 9). Figure 9 shows the confusion matrix of utilized methods of classification. The labeling (1, 2, and 3) illustrates MCI, AD, and HC, respectively. Based on the confusion matrix. The green values show the number of persons diagnosed correctly from 64 people. Based on Figure 9(a), from 64 patients who suffered from MCI, 51 of them (79.7%) are diagnosed correctly.

Moreover, for people with AD, the sensitivity of the KNN is 71.9% (46 out of 64 patients). Moreover, 62.5% (40 persons) of healthy samples are detected correctly. Based on the results, the lower row of the matrix illustrates the sensitivity of the methods for the classification of each category. Moreover, red %ages indicated miss rates of each class. The right column of the matrix represents the precision of the KNN technique. Based on the results, 63.7% of the MCI-diagnosed patients are classified correctly. Other precision values are depicted in the right column. Finally, the accuracy of the methods is presented in the lower-right corner of the matrix. Based on the result, the accuracy of KNN, SVM, and LDA approaches is 71.4%, 41.1%, and 43.8%, respectively. In other words, the accuracy of the KNN is higher than

TABLE 2: Comparison of the diagnosis methods used in this paper.

	Sensitivity			Precision			AUC	Accuracy
	MCI	AD	HC	MCI	AD	HC		
KNN	79.7%	71.9%	62.5%	63.7%	75.4%	78.4%	0.902	71.4%
SVM	9.4%	32.8%	81.3%	31.6%	47.7%	40.3%	0.593	41.1%
LDA	28.1%	53.1%	50.0%	45.0%	43.0%	48.8%	0.594	43.8%
Presented CNN	82.8%	89.1%	75.5%	81.5%	79.2%	87.3%	0.988	82.3%

that of the other methods. On the other hand, SVM and LDA methods represent results with very low accuracy. It can be noticed that the classification with two categories of patients and healthy people illustrates higher accuracy in comparison with three categories.

In this paper, a novel deep learning process for the classification of EEG signals is introduced based on a CNN. The architecture of the presented method is illustrated in Figure 10. Input layer includes the following:

(i) Seven features of 64 people for each category of MCI, AD, and HC

(ii) Input matrix 4D $[7 \times 1 \times 1 \times 192]$

(iii) Output matrix 1D $[1 \times 192]$

The labeling (1, 2, and 3) illustrates MCI, AD, and HC, respectively, output matrix. The process of the classification using the presented CNN architecture is shown in Figure 11. The maximum accuracy for the process reached almost 100%, and the loss value decreased to almost zero.

The confusion matrix of the presented method is indicated in Figure 12 without uncertainty value. The confident value of the accuracy for the CNN approach is 82.3%. In detail, 85% of the MCI patients are detected correctly; however, 89.1% of the AD and 75% of the normal sample are diagnosed correctly. For comparison of the resulted model of CNN with other KNN, SVM, and LDA techniques, the ROC plot is depicted in Figure 13. Regarding Figure 13, the value of FP rate versus TP rate is depicted based on output classification scores. Based on this graph, if the values would be higher for the TP rate and lower for the FP rate, it is better than the other graphs.

Moreover, the area under the curve (AUC) is one of the criteria for the classification method's performance analysis. Based on Figure 13, the presented CNN outperforms other methods, and the KNN method is in the following priority. The LDA and SVM reached low AUC values (see Table 2). To conclude, it can be seen that the presented CNN architecture is better and more accurate than other classification methods. Based on Table 2, the best method for diagnosing AD patients from EEG signals is the presented architecture's CNN approach.

5. Conclusion

In this article, the TD-PSD approach is used for EEG signal feature extraction from three groups of MCI, AD, and HC test samples. The final features used in three conventional classification methods are registered: KNN, SVM, and LDA, and the effects are recorded. Finally, the CNN architecture is provided for AD patient classification. The findings are indicated using performance measurement. Data were obtained with a signal duration of 300 seconds and a sampling frequency of 1024 and 256 samples per second. To minimize the EEG background artifacts, it obtains 180 seconds for each signal (i.e., from 60 to 240 seconds), and it is translated to 256 samples per second. The EEG ability of human samples belonging to three groups, including AD, MCI, and HC, has been summarized. First, the TD-PSD method is utilized to feature input EEG signals based on the procedure. The results of the extraction of features are the generation of seven value characteristics from 256 EEG samples. For classification methods, the output features are used in the next step. Methods used include KNN, SVM, and LDA. All the strategies are picked from effective methods in machine learning.

Based on the findings, 51 of the 64 patients with MCI (79.7%) were correctly diagnosed. In comparison, the KNN sensitivity for AD persons is 71.9% (46 out of 64 patients). In comparison, 62.5% (40 individuals) of sound samples are appropriately classified. Also, 63.7% of the MCI diagnosed patients are appropriately categorized based on the findings. In comparison, KNN, SVM, and LDA methods have a precision of 71.4%, 41.1%, and 43.8%, respectively. The precision of KNN, in other words, is greater than that of other processes. Then, a new EEG signal classification architecture is implemented that is focused on a CNN. For the CNN approach, the accurate meaning of accuracy is 82.3%. 85% of MCI cases are accurately detected in-depth, but 89.1% of the AD and 75% of the healthy population are correctly diagnosed. The presented CNN outperforms other approaches based on performance, and the KNN approach is the next target. The LDA and SVM were at low AUC values. For potential outcomes, it is recommended to modify feature extraction with another EEG signal justification for classification with a lower number of features and training using the design provided.

Data Availability

Data is extracted from the following ref: An integrated approach based on EEG signals processing combined with supervised methods to classify Alzheimer's disease patients.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] World Health Organization, *Alzheimer's Disease International. Dementia: A Public Health Priority*, vol. 2013, World Health Organization, Geneva, 2012.

[2] B. Duthey, "Background paper 6.11: Alzheimer disease and other dementias," *A Public Health Approach to Innovation*, vol. 6, pp. 1–74, 2013.

[3] M. Prince, A. Wimo, M. Guerchet et al., "The global impact of dementia: an analysis of prevalence, incidence, cost and trends," *Alzheimer's Disease International*, vol. 2015, 2015.

[4] E. E. Lee, B. Chang, S. Huege, and J. Hirst, "Complex clinical intersection: palliative care in patients with dementia," *The American Journal of Geriatric Psychiatry*, vol. 26, no. 2, pp. 224–234, 2018.

[5] C. R. Jack Jr., M. S. Albert, D. S. Knopman et al., "Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimer's & Dementia*, vol. 7, no. 3, pp. 257–262, 2011.

[6] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, "Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group∗ under the auspices of department of health and human services task force on Alzheimer's disease," *Neurology*, vol. 34, no. 7, pp. 939–944, 1984.

[7] J. Hort, J. T. O'Brien, G. Gainotti et al., "EFNS guidelines for the diagnosis and management of Alzheimer's disease," *European Journal of Neurology*, vol. 17, no. 10, pp. 1236–1248, 2010.

[8] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini-mental state: a practical method for grading the cognitive state of patients for the clinician," *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, 1975.

[9] A. J. Mitchell, "A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment," *Journal of Psychiatric Research*, vol. 43, no. 4, pp. 411–431, 2009.

[10] Z. S. Nasreddine, N. A. Phillips, V. Bédirian et al., "The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment," *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.

[11] P. S. Mathuranath, P. J. Nestor, G. E. Berrios, W. Rakowicz, and J. R. Hodges, "A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia," *Neurology*, vol. 55, no. 11, pp. 1613–1620, 2000.

[12] P. Amodio, H. Wenin, F. del Piccolo et al., "Variability of trail making test, symbol digit test and line trait test in normal people. A normative study taking into account age-dependent decline and sociobiological variables," *Aging Clinical and Experimental Research*, vol. 14, no. 2, pp. 117–131, 2002.

[13] K. I. Shulman, "Clock-drawing: is it the ideal cognitive screening test?," *International Journal of Geriatric Psychiatry*, vol. 15, no. 6, pp. 548–561, 2000.

[14] N. K. al-Qazzaz, S. H. B. M. D. Ali, S. A. Ahmad, K. Chellappan, M. S. Islam, and J. Escudero, "Role of EEG as biomarker in the early detection and classification of dementia," *The Scientific World Journal*, vol. 2014, Article ID 906038, 16 pages, 2014.

[15] T. J. Montine, C. H. Phelps, T. G. Beach et al., "National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practi-

[16] L. Wu, L. Wu, Y. Chen, and J. Zhou, "A promising method to distinguish vascular dementia from Alzheimer's disease with standardized low-resolution brain electromagnetic tomography and quantitative EEG," *Clinical EEG and Neuroscience*, vol. 45, no. 3, pp. 152–157, 2014.

[17] B. Dubois, H. H. Feldman, C. Jacova et al., "Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria," *The Lancet Neurology*, vol. 13, no. 6, pp. 614–629, 2014.

[18] D. Ferreira, L. Perestelo-Pérez, E. Westman, L. O. Wahlund, A. Sarría, and P. Serrano-Aguilar, "Meta-review of CSF core biomarkers in Alzheimer's disease: the state-of-the-art after the new revised diagnostic criteria," *Frontiers in Aging Neuroscience*, vol. 6, p. 47, 2014.

[19] B. Olsson, R. Lautner, U. Andreasson et al., "CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis," *The Lancet Neurology*, vol. 15, no. 7, pp. 673–684, 2016.

[20] M. Seeck, L. Koessler, T. Bast et al., "The standardized EEG electrode array of the IFCN," *Clinical Neurophysiology*, vol. 128, no. 10, pp. 2070–2077, 2017.

[21] S. Hu, Y. Lai, P. A. Valdes-Sosa, M. L. Bringas-Vega, and D. Yao, "How do reference montage and electrodes setup affect the measured scalp EEG potentials?," *Journal of Neural Engineering*, vol. 15, no. 2, article 026013, 2018.

[22] J. Wang, J. Barstein, L. E. Ethridge, M. W. Mosconi, Y. Takarae, and J. A. Sweeney, "Resting-state EEG abnormalities in autism spectrum disorders," *Journal of Neurodevelopmental Disorders*, vol. 5, no. 1, 2013.

[23] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, 2015.

[24] C. F. Muniz, A. V. Shenoy, K. L. O'Connor et al., "Clinical development and implementation of an institutional guideline for prospective EEG monitoring and reporting of delayed cerebral ischemia," *Journal of Clinical Neurophysiology*, vol. 33, no. 3, pp. 217–226, 2016.

[25] K. Nishida, M. Yoshimura, T. Isotani et al., "Differences in quantitative EEG between frontotemporal dementia and Alzheimer's disease as revealed by LORETA," *Clinical Neurophysiology*, vol. 122, no. 9, pp. 1718–1725, 2011.

[26] D.-H. Jeong, Y.-D. Kim, I.-U. Song, Y.-A. Chung, and J. Jeong, "Wavelet energy and wavelet coherence as EEG biomarkers for the diagnosis of Parkinson's disease-related dementia and Alzheimer's disease," *Entropy*, vol. 18, no. 1, p. 8, 2016.

[27] E. Neto, E. A. Allen, H. Aurlien, H. Nordby, and T. Eichele, "EEG spectral features discriminate between Alzheimer's and vascular dementia," *Frontiers in Neurology*, vol. 6, 2015.

[28] E. Neto, F. Biessmann, H. Aurlien, H. Nordby, and T. Eichele, "Regularized linear discriminant analysis of EEG features in dementia patients," *Frontiers in Aging Neuroscience*, vol. 8, 2016.

[29] S. J. Colloby, R. A. Cromarty, L. R. Peraza et al., "Multimodal EEG-MRI in the differential diagnosis of Alzheimer's disease and dementia with Lewy bodies," *Journal of Psychiatric Research*, vol. 78, pp. 48–55, 2016.

[30] H. Garn, C. Coronel, M. Waser, G. Caravias, and G. Ransmayr, "Differential diagnosis between patients with probable Alzheimer's disease, Parkinson's disease dementia,

or dementia with Lewy bodies and frontotemporal dementia, behavioral variant, using quantitative electroencephalographic features," *Journal of Neural Transmission*, vol. 124, no. 5, pp. 569–581, 2017.

[31] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*, Oxford University Press, New York, NY, USA, 2006.

[32] M. Cohen, *Analyzing Neural Time Series Data: Theory and Practice*, MIT Press, 2014.

[33] L. Sörnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, Academic Press, 2005.

[34] J. Dauwels, K. Srinivasan, M. Ramasubba Reddy et al., "Slowing and loss of complexity in Alzheimer's EEG: two sides of the same coin?," *International Journal of Alzheimer's Disease*, vol. 2011, article 539621, pp. 1–10, 2011.

[35] A. Alberdi, A. Aztiria, and A. Basarab, "On the early diagnosis of Alzheimer's disease from multimodal signals: a survey," *Artificial Intelligence in Medicine*, vol. 71, pp. 1–29, 2016.

[36] N. Malek, M. R. Baker, C. Mann, and J. Greene, "Electroencephalographic markers in dementia," *Acta Neurologica Scandinavica*, vol. 135, no. 4, pp. 388–393, 2017.

[37] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "Alzheimer's disease: models of computation and analysis of EEGs," *Clinical and EEG Neuroscience*, vol. 36, no. 3, pp. 131–140, 2005.

[38] B. Jelles, J. H. van Birgelen, J. P. Slaets, R. E. Hekster, E. J. Jonkman, and C. J. Stam, "Decrease of non-linear structure in the EEG of Alzheimer patients compared to healthy controls," *Clinical Neurophysiology*, vol. 110, no. 7, pp. 1159–1167, 1999.

[39] J. Jeong, S. Y. Kim, and S. H. Han, "Non-linear dynamical analysis of the EEG in Alzheimer's disease with optimal embedding dimension," *Electroencephalography and Clinical Neurophysiology*, vol. 106, no. 3, pp. 220–228, 1998.

[40] J. Jeong, J. H. Chae, S. Y. Kim, and S. H. Han, "Non-linear dynamic analysis of the EEG in patients with Alzheimer's disease and vascular dementia," *Journal of Clinical Neurophysiology*, vol. 18, no. 1, pp. 58–67, 2001.

[41] T. Takahashi, "Complexity of spontaneous brain activity in mental disorders," *Progress in Neuro-psychopharmacology & Biological Psychiatry*, vol. 45, pp. 258–266, 2013.

[42] T. Yagyu, J. Wackermann, M. Shigeta et al., "Global dimensional complexity of multi-channel EEG in mild Alzheimer's disease and age-matched cohorts," *Dementia and Geriatric Cognitive Disorders*, vol. 8, no. 6, pp. 343–347, 1997.

[43] N. Houmani, F. B. Vialatte, C. Latchoumane, J. Jeong, and G. Dreyfus, "Stationary epoch-based entropy estimation for early diagnosis of Alzheimer's disease," in *2013 IEEE Faible Tension Faible Consommation*, pp. 1–4, Paris, France, 2013.

[44] N. Houmani, F. B. Vialatte, and G. Dreyfus, "Epoch-based entropy for early screening of Alzheimer's disease," *International Journal of Neural Systems*, vol. 25, no. 8, article 1550032, 2015.

[45] D. Aba'solo, R. Hornero, P. Espino, D. Alvarez, and J. Poza, "Entropy analysis of the EEG background activity in Alzheimer's disease patients," *Physiological Measurement*, vol. 27, no. 3, pp. 241–253, 2006.

[46] T. J. De Bock, S. Das, M. Mohsin et al., "Early detection of Alzheimer's disease using non-linear analysis of EEG via Tsallis entropy," in *2010 Biomedical Sciences and Engineering Conference*, pp. 1–4, Oak Ridge, TN, USA, 2010.

[47] D. Aba'solo, R. Hornero, P. Espino, J. Poza, C. I. Sanchez, and R. de la Rosa, "Analysis of regularity in the EEG background activity of Alzheimer's disease patients with approximate entropy," *Clinical Neurophysiology*, vol. 116, no. 8, pp. 1826–1834, 2005.

[48] S. M. Pincus, "Approximate entropy as a measure of irregularity for psychiatric serial metrics," *Bipolar Disorders*, vol. 8, no. 5p1, pp. 430–440, 2006.

[49] J. Escudero, D. Aba'solo, R. Hornero, P. Espino, and M. Lopez, "Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy," *Physiological Measurement*, vol. 27, no. 11, pp. 1091–1106, 2006.

[50] D. Aba'solo, R. Hornero, and C. Gomez, "Analysis of EEG background activity in Alzheimer's disease patients with Lempel-Ziv complexity and central tendency measure," *Medical Engineering & Physics*, vol. 28, no. 4, pp. 315–322, 2006.

[51] W. Chen, Y. Y. Lam, C. P. Shen, H. Y. Sung, M. J. Chiu, and F. Lai, "Ultra-fast epileptic seizure detection using EMD based on multi-channel electroencephalogram," in *13th IEEE International Conference on BioInformatics and BioEngineering*, pp. 1–4, Chania, Greece, 2013.

[52] G. Pei, J. Wu, D. Chen et al., "Effects of an integrated neurofeedback system with dry electrodes: EEG acquisition and cognition assessment," *Sensors*, vol. 18, no. 10, p. 3396, 2018.

[53] G. Swapna, G. Swapna, G. Swapna, J. S. Suri, and J. S. Suri, "Automated EEG analysis of epilepsy: a review," *Knowledge-Based Systems*, vol. 45, pp. 147–165, 2013.

[54] T. Yan, W. Wang, T. Liu et al., "Increased local connectivity of brain functional networks during facial processing in schizophrenia: evidence from EEG data," *Oncotarget*, vol. 8, no. 63, pp. 107312–107322, 2017.

[55] G. Chen, W. Xie, T. D. Bui, and A. Krzyzak, "Automatic epileptic seizure detection in EEG using nonsubsampled wavelet–Fourier features," *Journal of Medical and Biological Engineering*, vol. 37, no. 1, pp. 123–131, 2017.

[56] A. Hamad, E. H. Houssein, A. E. Hassanien, and A. A. Fahmy, "A hybrid EEG signals classification approach based on grey wolf optimizer enhanced SVMs for epileptic detection," in *International Conference on Advanced Intelligent Systems and Informatics*, pp. 108–117, Cairo, Egypt, 2017.

[57] A. Subasi, J. Kevric, and M. Abdullah Canbaz, "Epileptic seizure detection using hybrid machine learning methods," *Neural Computing and Applications*, vol. 31, no. 1, pp. 317–325, 2017.

[58] B. Wang, T. Yan, S. Ohno, S. Kanazawa, and J. Wu, "Retinotopy and attention to the face and house images in the human visual cortex," *Experimental Brain Research*, vol. 234, no. 6, pp. 1623–1635, 2016.

[59] D. Ahmedt-Aristizabal, C. Fookes, K. Nguyen, S. Denman, S. Sridharan, and S. Dionisio, "Deep facial analysis: a new phase I epilepsy evaluation using computer vision," *Epilepsy & Behavior*, vol. 82, pp. 17–24, 2018.

[60] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Computers in Biology and Medicine*, vol. 100, pp. 270–278, 2018.

[61] R. Hussein, H. Palangi, R. Ward, and Z. J. Wang, *Epileptic Seizure Detection: A Deep Learning Approach*2018, https://arxiv.org/abs/1803.09848.

[62] N. Kulkarni and V. Bairagi, *EEG-Based Diagnosis of Alzheimer Disease: A Review and Novel Approach for Feature*

Extraction and Classification Techniques, Academic Press, Cambridge, 2018.

[63] M. Z. Parvez and M. Paul, "Seizure prediction using undulated global and local features," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 208–217, 2017.

[64] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, R. Wald, E. Muharemagic, and N. Seliya, "Deep learning applications and challenges in big data analytics," *Journal of Big Data*, vol. 2, no. 1, 2015.

[65] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[66] M. Coşkun, Ö. Yildirim, A. Uçar, and Y. Demir, "An overview of popular deep learning methods," *European Journal of Technic*, vol. 7, no. 2, pp. 165–176, 2017.

[67] J. Günther, P. M. Pilarski, G. Helfrich, H. Shen, and K. Diepold, "First steps towards an intelligent laser welding architecture using deep neural networks and reinforcement learning," *Procedia Technology*, vol. 15, pp. 474–483, 2014.

[68] Y. LeCun, B. Boser, J. S. Denker et al., "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.

[69] A. Uçar, Y. Demir, and C. Güzeliş, "Object recognition and detection with deep learning for autonomous driving applications," *Simulation*, vol. 93, no. 9, pp. 759–769, 2017.

[70] F. Beser, M. A. Kizrak, B. Bolat, and T. Yildirim, "Recognition of sign language using capsule networks," in *2018 26th IEEE signal processing and communications applications conference (SIU)*, Izmir, Turkey, 2018.

[71] V. Mnih, K. Kavukcuoglu, D. Silver et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[72] O. Yildirim, R. S. Tan, and U. R. Acharya, "An efficient compression of ECG signals using deep convolutional autoencoders," *Cognitive Systems Research*, vol. 52, pp. 198–211, 2018.

[73] Ö. Yildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification," *Computers in Biology and Medicine*, vol. 96, pp. 189–202, 2018.

[74] O. Yıldırım, P. Pławiak, R. S. Tan, and U. R. Acharya, "Arrhythmia detection using deep convolutional neural network with long duration ECG signals," *Computers in Biology and Medicine*, vol. 102, pp. 411–420, 2018.

[75] S. L. Oh, Y. Hagiwara, U. Raghavendra et al., "A deep learning approach for Parkinson's disease diagnosis from EEG signals," *Neural Computing and Applications*, vol. 32, no. 15, pp. 10927–10933, 2020.

[76] X. Hong, R. Lin, and C. Yang, "Predicting Alzheimer's disease using LSTM," *IEEE Access*, vol. 7, pp. 2169–3536, 2019.

[77] M. Aghili, S. Tabarestani, and M. Adjouadi, "Predictive modeling of longitudinal data for Alzheimer's disease diagnosis using RNNs," in *International Workshop on PRedictive Intelligence In MEdicine*, pp. 112–119, Granada, Spain, 2018.

[78] F. Razavi, M. J. Tarokh, and M. Alborzi, "An intelligent Alzheimer's disease diagnosis method using unsupervised feature learning," *Journal of Big Data*, vol. 6, no. 1, 2019.

[79] P. Padilla, M. López, J. M. Górriz, J. Ramírez, D. Salas-González, and I. Álvarez, "NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease," *IEEE Transactions on Medical Imaging*, vol. 31, no. 2, pp. 207–216, 2012.

[80] Q. Xin, S. Hu, S. Liu, X. Ma, H. Lv, and Y. D. Zhang, "Epilepsy EEG classification based on convolution support vector machine," *Journal of Medical Imaging and Health Informatics*, vol. 11, no. 1, pp. 25–32, 2021.

[81] I. Aliyu and C. G. Lim, "Selection of optimal wavelet features for epileptic EEG signal classification with LSTM," *Neural Computing and Applications*, pp. 1–21, 2021.

[82] T. Tuncer, "A new stable non-linear textural feature extraction method-based EEG signal classification method using substitution Box of the Hamsi hash function: Hamsi pattern," *Applied Acoustics*, vol. 172, article 107607, 2021.

[83] P. A. Cicalese, R. Li, M. B. Ahmadi et al., "An EEG-fNIRS hybridization technique in the four-class classification of Alzheimer's disease," *Journal of Neuroscience Methods*, vol. 336, article 108618, 2020.

[84] R. Ferri, C. Babiloni, V. Karami et al., "Stacked autoencoders as new models for an accurate Alzheimer's disease classification support using resting-state EEG and MRI measurements," *Clinical Neurophysiology*, vol. 132, no. 1, pp. 232–245, 2021.

[85] L. R. Trambaiolli, N. Spolaôr, A. C. Lorena, R. Anghinah, and J. R. Sato, "Feature selection before EEG classification supports the diagnosis of Alzheimer's disease," *Clinical Neurophysiology*, vol. 128, no. 10, pp. 2058–2067, 2017.

[86] S. Nobukawa, T. Yamanishi, S. Kasakawa, H. Nishimura, M. Kikuchi, and T. Takahashi, "Classification methods based on complexity and synchronization of electroencephalography signals in Alzheimer's disease," *Frontiers in Psychiatry*, vol. 11, 2020.

[87] N. N. Kulkarni and V. K. Bairagi, "Extracting salient features for EEG-based diagnosis of Alzheimer's disease using support vector machine classifier," *IETE Journal of Research*, vol. 63, no. 1, pp. 11–22, 2017.

[88] F. Vecchio, F. Miraglia, F. Alù et al., "Classification of Alzheimer's disease with respect to physiological aging with innovative EEG biomarkers in a machine learning implementation," *Journal of Alzheimer's Disease*, vol. 75, no. 4, pp. 1253–1261, 2020.

[89] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 3, pp. 306–310, 1970.

[90] R. N. Khushaba, M. Takruri, J. V. Miro, and S. Kodagoda, "Towards limb position invariant myoelectric pattern recognition using time-dependent spectral features," *Neural Networks*, vol. 55, pp. 42–58, 2014.

[91] V. Wellekens and C. Tyagi, "On desensitizing the Mel-Cepstrum to spurious spectral components for robust speech recognition," in *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005*, pp. 529–532, Philadelphia, PA, USA, 2005.

[92] D. G. Childers, D. P. Skinner, and R. C. Kemerait, "The cepstrum: a guide to processing," *Proceedings of the IEEE*, vol. 65, no. 10, pp. 1428–1443, 1977.

[93] J. He, D. Zhang, X. Sheng, S. Li, and X. Zhu, "Invariant surface EMG feature against varying contraction level for myoelectric control based on muscle coordination," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 874–882, 2014.

[94] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks,"

Advances in Neural Information Processing System, vol. 25, pp. 1097–1105, 2012.

[95] 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 110170, 2020.

[96] F. Milletari, N. Navab, and S. A. Ahmadi, "V-net: fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, Stanford, CA, USA, 2016.

[97] R. Zhang, L. Bai, J. Zhang, L. Tian, R. Li, and Z. Liu, "Convolutional LSTM networks for vibration-based defect identification of the composite structure," in *2020 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*, pp. 390–395, Xi'an, China, 2020.

[98] R. Liu, S. Wei, Y. Zhao, and Y. Yang, "Indexing of the CNN features for the large scale image search," *Multimedia Tools and Applications*, vol. 77, no. 24, pp. 32107–32131, 2018.

[99] S. H. Park, J. M. Goo, and C. H. Jo, "Receiver operating characteristic (ROC) curve: practical review for radiologists," *Korean Journal of Radiology*, vol. 5, no. 1, p. 11, 2004.

[100] H. H. Jasper, "The ten-twenty electrode system of the international federation," *Electroencephalography and Clinical Neurophysiology*, vol. 10, pp. 371–375, 1958.

Research Article

A New Measure of Pulse Rate Variability and Detection of Atrial Fibrillation Based on Improved Time Synchronous Averaging

Xiaodong Ding [iD], **Yiqin Wang** [iD], **Yiming Hao** [iD], **Yi Lv** [iD], **Rui Chen** [iD], **and Haixia Yan** [iD]

Shanghai Key Laboratory of Health Identification and Assessment, Laboratory of Traditional Chinese Medicine Four Diagnostic Information, Shanghai University of Traditional Chinese Medicine, Shanghai, China

Correspondence should be addressed to Haixia Yan; hjy2012ok@163.com

Background. Pulse rate variability monitoring and atrial fibrillation detection algorithms have been widely used in wearable devices, but the accuracies of these algorithms are restricted by the signal quality of pulse wave. Time synchronous averaging is a powerful noise reduction method for periodic and approximately periodic signals. It is usually used to extract single-period pulse waveforms, but has nothing to do with pulse rate variability monitoring and atrial fibrillation detection traditionally. If this method is improved properly, it may provide a new way to measure pulse rate variability and to detect atrial fibrillation, which may have some potential advantages under the condition of poor signal quality. *Objective*. The objective of this paper was to develop a new measure of pulse rate variability by improving existing time synchronous averaging and to detect atrial fibrillation by the new measure of pulse rate variability. *Methods*. During time synchronous averaging, two adjacent periods were regarded as the basic unit to calculate the average signal, and the difference between waveforms of the two adjacent periods was the new measure of pulse rate variability. 3 types of distance measures (Euclidean distance, Manhattan distance, and cosine distance) were tested to measure this difference on a simulated training set with a capacity of 1000. The distance measure, which can accurately distinguish regular pulse rate and irregular pulse rate, was used to detect atrial fibrillation on the testing set with a capacity of 62 (11 with atrial fibrillation, 8 with premature contraction, and 43 with sinus rhythm). The receiver operating characteristic curve was used to evaluate the performance of the indexes. *Results*. The Euclidean distance between waveforms of the two adjacent periods performs best on the training set. On the testing set, the Euclidean distance in atrial fibrillation group is significantly higher than that of the other two groups. The area under receiver operating characteristic curve to identify atrial fibrillation was 0.998. With the threshold of 2.1, the accuracy, sensitivity, and specificity were 98.39%, 100%, and 98.04%, respectively. This new index can detect atrial fibrillation from pulse wave signal. *Conclusion*. This algorithm not only provides a new perspective to detect AF but also accomplishes the monitoring of PRV and the extraction of single-period pulse wave through the same technical route, which may promote the popularization and application of pulse wave.

1. Introduction

The radial artery pulse wave is an important signal in health monitoring and disease diagnosis, which contains abundant physiological information. Different from the ECG signal which is often used to detect all kinds of arrhythmias [1], the radial artery pulse wave not only contains the information of heart rate and its variability which is widely used in smart watches and other wearable devices [2–4] but also can assist conventional methods to diagnose and monitor the occurrence and development of multiple common dis-

eases such as hypertension, diabetes, and coronary heart disease [5–10]. In addition, the acquisition of radial artery pulse wave is much more convenient than ECG. However, except for atrial fibrillation (AF) detection and other the applications closely related to pulse rate variability (PRV), most of the applications depend on the information contained in single-period pulse waves. Due to the susceptibility of wearable devices to interference, it is so difficult to collect valuable single-period pulse waves with existing wearable devices that the information contained in the single-period pulse waves is neglected by wearable device researchers. Even for PRV
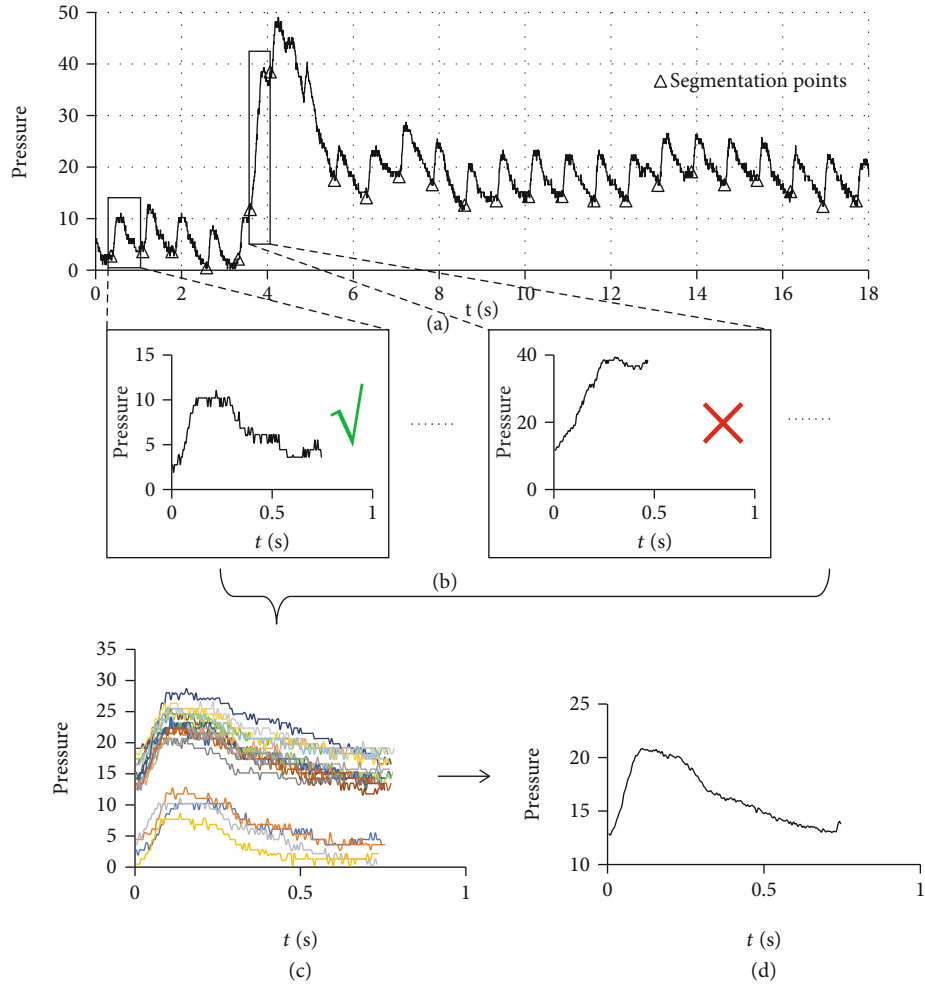
Figure 1: General steps of TSA combined with signal-period pulse wave quality evaluation method. (a) A pulse wave series was segmented into periods. (b) The signal quality of each segment was evaluated, and the abnormal segments were eliminated. (c) All the normal segments were synchronized with the starting points. (d) The noise was suppressed by averaging the synchronized signals.

monitoring and AF detection algorithms, it is imperative that subjects remain stationary during pulse wave acquisition. The application of pulse wave is restricted by the weak anti-interference ability.

Time synchronous averaging (TSA) is a widely used signal processing technique which enables periodic waveforms to be extracted from noisy signals [11, 12]. It is traditionally suited for the vibration analysis of mechanical systems which move circularly such as gearboxes. The noise of such signals can be effectively averaged out by gradually accumulating those portions of the signals that are synchronized with the fiducial points. Different from other noise reduction methods, TSA can effectively reduce all independent noise without considering frequency properties and threshold selection. Moreover, the signal-period pulse wave quality evaluation method [13] can be incorporated in TSA to identify and eliminate the seriously interfered periods. That is to say, we can select the less interfered periods from a pulse wave series with poor quality to complete TSA, rather than discarding the whole series (Figure 1). Similar algorithms have been applied to single-period pulse waveform extraction

[14, 15]. However, in existing applications, the pulse wave signal is assumed to be a strict periodic signal, and the starting point or the highest point of the waveform is used as the fiducial point for synchronization without discussing the basis of these steps. More importantly, the single-period pulse wave extracted by existing TSA method does not contain the PRV information. It has nothing to do with PRV monitoring and AF detection. This may be the reason why TSA is neglected by wearable device researchers.

However, if we take two adjacent periods as the basic unit to calculate the average signal, it can be expected that with the increasement of PRV, the waveform of the second period will be gradually distorted due to the misalignment (Figure 2). The difference between waveforms of the two adjacent periods obtained by TSA may provide a new measure of PRV. Compared with traditional PRV measures, it may have some potential advantages under the condition of poor signal quality. And this index mainly reflects the irregular changes of heart rate. For patients with premature contraction (PC) which usually have regular changes in heart rhythm, the difference between adjacent periods may not be large because
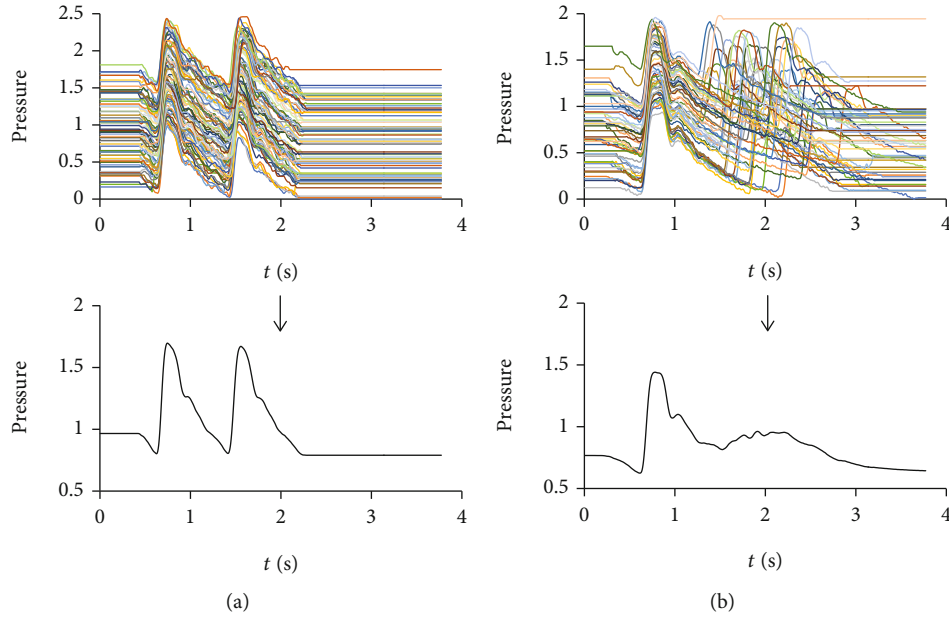
(a)

(b)

FIGURE 2: The average pulse wave with a basic unit of two adjacent periods. (a) The average pulse wave of a normal individual. The waveforms of the two adjacent periods are similar. (b) The average pulse wave of a patient with AF. The waveform of the second period is seriously distorted.

there are still a considerable number of second period waveforms are synchronous. The new index may effectively distinguish AF from PC and sinus rhythm (SR).

The objective of this paper was to develop a new measure of PRV by improving existing TSA and to detect atrial fibrillation by the new measure of pulse rate variability.

2. Methods

2.1. Data. In this study, the radial artery pulse wave signals were taken from 112 inpatients who had underwent an electrocardiographic (ECG) examination at Shanghai Shuguang Hospital between July 2019 and January 2020, including 11 cases with AF, 8 cases with PC, and 93 cases with SR. For each subject, a left radial artery pulse wave signal with a length of 60 seconds was taken by a wrist-type pulse wave monitor (type: Smart TCM-I, product by: Shanghai Asia & Pacific Computer Information System CO, Ltd, Shanghai, China) after the subject was either sitting or lying down for at least 5 min. ECG examination and pulse wave acquisition were performed on the same day but not simultaneously.

2.2. Preprocessing. The steps of preprocessing, including period segmentation and signal quality evaluation, are illustrated in Figure 3.

To segment the pulse wave series into periods, the derivative of the original signal was used to locate segmentation points by the threshold method (Figure 4). During threshold determination, each pulse wave series was segmented with 9 trial thresholds (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9). All the obtained segments were evaluated by a logistic regression model [13] which can divide the segments into normal segments and abnormal segments. The threshold with which
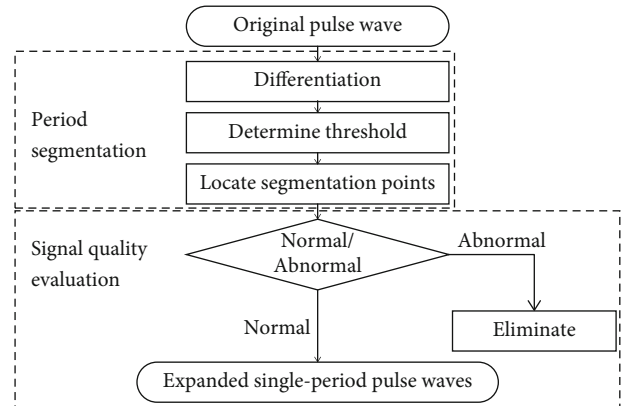

FIGURE 3: Steps of preprocessing. During period segmentation, the original pulse wave series were segmented into periods by threshold method. During signal quality evaluation, the segment obtained by period segmentation were divided into normal and abnormal segments by a logistic regression model. The abnormal segments were eliminated, and the range of normal segment was expanded by 50% on both sides to include the information of adjacent periods.

the maximum number of normal segments were obtained was selected for the next steps.

During signal quality evaluation, the segment obtained by period segmentation were divided into normal and abnormal segments by the same logistic regression model as used in threshold determination. The abnormal segments were eliminated, and the range of normal segment was expanded by 50% on both sides to prepare for the measurement of PRV (Figure 5).
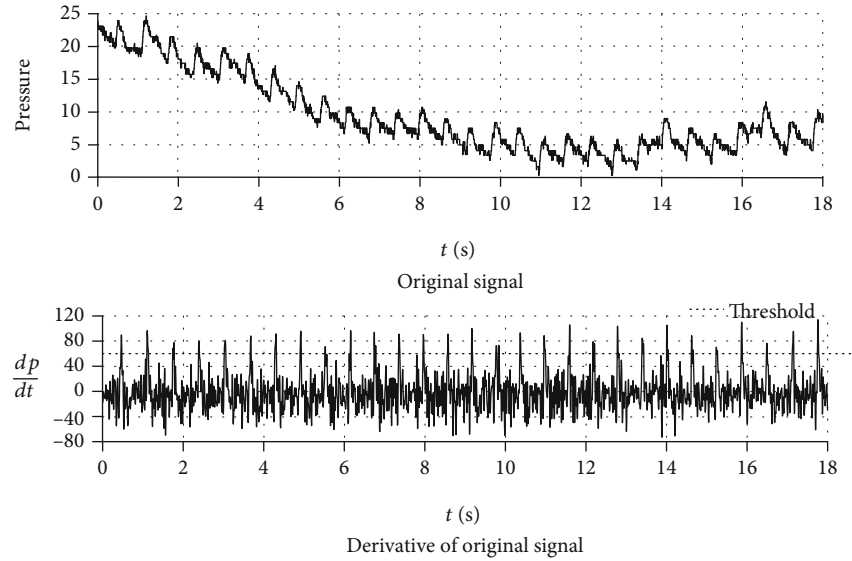
FIGURE 4: Pulse wave and its derivative with an applied threshold. The derivative of the original signal is almost entirely unaffected by baseline wander and shows clearer segmentation points. The first zero point of the derivative before each threshold point was defined as the period segmentation point, and the corresponding segments of the original signal between two adjacent period segmentation points were single-period waveforms.
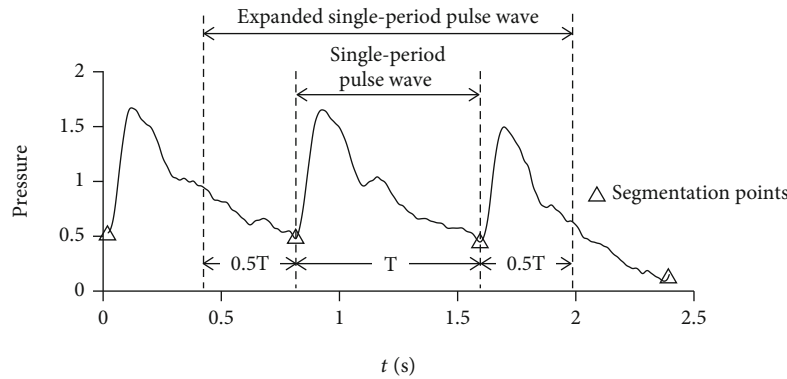

FIGURE 5: Expanded single-period pulse wave. The range of each segment was expanded by 50% on both sides to include the information of adjacent periods.

2.3. Time Synchronous Averaging. How to average the single-period pulse waves with different lengths in the same sequence and what is the appropriate fiducial point with the strongest anti-interference ability are questions that have not been fully discussed in current applications of TSA.

As shown in Figure 2, even in the pulse wave of a patient with AF, all the systoles have similar lengths and shapes, and the difference of cardiac cycle duration is mainly caused by the difference of diastolic duration. It is because the process of myocardial contraction and the state of arterial vessels are relatively stable for the same individual, and the duration of diastole does not significantly affect the left ventricular end-diastolic volume due to the low rate of left ventricular filling during late diastole. The initial condition and process of systole are basically stable. It is an appropriate averaging method to accumulate and average the preprocessed data without any stretching or compression, because most of the common time-

domain features except the duration of cardiac cycle and diastole are extracted from the pulse wave of systole.

Figure 4 shows a pulse wave series and its derivative. The derivative of the original signal is almost entirely unaffected by baseline wander and shows clearer segmentation points. The spikes of the derivative are formed by the periodic rapid ejections of blood from the left ventricle. Different from the starting point or the highest point of a period where the waveform is relatively gentle and easy to be distorted by external interference, the spikes of the derivative have stronger anti-interference ability because the change of pulse wave caused by rapid ejection of blood is more significant than that caused by external interference. Moreover, the QRS complex, which is the most frequently used heartbeat fiducial point to calculate the heart rate in ECG [16], is formed by the same cardiac event. Using the peak of the derivative in each period as the reference point, the calculated results may have better
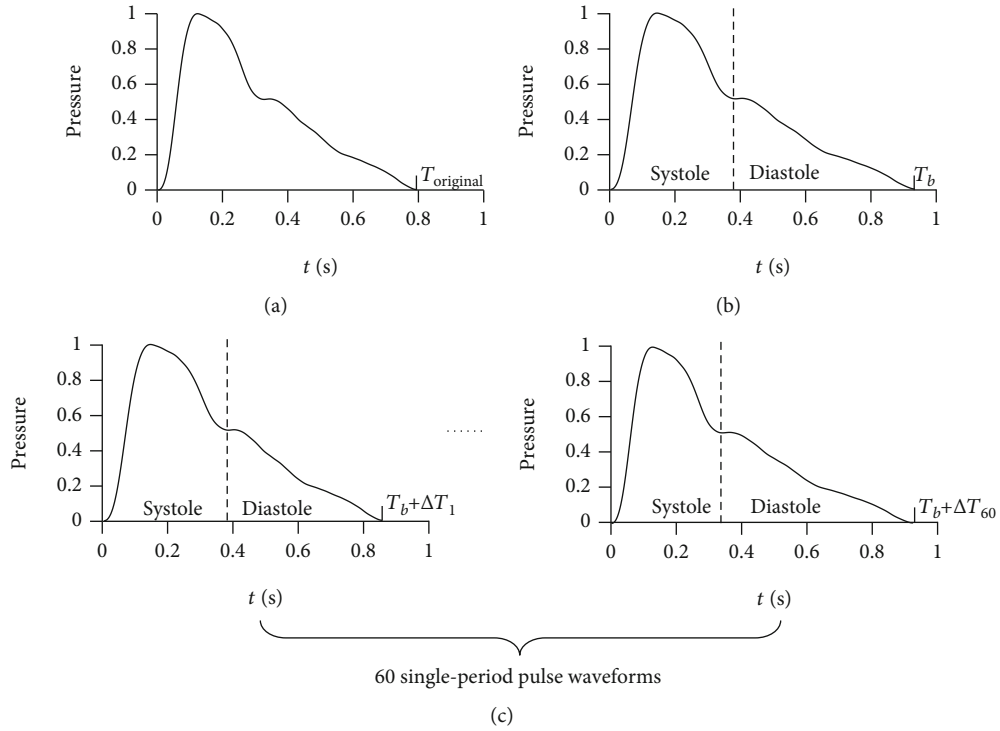
FIGURE 6: Steps to generate the simulated training set. (a) 50 single-period pulse waveforms were extracted from pulse wave signals of the 50 selected cases by TSA. (b) 1000 single-period pulse waveforms were generated by stretching or compressing the original single-period pulse waveform to make its length equal to T_b. (c) 60 single-period pulse waveforms were generated by stretching or compressing the corresponding single-period pulse waveform generated in (b) to make its length equal to $T_b + \Delta T$ separately. Only the waveforms of diastole were stretched or compressed in this step. The pulse wave series of each training sample were subsequently obtained by connecting the 60 single-period pulse waveforms end to end.
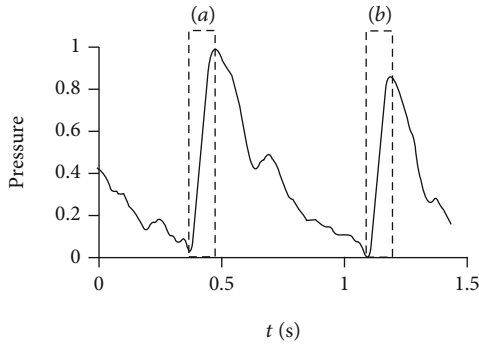

FIGURE 7: The first ascending limb and the second ascending limb of expanded single-period pulse wave. (a) The first ascending limb is defined as the data between the minimum value and the maximum value in the first half of the expanded single-period pulse waveform. (b) The second ascending limb is defined as the data starting from the minimum value between the maximum values of the first half and the second half of the expanded single-period pulse waveform and with the same length as the first ascending limb.

FIGURE 8: ROC curves of 3 candidate indexes on the training set. The AUC of D_e, D_m, and D_c are 0.857, 0.801, and 0.516, respectively.

comparability with the results of ECG. Therefore, the peak of the derivative is an appropriate fiducial point for synchronization of TSA.

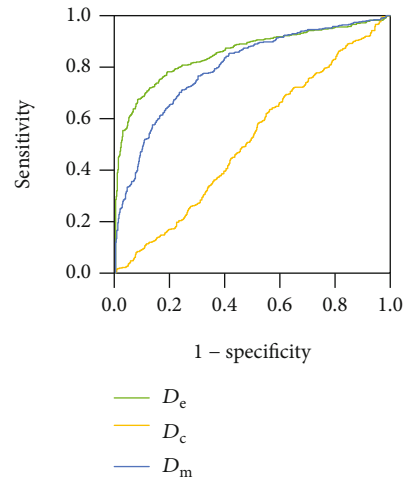Therefore, during TSA, the expanded single-period pulse waveforms obtained by preprocessing were synchronized with the maximum derivative value of each period, and all the waveforms from one pulse wave series were averaged directly without stretching or compression. After synchronizing, all the expanded single-period pulse waveforms were unified to the same length by filling with 0. If $X_i = \{x_1, x_2, \cdots, x_n\}$ was an expanded single-period pulse waveform and

Table 1: Comparison of D_e in different groups of testing set by Kruskal-Wallis test.

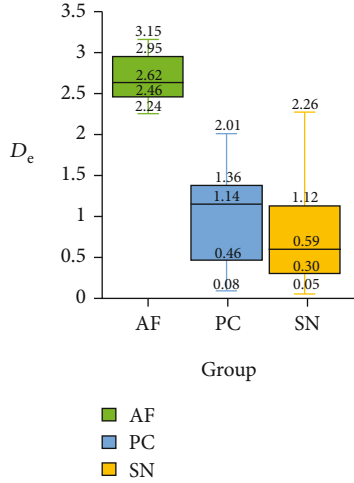| Groups	Test statistic	Standard error	p
SN-AF	32.049	6.096	<0.001
PC-AF	24.659	8.383	0.010
SN-PC	7.390	6.947	0.862

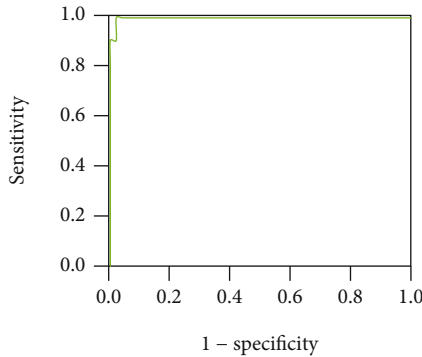Figure 9: Box-plot of D_e in different groups of testing set.

Figure 10: ROC curve of D_e to detect AF on the testing set. The AUC was 0.998, and the accuracy, sensitivity, and specificity were 98.39%, 100%, and 98.04%, respectively, with the threshold of 2.1.

N was the number of normal segments in the sequence, the average expanded single-period pulse waveform of the sequence was given by

$$X_a = \frac{1}{N}\sum_{i=1}^{N} X_i. \qquad (1)$$

2.4. Measure of PRV and Detection of AF. To find an effective index of PRV, 50 cases with SR were randomly selected from the data set to generate the simulated training set with a capacity of 1000. The testing set consisting of the other 62 cases (AF:11, PC:8, SR:43) was used to test the ability of the selected index to detect AF. The training set was generated according to a simple and commonly used identification cri-

terion for irregular heart rhythm—there is a variation of more than 0.16 seconds between the longest cardiac cycle duration and the shortest cardiac cycle duration [17]. The detailed steps to generate the simulated training set are as follows (Figure 6):

(1) 50 single-period pulse waveforms were extracted from pulse wave signals of the 50 selected cases by TSA

(2) Considering that the cardiac cycle duration is usually between 0.6 s and 1 s, for each single-period pulse waveform, 20 random numbers (denoted by T_b) which obey the uniform distribution U (0.6,1) were generated to simulate different cardiac cycle durations of different individuals. A total of 1000 base cardiac cycle durations were generated for the 1000 expected training samples

(3) 1000 single-period pulse waveforms were generated by stretching or compressing the original single-period pulse waveform to make its length equal to T_b. All the 1000 single-period pulse waveforms were randomly divided into arrhythmia group and control group with 500 waveforms in each group

(4) Each training sample consists of 60 cardiac cycles, and the duration of each cardiac cycle fluctuates around the base duration T_b. The duration of each cardiac cycle is given by $T = T_b + \Delta T$. In arrhythmia group, ΔT obeys the uniform distribution U (-0.09, 0.09). Whereas in control group, ΔT obeys the uniform distribution U (-0.07, 0.07). A sequence of 60 durations was generated for each training sample. In arrhythmia group, the variation between the maximum value and the minimum value of 60 durations is less than 0.18 but usually more than 0.16. Whereas in control group, the variation between the maximum value and the minimum value of 60 durations is less than 0.14. It is in accordance with the identification criterion for irregular heart rhythm

(5) 60 single-period pulse waveforms were generated by stretching or compressing the corresponding single-period pulse waveform generated in step 3 to make its length equal to the 60 durations separately. The pulse wave series of each training sample were subsequently obtained by connecting the 60 single-period pulse waveforms end to end. Considering that the systolic duration of an individual is almost constant, only the waveforms of diastole were stretched or compressed to satisfy the requirement of cardiac cycle durations in this step

After the training set was generated, the average expanded single-period pulse waveform of each training sample was extracted by TSA. The difference between the first ascending limb and the second ascending limb were tested to distinguish between arrhythmia group and control group. The first ascending limb is defined as the data between the minimum value and the maximum value in the first half

TABLE 2: Comparison of recent pulse-wave-based AF detection techniques.

Reference	Methods	Accuracy (%)
McManus DD, et al. (2013) [18]	RMSSD and Shannon entropy	96.76
Krivoshei L, et al. (2017) [19]	Shannon entropy and other IBI features	87.5
Fallet S, et al. (2019) [20]	Bagging decision tree based on IBI features	88.5
Kabutoya T, et al. (2019) [3]	Irregular heartbeat ratio	98.3
Kashiwa A, et al. (2019) [21]	IBI features	97.3
Zalabarria U, et al. (2020) [22]	ANN with foot point detection	93.68
Han D, et al. (2020) [23]	Random forest with Poincare plot	95.32
This paper	D_e between adjacent periods based on TSA	98.4

of the expanded single-period pulse waveform. The second ascending limb is defined as the data starting from the minimum value between the maximum values of the first half and the second half of the expanded single-period pulse waveform and with the same length as the first ascending limb (Figure 7). Considering that Euclidean distance (D_e), Manhattan distance (D_m), and cosine distance (D_c) are commonly used distance measures between two vectors, these 3 candidate indexes were tested on the training set to distinguish between arrhythmia group and control group. If $X_f = \{x_{f1}, x_{f2}, \cdots, x_{fn}\}$ and $X_s = \{x_{s1}, x_{s2}, \cdots, x_{sn}\}$ were the data of first ascending limb and the second ascending limb, respectively, D_e, D_m, and D_c were given by

$$D_e = \sqrt{\sum_{i=1}^{n} \left(x_{fi} - x_{si}\right)^2},$$

$$D_m = \sum_{i=1}^{n} \left|x_{fi} - x_{si}\right|, \tag{2}$$

$$D_c = \frac{\sum_{i=1}^{n} x_{fi} x_{si}}{\sqrt{\sum_{i=1}^{n} x_{fi}^2} \sqrt{\sum_{i=1}^{n} x_{si}^2}}.$$

The receiver operating characteristic (ROC) curve was used to evaluate the performance of the indexes. The index with the maximum area under ROC curve (AUC) was selected to detect AF on the testing set. The distribution of the selected index in different groups of the testing set was compared by Kruskal-Wallis test. And the AF identification performance on the testing set was evaluated by ROC curve.

3. Results

3.1. Performance of Candidate Indexes on Training Set.
The ROC curves of 3 candidate indexes on the training set are shown in Figure 8. The AUC of D_e, D_m, and D_c were 0.857, 0.801, and 0.516, respectively. Both D_e and D_m can effectively identify irregular pulse rhythm, and D_e performed best in this task.

3.2. Comparison of D_e in Different Groups of Testing Set.
The comparison result of D_e in different groups of testing set by Kruskal-Wallis test is shown in Table 1. And the box-plot
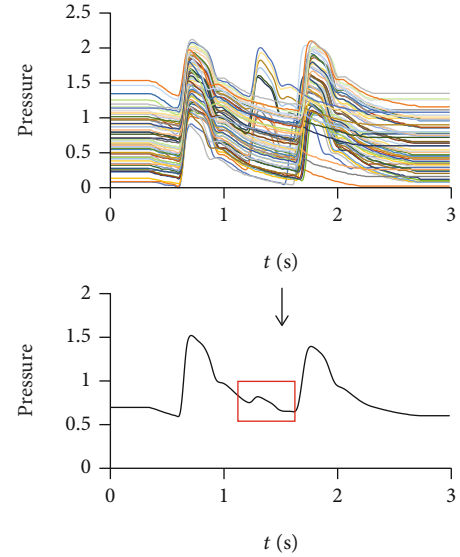

FIGURE 11: A typical average pulse wave of PC. There are still enough synchronized second ascending limbs in pulse wave of PC to form a similar average waveform with the first ascending limbs. The other second ascending limbs will form the bulge in the red box.

of D_e in different groups of testing set is shown in Figure 9. The result indicated that D_e in AF group is significantly higher than that of the other two groups, and there was no significant difference between the PC and SN group. Therefore, D_e can be used as an indicator to detect AF.

3.3. Performance of D_e to Detect AF on Testing Set.
The ROC curve of D_e to identify AF on the testing set are shown in Figure 10. The AUC was 0.998, and the accuracy, sensitivity, and specificity were 98.39%, 100%, and 98.04%, respectively, with the threshold of 2.1. D_e can effectively detect AF from pulse wave signals.

3.4. Comparison with Other Works.
With the popularity of wearable devices, the research of AF detection based on pulse wave is increasing in recent years (Table 2). However, except Shannon entropy, most of the features used to detect AF are based on the interbeat interval (IBI) series, which makes the accurate calculation of the cardiac cycle duration a prerequisite for AF detection. Therefore, the sensitivity to external interference has become a common weakness of these

studies. As indicated in Table 2, the method proposed in this paper is one of the most accurate methods. And it does not rely on IBI series, consequently, and may have stronger anti-interference ability.

4. Discussion

In this paper, we propose a new measure of PRV based on TSA. It was discovered that this new index can effectively detect AF from pulse wave signals. It can not only be applied to the seriously interfered signal by combining with single-period pulse wave quality evaluation method, but also extract a high-quality single-period pulse waveform at the same time, which can be used in other pulse wave-related applications. In addition, it can distinguish AF and PC, which has long been a problem in the identification of AF [21].

Figure 11 shows a typical average pulse wave of PC. Although the PRV of patient with PC is large, the change of its cardiac cycle durations is usually regular. Therefore, there are still enough synchronized second ascending limbs to form a similar average waveform with the first ascending limbs. The other second ascending limbs will form the bulge in the red box of Figure 9. This feature is usually located in the diastolic of the average waveform. The diastolic pulse wave of healthy people usually decreases gradually without obvious features. This feature may be used to detect PC in the future.

In summary, the new index provides a new perspective to measure PRV and to detect AF. Moreover, it accomplishes the monitoring of PRV and the extraction of single-period pulse wave through the same technical route, which may promote the popularization and application of pulse wave. However, this study also has limitations: (1) the sample size is so limited that we had to use simulated data instead of real clinical data to screening candidate indexes. Therefore, D_e may not be the best choice for real clinical data. (2) The ECG and pulse wave are not collected simultaneously, which may lead to incorrect label. (3) The anti-interference ability has not been verified because the new index was not tested on seriously interfered data set.

In the future, we hope to improve this algorithm by collecting more real clinical data and screening more distance measures. In addition, it has been discovered that there is a unique characteristic on the average pulse wave of PC. It is also one of the future research directions to develop an automatic PC detection algorithm based on this characteristic.

Abbreviations

PRV: Pulse rate variability
AF: Atrial fibrillation
TSA: Time synchronous averaging
PC: Premature contraction
SN: Sinus rhythm
ROC: Receiver operating characteristic
AUC: Area under curve
ECG: Electrocardiographic

IBI: Interbeat interval.

Data Availability

The original data used to support the findings of this study are available at https://github.com/Xiaodong-Ding/AFDetect.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. Naz, J. H. Shah, M. A. Khan, M. Sharif, M. Raza, and R. Damaševičius, "From ECG signals to images: a transformation based approach for deep learning," *PeerJ Computer Science*, vol. 7, article e386, 2021.

[2] J. M. Raja, C. Elsakr, S. Roman et al., "Apple watch, wearables, and heart rhythm: where do we stand?," *Annals of translational medicine*, vol. 7, no. 17, p. 417, 2019.

[3] T. Kabutoya, S. Takahashi, T. Watanabe et al., "Diagnostic accuracy of an algorithm for detecting atrial fibrillation in a wrist-type pulse wave monitor," *The Journal of Clinical Hypertension*, vol. 21, no. 9, pp. 1393–1398, 2019.

[4] M. Dörr, V. Nohturfft, N. Brasier et al., "The WATCH AF trial: SmartWATCHes for detection of atrial fibrillation," *JACC: Clinical Electrophysiology*, vol. 5, no. 2, pp. 199–208, 2019.

[5] N. G. Moura, I. Cordovil, and A. S. Ferreira, "Traditional Chinese medicine wrist pulse-taking is associated with pulse waveform analysis and hemodynamics in hypertension," *Journal of Integrative Medicine*, vol. 14, no. 2, pp. 100–113, 2016.

[6] Y. Hao, F. Cheng, M. Pham et al., "A noninvasive, economical, and instant-result method to diagnose and monitor type 2 diabetes using pulse wave: case-control study," *JMIR Mhealth Uhealth*, vol. 7, no. 4, article e11959, 2019.

[7] C. W. Chang, K. M. Liao, Y. T. Chang, S. H. Wang, Y. C. Chen, and G. C. Wang, "Fourth harmonic of radial pulse wave predicts adverse cardiac events in asymptomatic patients with type 2 diabetes," *Journal of Diabetes and its Complications*, vol. 33, no. 6, pp. 413–416, 2019.

[8] F. F. Li, R. Sun, S. Xue et al., "Pulse signal analysis of patients with coronary heart diseases using Hilbert-Huang transformation and time-domain method," *Chinese journal of integrative medicine*, vol. 21, no. 5, pp. 355–360, 2015.

[9] J. Chen, H. Huang, W. Hao, and J. Xu, "A machine learning method correlating pulse pressure wave data with pregnancy," *International journal for numerical methods in biomedical engineering*, vol. 36, no. 1, p. e3272, 2020.

[10] Z. D. Liu, J. K. Liu, B. Wen, Q. Y. He, Y. Li, and F. Miao, "Cuffless blood pressure estimation using pressure pulse wave signals," *Sensors*, vol. 18, no. 12, p. 4227, 2018.

[11] P. D. McFadden, "A revised model for the extraction of periodic waveforms by time domain averaging," *Mechanical Systems and Signal Processing*, vol. 1, no. 1, pp. 83–95, 1987.

[12] F. Combet and L. Gelman, "An automated methodology for performing time synchronous averaging of a gearbox signal without speed sensor," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2590–2606, 2007.

[13] X. Ding, F. Cheng, R. Morris, C. Chen, and Y. Wang, "Machine learning-based signal quality evaluation of single-period radial artery pulse waves: model development and validation," *JMIR medical informatics*, vol. 8, no. 6, article e18134, 2020.

[14] D. Wang and G. Lu, "Period segmentation for wrist pulse signal based on adaptive cascade thresholding and machine learning," in *2014 International Conference on Medical Biometrics*, Shenzhen, China, 2014.

[15] C. Xia, Y. Li, J. Yan et al., "A practical approach to wrist pulse segmentation and single-period average waveform estimation," in *2008 International Conference on BioMedical Engineering and Informatics*, Sanya, China, 2008.

[16] J. Koenig and J. F. Thayer, "Sex differences in healthy human heart rate variability: a meta-analysis," *Neuroscience and Biobehavioral Reviews*, vol. 64, pp. 288–310, 2016.

[17] G. Ramesh, D. Satyanarayana, and M. Sailaja, "Automatic detection of cardiac arrhythmia through ECG signal analysis: a review," *IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)*, vol. 7, no. 6, pp. 1–12, 2017.

[18] D. D. McManus, J. Lee, O. Maitas et al., "A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation," *Heart Rhythm*, vol. 10, no. 3, pp. 315–319, 2013.

[19] L. Krivoshei, S. Weber, T. Burkard et al., "Smart detection of atrial fibrillation," *Europace*, vol. 19, no. 5, pp. 753–757, 2016.

[20] S. Fallet, M. Lemay, P. Renevey, C. Leupi, E. Pruvot, and J. M. Vesin, "Can one detect atrial fibrillation using a wrist-type photoplethysmographic device?," *Medical & Biological Engineering & Computing*, vol. 57, no. 2, pp. 477–487, 2019.

[21] A. Kashiwa, F. Koyama, K. Miyamoto et al., "Performance of an atrial fibrillation detection algorithm using continuous pulse wave monitoring," *Annals of Noninvasive Electrocardiology*, vol. 24, no. 2, article e12615, 2019.

[22] U. Zalabarria, E. Irigoyen, and A. Lowe, "Diagnosis of atrial fibrillation based on arterial pulse wave foot point detection using artificial neural networks," *Computer Methods and Programs in Biomedicine*, vol. 197, p. 105681, 2020.

[23] D. Han, S. K. Bashar, F. Zieneddin et al., "Digital image processing features of smartwatch photoplethysmography for cardiac arrhythmia detection," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 4071–4074, Montreal, QC, Canada, 2020.