

Intelligent Decision Support Systems for Complex Healthcare Applications

Lead Guest Editor: K. Shankar

Guest Editors: Gyanendra Prasad Joshi and Mohamed Elhoseny





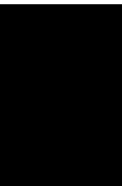
Intelligent Decision Support Systems for Complex Healthcare Applications

Journal of Healthcare Engineering

Intelligent Decision Support Systems for Complex Healthcare Applications

Lead Guest Editor: K. Shankar

Guest Editors: Gyanendra Prasad Joshi and
Mohamed Elhoseny



Copyright © 2023 Hindawi Limited. All rights reserved.

This is a special issue published in "Journal of Healthcare Engineering." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Associate Editors

Xiao-Jun Chen , China
Feng-Huei Lin , Taiwan
Maria Lindén, Sweden

Academic Editors





Cherif Adnen, Tunisia
Saverio Affatato , Italy
Óscar Belmonte Fernández, Spain
Sweta Bhattacharya , India
Prabadevi Boopathy , India
Weiwei Cai, USA
Gin-Shin Chen , Taiwan
Hongwei Chen, USA
Daniel H.K. Chow, Hong Kong
Gianluca Ciardelli , Italy
Olawande Daramola, South Africa
Elena De Momi, Italy
Costantino Del Gaudio , Italy
Ayush Dogra , India
Luobing Dong, China
Daniel Espino , United Kingdom
Sadiq Fareed , China
Mostafa Fatemi, USA
Jesus Favela , Mexico
Jesus Fontecha , Spain
Agostino Forestiero , Italy
Jean-Luc Gennisson, France
Badicu Georgian , Romania
Mehdi Gheisari , China
Luca Giancardo , USA
Antonio Gloria , Italy
Kheng Lim Goh , Singapore
Carlos Gómez , Spain
Philippe Gorce, France
Vincenzo Guarino , Italy
Muhammet Gul, Turkey
Valentina Hartwig , Italy
David Hewson , United Kingdom
Yan Chai Hum, Malaysia
Ernesto Iadanza , Italy
Cosimo Ieracitano, Italy

Giovanni Improta , Italy
Norio Iriguchi , Japan
Mihajlo Jakovljevic , Japan
Rutvij Jhaveri, India
Yizhang Jiang , China
Zhongwei Jiang , Japan
Rajesh Kaluri , India
Venkatachalam Kandasamy , Czech Republic
Pushpendu Kar , India
Rashed Karim , United Kingdom
Pasi A. Karjalainen , Finland
John S. Katsanis, Greece
Smith Khare , United Kingdom
Terry K.K. Koo , USA
Srinivas Koppu, India
Jui-Yang Lai , Taiwan
Kuruva Lakshmanna , India
Xiang Li, USA
Lun-De Liao, Singapore
Qiu-Hua Lin , China
Aiping Liu , China
Zufu Lu , Australia
Basem M. ElHalawany , Egypt
Praveen Kumar Reddy Maddikunta , India
Ilias Maglogiannis, Greece
Saverio Maietta , Italy
M.Sabarimalai Manikandan, India
Mehran Moazen , United Kingdom
Senthilkumar Mohan, India
Sanjay Mohapatra, India
Rafael Morales , Spain
Mehrbakhsh Nilashi , Malaysia
Sharnil Pandya, India
Jialin Peng , China
Vincenzo Positano , Italy
Saeed Mian Qaisar , Saudi Arabia
Alessandro Ramalli , Italy
Alessandro Reali , Italy
Vito Ricotta, Italy
Jose Joaquin Rieta , Spain
Emanuele Rizzuto , Italy


Dinesh Rokaya, Thailand
Sébastien Roth, France
Simo Saarakkala , Finland
Mangal Sain , Republic of Korea
Nadeem Sarwar, Pakistan
Emiliano Schena , Italy
Prof. Asadullah Shaikh, Saudi Arabia
Jiann-Shing Shieh , Taiwan
Tiago H. Silva , Portugal
Sharan Srinivas , USA
Kathiravan Srinivasan , India
Neelakandan Subramani, India
Le Sun, China
Fabrizio Taffoni , Italy
Jinshan Tang, USA
Ioannis G. Tollis, Greece
Ikram Ud Din, Pakistan
Sathishkumar V E , Republic of Korea
Cesare F. Valenti , Italy
Qiang Wang, China
Uche Wejinya, USA
Yuxiang Wu , China
Ying Yang , United Kingdom
Elisabetta Zanetti , Italy
Haihong Zhang, Singapore
Ping Zhou , USA

Contents







Lung Diseases Detection Using Various Deep Learning Algorithms

M. Jasmine Pemeena Priyadarsini, Ketan kotecha , G. K. Rajini, K. Hariharan, K. Utkarsh Raj, K. Bhargav Ram, V. Indragandhi , V. Subramaniaswamy , and Sharnil Pandya 
Research Article (13 pages), Article ID 3563696, Volume 2023 (2023)



An Ensemble of Deep Learning Enabled Brain Stroke Classification Model in Magnetic Resonance Images

Ala' A. Eshmawi, Mashael Khayyat, Abeer D. Algarni , and Inès Hilali-Jaghdam
Research Article (11 pages), Article ID 7815434, Volume 2022 (2022)





Machine Learning with Quantum Seagull Optimization Model for COVID-19 Chest X-Ray Image Classification

Mahmoud Ragab , Samah Alshehri , Nabil A. Alhakamy , Wafaa Alsaggaf , Hani A. Alhadrami , and Jaber Alyami 
Research Article (13 pages), Article ID 6074538, Volume 2022 (2022)

Arithmetic Optimization with RetinaNet Model for Motor Imagery Classification on Brain Computer Interface

Areej A. Malibari, Fahd N. Al-Wesabi , Marwa Obayya, Mimouna Abdullah Alkhonaini, Manar Ahmed Hamza , Abdelwahed Motwakel, Ishfaq Yaseen, and Abu Sarwar Zamani
Research Article (11 pages), Article ID 3987494, Volume 2022 (2022)


Optimal Deep Learning Enabled Prostate Cancer Detection Using Microarray Gene Expression

Abdulrhman M. Alshareef , Raed Alsini , Mohammed Alsieni , Fadwa Alrowais, Radwa Marzouk, Ibrahim Abunadi, and Nadhem Nemri 
Research Article (12 pages), Article ID 7364704, Volume 2022 (2022)






Estimation and Prediction of Hospitalization and Medical Care Costs Using Regression in Machine Learning

Ahmed I. Taloba , Rasha M. Abd El-Aziz , Huda M. Alshanbari , and Abdal-Aziz H. El-Bagoury 
Research Article (10 pages), Article ID 7969220, Volume 2022 (2022)





Optimal Deep-Learning-Enabled Intelligent Decision Support System for SARS-CoV-2 Classification

Ashit Kumar Dutta, Nasser Ali Aljarallah, T. Abirami, M. Sundarajan, Seifedine Kadry, Yunyoung Nam , and Chang-Won Jeong
Research Article (14 pages), Article ID 4130674, Volume 2022 (2022)

Fusion-Based Deep Learning with Nature-Inspired Algorithm for Intracerebral Haemorrhage Diagnosis

Nada M. Alfaer , Hassan M. Aljohani , Sayed. Abdel-Khalek , Abdulaziz S. Alghamdi , and Romany F. Mansour 
Research Article (12 pages), Article ID 4409336, Volume 2022 (2022)

Metaheuristics with Deep Learning-Enabled Parkinson's Disease Diagnosis and Classification Model

Adel A. Bahaddad , Mahmoud Ragab , Ehab Bahaudien Ashary , and Eied M. Khalil 
Research Article (14 pages), Article ID 9276579, Volume 2022 (2022)

Research Article

Lung Diseases Detection Using Various Deep Learning Algorithms

M. Jasmine Pemeena Priyadarsini,¹ Ketan kotecha^{2,3} , G. K. Rajini,⁴ K. Hariharan,¹ K. Utkarsh Raj,¹ K. Bhargav Ram,¹ V. Indragandhi⁴ , V. Subramaniaswamy⁵ , and Sharnil Pandya⁶ 

¹School of Electronics Engineering, Vellore Institute of Technology, Vellore 632014, India

²Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International (Deemed University), Pune, India

³School of Mathematical Sciences, Sunway University, Subang Jaya, Malaysia

⁴School of Electrical Engineering, Vellore Institute of Technology, Vellore 632014, India

⁵School of Computing, SASTRA Deemed University, Thanjavur, India

⁶Computer Science and Media Technology Department, Faculty of Technology, Linnaeus University, P G Vejdes väg 351 95, Växjö, Sweden

Correspondence should be addressed to Sharnil Pandya; sharnil.pandya@lnu.se

Received 24 March 2022; Revised 17 August 2022; Accepted 24 November 2022; Published 3 February 2023

Academic Editor: Ayush Dogra

Copyright © 2023 M. Jasmine Pemeena Priyadarsini et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The primary objective of this proposed framework work is to detect and classify various lung diseases such as pneumonia, tuberculosis, and lung cancer from standard X-ray images and Computerized Tomography (CT) scan images with the help of volume datasets. We implemented three deep learning models namely Sequential, Functional & Transfer models and trained them on open-source training datasets. To augment the patient's treatment, deep learning techniques are promising and successful domains that extend the machine learning domain where CNNs are trained to extract features and offers great potential from datasets of images in biomedical application. Our primary aim is to validate our models as a new direction to address the problem on the datasets and then to compare their performance with other existing models. Our models were able to reach higher levels of accuracy for possible solutions and provide effectiveness to humankind for faster detection of diseases and serve as best performing models. The conventional networks have poor performance for tilted, rotated, and other abnormal orientation and have poor learning framework. The results demonstrated that the proposed framework with a sequential model outperforms other existing methods in terms of an F1 score of 98.55%, accuracy of 98.43%, recall of 96.33% for pneumonia and for tuberculosis F1 score of 97.99%, accuracy of 99.4%, and recall of 98.88%. In addition, the functional model for cancer outperformed with an accuracy of 99.9% and specificity of 99.89% and paves way to less number of trained parameters, leading to less computational overhead and less expensive than existing pretrained models. In our work, we implemented a state-of-the-art CNN with various models to classify lung diseases accurately.

1. Introduction

Lungs play a vital role in the human system, which performs expansion and relaxation to bring in oxygen and take out carbon dioxide. Lung diseases are respiratory diseases that affect the various organs and tissues associated with breathing, leading to airway diseases, lung tissue diseases, and lung circulation diseases. Some of the respiratory diseases like common cold and influenza cause mild discomfort and hindrance while others like pneumonia, tuberculosis

and lung cancer are life-threatening and cause severe acute respiratory problems [1].

According to a research study done by the Forum of International Respiratory Societies called "The Global Impact of Respiratory Disease," 10.4 million people suffered mild or severe symptoms of tuberculosis, and 1.4 million of those affected died as per the survey reported [2]. Lung cancer kills an astounding number of people every year. More than 1.6 million people were reported to have died in the year the survey was carried out. Pneumonia is one of the

top respiratory diseases and 1.23 million children under the age of 5 died due to pneumonia according to the Johns Hopkins Bloomberg School of Public Health report titled "Pneumonia and Diarrhea Progress Report 2020" [3]. Detection of the abovementioned diseases at the early stages of infection can drastically increase the chances of survival and can prevent human casualties. Chest X-ray images and CT scans are common examinations that determine the presence of these diseases [4]. The presence of trained professionals is required to examine the scanned images and determine the infections. According to the Union Health Ministry data statistics there exists a shortfall of 76.1 percent of physicians at the Community Health Centres (CHCs) in rural areas. To overcome this, deep learning techniques are implemented, which pave the way for a new strategy.

Deep learning is a branch of machine learning that provides state-of-the-art accuracy and is a subset of the artificial intelligence with representation learning. This tool has drawn attention in recent times due to their ability to read image data, process them, and provide results based on the previously trained data [5]. Deep learning models can learn features and patterns from dataset images and use the learned features to classify new test images that have not been previously visualized by the model.

Numerous works have already been conducted by researchers around the world and have led to promising results. These works can help support existing methods or open pathways to new ones that could not have been possible. These advancements can help in quick and accurate detection as well as classification of diseases and provide quick support to obtain impressive results to eliminate deadly infectious diseases.

The rest of the manuscript is structured as follows: In Section 2, we introduce the basic foundation of the convolution neural networks. Section 3 describes the architecture of the proposed model. Section 4 discusses the implementation of the proposed CNN technique and the experimental results. Finally, we conclude our work with a summary and future directions in Section 5.

2. Related Work

One of the best techniques currently used in medical image analysis are CNNs, which have a remarkable efficiency in classifying the images. Some of the contemporary CNN models are Pre-Trained, Functional, and Sequential, which are reviewed in the forthcoming sections.

Liu et al. proposed three different types for the applications of CNN-trained models in tuberculosis detection. In all these three methods, features are extracted by the CNN architectures and are trained by the support vector machine (SVM), and in the second proposal, features are extracted from coreference resolution (CR) and are trained in the SVM classifier. In the third proposal, these two proposals are combined together to create an ensemble of the classifiers. The Montgomery dataset has a total of 138 X-ray images and the Shenzhen dataset has a total of 662 X-ray images. These trained models help reduce the processing time but provide low accuracy, which is not conducive and cannot be implemented in medical diagnosis.

Amit Kumar Jaiswal, Prayag Tiwari, Sachin Kumar, and Deepak Gupta proposed a method called mask RCNN. It is a deep neural network model that can extract two types of features: global and local. Pixel-wise division is carried out and this method is expected to have a better performance evaluated on the radiograph dataset. This technique highlights the infected regions and provides a heat map for better understanding for people looking at the results. But they have ensembled ResNet50 and ResNet101 (Mask RCNN models) but achieved less biased results than expected and require more GPU processing power to train.

Elshennawy and Ibrahim, presented on four different models. Among these four models, CNN and LSTM-CNN started from the beginning and the other two are pre-trained models and the specific models used are ResNet152v2 and MobileNetV2. They formulated to create from the ground, a deep learning neural network model, which could diagnose pneumonia symptoms using chest X-ray images, which has pneumonia [6]. Some of the disadvantages are that it has a humongous architecture with hundreds of millions of trainable parameter weights [7, 8]. This type of model requires high computing and processing power.

Various deep learning techniques, Naik and Edla [9] developed a lung nodule classification and identification model for computed tomography (CT) images. The CT scans required a computer-aided detection system for categorizing the lung nodule into benign and malignant types, along with the highest level of accuracy to protect from a delay in diagnosis. The deep learning approaches used to categorize the lung nodule have positive outcomes compared to other methods. When the mutations were implemented in the deep learning architecture, the accuracy of the classification system increased rapidly. The deep learning method was used to specify the new impacts in nodule classification and also recognized the preliminary stage of a malignant lesion [10].

3. Proposed Methodology

This section discusses the datasets used, the preprocessing, the data augmentation methods, and the various algorithms used. The workflow of the proposed technique is presented in a flowchart form in Figure 1.

3.1. Datasets. All the datasets used in this work are from open-source datasets published on the website "Kaggle."

The pneumonia dataset published by Paul Mooney contains 5,856 frontal chest X-ray images, 1,583 images of the dataset are of people with no abnormalities in their lungs, and 4,273 images predict some abnormalities and symptoms of pneumonia.

The tuberculosis dataset published by Scott Mader has 662 frontal X-rays. These images were collected by physicians in the Guangdong Hospital, Shenzhen, China. Hence, this dataset is commonly known as the Shenzhen dataset. It contains 326 images, which contains lung images of healthy persons and in turn contains 336 images that are infected by tuberculosis.

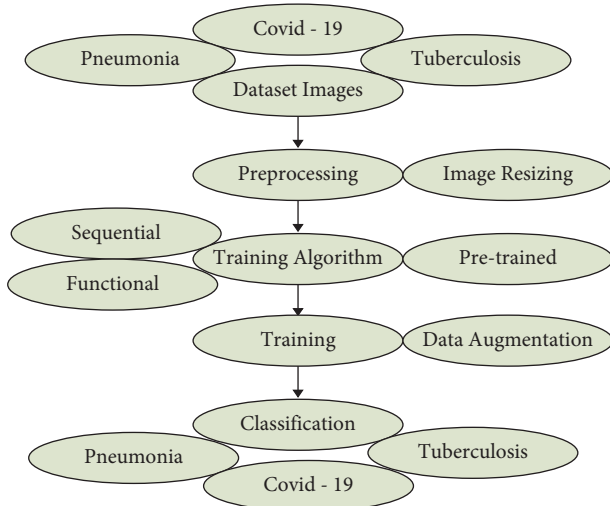


FIGURE 1: Workflow of the classification model.

The cancer dataset published by Mohamed Hany has 907 lung CT-scan images, 215 images of the dataset are of people with no signs of cancer, and 692 images of the dataset are of people with cancer. The dataset contains 3 types of cancer images: adenocarcinoma, large cell carcinoma, and squamous cell carcinoma. Figure 2 shows few sample images from the CT-scan dataset.

3.1.1. Preprocessing and Data Augmentation. The images present in the datasets are of different resolutions. However, the CNN models require images to be of one specified size. Hence, all the images in the dataset were resized to 224×224 . Lowering the input image size helps process a faster execution of images and thus, makes the model faster for the specific associated task.

Data augmentation is a common support method used to significantly increase the training data volume by introducing slight variations of an image in each training epoch. The variations used in this work are horizontal flip, zoom, shear, rotation, and rescale. This technique is essential to get high levels of accuracy as the CNN model is able to train on more data than originally present in the dataset. Figure 3 shows the variations that can be created from one sample image.

3.2. Deep Learning Algorithms. In recent times, a dataset of medical images has been available in the Kaggle repository. In this paper, this dataset has been implemented using the novel models of CNN, namely, sequential and functional models, combining CNN and data augmentation. Three different model algorithms were deployed in this proposed work. These are explained in detail in the following subsections.

3.2.1. Sequential Model. The sequential model is a model in which layers are stacked to form a sequential order. The input is passed through all the layers in the order in which the layers are stacked. Features are learnt at each and every

layer and more deeper into the layer, the model is capable of distinguishing the infected areas and noninfected areas from the chest X-rays [11].

The proposed sequential model has five convolutional layers with the number of filters increasing as it proceeds deeper into the network [12]. The alpha parameter was set to 0.66. Leaky ReLU allows a small gradient to pass through, while ReLU completely removes any gradient when the unit is not active. In addition, max pooling was carried out after each activation. Adam optimizer and learning rate of 0.0001 was employed. The block diagram of the sequential model is presented in Figure 4.

3.2.2. Functional Model. The functional model has more flexibility than the other algorithms. It can form connections between any two layers contrary to the others and progress in a linear fashion. This allows us to create more complicated and sophisticated networks [13]. The input goes through the first layer and then proceeds along the designed architecture. This method also trains from the beginning, contrary to the pretrained model.

The proposed functional model has two convolution layers of 7×7 window and another with 1×1 on top of 3×3 window as presented in Figure 5. The input is passed through both convolution layers separately and then the output from both layers is appended and then passed to five 3×3 convolution layers. The Adam optimizer with learning rate = 0.0001 was employed.

3.2.3. Pretrained Model. This is the easiest and most commonly used model for image classification. Instead of training a model from the beginning, this technique uses already trained weights on a large dataset of images to classify the required images [14, 15]. This technique is also called transfer learning as previously learned weights are transferred and used for classification. Generally, this model takes less time to train and produces better results and accuracy.

The pretrained model used here is VGG-16, a convolutional neural network (CNN), famous for high accuracy and achieved the top 5 accuracies in the ImageNet competition with an accuracy of 97.7%.

4. Results and Discussion

The various models were trained, their accuracies and losses were plotted, and the test accuracy was obtained and compared with other research works for lung disease detection with CNN [16, 17]. The performance metrics involved in this proposed work are accuracy, precision, recall, and F1 score.

- (i) Accuracy represents the number of correctly classified data instances over the total number of data instances.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (1)$$

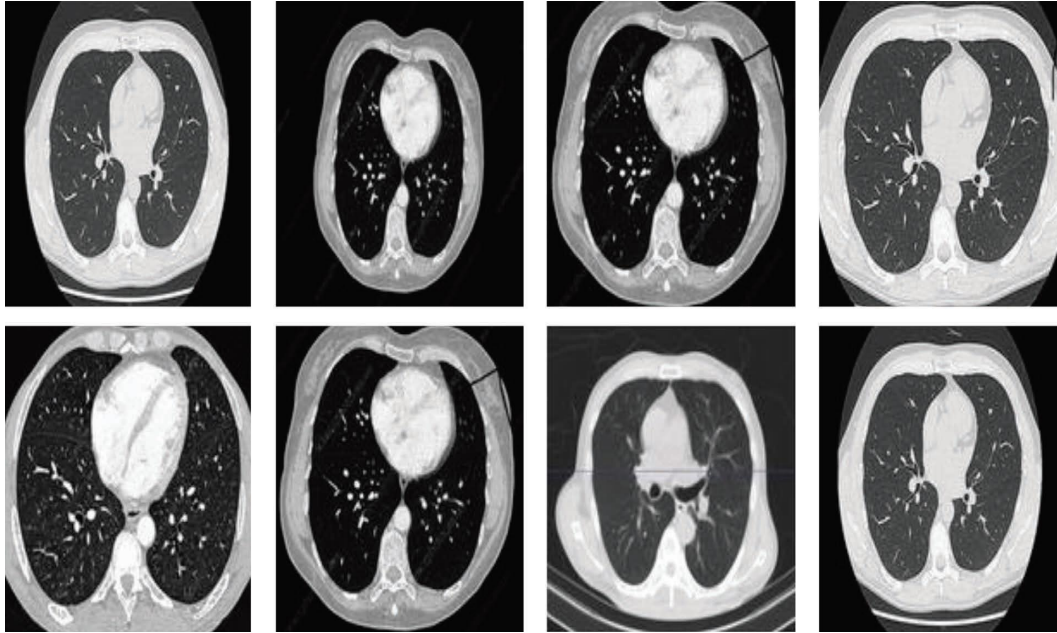


FIGURE 2: Chest-CT scan images (source: kaggle).

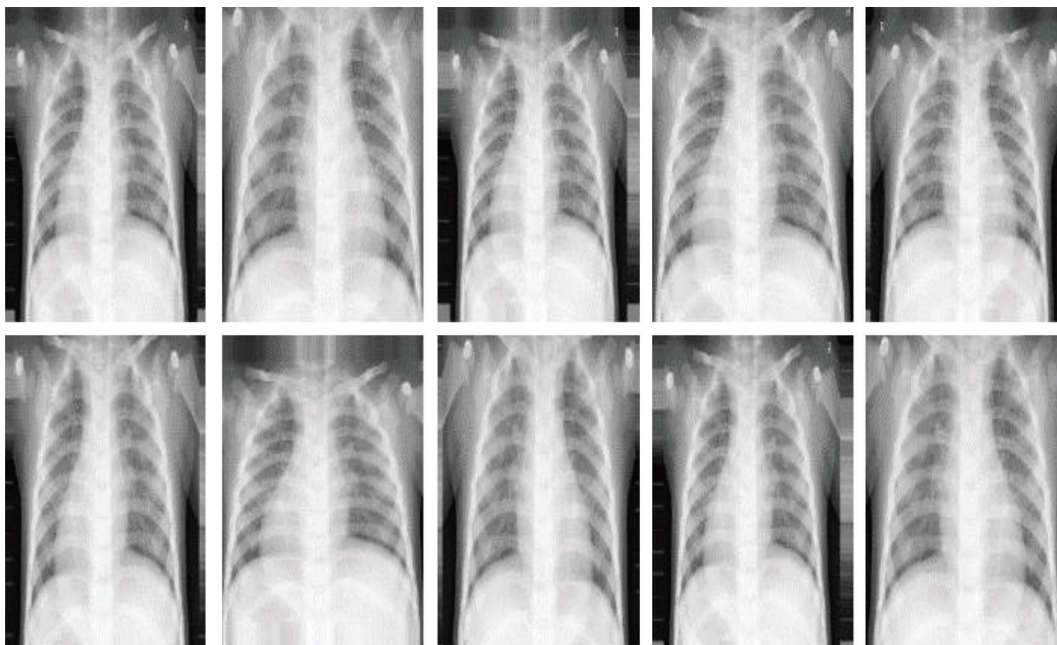


FIGURE 3: Variations of a chest X-ray image.

where, true positive is abbreviated as TP, true negative as TN, false positive as FP, and false negative as FN.

- (ii) Precision should ideally be 1 (high) for a good classifier. Precision becomes 1 only when the numerator and denominator are equal, i.e, $TP = TP + FP$, this also means FP is zero. As FP increases, the value of the denominator becomes greater than the numerator and the precision value decreases.

$$\text{precision} = \frac{TP}{TP + FP}. \quad (2)$$

- (iii) Recall is also known as sensitivity or true positive rate and is defined as follows:

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (3)$$

- (iv) F1-score is a metric that takes into account both precision and recall and is defined as follows:

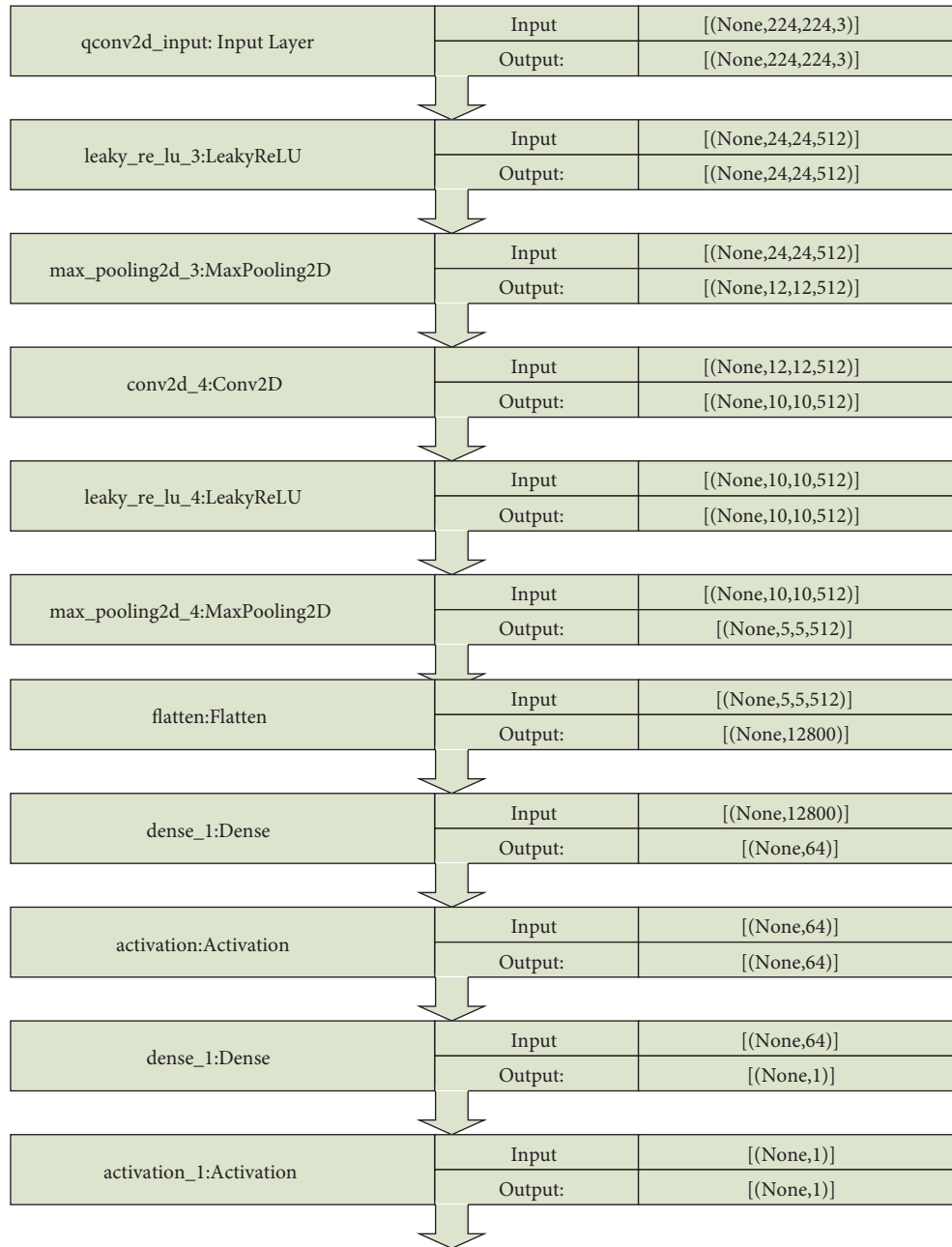


FIGURE 4: Block diagram of sequential model.

$$F1 \text{ score} = 2 * \frac{\text{precision} * \text{Recall}}{\text{precision} + \text{Recall}} \quad (4)$$

4.1. Sequential Model for Pneumonia. In medical diagnostics, it is common to analyze the classifier performance using sensitivity (true positive rate) and specificity (true negative rate) instead of accuracy [6]. To assess the overall classification F1 score is computed [7, 18]. From the dataset of 5,856 chest X-ray images, 2,000 images were used for training of which 1,000 images were of normal chest X-rays and the other 1,000 images were of pneumonia-infected chest X-rays.

The model was trained for 50 epochs. Figure 6 shows the increase in accuracy as the model trains with trained set images and Figure 7 shows that the loss encountered with this model is less. The accuracy starts from 75% and gradually increases to 90% with 10 epochs.

After training, the model was used to predict the labels of test images that were not known by the model during training. The test image set had 583 images of normal chest X-rays and 3,273 images of pneumonia-infected chest X-rays. Table 1 provides accuracy of our model with the existing works related to pneumonia and found that our model outperforms other existing works. The model predicted the labels accurately for 533 images from 583 normal

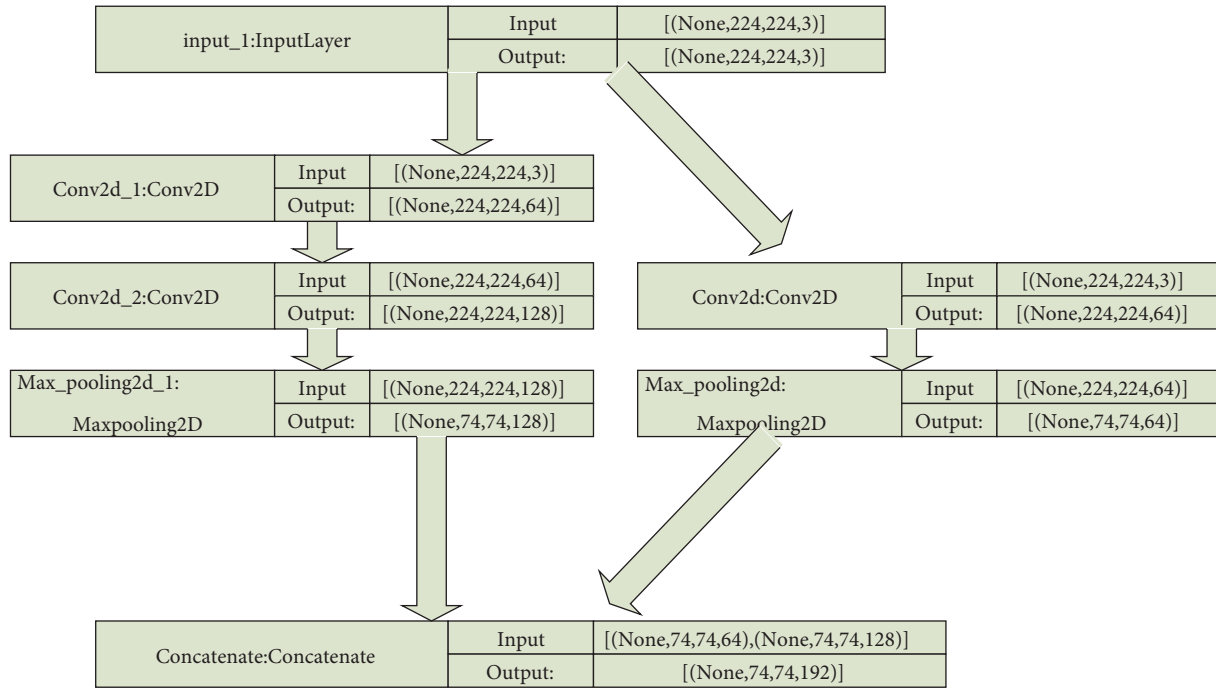


FIGURE 5: Block diagram of functional model.

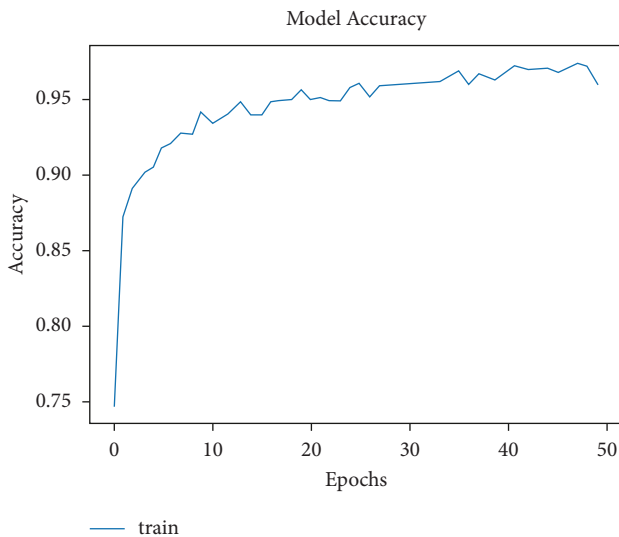


FIGURE 6: Model accuracy vs. epochs for pneumonia (sequential).

CXR images and 3,070 images from 3,273 pneumonia-infected CXR images.

4.2. *Sequential Model for Tuberculosis.* The tuberculosis dataset has a grand total of 662 chest X-ray images. Among 662 images, 285 images of normal chest X-rays and 292 images of tuberculosis-infected chest X-rays were used for training. As depicted in Figure 8, the tuberculosis model started with a very low accuracy of 50%. After training for around 100 epochs, the model accuracy value of 97% was obtained.

Numerous works have already been carried out by researchers around the world and have led to promising

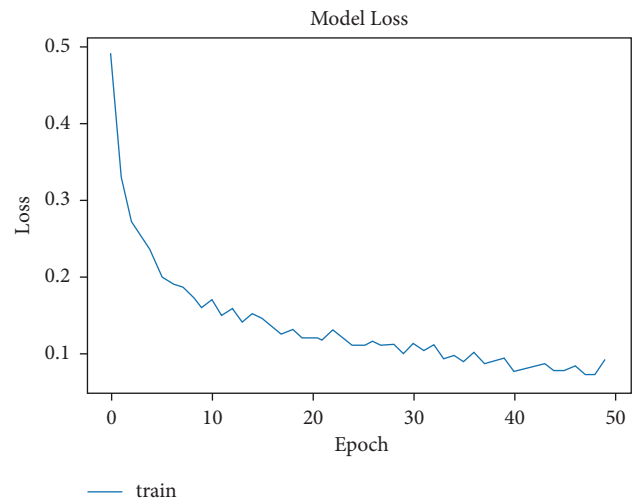


FIGURE 7: Model loss vs. epochs for pneumonia (sequential).

results. These works can help support the existing methods or open pathways to new methods, which could not have been possible before [8, 24]. These advancements can help in faster and accurate detection, as well as classification of diseases and provide support to obtain impressive results to eliminate deadly infectious diseases.

The model was used to predict the labels for test images. The test image set had a total of 85 images of which 41 were of normal and 44 were of tuberculosis-infected. The model predicted 37 images of normal and 39 images of tuberculosis-infected accurately as presented in Figure 9. Table 2 provides an accuracy of our model with existing works related to tuberculosis and finds to be superior when compared to other existing works.

TABLE 1: Comparison of our model with existing works related to pneumonia.

Related work	Dataset	Techniques	F1score	Accuracy	Recall	Precision	Specificity
Sourab and Kabir [19]	Mendeley datav2 (5856 images)	CNN	96	94.55	94	99	85
Gayathri [20]	GitHub (783 images)	CNN	95.63	95.78	—	95.63	95
Vrbanić and Podgorelec [21]	ImageNet	Ensemble	95.01	91.98	93.8	96.38	90
Fernandes et al. [22]	CXR dataset (3883 images)	CNN	93.2	95.7	91.2	95.1	97.5
Hasija et al. [23]	—	CNN	98.52	98.38	—	—	99.81
Our proposed model	Chest X-ray (1000 images)	CNN-sequential model	98.55	98.4375%	96.33%	99%	99.89%

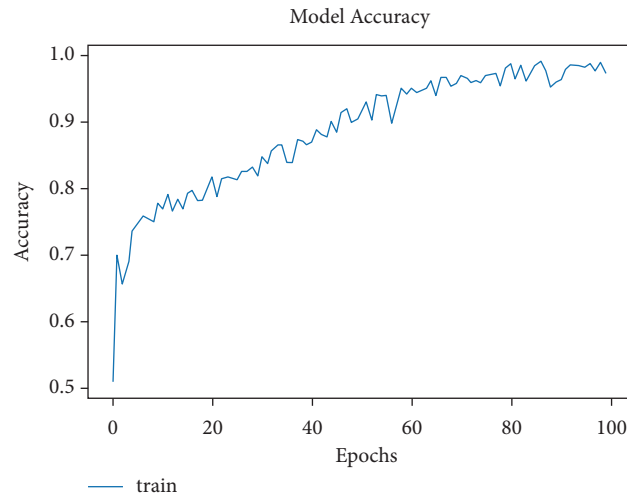


FIGURE 8: Model accuracy vs. epochs for tuberculosis.

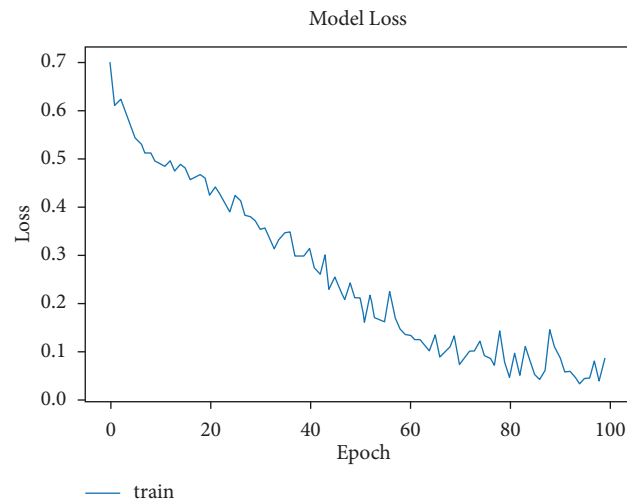


FIGURE 9: Model loss vs. epochs for tuberculosis.

4.3. Functional Model for Cancer. The dataset has a total of 907 lung CT-scan images, 215 images of people with no signs of cancer, and 692 images of people infected with cancer were used for training the model [30, 31]. The model was trained for 100 epochs. As seen in Figure 10, the model started with an accuracy of 70% and increased to 90% in about 10 epochs.

The model was presented to predict for the test images. The test dataset had a total of 278 images, of which 224 were cancer infected and 54 were normal. The model accurately predicted 54 images of normal and 204 images of cancer infected, and the loss is shown in Figure 11.

Table 3 depicts the accuracy of our model with existing works related to cancer and finds to be extraordinary when compared to other existing works.

4.4. Functional Model for Pneumonia. The dataset used for the functional model is the one that was utilized in sequential

model. The model accuracy starts from around 81% and rapidly increases to 90% in less than 5 epochs.

4.5. Pretrained Model for Pneumonia. Figure 12 shows the model accuracy gradual improvement for the pneumonia disease with the functional model. The same dataset was used for this model, i.e., from Paul Mooney with 5,856 images of which 1,000 are normal X-rays and other 1,000 are infected chest X-rays. As the model has already been trained before, the starting accuracy is very good. There is a minor improvement after training for 15 epochs as is evident in Figure 13.

Figure 14 shows that the initial model loss is low as compared to the other models, hence, there is no continuous progress like that in the sequential and functional models.

Figure 15 shows that the pretrained models are easy to train and that the loss gradually decreases as they have previously been trained on various datasets.

TABLE 2: Accuracy of our model with existing works related to tuberculosis.

Authors	Dataset	Techniques	F1-score	Accuracy	Recall	Precision
Momeny et al. [25]	Annotated (1078 images)	CNN	89	93	85	92
Lopes and Valiati [26]	ImageNet (10000 images)	CNN	—	76	—	—
Sineglazov et al. [27]	National institute of phthisiology and pulmonology (9311 slices)	CNN	99.04	99	97.6	98.34
Mamalakis et al. [28]	CXR dataset (3883 images)	ResNet CNN	81.64	71	91	78.6
Duong et al. [29]	(i) Montgomery county & CXR dataset (138 images) (ii) Shenzhen dataset (662 images)	CNN	97.9	99	—	—
Our proposed model	Chest X-ray (662 images)	CNN	97.99	99.4%	98.8	98.55

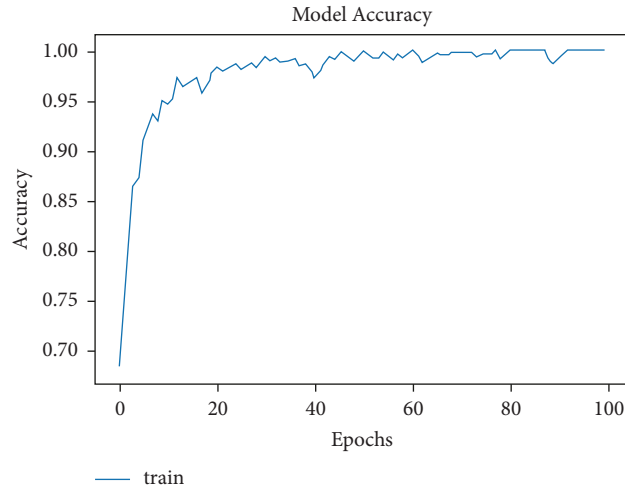


FIGURE 10: Model accuracy vs. epochs for cancer.

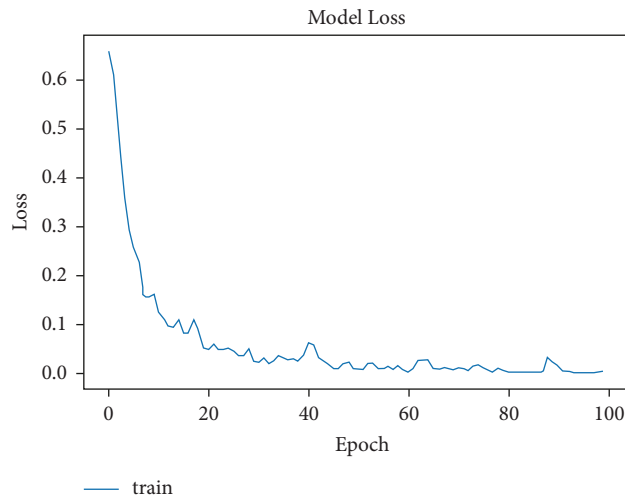


FIGURE 11: Model loss vs. epochs for cancer.

TABLE 3: Accuracy of our model with existing works related to cancer.

Authors	Dataset	Techniques	Accuracy	Specificity	Sensitivity
Lee et al. [32]	Annotated dataset-87	CNN	92.5	—	—
Tomassini et al. [33]	Planar data	CNN	74	—	81%
Wei et al. [34]	Annotated dataset-500 images	CNN	99.3	98.31	100
Desai and Shah [35]	Annotated dataset-1000 images	MLP	91.92	92.3	91
Hassantabar et al. [36]	Annotated dataset-682 images	CNN	93.20	99.71	96.09
Our proposed model	278 images	CNN	99.9	99.89	100

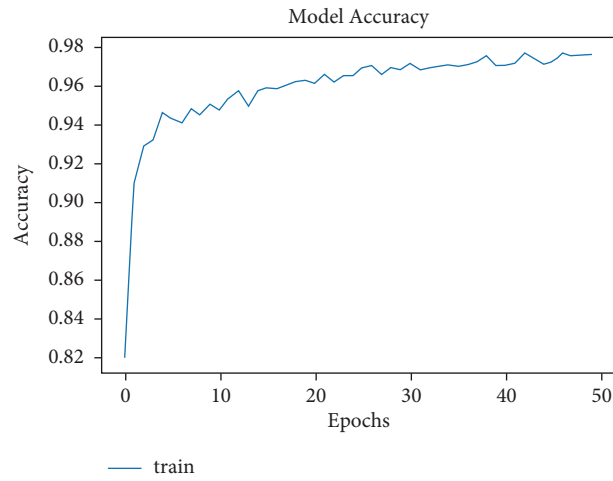


FIGURE 12: Model accuracy vs. epochs for pneumonia (functional).

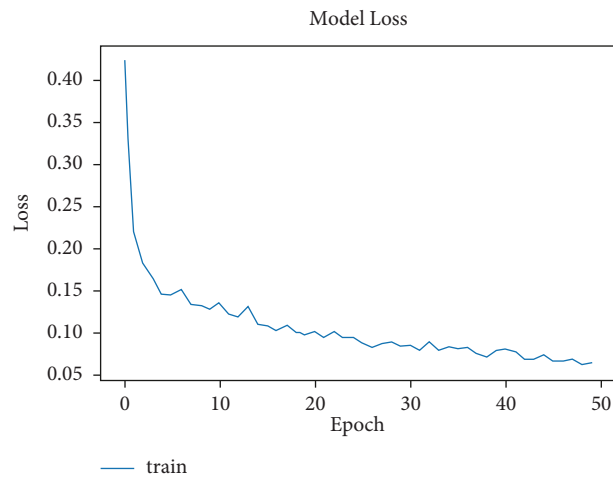


FIGURE 13: Model loss vs. epochs for pneumonia (functional).

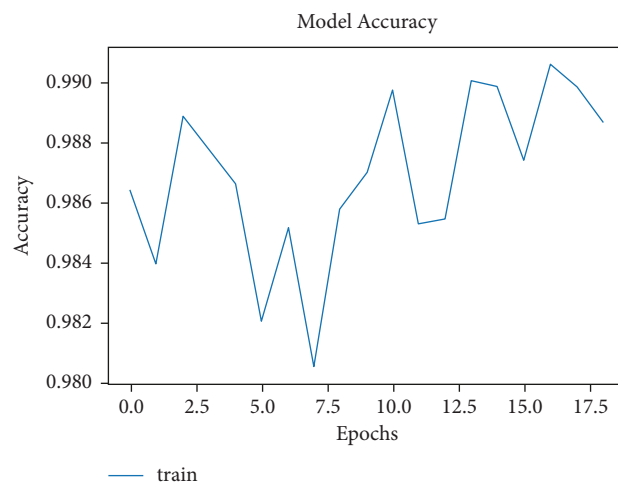


FIGURE 14: Model accuracy vs. epochs for pneumonia (pretrained).

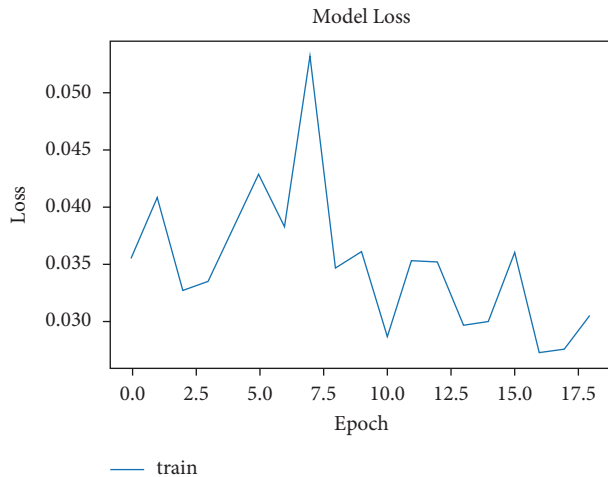


FIGURE 15: Model loss vs. epochs for pneumonia (pretrained).

5. Conclusion and Future Work

We have proposed three different architecture models of CNNs, which were used to train on various lung diseases that are available in the open-source dataset. The trained models were used to predict the labels of some test images that were not visualized by the models. The results of the proposed models performed better than other related works. The results obtained through this framework with a sequential model outperform other existing methods in terms of F1 score, accuracy and recall for pneumonia and for tuberculosis. In addition, the functional model for cancer outperformed with accuracy and specificity, and it requires less computation cost and time. In future, varying the optimizers, learning rate, and introduction of more data augmentation could potentially lead to further improvements in the classification accuracy of the proposed CNN models. Early stopping techniques will likely provide further insights into diagnosing lung diseases that can be passed down to avoid overfitting.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

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

References

- [1] WHO, *Tuberculosis*, World Health Organization, Mexico, UK, 2018.
- [2] A. A. Cruz, *Global Surveillance, Prevention and Control of Chronic Respiratory Diseases: A Comprehensive Approach*, World Health Organization, Mexico, UK, 2007.
- [3] International Vaccine Access Center Johns Hopkins Bloomberg School of Public Health, *Pneumonia and Diarrhea Progress Report 2020*, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA, 2020.
- [4] D. Shen, G. Wu, and H. I. Suk, "Deep learning in medical image analysis," *Annual Review of Biomedical Engineering*, vol. 19, no. 1, pp. 221–248, 2017.
- [5] J. Ma, Y. Song, X. Tian, Y. Hua, R. Zhang, and J. Wu, "Survey on deep learning for pulmonary medical imaging," *Frontiers of Medicine*, vol. 14, pp. 450–469, 2020.
- [6] N. M. Elshennawy and D. M. Ibrahim, "Deep-pneumonia framework using deep learning models based on chest X-ray images," *Diagnostics*, vol. 10, no. 9, p. 649, 2020.
- [7] Xi Ouyang, J. Huo, L. Xia et al., "Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2595–2605, 2020.
- [8] M. M. Ahsan, T. E. Alam, T. Trafalis, and P. Huebner, "Deep MLP-CNN model using mixed-data to distinguish between COVID-19 and non-COVID-19 patients," *Symmetry*, vol. 12, no. 9, p. 1526, 2020.
- [9] A. Naik and D. R. Edla, "Lung nodule classification on computed tomography images using deep learning," *Wireless Personal Communications*, vol. 116, pp. 655–690, 2021.
- [10] M. Kanipriya, C. Hemalatha, N. Sridevi, S. SriVidhya, and S. Jany Shabu, "An improved capuchin search algorithm optimized hybrid CNN-LSTM architecture for malignant lung nodule detection," *Biomedical Signal Processing and Control*, vol. 78, no. 2022, Article ID 103973, 2022.
- [11] L. Li, L. Qin, Z. Xu et al., "Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy," *Radiology*, vol. 296, no. 2, pp. E65–E71, 2020.
- [12] S. G. Armato, G. McLennan, L. Bidaut et al., "The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on CT scans," *Medical Physics*, vol. 38, no. 2, pp. 915–931, 2011.
- [13] X. W. Gao, C. James-reynolds, and E. Currie, "Analysis of tuberculosis severity levels from CT pulmonary images based on enhanced residual deep learning architecture," *Neuro-computing*, vol. 392, 2019.
- [14] N. Alsharman and I. Jawarneh, "GoogleNet CNN neural network towards chest CT coronavirus medical image classification," *Journal of Computer Science*, vol. 16, no. 5, pp. 620–625, 2020.
- [15] D. Singh, V. Kumar, Vaishali, and M. Kaur, *Classification of COVID-19 Patients from Chest CT Images Using Multi-Objective Differential Evolution-Based Convolutional Neural Networks*, European Journal of Clinical Microbiology & Infectious Diseases, Berlin, Germany, 2020.
- [16] Y. Rivenson, Z. Gorocs, H. Gunaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," *Optica*, vol. 4, pp. 1437–1443, 2017.
- [17] J. A. Quinn, R. Nakasi, P. K. B. Mugagga, P. Byanyima, W. Lubega, and A. Andama, "Deep convolutional neural networks for microscopy-based point of care diagnostics," in *Proceedings of the Machine Learning for Healthcare Conference PMLR*, August 2016, <https://arxiv.org/abs/1608.02989>.
- [18] K. Kuan, M. Ravaut, G. Manek et al., "Deep Learning for Lung Cancer Detection: Tackling the Kaggle Data Science Bowl 2017 challenge," 2017, <https://arxiv.org/abs/1705.09435>.
- [19] S. Y. Sourab and M. A. Kabir, "A comparison of hybrid deep learning models for pneumonia diagnosis from chest radiograms," *Sensors International*, vol. 3, Article ID 100167, 2022.
- [20] J. L. Gayathri, "A computer-aided diagnosis system for the classification of COVID-19 and non-COVID-19 pneumonia on chest X-ray images by integrating CNN with sparse

- autoencoder and feed forward neural network,” *Computers in Biology and Medicine*, vol. 141, Article ID 105134, 2022.
- [21] G. Vrbančić and V. Podgorelec, “Efficient ensemble for image-based identification of Pneumonia utilizing deep CNN and SGD with warm restarts,” *Expert Systems with Applications*, vol. 187, Article ID 115834, 2022.
- [22] V. Fernandes, G. B. Junior, A. C. de Paiva, A. C. Silva, and M. Gattass, “Bayesian convolutional neural network estimation for pediatric pneumonia detection and diagnosis,” *Computer Methods and Programs in Biomedicine*, vol. 208, Article ID 106259, 2021.
- [23] S. Hasija, P. Akash, M. Bhargav Hemanth, A. Kumar, and S. Sharma, “A novel approach for detection of COVID-19 and Pneumonia using only binary classification from chest CT-scans,” *Neuroscience Informatics*, vol. 2, Article ID 100069, 2022.
- [24] R. Golan, C. Jacob, and J. Denzinger, “Lung nodule detection in CT images using deep convolutional neural networks,” in *Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN)*, IEEE, Vancouver, BC, Canada, July 2016.
- [25] M. Momeny, A. A. Neshat, A. Gholizadeh et al., “Greedy Autoaugmentation for classification of mycobacterium tuberculosis image via generalized deep CNN using mixed pooling based on minimum square rough entropy,” *Computers in Biology and Medicine*, vol. 141, Article ID 105175, 2022.
- [26] U. K. Lopes and J. Valiati, “Pre-trained convolutional neural networks as feature extractors for tuberculosis detection,” *Computers in Biology and Medicine*, vol. 89, pp. 135–143, 2017.
- [27] V. Sineglazov, K. Riazanovskiy, A. Klanovets, E. Chumachenko, and N. Linnik, “Intelligent tuberculosis activity assessment system based on an ensemble of neural networks,” *Computers in Biology and Medicine*, vol. 147, Article ID 105800, 2022.
- [28] M. Mamalakis, A. J. Swift, B. Vorselaars et al., “DenResCov-19: a deep transfer learning network for robust automatic classification of COVID-19, pneumonia, and tuberculosis from X-rays,” *Computerized Medical Imaging and Graphics*, vol. 94, Article ID 102008, 2021.
- [29] L. T. Duong, N. H. Le, T. B. Tran, V. M. Ngo, and P. T. Nguyen, “Detection of tuberculosis from chest X-ray images: boosting the performance with vision transformer and transfer learning,” *Expert Systems with Applications*, vol. 184, Article ID 115519, 2021.
- [30] S. N. Kumar, A. L. Fred, P. Padmanabhan, B. Gulyas, H. Ajay kumar, and L. R. Jonisha Miriam, “Deep learning algorithms in medical image processing for cancer diagnosis: overview, challenges and future,” *Deep Learning for Cancer Diagnosis*, pp. 37–66, 2021.
- [31] M. M. Ahsan, T. E. Alam, T. Trafalis, and P. Huebner, “Deep MLP-CNN model using mixed-data to distinguish between COVID-19 and Non-COVID-19 patients,” *Symmetry*, vol. 12, p. 1526, 2020.
- [32] A. L. S. Lee, C. C. K. To, A. L. H. Lee, J. J. X. Li, and R. C. K. Chan, “Model architecture and tile size selection for convolutional neural network training for non-small cell lung cancer detection on whole slide images,” *Informatics in Medicine Unlocked*, vol. 28, Article ID 100850, 2022.
- [33] S. Tomassini, N. Falcionelli, P. Sernani, L. Burattini, and A. F. Dragoni, “Lung nodule diagnosis and cancer histology classification from computed tomography data by convolutional neural networks: a survey,” *Computers in Biology and Medicine*, vol. 146, Article ID 105691, 2022.
- [34] J. Wei, R. Zhu, H. Zhang, P. Li, A. Okasha, and A. K. Muttar, “Application of PET/CT image under convolutional neural network model in postoperative pneumonia virus infection monitoring of patients with non-small cell lung cancer,” *Results in Physics*, vol. 26, Article ID 104385, 2021.
- [35] M. Desai and M. Shah, “An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN),” *Clinical eHealth*, vol. 4, pp. 1–11, 2021.
- [36] S. Hassantabar, M. Ahmadi, and A. Sharifi, “Diagnosis and detection of infected tissue of COVID-19 patients based on lung X-ray image using convolutional neural network approaches,” *Chaos, Solitons & Fractals*, vol. 140, Article ID 110170, 2020.

Research Article

An Ensemble of Deep Learning Enabled Brain Stroke Classification Model in Magnetic Resonance Images

Ala' A. Eshmawi,¹ Mashael Khayyat,² Abeer D. Algarni ,³ and Inès Hilali-Jaghdam⁴

¹University of Jeddah, College of Computer Science and Engineering, Department of Cybersecurity, Jeddah, Saudi Arabia

²University of Jeddah, College of Computer Science and Engineering, Department of Information Systems and Technology, Jeddah, Saudi Arabia

³Department of Information Technology, College of Computer and Information Sciences,

Princess Nourah Bint Abdulrahman University, P.O.Box 84428, Riyadh 11671, Saudi Arabia

⁴Department of Computer Science and IT, College of Community, Princess Nourah Bint Abdulrahman University, P.O.Box 84428, Riyadh 11671, Saudi Arabia

Correspondence should be addressed to Abeer D. Algarni; adalqarni@pnu.edu.sa

Received 30 December 2021; Revised 27 January 2022; Accepted 4 February 2022; Published 18 November 2022

Academic Editor: K. Shankar

Copyright © 2022 Ala' A. Eshmawi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Brain stroke is a major cause of global death and it necessitates earlier identification process to reduce the mortality rate. Magnetic resonance imaging (MRI) techniques is a commonly available imaging modality used to diagnose brain stroke. Presently, machine learning (ML) and deep learning (DL) models can be extremely utilized for disease detection and classification processes. Amongst the available approaches, the convolutional neural network (CNN) models have been widely used for computer vision and image processing issues such as ImageNet, facial detection, and digit classification. In this article, a novel computer aided diagnosis (CAD) based brain stroke detection and classification (CAD-BSDC) model has been developed for MRI images. The proposed CAD-BSDC technique aims in classifying the provided MR brain image as normal or abnormal. The CAD-BSDC technique involves different subprocesses such as preprocessing, feature extraction, and classification. Firstly, the input image undergoes preprocessing using adaptive thresholding (AT) technique for improving the image quality. Followed by, an ensemble of feature extractors such as MobileNet, CapsuleNet, and EfficientNet models are used. Besides, the hyperparameter tuning of the deep learning models takes place using the improved dragonfly optimization (IDFO) algorithm. Moreover, satin bowerbird optimization (SBO) based stacked autoencoder (SAE) is used for the classification of brain stroke. The design of optimal SAE using the SBO algorithm shows the novelty of the work. The performance of the presented technique was validated utilizing benchmark dataset which includes T2-weighted MR brain image collected from the axial axis with size of 256×256 . The simulation outcomes indicated the promising efficiency of the proposed CAD-BSDC technique over the latest state of art approaches in terms of various performance measures.

1. Introduction

Strokes are the 3rd most common cause of death around the world as per the report of the world health organization (WHO). With around 87% the more common type are ischemic strokes, caused by disturbance in the brain blood supply. Measuring volume and lesion location could assist diagnoses and guide treatment decisions [1]. Moreover, lesion classification plays a significant role in cognitive

neuroscience research. This frequently includes an anatomical analysis, where brain area is associated with the neurological deficit that requires manual examinations of massive stroke image databases. Hence, an automated methodology for segmenting ischemic lesions in brain images is extremely needed [2, 3]. Ischemic stroke lesions undergo several developmental stages. The moment of partial or total loss of blood supply to the infected brain areas is called the onset of the stroke and marks the beginning of

the hyperacute stage. At onset, the infection areas are separated into a core of infarcted tissue and a surrounding penumbra of under perfused, however possibly salvageable tissue i.e., partially provided by collateral blood flow [4]. The stroke developmental phase is related to amount of cell death and reconstructing mechanism which affects visibility of the stroke area in magnetic resonance imaging (MRI), mostly by the migration of water molecules [5].

Brain imaging methodologies, namely, computed tomography (CT) and MRI are very supportive for a physician for starting the early screening of the patients [6]. Also, there are several imaging modalities for analyzing brain that might involve diffuse optical imaging, X-ray imaging, positron emission tomography, magnetoencephalography, and functional MRI [7]. But this imaging technique requires well-trained operators and higher operating costs, thus many of these imaging methods may not be presented in every hospital and clinic. Image classification is broadly utilized in medical imaging systems [8]. But the classification method outcomes must be closer to the manual diagnoses.

This study introduces a novel computer aided diagnosis (CAD) based brain stroke detection and classification (CAD-BSDC) model on MRI images. The proposed CAD-BSDC technique involves preprocessing using adaptive thresholding (AT) technique to improve the image quality. In addition, an ensemble of feature extractors such as MobileNet, CapsuleNet, and EfficientNet models are used. Besides, the hyperparameter tuning of the deep learning models takes place using the improved dragonfly optimization (IDFO) algorithm. Furthermore, satin bowerbird optimization (SBO) based stacked autoencoder (SAE) is used to classify the MR brain image as normal or abnormal. The experimental result analysis of the CAD-BSDC technique takes place utilizing benchmark dataset which comprises T2-weighted MR brain images.

2. Literature Review

Currently, deep learning (DL) method was widely utilized as a classification system since it calculates features automatically within the convolution layer of the deep system [9]. The major benefit of utilizing DL method is that it outperforms other traditional methodologies for the classification of images. Several DL methodologies have existed like deep belief nets (DBN), RNNs, LSTM, and so on. Amongst this method, convolutional neural network (CNN) was widely employed in medical image processing and computer vision challenges such as house numbers digit classification, ImageNet, patch classification from medical images, face recognition, and so on [10].

Nishio et al. [11] evaluated and developed an automated acute ischemic stroke (AIS) detection method including 2-phase DL models. Next, the 2-phase method implemented the AIS recognition system in the testing set. To evaluate the detection outcomes, a board-certified radiologist assessed the testing set head CT image with and without help of detection system [12]. Hilbert et al. [13] examined DL methods to build model to directly forecast better reperfusion afterward endovascular treatment (EVT) and better

functional outcomes using CT images. This model does not need image annotation and is faster to calculate. The study compared DL to ML methods using conventional radiological image biomarkers. Pan et al. [14] investigated a new method based mainly on DL-ResNet for detecting infarct cores on non-contrast CT images and enhancing the performance of acute ischemic stroke diagnoses. They endlessly enrolled magnetic resonance diffusion weighted image (MR-DWI) confirmed first-episode ischemic stroke patients. Next, utilize decision curve analysis (DCA) model for analyzing the values of this technique in medical settings.

Zhang et al. [15] introduced a DL method that leverages MRI diffusion series for classifying TSS based medically validated threshold. Also, the study presented an intra-domain task-adoptive transfer learning technique that includes model training on simple medical tasks (stroke recognition) and refined the method with distinct binary thresholds of TSS. Wang et al. [16] evaluated and developed a DL based method to assist the selection of appropriate patients with acute ischemic stroke for endovascular treatment-based 3D pseudo-continuous arterial spin labeling (pCASL). The DL and six ML methods have been trained by using 10-fold CV.

3. The Proposed Model

In this study, a new CAD-BSDC model has been developed for MRI images for classifying them into normal or abnormal. The CAD-BSDC technique involves different subprocesses such as AT based preprocessing, ensemble of feature extraction, IDFO-based hyperparameter tuning, SAE-based classification, and SBO-based parameter tuning. Figure 1 illustrates the overall process of CAD-BSDC technique.

3.1. Image Preprocessing Using at Technique. At the primary level, the AT technique is applied on MRI images to remove the noise and enhance the quality. It is an effective method to determine the infected regions by the use of thresholding concept. In the AT technique, the investigation of the MRI images takes place for the distributed pixel intensities and the threshold value is chosen. In this case, the input MRI image can be denoted as $g(x, y)$, I implies the threshold value, and the final image can be defined as $f(x, y)$. It can be mathematically defined as follows [17]:

$$f(x, y) = \begin{cases} 1, & g(x, y) \geq I, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

3.2. Ensemble of Feature Extraction Approaches. During the feature extraction process, the ensemble of feature extractors namely MobileNet, CapsuleNet, and EfficientNet models are used. The DL is a type of CNN and has extremely utilized for images [18]. In recent times, DL was extremely utilized in the analysis of several medicinal diseases. Also, several researchers are developed by analysis of skin disease utilizing DL. The DL has several linked layers using distinct weight as

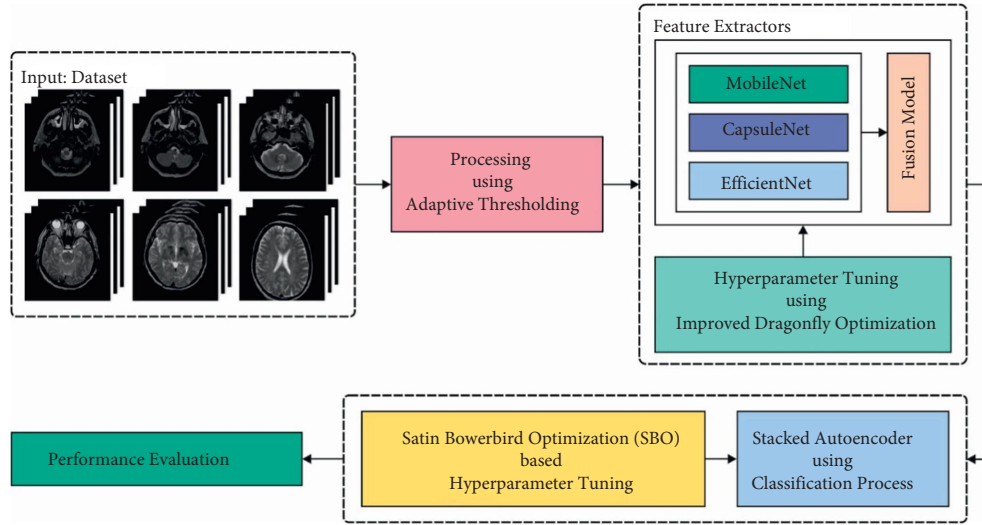


FIGURE 1: Overall process of CAD-BSDC technique.

well as activation functions. A fundamental DL technique involves convolution, pooling, and connected layers. Many activation functions were utilized for adjusting the weight. During this case, EfficientNetB3 was utilized for glaucoma detection. The EfficientNetB3 is current, cost-efficient, and robust technique established by scaling 3 parameters namely depth, width, and resolution [19]. An EfficientNetB3 method with noisy-student weight was utilized from scenarios I and III to the transfer learning (TL) procedure, but “isicall_eff3_weights” weight is utilized as pretrained to scenarios II and IV. The GlobalAveragePooling2D layers were added to all scenarios for generalizing the optimum model. The amount of parameters is decreased. Also, the rectified linear unit (ReLU) activation function was utilized with 3 dense and 2 dropout layers. The resultant layer has several outcome units to multiclass classification utilizing the softmax activation functions. Figure 2 demonstrates the structure of CapsNet.

The MobileNet [20] has lesser framework, minimum computation, and superior precision that is utilized to mobile terminal and embedding devices. According to depthwise separable convolutional, MobileNets utilize 2 global hyperparameters for keeping a balance amongst efficacy and accuracy. The basic concept of MobileNet is decomposition of convolutional kernel. With utilizing depthwise separable convolutional, the typical convolutional was decomposed as to depthwise convolutional and pointwise convolutional with convolutional kernels. The depthwise convolutional filter execute convolutional for all channels, and convolutional was utilized for combining the outcomes of depthwise convolutional layer. During this technique, N typical convolutional kernel. The typical convolution filters integrate an input as to a novel group of outputs, but the depthwise separable convolutional separates the inputs as to 2 layers, one to filter and another to merge. The MobileNetV2 establishes novel components with inverted remaining framework.

In order to compensate for shortcomings of CNN, the network framework named the CapsNet was presented [21]. The CapsNet is a deep network technique involving capsules. The capsule was comprised of a set of neurons. The activation neuron signifies the features of modules from the objects. All the capsules are responsible to determine a single module from the object, and every capsule jointly defines the entire framework of objects. Conversely, for any DNNs (for instance, DBN), this framework preserves object modules and spatial data. Related to CNN, the CapsNet was comprised of multi-layer networks.

3.3. Hyperparameter Tuning Using IDFO Algorithm. For optimally adjusting the hyperparameters of the DL models, the IDFO algorithm is applied. The DFO method was coined by Mirjalili at Griffith University in 2016 [22]. This method is a meta-heuristic approach-based SI is stimulated by dynamic as well as static behaviors of dragonflies in nature. There are 2 primary phases of optimization: exploitation and exploration. These two stages are modelled by dragonflies, either statically or dynamically searching for food or avoiding the enemy. The 2 further behaviors are added to these three fundamental behavior in *DA*: move to the food and avoid the enemy. Thus, once each individual moves to food source (equation (5)), they need to avoid the enemy simultaneously (equation (6)).

$$\begin{aligned}
 S_i &= -\sum_{j=1}^N X - X_j, \\
 A_i &= \frac{\sum_{j=1}^N V_j}{N}, \\
 C_i &= \frac{\sum_{j=1}^N X_j}{N} - X, \\
 F_i &= X^+ - X, \\
 E_i &= X^- + X,
 \end{aligned} \tag{2}$$

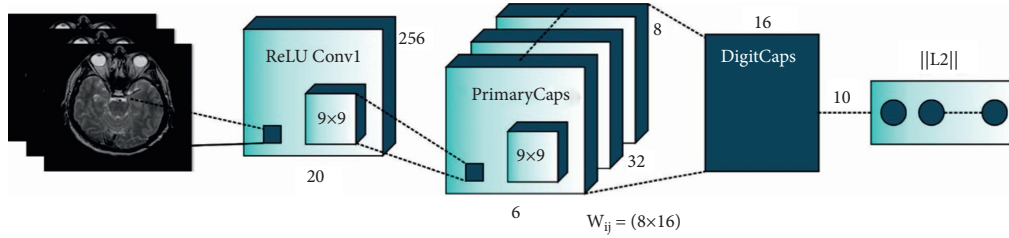


FIGURE 2: Structure of CapsNet.

where, X indicates the immediate location of the individual, in which X_j indicates the immediate location of j^{th} individual. N characterizes the amount of neighboring individuals, in which y_j shows the speed of j^{th} neighboring individual. X^+ and X^- denotes the position of the food and enemy source, correspondingly [23]. The overall steps of the DFO algorithm are given in 1.

To upgrade the place of artificial dragonflies in the searching space and simulate the motion, two vectors are taken into consideration: position (X) and step (X). The step vector considers as speed, shows the direction of dragonfly motion (equation (7)). Then estimating the step vectors, the position vector is upgraded (equation (8)):

$$\begin{aligned} \nabla X_{t+1} &= (sS_i + aA_i + cC_i + fF_i + eE_i) + w\nabla X_t, \\ X_{t+1} &= X_t + \nabla X_{t+1}. \end{aligned} \quad (3)$$

In which, and f , e , w , and t represents the food factor, enemy factor, inertia coefficient, iteration number, correspondingly and the, a , and c indicates separation, alignment, and cohesion coefficient, correspondingly. This coefficient and the abovementioned factor enable to implementation of exploitative and exploratory behaviors. In the IDFO algorithm, the traditional DFO algorithm is integrated into the flower authorization algorithm, we set the value ranges $[S_{min}, S_{max}]$. To efficiently evade the situation where the dragonfly collectively gathers in the first phase, the uniform distribution is utilized for implementing random initialization process on all the dimensions [24],

$$x_{id}^0 = S_{min} + (S_{max} - S_{min}). \quad (4)$$

Afterward preprocessing, the 2 procedures are further merged for guiding the dragonfly to fly to an optimal location.

3.4. Image Classification Using SBO-SAE Model. Finally, the SBO-SAE model can be employed for the classification of MRI images. The SAE is developed based on the concept of auto encoder (AE). In SAE model, the encoding part of the AE is stacked together, i.e., the input of initial layer of an AE model is actual data and the input of lower layer is hidden layer data. Lastly, a classification model is appended to the network [25]. The training process of the SAE model involves the pretraining and the inverse fine-tuning procedure. It makes use of a huge quantity of unlabeled data for unsupervised learning, independently extracted the features,

and utilizes the labelled data for inverse fine tuning of the network. For boosting the performance of the SAE technique, the weight and bias values are optimally chosen by the SBO algorithm.

SBO technique begins generating a primary uniform arbitrary population that contains a group of places to bower [26]. All positions ($pop(i).Pos$) are determined to the parameter which is supposed that optimize as written in equation (6). It could be noticeable the value of primary population lie among the existing minimal as well as maximal limit of optimizing parameters.

$$\begin{aligned} pop(i) \cdot Pos &= rand(1, n_{var}) \cdot (Var_{Max} - Var_{Min}) \\ &+ Var_{Min}, \quad \forall i \in n_{pop}. \end{aligned} \quad (5)$$

Comparatively, same as ABC, the probability of fascinating of male/female ($Prob_i$) to bower was calculated as follows.

$$\begin{aligned} Prob_i &= \frac{cost_i}{\sum_{k=1}^{n_{pop}} cost_i}, \quad \forall i \in n_{pop}, \\ cost_i &= \begin{cases} \frac{1}{1 + f(x_i)}, & f(x_i) \geq 0, \\ 1 + |f(x_i)|, & f(x_i) < 0. \end{cases} \end{aligned} \quad (6)$$

Same as other evolutionary dependent upon optimizer, elitism was utilized for storing an optimum solution(s) at all iterations of optimized procedure. In the mating season, males like every other bird utilize its drives for building and decorating the bower. Noticeably, older and experienced males are appealed further attention of others to their bower. Conversely, this bower has further fitness than the other bower. During the SBO processes, the place of an optimum bower created by bird processes, the place of an optimum bower created by bird was estimated as elite of k^{th} iteration ($x_{elite,k}$) that is maximum fitness and is capable of affecting the other places. In all iterations, a novel modification at some bower was computed dependent upon equation demonstrated in

$$\chi_{ik}^{new} = \chi_{i,k}^{old} + \beta_k \left[\left(\frac{\chi_{jk} + x_{elite,k}}{2} \right) - \chi_{i,k}^{old} \right]. \quad (7)$$

It can be worth maintaining that roulette wheel selective process was utilized for picking up bower with superior probability (x_{jk}). In SBO, Parameter β_k defines the count of steps for selecting target bowers that are calculated to all variables and modified based on


```

Parameter Initialization
Assume maximum number of hyperspheres
Represent the archive size
While the termination criteria is not fulfilled
    Determine the objective value of every dragonfly
    Compute nondominated solution
    Upgrade the archives based on the attained nondominated solution
    If the archive is occupied
        Execute the archive maintenance scheme to discard the present archive member
    Append the newly attained solutions to the archive
    End if
    If any of the newly appended solutions to the archive is situated external the hyper sphere
    Upgrade and reposition the hypersphere for covering the newly obtained solutions
    End if
    Elect a food source from archive
    Elect an enemy from archive
    Upgrade step vector by the use of  $T(\Delta x) = |\Delta x / \sqrt{\Delta x^2 + 1}|$ 
    Upgrade position vector by

$$X_{t+1} = \begin{cases} -X_t & r < T(\Delta x_{t+1}) \\ X_t & r \geq T(\Delta x_{t+1}) \end{cases}$$

    Ensure and adjust the newly obtained position depending upon the boundary variable
End while

```

ALGORITHM 1: Pseudocode of DFO Algorithm.

```

Input: population  $\vec{P}_{sp}$ 
Output: optimal searching agent,  $\vec{P}_{bst}$ 
Procedure SOA
    Parameter Initialization:  $C_A$  and  $C_B$ 
    Determine the fitness of every search agent
     $\vec{P}_{bst} \leftarrow$  optimal searching agent
    While ( $z < \text{Max}_{\text{iterations}}$ ) do
        for every searching agent do
            Upgrade the location of searching agent
        end for
        Upgrade variables  $C_A$  and  $C_B$ 
        Determine fitness value of all searching agents
        Upgrade  $\vec{P}_{bst}$  when better solution exist over earlier optimum solution
         $z \leftarrow z + 1$ 
    End while
    Return  $\vec{P}_{bst}$ 
End procedure

```

ALGORITHM 2: Pseudocode of SBO Algorithm.

$$\beta_k = \frac{\alpha}{1 + \text{Prop}_i} \quad (8)$$

Arbitrary modifies were executed to x_{ik} with specific probability, where normal distribution (N) has been utilized with average of x_{ik}^{old} and variance of σ as stated in equation.

$$X_{ik}^{\text{new}} \sim X_{ik}^{\text{old}} + \sigma \cdot N(0, 1), \quad (9)$$

$$\sigma = Z \cdot (\text{Var}_{\text{Max}} - \text{Var}_{\text{Min}}).$$

Finally, all the cycle is an old population and population attained in modifies as aforementioned were evaluated, integrated, sorted and novel population was created. The pseudocode of SBO algorithm is given in 2.

The SBO approach develops a FF for attaining enhanced classification efficiency. It defines a positive integer for representing the optimum efficiency of candidate solution. During this case, the minimized classification error rate was regarded as FF is provided in equation (10). The optimum

TABLE 1: Dataset descriptions.

Categories of MR brain stroke images	Class labels	Dimensions	Number of images
Acute (speech arrest)	Class 0	256 * 256	25
Cerebral haemorrhages	Class 1	256 * 256	25
Fatal stroke	Class 2	256 * 256	24
Normal images	Class 3	256 * 256	54
Subacute stroke (hesitating speech)	Class 4	256 * 256	26
Subacute stroke (loss of sensation)	Class 5	256 * 256	24

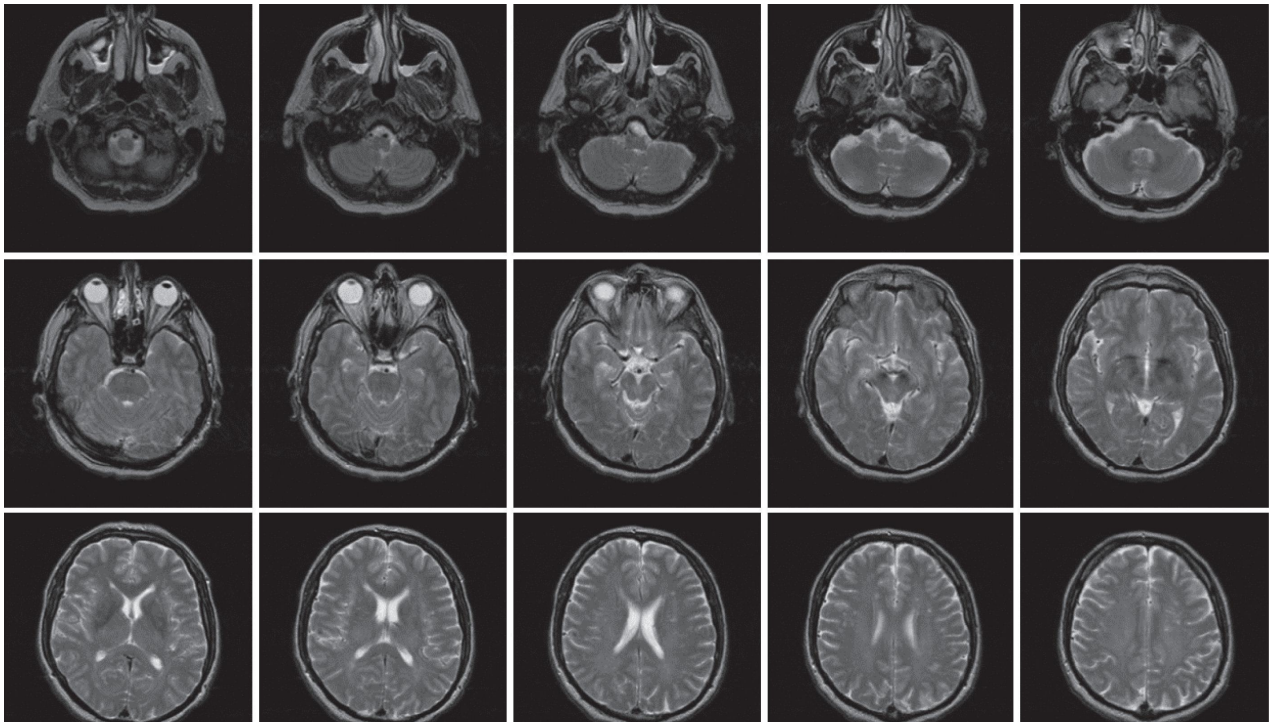


FIGURE 3: Sample test images.

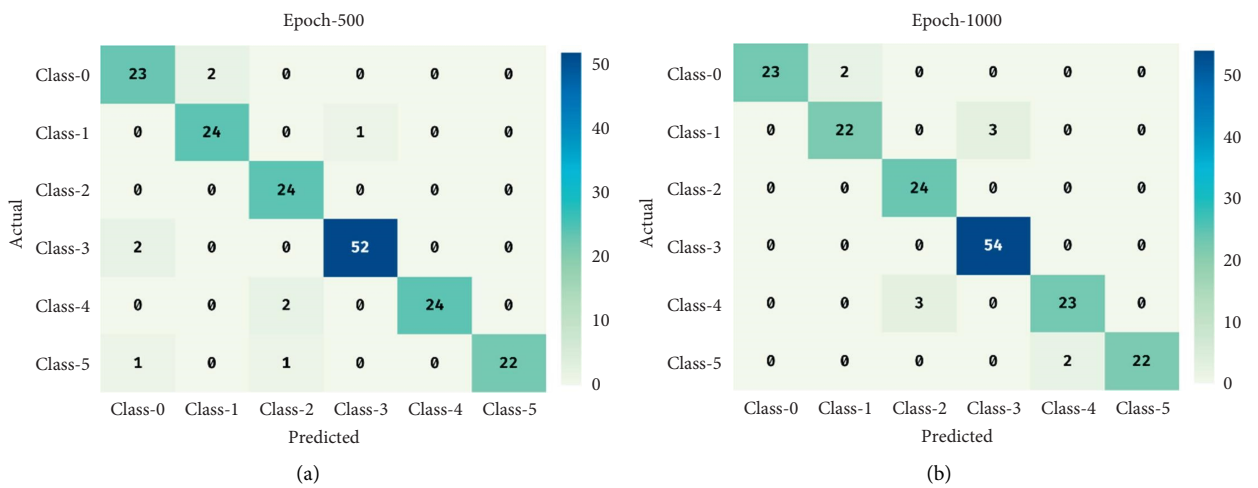


FIGURE 4: Continued.

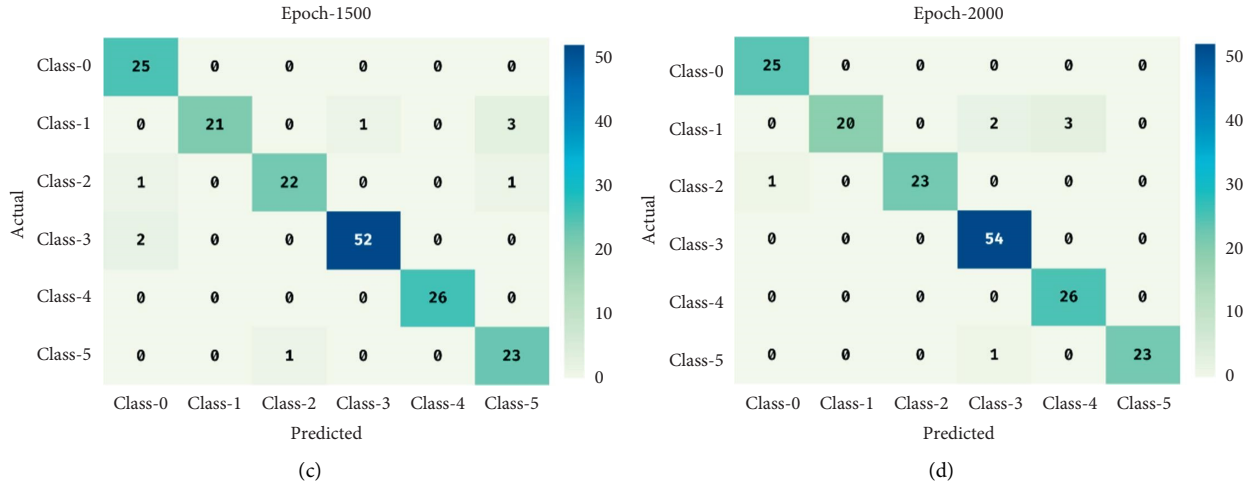


FIGURE 4: Confusion matrix of CAD-BSDC technique.

TABLE 2: Result analysis of CAD-BSDC technique with different classes and epoch counts.

Classes	Sensitivity	Specificity	Accuracy	Precision	F-score	MCC
Epoch-500						
Class 0	92.00	98.04	97.19	88.46	90.20	88.58
Class 1	96.00	98.69	98.31	92.31	94.12	93.16
Class 2	100.00	98.05	98.31	88.89	94.12	93.36
Class 3	96.30	99.19	98.31	98.11	97.20	96.00
Class 4	92.31	100.00	98.88	100.00	96.00	95.45
Class 5	91.67	100.00	98.88	100.00	95.65	95.13
Average	94.71	99.00	98.31	94.63	94.55	93.61
Epoch-1000						
Class 0	92.00	100.00	98.88	100.00	95.83	95.30
Class 1	88.00	98.69	97.19	91.67	89.80	88.19
Class 2	100.00	98.05	98.31	88.89	94.12	93.36
Class 3	100.00	97.58	98.31	94.74	97.30	96.15
Class 4	88.46	98.68	97.19	92.00	90.20	88.58
Class 5	91.67	100.00	98.88	100.00	95.65	95.13
Average	93.35	98.83	98.13	94.55	93.82	92.78
Epoch-1500						
Class 0	100.00	98.04	98.31	89.29	94.34	93.56
Class 1	84.00	100.00	97.75	100.00	91.30	90.48
Class 2	91.67	99.35	98.31	95.65	93.62	92.67
Class 3	96.30	99.19	98.31	98.11	97.20	96.00
Class 4	100.00	100.00	100.00	100.00	100.00	100.00
Class 5	95.83	97.40	97.19	85.19	90.20	88.77
Average	94.63	99.00	98.31	94.71	94.44	93.58
Epoch-2000						
Class 0	100.00	99.35	99.44	96.15	98.04	97.74
Class 1	80.00	100.00	97.19	100.00	88.89	88.02
Class 2	95.83	100.00	99.44	100.00	97.87	97.58
Class 3	100.00	97.58	98.31	94.74	97.30	96.15
Class 4	100.00	98.03	98.31	89.66	94.55	93.75
Class 5	95.83	100.00	99.44	100.00	97.87	97.58
Average	95.28	99.16	98.69	96.76	95.75	95.13

TABLE 3: Overall Classification result analysis of CAD-BSDC technique.

No. of epochs	Sensitivity	Specificity	Accuracy	Precision	F-score	MCC
Epoch-500	94.71	99.00	98.31	94.63	94.55	93.61
Epoch-1000	93.35	98.83	98.13	94.55	93.82	92.78
Epoch-1500	94.63	99.00	98.31	94.71	94.44	93.58
Epoch-2000	95.28	99.16	98.69	96.76	95.75	95.13
Average	94.49	99.00	98.36	95.16	94.64	93.78

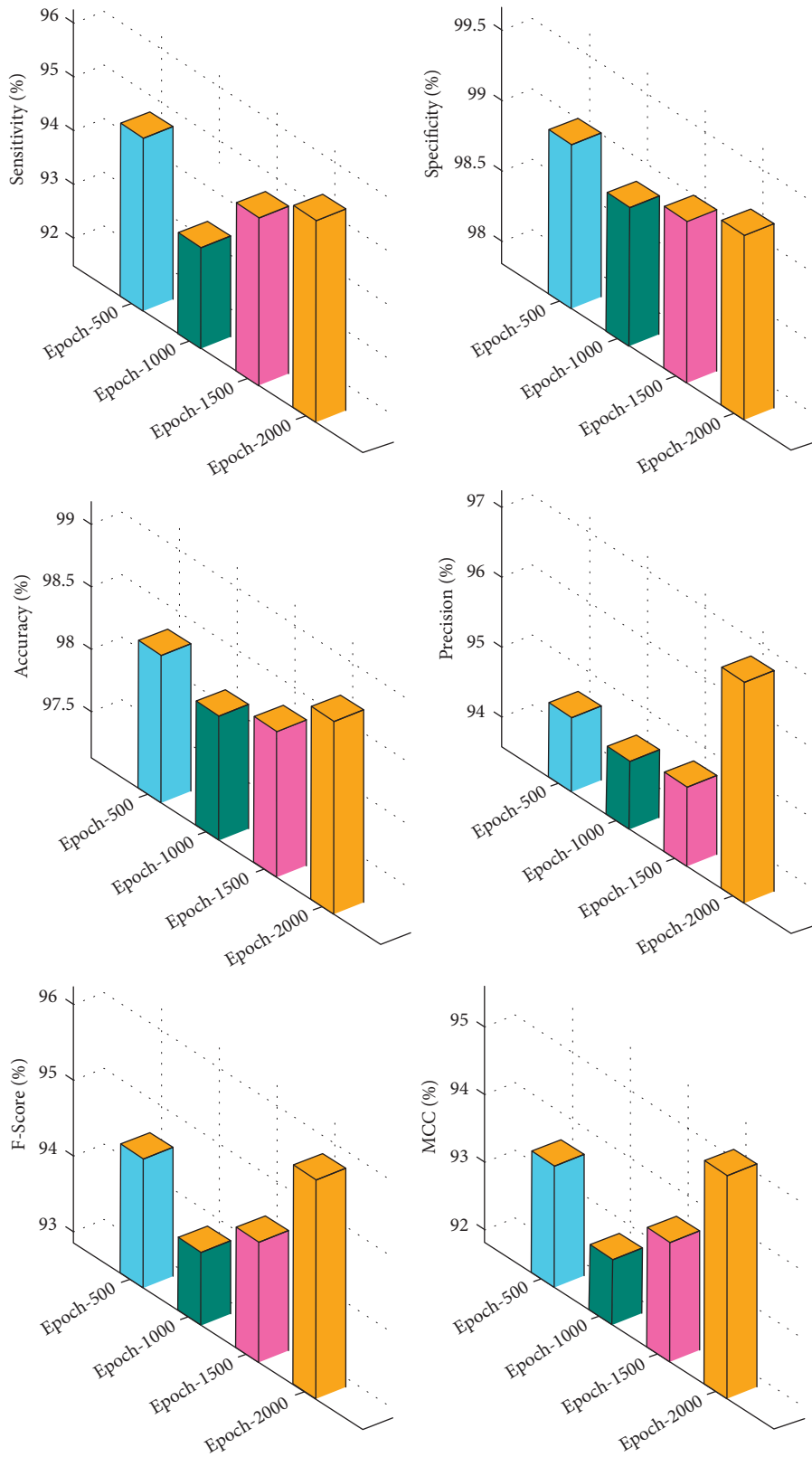


FIGURE 5: Overall classification result analysis of CAD-BSDC technique.

TABLE 4: Comparative analysis of CAD-BSDC technique with existing approaches.

Methods	Sensitivity	Specificity	F-measure	Accuracy
CAD-BSDC	94.49	99.00	96.64	98.36
SIFT-DT model	91.04	98.23	91.91	97.25
SURF-DT model	78.10	95.11	80.15	92.62
EM-PSORF	92.50	95.80	95.17	93.40
EM-PSOSVM	90.40	94.60	93.44	93.20
FODPSO-RF	89.40	93.70	93.54	92.40
FODPSO-SVM	88.30	92.60	91.69	91.70

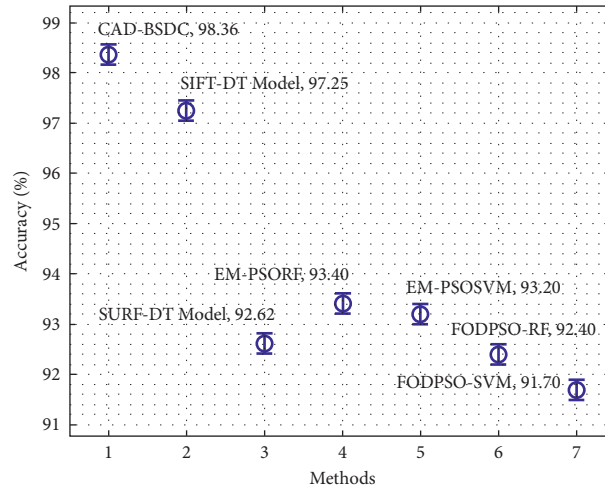


FIGURE 6: Accuracy analysis of CAD-BSDC technique with existing approaches.

solutions have reduced error rates and the worst solutions obtain an enhanced error rate.

$$\begin{aligned} \text{fitness}(x_i) &= \text{ClassifierErrorRate}(x_i), \\ &= \frac{\text{number of misclassified instances}}{\text{total number of instances}} * 100. \end{aligned} \quad (10)$$

4. Experimental Validation

The performance validation of the CAD-BSDC technique takes place using the benchmark dataset [27], which contains MRI images under six distinct classes. The details relevant to the dataset are given in Table 1. Figure 3 shows the sample MRI images.

Figure 4 shows the confusion matrices offered by the CAD-BSDC technique on the classification of brain stroke. The figure shows that the CAD-BSDC technique has effectually identified distinct classes of brain stroke. For instance, under 500 epochs, the CAD-BSDC technique has categorized 23 images in class 0, 24 images under class 1, 24 images in class 2, 52 images under class 3, 24 images in class 4, and 22 images in class 5, respectively. Simultaneously, in 1000 epochs, the CAD-BSDC approach has classified 23 images under class 0, 22 images in class 1, 24 images in class 2, 54 images under class 3, 23 images in class 4, and 22 images in class 5 correspondingly. Furthermore, under 1500 epochs, the CAD-BSDC methodology has classified 25

images in class 0, 21 images under class 1, 22 images in class 2, 52 images in class 3, 26 images in class 4, and 23 images in class 5 correspondingly. Furthermore, under 2000 epochs, the CAD-BSDC system has categorized 25 images in class 0, 20 images in class 1, 23 images in class 2, 54 images under class 3, 26 images in class 4, and 23 images in class 5 correspondingly.

Table 2 offers a detailed classification result analysis of the CAD-BSDC technique under various classes and epoch counts. The results ensured the effective performance of the CAD-BSDC technique interms of different measures.

Table 3 and Figure 5 depict the overall classification result analysis of the CAD-BSDC technique under varying epochs. The results show that the CAD-BSDC technique has resulted in improved classification results. For instance, with 500 epochs, the CAD-BSDC technique has resulted in the $sens_y$, $spec_y$, $accu_y$, $prec_n$, F_{score} , and MCC of 94.71%, 99%, 98.31%, 94.63%, 94.55%, and 93.61%, respectively. As well as, with 1000 epochs, the CAD-BSDC process has resulted in $sens_y$, $spec_y$, $accu_y$, $prec_n$, F_{score} , and MCC of 93.35%, 98.83%, 98.13%, 94.55%, 93.82%, and 92.78% correspondingly. Furthermore, with 1500 epochs, the CAD-BSDC method has resulted to the $sens_y$, $spec_y$, $accu_y$, $prec_n$, F_{score} , and MCC of 94.63%, 99%, 98.31%, 94.71%, 94.44%, and 93.58% correspondingly. Finally, with 2000 epochs, the CAD-BSDC approach has resulted in the $sens_y$, $spec_y$, $accu_y$, $prec_n$, F_{score} , and MCC of 95.28%, 99.16%, 98.69%, 96.76%, 95.75%, and 95.13% correspondingly.

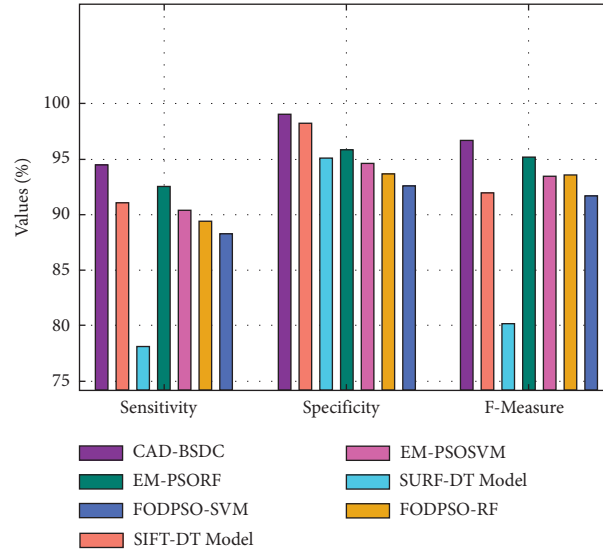


FIGURE 7: Comparative analysis of CAD-BSDC technique with existing methods.

In order to ensure the improvements of the CAD-BSDC technique, a comprehensive comparison study is made in Table 4 [28].

Figure 6 investigates the comparative accuracy analysis of the CAD-BSDC with recent methods on the test dataset. The figure demonstrated that the FODPSO-SVM technique has accomplished ineffectual outcomes with the least values of accuracy. In line with, the SURF-DT and FODPSO-RF techniques have obtained slightly increased values of accuracy. Followed by, the EM-PSORF and EM-PSOSVM techniques have reached moderately improved accuracy values. Though the SIFT-DT technique has reached near optimal accuracy of 97.25%, the CAD-BSDC technique has accomplished maximum accuracy of 98.36%.

Figure 7 explores the comparative $sens_y$, $spec_y$, and $F_{measure}$ analysis of the CAD-BSDC with current methodologies on the test dataset. The figure illustrates that the FODPSO-SVM approach has achieved ineffectual outcomes with the least values of $sens_y$, $spec_y$, and $F_{measure}$. In line with, the SURF-DT and FODPSO-RF methodologies have attained slightly improved values of $sens_y$, $spec_y$, and $F_{measure}$. Then, the EM-PSORF and EM-PSOSVM algorithms have reached better $sens_y$, $spec_y$, and $F_{measure}$ values. Although the SIFT-DT model has reached to near optimum $sens_y$, $spec_y$, and $F_{measure}$ of 91.04% 98.23%, and 91.91%, the CAD-BSDC model has attained maximal $sens_y$, $spec_y$, and $F_{measure}$ of 94.49%, 99%, and 96.64%.

The above mentioned tables and figures demonstrated that the CAD-BSDC technique has showcased superior performance over the other techniques.

5. Conclusion

In this study, a new CAD-BSDC model has been developed for MRI images for classifying them into normal or abnormal. The CAD-BSDC technique involves different subprocesses such as AT based preprocessing, ensemble of

feature extraction, IDFO-based hyperparameter tuning, SAE based classification, and SBO based parameter tuning. The experimental result analysis of the CAD-BSDC technique takes place utilizing benchmark dataset which includes T2-weighted MR brain images. The simulation outcomes indicated the promising efficiency of the proposed CAD-BSDC technique over the latest state of art approaches in terms of various performance measures. Thus, the CAD-BSDC technique can be realized in a real time environment to aid physicians. As a part of future extension, the classification performance of the CAD-BSDC technique can be enhanced by the use of DL-based segmentation approaches.

Data Availability

Data sharing not applicable to this article as no datasets were generated during the current study.

Ethical Approval

This article does not contain any studies with human participants performed by any of the authors.

Consent

Not applicable.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgments

This work was supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R51), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

- [1] A. Zotin, M. Kurako, Y. Hamad, and S. Kirillova, "Edge detection in MRI brain tumor images based on fuzzy C-means clustering," *Procedia Computer Science*, vol. 126, pp. 1261–1270, 2018.
- [2] S. Cheon, J. Kim, and J. Lim, "The use of deep learning to predict stroke patient mortality," *International Journal of Environmental Research and Public Health*, vol. 16, no. 11, p. 1876, 2019.
- [3] N. M. Murray, M. Unberath, G. D. Hager, and F. K. Hui, "Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review," *Journal of Neurointerventional Surgery*, vol. 12, no. 2, pp. 156–164, 2020.
- [4] O. Öman, T. Mäkelä, E. Salli, S. Savolainen, and M. Kangasniemi, "3D convolutional neural networks applied to CT angiography in the detection of acute ischemic stroke," *European radiology experimental*, vol. 3, no. 1, pp. 8–11, 2019.
- [5] R. Gupta, S. P. Krishnam, P. W. Schaefer, M. H. Lev, and R. G. Gonzalez, "An east coast perspective on artificial intelligence and machine learning," *Neuroimaging Clinics of North America*, vol. 30, no. 4, pp. 467–478, 2020.
- [6] A. Clèrigues, S. Valverde, J. Bernal, J. Freixenet, A. Oliver, and X. Lladó, "Acute ischemic stroke lesion core segmentation in CT perfusion images using fully convolutional neural networks," *Computers in Biology and Medicine*, vol. 115, Article ID 103487, 2019.
- [7] C. M. J. M. Dourado Jr, S. P. P. da Silva, R. V. M. da Nóbrega, P. P. R. Filho, and V. H. C. de Albuquerque, "Deep learning IoT system for online stroke detection in skull computed tomography images," *Computer Networks*, vol. 152, pp. 25–39, 2019.
- [8] D. S. Liebeskind, "Artificial intelligence in stroke care: deep learning or superficial insight?" *EBioMedicine*, vol. 35, pp. 14–15, 2018.
- [9] G. S. Saragih, Z. Rustam, D. Aldila, R. Hidayat, R. E. Yunus, and J. Pandelaki, "Ischemic stroke classification using random forests based on feature extraction of convolutional neural networks," *International Journal of Advanced Science, Engineering and Information Technology*, vol. 10, no. 5, 2020.
- [10] Y. Mokli, J. Pfaff, D. P. Dos Santos, C. Herweh, and S. Nagel, "Computer-aided imaging analysis in acute ischemic stroke—background and clinical applications," *Neurological Research and Practice*, vol. 1, no. 1, pp. 1–13, 2019.
- [11] M. Nishio, S. Koyasu, S. Noguchi et al., "Automatic detection of acute ischemic stroke using non-contrast computed tomography and two-stage deep learning model," *Computer Methods and Programs in Biomedicine*, vol. 196, Article ID 105711, 2020.
- [12] N. Takahashi, Y. Shinohara, T. Kinoshita et al., "Computerized identification of early ischemic changes in acute stroke in noncontrast CT using deep learning," in *Proceeding of the Medical Imaging 2019: Computer-Aided Diagnosis*, vol. 10950, March 2019, Article ID 109503A.
- [13] A. Hilbert, L. A. Ramos, H. J. A. van Os et al., "Data-efficient deep learning of radiological image data for outcome prediction after endovascular treatment of patients with acute ischemic stroke," *Computers in Biology and Medicine*, vol. 115, Article ID 103516, 2019.
- [14] J. Pan, G. Wu, J. Yu, D. Geng, J. Zhang, and Y. Wang, "Detecting the early infarct core on non-contrast CT images with a deep learning residual network," *Journal of Stroke and Cerebrovascular Diseases*, vol. 30, no. 6, Article ID 105752, 2021.
- [15] H. Zhang, J. S. Polson, K. Nael et al., "Intra-domain task-adaptive transfer learning to determine acute ischemic stroke onset time," *Computerized Medical Imaging and Graphics*, vol. 90, Article ID 101926, 2021.
- [16] K. Wang, Q. Shou, S. J. Ma et al., "Deep learning detection of penumbral tissue on arterial spin labeling in stroke," *Stroke*, vol. 51, no. 2, pp. 489–497, 2020.
- [17] D. S. Tan, W.-Y. Chen, and K.-L. Hua, "Deep demosaicking: adaptive image demosaicking via multiple deep fully convolutional networks," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2408–2419, 2018.
- [18] S. Akter, F. M. J. M. Shamrat, S. Chakraborty, A. Karim, and S. Azam, "COVID-19 detection using deep learning algorithm on chest X-ray images," *Biology*, vol. 10, no. 11, p. 1174, 2021.
- [19] B. Koonce, "Efficient net," in *Convolutional Neural Networks with Swift for Tensorflow*, pp. 109–123, Apress, Berkeley, CA, 2021.
- [20] W. Wang, Y. Li, T. Zou, X. Wang, J. You, and Y. Luo, "A novel image classification approach via dense-mobilenet models," *Mobile Information Systems*, vol. 2020, Article ID 7602384, 2020.
- [21] E. Goceri, "CapsNet topology to classify tumours from brain images and comparative evaluation," *IET Image Processing*, vol. 14, no. 5, pp. 882–889, 2020.
- [22] S. Mirjalili, "Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems," *Neural Computing & Applications*, vol. 27, no. 4, pp. 1053–1073, 2016.
- [23] Ç.İ. Acı and H. Gülcan, "A modified dragonfly optimization algorithm for single-and multiobjective problems using brownian motion," *Computational intelligence and neuroscience*, vol. 2019, Article ID 6871298, 8 pages, 2019.
- [24] F. Liu, "Path recognition of the regional education expansion based on improved dragonfly algorithm," *Mathematical Problems in Engineering*, vol. 2021, Article ID 9928020, 11 pages, 2021.
- [25] H. Pan, W. Tang, J. J. Xu, and M. Binama, "Rolling bearing fault diagnosis based on stacked autoencoder network with dynamic learning rate," *Advances in Materials Science and Engineering*, vol. 2020, Article ID 6625273, 2020.
- [26] S. H. Samareh Moosavi and V. Khatibi Bardsiri, "Satin bowerbird optimizer: a new optimization algorithm to optimize ANFIS for software development effort estimation," *Engineering Applications of Artificial Intelligence*, vol. 60, pp. 1–15, 2017.
- [27] K. A. Johnson and J. Alex Becker, "The whole brain atlas," 2000, <https://www.med.harvard.edu/AANLIB/>.
- [28] S. Manikandan and P. Dhanalakshmi, "Computer aided diagnosis model for brain stroke classification in MRI images using machine learning algorithms," *International Journal of Advanced Science and Technology*, vol. 29, no. 5, pp. 7976–7990, 2020.

Research Article

Machine Learning with Quantum Seagull Optimization Model for COVID-19 Chest X-Ray Image Classification

Mahmoud Ragab ^{1,2,3} **Samah Alshehri** ⁴ **Nabil A. Alhakamy** ^{5,6,7} **Wafaa Alsaggaf** ¹
Hani A. Alhadrami ^{8,9,10} and **Jaber Alyami** ^{11,12}

¹Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

²Centre of Artificial Intelligence for Precision Medicines, King Abdulaziz University, Jeddah 21589, Saudi Arabia

³Mathematics Department, Faculty of Science, Al-Azhar University, Naser City 11884, Cairo, Egypt

⁴Pharmacy Practice Department, Faculty of Pharmacy, King Abdulaziz University, Jeddah, Saudi Arabia

⁵Pharmaceutics Department, Faculty of Pharmacy, King Abdulaziz University, Jeddah, Saudi Arabia

⁶Center of Excellence for Drug Research and Pharmaceutical Industries, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁷Mohamed Saeed Tamer Chair for Pharmaceutical Industries, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁸Medical Laboratory Technology Department, Faculty of Applied Medical Sciences, King Abdulaziz University, P.O. BOX 80402, Jeddah 21589, Saudi Arabia

⁹Molecular Diagnostic Lab, King Abdulaziz University Hospital, King Abdulaziz University, P.O. BOX 80402, Jeddah 21589, Saudi Arabia

¹⁰Special Infectious Agent Unit, King Fahd Medical Research Center, King Abdulaziz University, P.O. BOX 80402, Jeddah 21589, Saudi Arabia

¹¹Diagnostic Radiology Department, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia

¹²Imaging Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Correspondence should be addressed to Mahmoud Ragab; mragab@kau.edu.sa

Received 21 January 2022; Revised 15 February 2022; Accepted 1 March 2022; Published 30 March 2022

Academic Editor: K. Shankar

Copyright © 2022 Mahmoud Ragab et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Early and accurate detection of COVID-19 is an essential process to curb the spread of this deadly disease and its mortality rate. Chest radiology scan is a significant tool for early management and diagnosis of COVID-19 since the virus targets the respiratory system. Chest X-ray (CXR) images are highly useful in the effective detection of COVID-19, thanks to its availability, cost-effective means, and rapid outcomes. In addition, Artificial Intelligence (AI) techniques such as deep learning (DL) models play a significant role in designing automated diagnostic processes using CXR images. With this motivation, the current study presents a new Quantum Seagull Optimization Algorithm with DL-based COVID-19 diagnosis model, named QSGOA-DL technique. The proposed QSGOA-DL technique intends to detect and classify COVID-19 with the help of CXR images. In this regard, the QSGOA-DL technique involves the design of EfficientNet-B4 as a feature extractor, whereas hyperparameter optimization is carried out with the help of QSGOA technique. Moreover, the classification process is performed by a multilayer extreme learning machine (MELM) model. The novelty of the study lies in the designing of QSGOA for hyperparameter optimization of the EfficientNet-B4 model. An extensive series of simulations was carried out on the benchmark test CXR dataset, and the results were assessed under different aspects. The simulation results demonstrate the promising performance of the proposed QSGOA-DL technique compared to recent approaches.

1. Introduction

Coronavirus disease (COVID-19) should be diagnosed in early stages in order to reduce the spread of virus and prevent further complications. With the increasing spread of COVID-19 cases, on a day-to-day basis across the globe, the limitation of the present diagnosis tool imposes challenges in managing and curbing the outbreak. Global researchers have conducted vigorous research to develop efficient diagnosis procedures and speed up the development of treatments and vaccine [1]. In general, three diagnostic procedures are widely employed such as medical imaging, blood tests, and viral tests [2]. One of the most widely employed viral tests that is identified as the gold-standard for detecting COVID-19 is Reverse Transcription Polymerase Chain Reaction (RT-PCR) which is employed as first-line screening tool. However, a number of researchers found that the experiment results achieved a sensitivity between 50 and 62% only [3]. This reveals the fact that first RT-PCR results could be attained as negative also. Therefore, in order to validate the accuracy of the experimental diagnoses, many RT-PCR experiments are conducted on a 14-day period of observation. In other words, an RT-PCR negative result for a suspicious COVID-19 case is taken into account as True Negative, if there is no positive RT-PCR result after running screening tests during the 14-day period of observation. Practically, this frustrates the patient diagnosed with COVID-19 and stresses the already-exhausted healthcare infrastructure of most of the nations due to lack of sufficient RT-PCR kits and qualified personnel [4].

As per the literature, chest X-rays (CXR) were employed as a first-line diagnosis tool in Italy and several countries [5]. Radiology scans can be run to detect the condition of the lungs and the patient's distinct phase of recovery/illness in an efficient manner [6]. Radiologists have observed a range of abnormalities present in radiology scan reports of COVID-19 patients. In recent years, deep learning, observed as the fundamental component of enhancing Artificial Intelligence technology, was stated to have considerable diagnosis accuracy, in medical imaging, for automated diagnosis of lung disease. It exceeded human level performances on ImageNet classification tasks with one million images to train in 2015 and further displayed dermatologist level performances on the classification of skin lesions in 2017. It further produced outstanding results in terms of screening lung cancer in 2019 [7].

In general, a radiologist's manual screening process may bring bias or wrong diagnoses and increases the possible risk of lost diagnoses for minuscule lesions. Therefore, health professionals such as radiotherapists gain excellent benefit out of emerging Artificial Intelligence (AI) methods in computer-aided COVID-19 diagnostics. Artificial Intelligence (AI) and advanced software, in the field of healthcare image analyses, have directly assisted the healthcare professionals in fighting this novel coronavirus. These systems offer effective and high-quality diagnosis result and drastically reduces manpower requirement [8]. Recently, machine learning and deep learning, the two main fields of AI, have forayed into healthcare applications commonly. Deep

learning-based support system is established in the diagnosis of COVID-19 using X-ray and CT scan samples. Few schemes have been proposed according to the pretrained models using transfer learning, whereas some methods have been presented with a personalized network [9]. Data science and machine learning, though being different domains, have been brought together and are dynamically employed in different stages such as prognosis, diagnosis, outbreak forecasting, and prediction for COVID-19. However, almost all of the DL-based techniques, used in disease diagnosis, require annotating the lesion, particularly for the disease diagnoses in CT volume. Annotating the lesion of COVID-19 incurs heavy cost, time, and effort for the radiotherapist which prevents efficient curbing of the disease. COVID-19 has rapidly spread to global nations, and there is a huge shortage for radiotherapists. Therefore, conducting COVID-19 diagnosis using DL models is of great significance for the community.

The current study focuses on the design of a new Quantum Seagull Optimization Algorithm with DL-based COVID-19 diagnosis model, named QSGOA-DL technique. Besides, the proposed QSGOA-DL technique involves the design of EfficientNet-B4 as a feature extractor, whereas the hyperparameter optimization process is carried out by the QSGOA technique. Moreover, the classification process is performed by a multilayer extreme learning machine (MELM) model. In order to showcase the supremacy of the proposed QSGOA-DL technique, a wide range of experimental analyses was conducted on benchmark test CXR dataset and the results were assessed under several aspects.

The rest of the paper is organized as follows: Section 2 reviews the literature; Section 3 discusses the proposed model; Section 4 validates the performance of the proposed model; at last, Section 5 concludes the study.

2. Related Works

Roy et al. [10] presented a new deep network acquired from the spatial transformer network. This network can predict the disease's seriousness rate concurrently based on input frames and offer positioning of pathological artefact in a weakly supervised manner. Additionally, the authors presented a novel methodology according to the uninorm for aggregation of efficient frame scores at a video level. At last, advanced deep methods were validated to estimate the pixel-level segmentation of COVID-19 imaging biomarker. In [11], a matrix profile technique was presented to detect the abnormalities in CT scan image through two stages. Abnormality Severity Score (CT-SS) was evaluated, and the variance of CT-SS between the COVID-19 CT image and non-COVID-19 CT image was examined. A sparse abnormality mask was evaluated and used for penalizing the pixel value of all the images. The abnormality-weighted images were utilized later for training the benchmark DenseNet DL model to differentiate COVID-19 CT from non-COVID-19 CT image. In this study, the authors applied the VGG19 model as a baseline model for comparison purposes.

Sakib et al. [12] proposed a feasible and effective DL-CRC framework for distinguishing COVID-19 from other

abnormalities (for example, pneumonia) and usual case with high precision. Exclusive datasets were developed from four open sources with PA chest sight of X-ray information for pneumonia, COVID-19, and usual case. The presented DL-CRC frameworks leveraged the DARI model for COVID-19 data by adaptively using GAN and GAD models. Kaur et al. [13] proposed expert models on the basis of deep feature and PF-BAT enhanced PF-FKNN classifiers to diagnose the novel coronavirus. In the presented method, the feature is extracted from the FC layer of transfer-learned MobileNetv2 and FKNN training. The hyperparameter of FKNN is fine-tuned with the help of PF-BAT algorithm.

Singh and Singh [14] proposed an automatic approach to diagnose COVID-19 from chest X-ray images. The study proposed an enhanced depth-wise CNN model to analyze the chest X-ray image. Wavelet decompositions were used in this study to integrate multiresolution analyses in the network. The frequency subbands, attained from the input image, were fed into the network to identify the disease. The networks were developed to predict the class of input image as either COVID-19 or normal or viral pneumonia.

Li et al. [15] proposed a new method for efficient and effective training of COVID-19 classification network with less number of COVID-19 CT exams and a record of negative samples. Specifically, new self-supervised learning methods were introduced to extract the features from negative sample and COVID-19-positive samples. Next, two types of soft labels (“diversity” and “difficulty”) were made for a negative sample by calculating the earth mover distance between COVID-19 features and negative samples, where the data “value” of the negative sample could be measured. Shamsi et al. [16] presented a deep uncertainty-aware TL architecture for COVID-19 recognition using healthcare image. Four common CNNs, including InceptionResNetV2, VGG16, ResNet50, and DenseNet121, were initially used in this study to extract the deep features from CT and X-ray images. Later, feature extraction was accomplished using distinct ML and statistical modelling methods to identify COVID-19 cases.

Wu et al. [17] developed a new JCS system to execute explainable and real-time COVID-19 chest CT diagnoses. In order to train these JCS systems, the authors created a large-scale COVID-19 Segmentation and Classification (COVID-CS) dataset containing 144,167 chest CT images collected from 400 COVID-19 persons and 350 negative samples. A total of 3,855 chest CT images, collected from 200 persons, were annotated to fine-grained pixel-level label of opacification, i.e., improved attenuation of lung parenchyma. Han et al. [18] proposed an AD3D-MIL model in which a person-level label is allocated to a 3D chest CT scan image that is viewed as a bag of instance. AD3D-MIL could semantically create deep 3D instances by following the probably diseased region. Furthermore, AD3D-MIL employs an attention-based pooling method for 3D instances so as to provide insight to every instance that contributes toward bag labels. Finally, AD3D-MIL learns Bernoulli’s distribution of bag-level label for easily available learning.

3. The Proposed Model

In this study, a novel QSGOA-DL technique is presented to detect and classify COVID-19 using CXR images. The presented QSGOA-DL technique encompasses different operational stages such as preprocessing, EfficientNet-B4-based feature extraction, QSGO-based hyperparameter optimization, and MELM-based classification. Figure 1 illustrates all the processes involved in the proposed QSGOA-DL model. The design of QSGO technique assists in optimal selection of hyperparameter values of EfficientNet-B4 model.

3.1. Preprocessing. In the presented model, the images undergo preprocessing through two ways such as data augmentation and image resizing. The augmentation technique generates the perturbed versions of the available images. Scaling, rotations, and other affine conversions are commonly used herewith. It is generally carried out to increase the size of the dataset and provide effective training to the deep learning model on different types of images. Besides, the 2D array (x -axis and y -axis) of the image of X-data (size of 512×512) is normalized for pixel values between 0 and 255 and stored from PNG format with the help of OpenCV library. All the preprocessed images measure 512×512 and have three channels.

3.2. EfficientNet-B4-Based Feature Extraction. In this stage, the preprocessed CXR images are passed onto EfficientNet-B4 technique and generate a useful set of feature vectors. Here, the CNN is directed towards an acyclic graph. This network is able to learn extremely nonlinear functions too. Neurons are the fundamental unit inside a CNN. All the layers, in a CNN, are made up of many neurons. These neurons are hooked together, i.e., the output of neuron from layer l becomes the input of neuron at layers $l + 1$, as given in the following equation:

$$a^{(l+1)} = f(W^{(l)}a^{(l)} + b^{(l)}), \quad (1)$$

where $W^{(l)}$ represents the weight matrix of layers l , $b^{(l)}$ denotes bias term, and f indicates the activation function. The activation for layer l is represented as $a^{(l)}$. In order to train a CNN, it is important to learn W and b for all the layers, so the cost functions are minimized [19]. Generally, assume a training set $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$ with m training example; weight W and bias b should be defined since they minimize the cost, i.e., the differences between the preferred output y and actual output $f_{W,b}(x)$. The cost functions for individual training examples are determined as follows:

$$J(W, b; x, y) = \frac{1}{2} \|h_{W,b}(x) - y\|^2, \quad (2)$$

where $h(x)$ represents the activation of final layer. Minimization process is iteratively performed by following the gradient descent method. This method involves the computation of partial derivatives of cost functions with regard

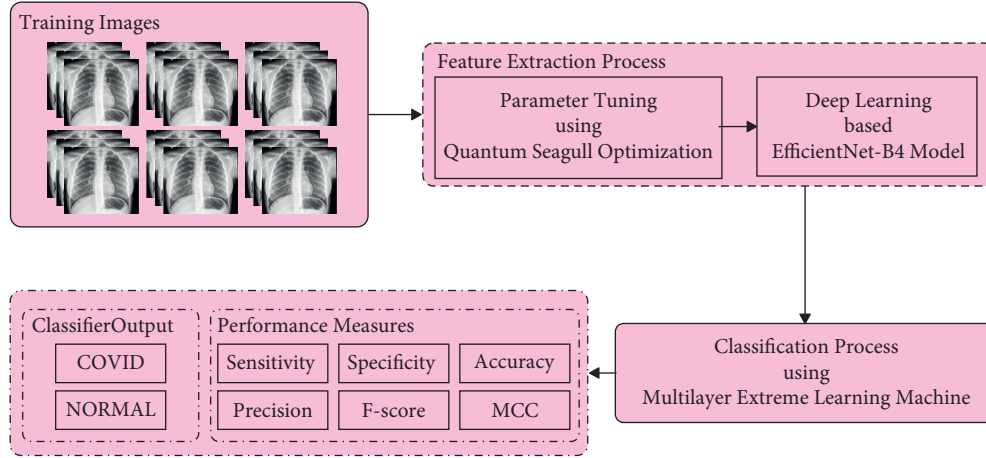


FIGURE 1: Overall process of the QSGOA-DL model.

to weight and updates the weight consequently. A single iteration of gradient descent updates the variables W and b as follows:

$$W^{(l)} = W^{(l)} - \alpha \frac{\partial}{\partial W^{(l)}} J(W, b),$$

$$b^{(l)} = b^{(l)} - \alpha \frac{\partial}{\partial b^{(l)}} J(W, b).$$
(3)

The BP model is employed in the computation of a partial derivative of cost function. Each FC has a hidden unit interconnected to each input unit. This increases the numbers of connections to extreme levels, while at the same time, it can also handle high-dimension information such as images. When the image size is assumed to be its dimensions, then the process of interconnecting every input pixel to all the neurons incurs heavy computation cost. An image as small as 100×100 pixel requires $10^4 \times N$ connection at the input layer, in which N represents the number of neurons at the initial layer. The convolution layer allows the construction of a sparse connection by assigning parameters through neurons. In comparison with the FC layer, the convolution layer has fewer parameters. So, it can be trained easily. It is derived at the cost of small reduction in the performance. The widely employed CNN for image detection includes convolutional and FS layers too. This network is frequently called as a deep network.

In DL training procedure models, expansion of network width, intensification of network depth, and improvement of input image solutions are the most widely employed methods to improve the precision of the models. Even though previous works such as ResNet and WideResNet proved the supremacy of the abovementioned approaches, it is important to balance each dimension in network resolution or width or depth so that the balance could be attained by scaling all the dimensions at a constant ratio. Tan presented the EfficientNet models that could produce appropriate effects on the extension of resolution, depth, and width of the networks and later attain a better performance. Initially, the researchers could portray CNN as a function:

$Y_i = F_i(X_i)$, in which F_i represents the operator (op), Y_i indicates the tensor of output, and X_i signifies the input tensor of shape $\langle H_i, W_i, C_i \rangle$, where C_i , H_i , and W_i denote the numbers of channels of an input image, height, and width. A CNN could be determined as a sequence of layers: $\text{Net} = F_k \circ \dots \circ F_2 \circ F_1(X_1) = \circ_{j=1 \dots k} F_j(X_1)$. In actual application procedure, the CNN layers are generally employed at many phases, where every phase uses a similar network framework. Hence, it is determined as follows [20]:

$$\text{Net} = \circ_{i=1 \dots s} F_i^{L_i}(X_{\langle H_i, W_i, C_i \rangle}),$$
(4)

where $F_i^{L_i}$ represents the layer F_i which is continued L_i time in a phase i and $\langle H_i, W_i, C_i \rangle$ represents the height, width, and numbers of channels of input tensor X of a layer i . Next, the standard CNN design mostly focuses on identifying an optimum layer framework F_i . However, according to the predetermined F_i baseline network framework, model scaling mostly extends the resolution (H_i, W_i), length (L_i), and width (C_i) of the networks. In the meantime, model scaling overcomes the implementation problems for a novel resource constraint by setting F_i . They could also examine L_i, C_i, H_i , and W_i distinctly for all the layers because it is a sample design space. EfficientNet stresses that each layer should be uniformly scaled by a constant ratio to reduce the design space. The target is to considerably enhance the precision of the models in the provided resource constraint environment since it is considered as an optimization problem:

$$\max_{d,w,r} \text{Accuracy}(\text{Net}(d, w, r)),$$

$$\text{s. t. Net}(d, w, r) = \circ_{j=1 \dots k} \hat{F}_i^{d, \hat{L}_i} \left(X_{\langle r \cdot \hat{H}_i, r \cdot \hat{W}_i, w \cdot \hat{C}_i \rangle} \right),$$
(5)

$$\text{Memory}(\text{Net}) \leq \text{target_memory},$$

$$\text{FLOPS}(\text{Net}) \leq \text{target_flops},$$

where w , d , and r represent the coefficients employed to scale the width, depth, and resolution of the network; $\hat{F}_i, \hat{L}_i, \hat{H}_i, \hat{W}_i$, and \hat{C}_i represent the predetermined parameters in the baseline network. Next, a novel compound scaling

technique, using a compound coefficient ϕ , is employed for uniform expansion of depth and width of the network as follows:

$$\begin{aligned} \text{depth: } d &= \alpha^\phi, \\ \text{width: } w &= \beta^\phi, \\ \text{resolution: } r &= \gamma^\phi, \\ \text{s. t. } \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2, \\ \alpha \geq 1, \beta \geq 1, \gamma &\geq 1, \end{aligned} \quad (6)$$

where α , β , and γ are constants. Amongst others, ϕ represents a stated value that determines how much resource is valid for expanding the models, whereas α , β , and γ determine the allocation method of extra resources to resolution, width, and depth of the network correspondingly. Also, there is a certain relationship between the FLOPS of a standard convolutional op and d , w^2 , and r^2 . When the depth of network doubles, then FLOPS doubles as well. However, when the network resolution/width doubles, FLOPS quadruples. Since convolutional ops frequently control the computational costs in the CNN, the CNN is expanded with equation (7) which accurately increases the overall FLOPS as $(\alpha \cdot \beta^2 \cdot \gamma^2)^\phi$. At last, scaling models does not alter the layer operator F_i in the predetermined baseline networks. Therefore, it is crucial to have a baseline network in place. EfficientNet, a novel mobile-size baseline network, is proposed with multiobjective neural framework which enhances both FLOPS and accuracy. The fundamental component consists of squeeze and excitation optimization and mobile-inverted bottleneck MBConv.

3.3. Hyperparameter Optimization. The QSGOA technique is deployed for optimal selection of hyperparameters involved in the EfficientNet-B4 model. In line with this, the performance gets boosted. Seagulls (scientific term: *Larus minutus*) are one amongst the coastal birds that started inhabiting the planet before 30 million years. They exist nearly everywhere in the world. With large wings, seagulls' hind legs have evolved so that they can travel in water too. Though fish is cited as the major food source for seagulls, they also consume amphibians, reptiles, moles, earthworms, and insects. In other terms, seagulls are omnivorous. They are considered as intelligent birds, while the average life span of seagulls is between 10 and 15 years. Generally, they live as a swarm and have a unique behaviour at the time of migration.

Migration is the movement of birds to the south during fall and to north during moving/spring from the ground to the height or from coast-coast to endure the winter condition and get wealthy food source with adequate amount of ease. This migration phenomenon of seagulls, which is a seasonal behaviour, is taken into account since they migrate everywhere to achieve a wide range of food sources to gain sufficient energy [21]. The procedure is given as follows:

- (i) Migration starts when swarms of seagulls started travelling towards north/south. In order to evade

collision, their primary position is made distinct from one another.

- (ii) One of the benefits from this swarm's experience is that they attempt to travel in the direction of optimal survival so as to achieve the minimum cost value.

In general, seagulls attack the migrating birds on the sea. This phenomenon occurs as a spiral-shaped behaviour at the time of attack. Seagull models for SGO are deliberated through the following points. The migration behaviour simulates the mobility of seagull swarms towards the position. For this purpose, three conditions must be fulfilled.

Collision avoidance: in order to evade the collisions amongst the neighboring seagulls, the models are determined as further parameter A to update the novel position of the deliberated seagull (search agents):

$$\begin{aligned} \vec{P} &= A \times \vec{p}_c(i), \\ i &= 0, 1, 2, \dots, \text{Max}(i), \end{aligned} \quad (7)$$

where \vec{P}_N describes the location that avoids colliding with other search agents, $\vec{p}_c(i)$ represents the location of the candidates in their current iteration (i), and A describes the movement behaviour of searching agents in their searching region which is also modelled as follows:

$$A = f_c - \left(i \times \left(\frac{f_c}{\text{Max}(i)} \right) \right), \quad (8)$$

where i describes the iteration and f_c represents the frequency control of parameter A in the range of $[0, f_c]$.

- (i) With another neighbors' experience: after avoiding the collision from the neighbor, the candidate progresses in the direction of optimal neighbors (optimal solutions).

$$\vec{d}_e = B \times \left(\vec{P}(i) - \vec{p}_c(i) \right), \quad (9)$$

where \vec{d}_e describes the position $\vec{p}_c(i)$ of candidate towards an optimal fitness candidate $\vec{p}_b(i)$. The coefficient B is an arbitrary value which makes the trade-off between exploration and exploitation phases. B is attained as follows:

$$B = 2 \times A^2 \times R. \quad (10)$$

Let R describe the arbitrary values between zero and one.

- (ii) Migration towards optimal solutions (search agents): at last, search agents upgrade their location according to the optimal solutions as follows:

$$\vec{D}_e = \left| \vec{P}_N + \vec{d}_e \right|, \quad (11)$$

where \vec{D}_e describes the variance between optimal costs and seagulls.

At the time of migration, seagulls change the attack speed and angle frequently. The location of seagulls can be

retained in the air by using their wings and weight. During attack procedure, the seagull moves in a spiral direction in air in x , y , and z plane by

$$\begin{aligned} X &= r \times \cos(t), \\ y &= r \times \sin(t), \\ Z &= r \times t, \end{aligned} \quad (12)$$

where t describes arbitrary values in the range between 0 and 2π and r denotes the radius of spiral turn as per the following formula:

$$r = \alpha \times e^{\beta t}, \quad (13)$$

where e describes the natural logarithm base and α and β represent the shapes of the spiral. The novel positions of the seagull are upgraded as follows:

$$\vec{P}_c(i) = \left(\vec{D}_e \times \hat{x} \times \hat{y} \times \hat{z} \right) + \vec{P}_b(i), \quad (14)$$

where $\vec{P}_c(i)$ keeps the optimal result. In order to improve the exploration abilities of SGO algorithm, QSGOA is designed including quantum computing.

Bit is the smallest unit of data from digital computers which demonstrates either 0 or 1 at a particular time, while Q-bit or quantum bit has achieved minimum unit of data from quantum computing. All Q-bits are capable to exist in the range of 0, 1, or a group of combined states simultaneously. This is named as superposition. Q-bit is referred to as a pair of numbers (α, β) , in which the values of $|\alpha|^2$ and $|\beta|^2$ signify the probabilities of determining the Q-bit from the states 0 and 1 correspondingly. The state of Q-bit is projected as follows:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle. \quad (15)$$

All the Q-bits must fulfill the normalization formula given as follows:

$$|\alpha|^2 + |\beta|^2 = 1. \quad (16)$$

In quantum computer, a separate q is signified as the order of n Q-bits as follows [22]:

$$q = [q_1, q_2, \dots, q_n] = \left[\begin{array}{c|c|c} \alpha_1 & \alpha_1 & \dots & \alpha_n \\ \beta_1 & \beta_2 & & \beta_n \end{array} \right]. \quad (17)$$

When a quantum state's performance is detected, it collapses toward the single state. The observation procedure of Q-bit i is carried out as follows:

If $\text{rand.}(0, 1) < (\alpha_i)^2$

Then. $f_i = 0$

Else. $f_i = 1$

In quantum computer, the order of quantum functions is implemented to update the values to Q-bits from all the individuals. This results in adherence of the upgraded Q-bits as in equation (20). Q-gate is the most quantum function to update Q-bits. There exist different Q-gates such as NOT gate, controlled NOT gate, rotation gate, Hadamard gate, x -gate, y -gate, and z -gate. In major analysis, the rotation Q-gate is utilized over other Q-gates. The rotation Q-gate $U(\Delta\theta_i)$ can be determined as follows:

$$U(\Delta\theta_i) = \begin{bmatrix} \cos(\Delta\theta_i) & -\sin(\Delta\theta_i) \\ \sin(\Delta\theta_i) & \cos(\Delta\theta_i) \end{bmatrix}, \quad (18)$$

where $\Delta\theta_i$ refers to the rotation angle of Q-bit i near 0/1 state. The state of Q-bit i at time t gets upgraded as follows:

$$\begin{bmatrix} \alpha_i(t+10) \\ \beta_i(t+1) \end{bmatrix} = U(\Delta\theta_i) \begin{bmatrix} \alpha_i(t) \\ \beta_i(t) \end{bmatrix}. \quad (19)$$

3.4. Image Classification. In this final stage, the derived set of features is fed into MELM classifier to allot appropriate class labels to the test CXR images. In the basic forms of SLFN, Huang et al. presented ELM to enhance the training speed of the work and later extended the hypotheses of ELM from neurons hidden node to another hidden node. Sample training can be expressed by $\{x_i, t_i\}_{i=1}^n$, where n represents the training sample, x_i indicates the input of i th sample using m dimension. Furthermore, t_i denotes the output of i th instance. Later, the input vector x is assumed to be the output of SLFN using L hidden node, and it is expressed as follows:

$$\begin{aligned} f(x) &= \sum_{i=1}^L \beta_i h_i(x) \\ &= h^T(x)\beta, \end{aligned} \quad (20)$$

where $h(x) = [h_1(x) \dots h_L(x)]^T$ represents the hidden output and $\beta = [\beta_1 \dots \beta_L]^T$ indicates the output weight. Given the output of n training, the sample could be estimated by zero error and is given as follows:

$$H\beta = t, \quad (21)$$

where $H = [h(x_1) \dots h(x_n)]^T$ signifies the hidden output matrix [23]. The output weight β solutions involve a linear formula, while such solutions might be equal to mitigation of training errors, namely, $\min H\beta - t$. The optimum approximation of output weight might be denoted as Moore-Penrose generalized inverse H^\dagger :

$$\hat{\beta} = H^\dagger t. \quad (22)$$

In general, orthogonal projection is employed to resolve the generalized inverse H^\dagger . If $H^T H$ is nonsingular, $H^\dagger = (H^T H)^{-1} H^T$, or if HH^T is nonsingular, $H^\dagger = H^T (HH^T)^{-1}$.

MELM is a multilayer NN in which multi-ELM-AEs are stacked together, where $X^{(i)} = [x_1^{(i)}, \dots, x_n^{(i)}]$; let $x_k^{(i)}$ be the i th data depiction for input x_k , $k = 1$ to n . Assume $\Lambda^{(i)} = [\lambda_1^{(i)}, \dots, \lambda_n^{(i)}]$ denotes the i th transformation matrix, in which $\lambda_k^{(i)}$ denotes the transformation vectors employed in depiction learning regarding $x_k^{(i)}$. Based on this, B replaces with $\Lambda^{(i)}$, where T is replaced by $X^{(i)}$ correspondingly [24]:

$$H^{(i)} \Lambda^{(i)} = X^{(i)}. \quad (23)$$

Let $H^{(i)}$ be the output matrix of i th hidden layer with regard to $X^{(i)}$, and $\Lambda^{(i)}$ is resolved as follows:

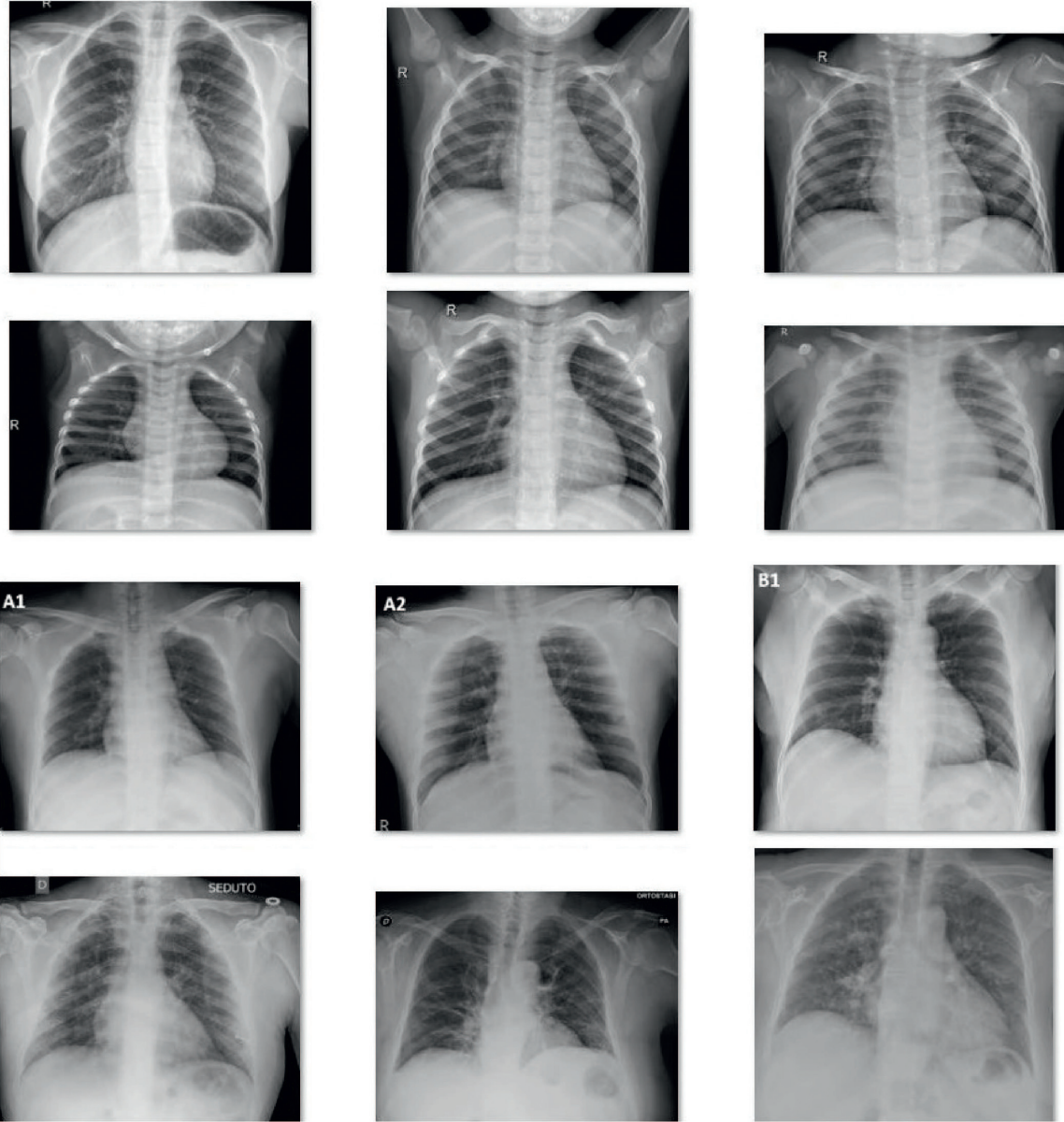


FIGURE 2: Sample images.

$$\Lambda^{(i)} = (H^{(i)})^T \left(\frac{I}{C} + H^{(i)} (H^{(i)})^T \right)^{-1} X^{(i)}. \quad (24)$$

Next,

$$X^* = g \left(X^{(i)} (\Lambda^{(i)})^T \right), \quad (25)$$

where χ^* represents the final depiction of $X^{(1)}$. X^* is employed as the hidden layer outputs to estimate the output weights β^* and β^* which are evaluated by

$$\beta^* = (X^*)^\dagger, \quad (26)$$

$$T = (X^*)^T \left(\frac{I}{C} + X^* (X^*)^T \right)^{-1} T.$$

4. Results and Discussion

The proposed model was simulated using Python 3.6.5 tool on a benchmark CXR image dataset [25]. The results were investigated under varying sizes of training and testing datasets. Figure 2 illustrates a few sample images considered for the study.

Figure 3 portrays the confusion matrices generated by the QSGOA-DL technique on test data with different training/testing data. Figure 3(a) depicts the confusion matrix produced by the proposed QSGOA-DL technique on training/testing of 80:20. The figure exhibits that the QSGOA-DL technique classified 3218 images as COVID-19 and 3219 images as healthy samples. Meanwhile, Figure 3(b) showcases the confusion matrix developed by QSGOA-DL

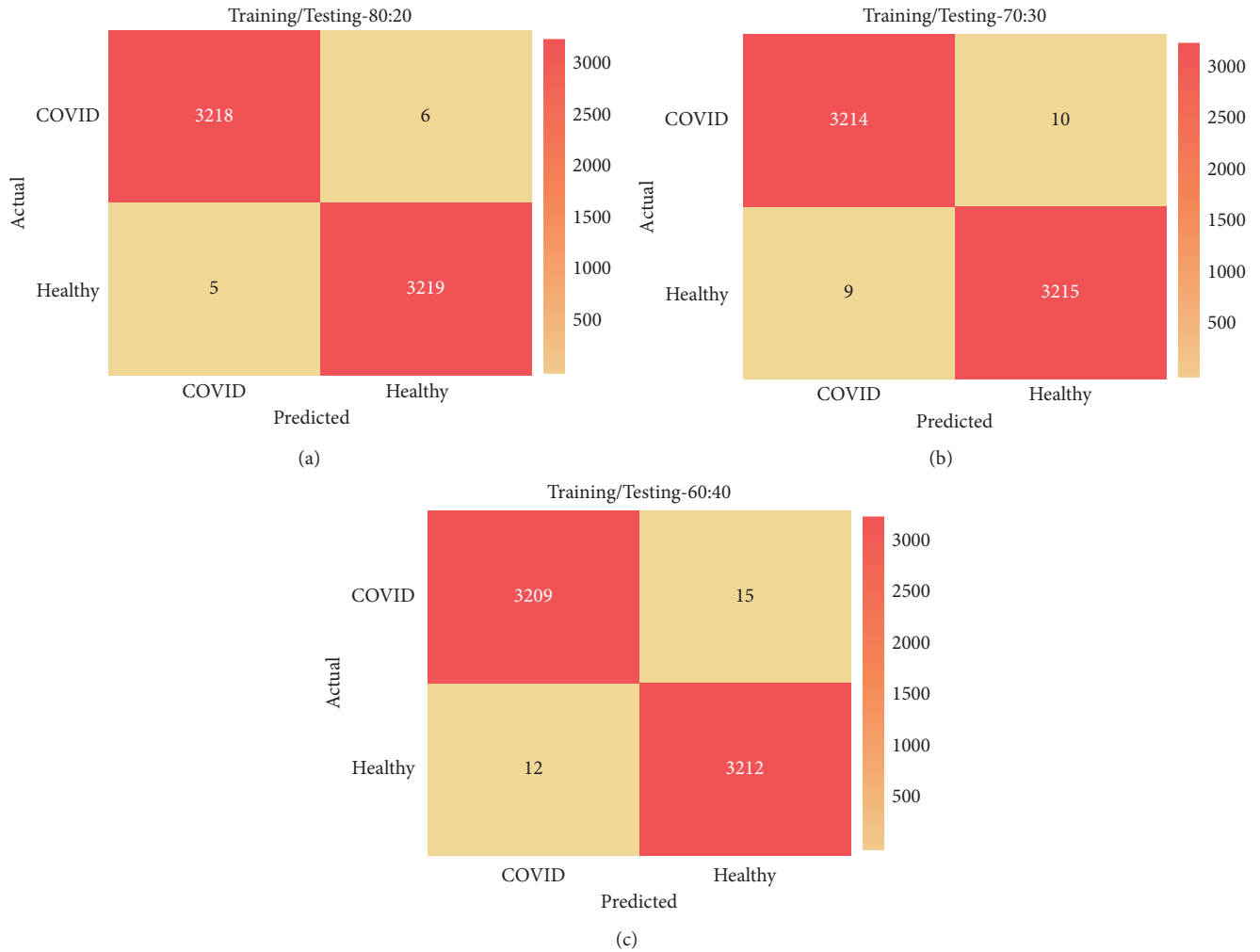


FIGURE 3: Confusion matrix analysis results of the QSGOA-DL model.

TABLE 1: Results of the analysis of QSGOA-DL model against different training/testing datasets.

Measures	Precision	Sensitivity	Specificity	Accuracy	F-score	MCC
Training/testing (80 : 20)	0.9984	0.9981	0.9984	0.9983	0.9983	0.9966
Training/testing (70 : 30)	0.9972	0.9969	0.9972	0.9971	0.9971	0.9941
Training/testing (60 : 40)	0.9963	0.9953	0.9963	0.9958	0.9958	0.9916
Average	0.9973	0.9968	0.9973	0.9971	0.9971	0.9941

manner on training/testing of 70 : 30. The figure shows that the QSGOA-DL algorithm outperformed compared to others and classified 3214 images as COVID-19 and 3215 images as healthy ones. Eventually, Figure 3(c) illustrates the confusion matrix generated by the QSGOA-DL algorithm on training/testing of 60 : 40. The figure demonstrates that the proposed QSGOA-DL methodology classified 3209 images as COVID-19 and 3212 images as healthy.

Table 1 shows the overall classification results attained by the QSGOA-DL technique under different training/testing data sizes. The results demonstrate that the proposed QSGOA-DL technique accomplished the maximum classification outcomes on all training/testing

sizes. For instance, with a training/testing data size of 80 : 20, the QSGOA-DL technique resulted in a precision of 0.9984, sensitivity of 0.9981, specificity of 0.9984, accuracy of 0.9983, F-score of 0.9983, and MCC of 0.9966. Moreover, with a training/testing data size of 70 : 30, QSGOA-DL manner resulted in a precision of 0.9972, sensitivity of 0.9969, specificity of 0.9972, accuracy of 0.9971, F-score of 0.9971, and MCC of 0.9941. Furthermore, with a training/testing data size being 60 : 40, the proposed QSGOA-DL method produced a precision of 0.9963, sensitivity of 0.9953, specificity of 0.9963, accuracy of 0.9958, F-score of 0.9958, and MCC of 0.9916.

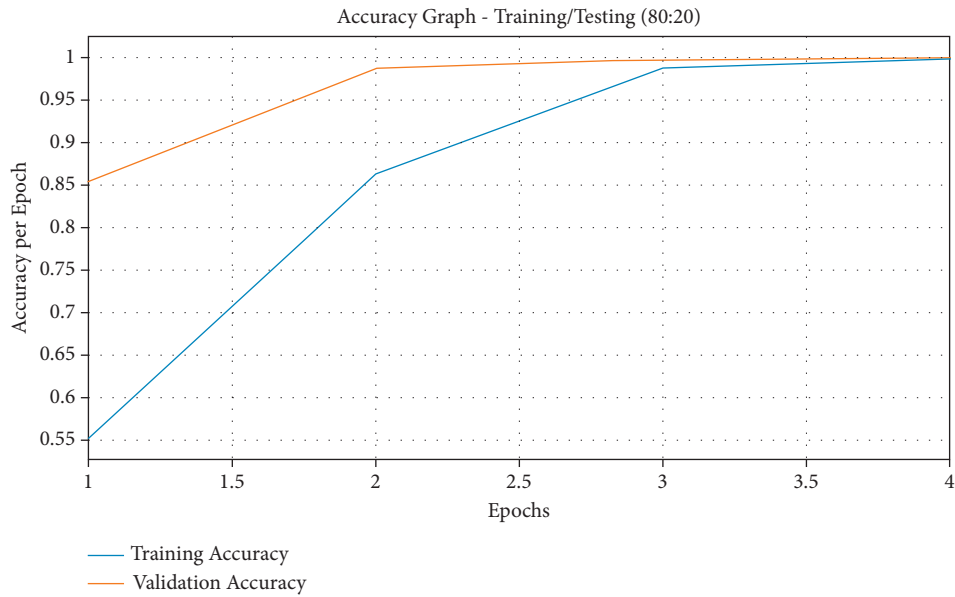


FIGURE 4: Accuracy graph analysis of the QSGOA-DL model on training/testing (80 : 20).

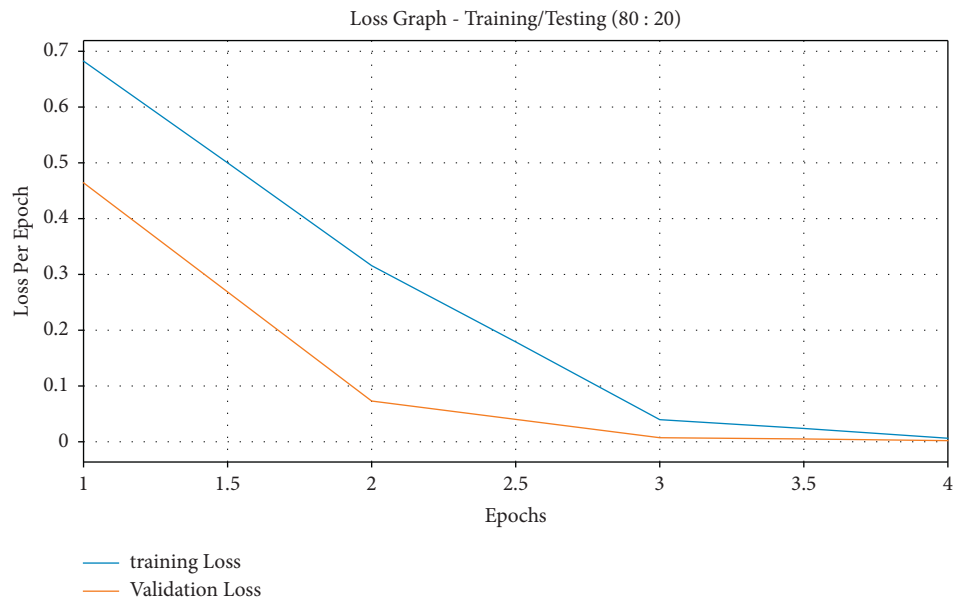


FIGURE 5: Loss graph analysis of the QSGOA-DL model on training/testing (80 : 20).

Figure 4 illustrates the accuracy graph plotted based on the results from the QSGOA-DL technique on the applied training/testing data size of 80 : 20. The figure reports that both training and testing accuracies got increased with an increase in epoch count. It got saturated after a maximum epoch count. It is also observed that the training accuracy got considerably higher than the testing accuracy.

Figure 5 exemplifies the loss graph plotted on the basis of results from the QSGOA-DL technique on the applied training/testing data size of 80 : 20. The figure states that both training and testing losses got heavily reduced with an increase in epoch count and got saturated after a maximum

epoch count. It is noticed that the training loss is lower than the testing accuracy.

Figure 6 showcases the accuracy graph plotted based on QSGOA-DL method results on the applied training/testing of 70 : 30. The figure describes that both training and testing accuracy values got increased with an increase in epoch count and got saturated after a maximal epoch count. It is also detected that the training accuracy got significantly enhanced to the testing accuracy.

Figure 7 demonstrates the loss graph plotted based on the analysis results of QSGOA-DL method on the applied training/testing of 70 : 30. The figure indicates that both

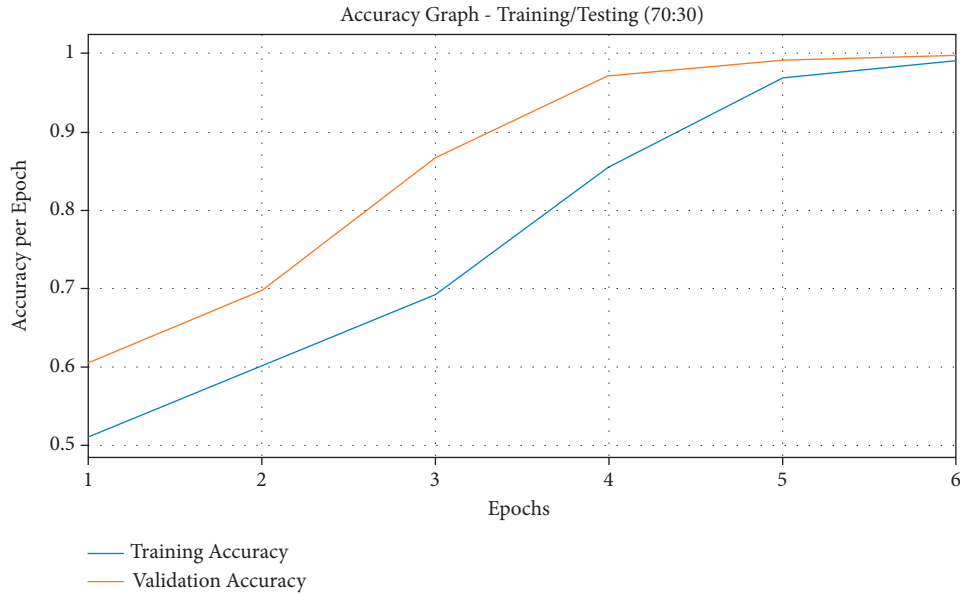


FIGURE 6: Accuracy analysis results of the QSGOA-DL model on training/testing (70 : 30).

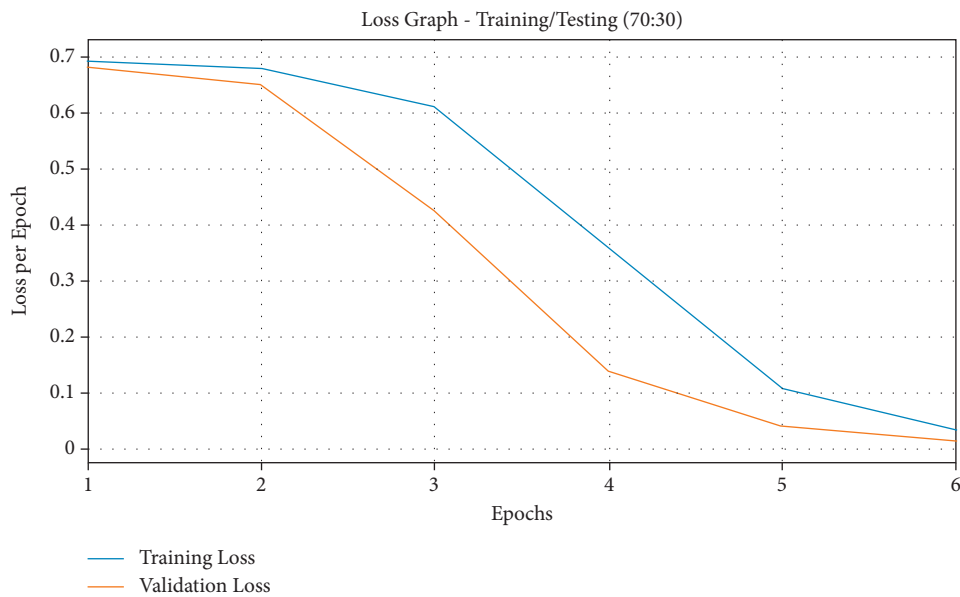


FIGURE 7: Loss analysis results of the QSGOA-DL model on training/testing (70 : 30).

training and testing losses got increased with a higher epoch count and got saturated after a superior epoch count. It is also observed that the training loss was lesser than the testing accuracy.

Figure 8 demonstrates the results from accuracy graph analysis of QSGOA-DL algorithm on the applied training/testing of 60 : 40. The figure states that both training and testing accuracy values get enhanced with an increase in epoch count and attained saturation after a high epoch count. From the results, it can be inferred that the training accuracy is noticeably superior to the testing accuracy. Figure 9 represents the loss graph analysis plot for the

presented QSGOA-DL technique on applied training/testing of 60 : 40. The figure showcases that both training and testing losses turn into minimum value with a superior epoch count and gets saturated after an increased epoch count. It can be observed that the training loss got established and was lesser than the testing accuracy.

Finally, a detailed comparative study was conducted between the proposed QSGOA-DL technique and other recent approaches, and the results are shown in Table 2 and Figures 10 and 11 [26]. By examining the results in terms of precision, it is evident that DHL-2, ResNet-1, and ResNet-2 techniques attained a minimal precision of 97%, 97%, and

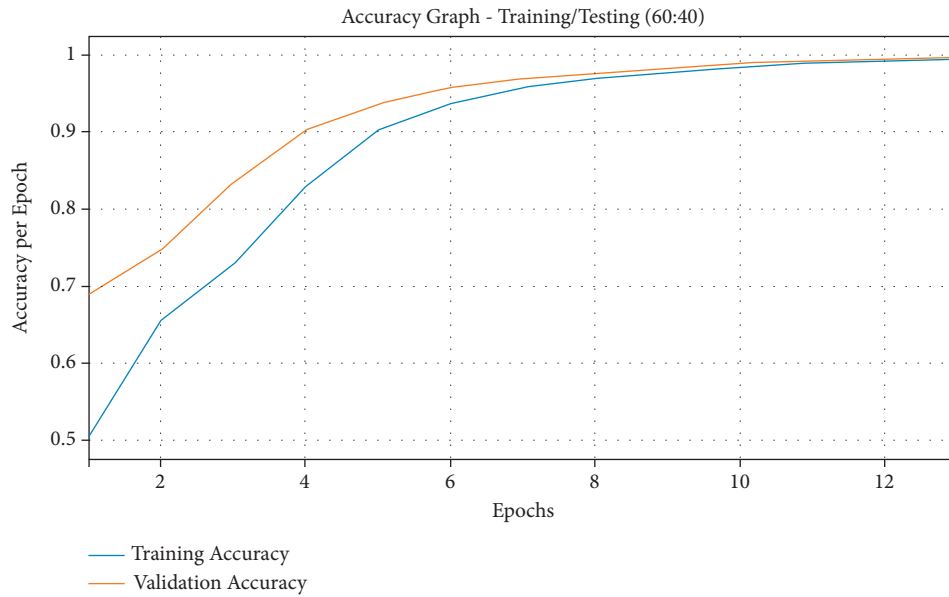


FIGURE 8: Accuracy graph analysis of the QSGOA-DL model on training/testing (60:40).

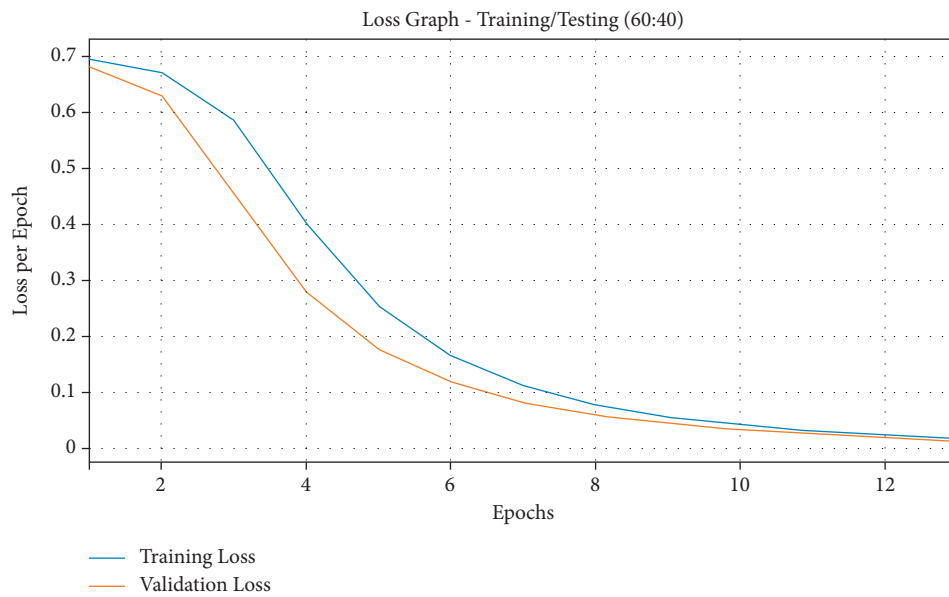


FIGURE 9: Loss graph analysis of the QSGOA-DL model on training/testing (60:40).

97%, respectively. Likewise, DHBL, DHL-1, and TL-ResNet-2 techniques accomplished moderate precision values of 98%, 98%, and 98%, respectively. Though TL-RENet-1 produced a near-optimal precision of 99%, the proposed QSGOA-DL technique gained a high precision of 99.80%.

Besides, with respect to sensitivity, it is clear that the models such as TL-RENet-1, ResNet-1, and ResNet-2 have obtained the least possible sensitivity of 97%, 97%, and 97%, respectively. Likewise, DHL-1, TL-ResNet-2, and DHL-2 techniques have accomplished moderate sensitivity values of 98%, 98%, and 99%, respectively. However, DBHL produced a near-optimal sensitivity of 99%, whereas the presented QSGOA-DL methodology attained a superior sensitivity of 99.80%. At the same time, by examining the results in terms

of specificity, DHL-2, ResNet-1, and ResNet-2 techniques attained the least specificity values, namely, 97%, 97%, and 97%, respectively. In line with this, DHBL, DHL-1, and TL-ResNet-2 systems accomplished moderate specificity values of 98%, 98%, and 98%, respectively. TL-RENet-1 achieved a near-optimal specificity of 99%, while the projected QSGOA-DL algorithm reached the maximum specificity of 99.80%.

On the other hand, by inspecting the results in terms of accuracy, ResNet-1, ResNet-2, and TL-RENet-1 methods attained the least accuracy values of 97.21%, 97.21%, and 98.06%, respectively. Likewise, TL-ResNet-2, DHL-1, and DHL-2 methodologies too accomplished moderate accuracy values of 98.14%, 98.14%, and 98.29%, respectively. Though

TABLE 2: Comparative analysis results of the QSGOA-DL model with different measures.

Methods	Precision	Sensitivity	Specificity	Accuracy	F-score	MCC
DBHL	98.00	99.00	98.00	98.53	98.00	97.00
DHL-2	97.00	99.00	97.00	98.29	98.00	97.00
DHL-1	98.00	98.00	98.00	98.14	98.00	96.00
ResNet-2	97.00	97.00	97.00	97.21	97.00	94.00
TL-ResNet-2	98.00	98.00	98.00	98.14	98.00	96.00
ResNet-1	97.00	97.00	97.00	97.21	97.00	94.00
TL-RENet-1	99.00	97.00	99.00	98.06	98.00	96.00
QSGOA-DL	99.80	99.80	99.80	99.83	99.80	99.70

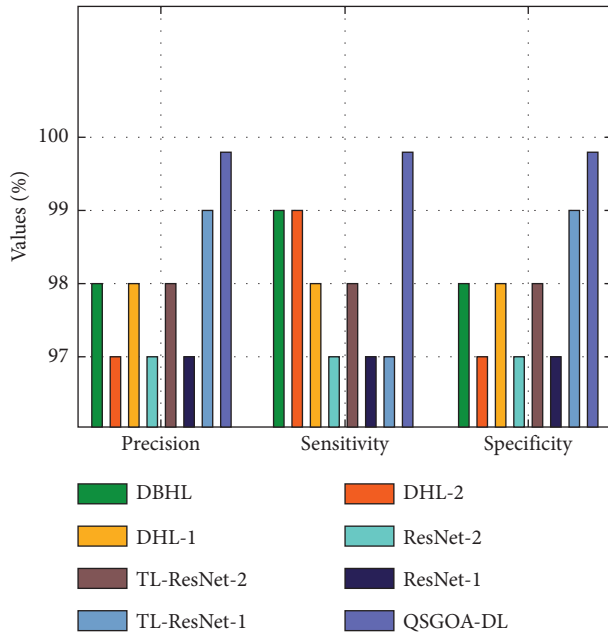


FIGURE 10: Comparative analysis results of the QSGOA-DL model under different measures.

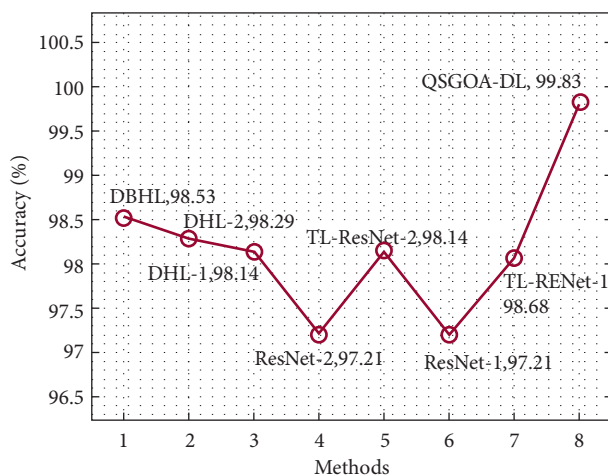


FIGURE 11: Accuracy analysis results of the QSGOA-DL model against existing approaches.

DBHL resulted in a near-optimal accuracy of 98.53%, the proposed QSGOA-DL manner accomplished a superior accuracy of 99.83%. The abovementioned results imply that

the proposed QSGOA-DL technique outperformed the existing methods with a maximum precision of 99.80%, sensitivity of 99.80%, specificity of 99.80%, accuracy of 99.83%, F-score of 99.80%, and MCC of 99.70%. Therefore, the proposed model can be utilized as a proper tool to diagnose COVID-19 using CXR images.

5. Conclusion

In this study, a novel QSGOA-DL technique is presented to detect and classify COVID-19 using CXR images. The proposed QSGOA-DL technique encompasses different operational stages such as preprocessing, EfficientNet-B4-based feature extraction, QSGO-based hyperparameter optimization, and MELM-based classification. The design of QSGO technique assists in the optimal selection of hyperparameter values of EfficientNet-B4 model. In order to showcase the supremacy of the proposed QSGOA-DL technique, a wide range of experimental analyses was conducted on benchmark test CXR dataset. The results were assessed under several aspects. The simulation results demonstrate the promising performance of QSGOA-DL technique than the existing approaches. In future, the performance of QSGOA-DL technique can be validated using computed tomography (CT) scan images in the diagnosis of COVID-19.

Data Availability

Data sharing is not applicable to this article as no datasets were generated during the current study.

Ethical Approval

This article does not contain any studies with human participants performed by any of the authors.

Consent

Not applicable.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgments

The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education, Saudi Arabia, for funding this research work through the project no. (IFPRC-215-249-2020), and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

References

- [1] M. M. Islam, F. Karray, R. Alhaji, and J. Zeng, "A review on deep learning techniques for the diagnosis of novel coronavirus (covid-19)," *IEEE Access*, vol. 9, pp. 30551–30572, 2021.
- [2] D. Ji, Z. Zhang, Y. Zhao, and Q. Zhao, "Research on classification of covid-19 chest x-ray image modal feature fusion based on deep learning," *Journal of Healthcare Engineering*, vol. 2021, Article ID 6799202, 2021.
- [3] M. Mahin, S. Tonmoy, R. Islam, T. Tazin, M. Monirujjaman Khan, and S. Bourouis, "Classification of COVID-19 and pneumonia using deep transfer learning," *Journal of Healthcare Engineering*, vol. 2021, Article ID 3514821, 2021.
- [4] X. Li, W. Tan, P. Liu, Q. Zhou, and J. Yang, "Classification of COVID-19 chest CT images based on ensemble deep learning," *Journal of Healthcare Engineering*, vol. 2021, Article ID 5528441, 2021.
- [5] H. Y. F. Wong, H. Y. S. Lam, A. H. T. Fong et al., "Frequency and distribution of chest radiographic findings in patients positive for COVID-19," *Radiology*, vol. 296, 2020.
- [6] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Computers in Biology and Medicine*, vol. 121, 2020.
- [7] X. Wang, X. Deng, Q. Fu et al., "A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2615–2625, 2020.
- [8] D. Ardila, A. P. Kiraly, S. Bharadwaj et al., "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography," *Nature Medicine*, vol. 25, no. 6, pp. 954–961, 2019.
- [9] A. Paszke, F. Massa, A. Lerer et al., "Pytorch: an imperative style, high-performance deep learning library," *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, pp. 8024–8035, 2019, <https://arxiv.org/abs/1912.01703>.
- [10] S. Roy, W. Menapace, S. Oei et al., "Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2676–2687, 2020.
- [11] Q. Liu, C. K. Leung, and P. Hu, "A two-dimensional sparse matrix profile DenseNet for COVID-19 diagnosis using chest CT images," *IEEE Access*, vol. 8, pp. 213718–213728, 2020.
- [12] S. Sakib, T. Tazrin, M. M. Fouda, Z. M. Fadlullah, and M. Guizani, "DL-CRC: deep learning-based chest radiograph classification for COVID-19 detection: a novel approach," *IEEE Access*, vol. 8, pp. 171575–171589, 2020.
- [13] T. Kaur, T. K. Gandhi, and B. K. Panigrahi, "Automated diagnosis of COVID-19 using deep features and parameter free BAT optimization," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 9, pp. 1–9, 2021.
- [14] K. K. Singh and A. Singh, "Diagnosis of COVID-19 from chest X-ray images using wavelets-based depthwise convolution network," *Big Data Mining and Analytics*, vol. 4, no. 2, pp. 84–93, 2021.
- [15] Y. Li, D. Wei, J. Chen et al., "Efficient and effective training of covid-19 classification networks with self-supervised dual-track learning to rank," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2787–2797, 2020.
- [16] A. Shamsi, H. Asgharnezhad, S. S. Jokandan et al., "An uncertainty-aware transfer learning-based framework for COVID-19 diagnosis," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 4, pp. 1408–1417, 2021.
- [17] Y.-H. Wu, S.-H. Gao, J. Mei et al., "Jcs: an explainable covid-19 diagnosis system by joint classification and segmentation," *IEEE Transactions on Image Processing*, vol. 30, pp. 3113–3126, 2021.
- [18] Z. Han, B. Wei, Y. Hong et al., "Accurate screening of COVID-19 using attention-based deep 3D multiple instance learning," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2584–2594, 2020.
- [19] F. Husain, B. Dellen, and C. Torras, "Scene understanding using deep learning," *Handbook of Neural Computation*, pp. 373–382, Academic Press, 2017.
- [20] P. Zhang, L. Yang, and D. Li, "EfficientNet-B4-Ranger: a novel method for greenhouse cucumber disease recognition under natural complex environment," *Computers and Electronics in Agriculture*, vol. 176, Article ID 105652, 2020.
- [21] Y. Che and D. He, "A Hybrid Whale Optimization with Seagull Algorithm for Global Optimization Problems," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6639671, 2021.
- [22] F. Barani and H. Nezamabadi-pour, "Bqibc: a new quantum-inspired artificial bee colony algorithm for binary optimization problems," *Journal of AI and Data Mining*, vol. 6, no. 1, pp. 133–143, 2018.
- [23] J. Lu, J. Huang, and F. Lu, "Distributed kernel extreme learning machines for aircraft engine failure diagnostics," *Applied Sciences*, vol. 9, no. 8, p. 1707, 2019.
- [24] J. Li, B. Xi, Q. Du, R. Song, Y. Li, and G. Ren, "Deep kernel extreme-learning machine for the spectral-spatial classification of hyperspectral imagery," *Remote Sensing*, vol. 10, no. 12, p. 2036, 2018.
- [25] Github, "COVID ChestXray-dataset," 2020, <https://github.com/ieee8023/covid-chestxray-dataset>.
- [26] S. H. Khan, A. Sohail, A. Khan et al., "COVID-19 detection in chest X-ray images using deep boosted hybrid learning," *Computers in Biology and Medicine*, vol. 137, Article ID 104816, 2021.

Research Article

Arithmetic Optimization with RetinaNet Model for Motor Imagery Classification on Brain Computer Interface

Areej A. Malibari,¹ Fahd N. Al-Wesabi¹,² Marwa Obayya,³ Mimouna Abdullah Alkhonaini,⁴ Manar Ahmed Hamza¹,⁵ Abdelwahed Motwakel,⁵ Ishfaq Yaseen,⁵ and Abu Sarwar Zamani⁵

¹Department of Industrial and Systems Engineering, College of Engineering, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

²Department of Computer Science, College of Science & Art at Mahayil, King Khalid University, Saudi Arabia

³Department of Biomedical Engineering, College of Engineering, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

⁴Department of Computer Science, College of Computer and Information Sciences, Prince Sultan University, Saudi Arabia

⁵Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam Bin Abdulaziz University, AlKharj, Saudi Arabia

Correspondence should be addressed to Manar Ahmed Hamza; ma.hamza@psau.edu.sa

Received 12 January 2022; Revised 9 February 2022; Accepted 2 March 2022; Published 24 March 2022

Academic Editor: K. Shankar

Copyright © 2022 Areej A. Malibari et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Brain Computer Interface (BCI) technology commonly used to enable communication for the person with movement disability. It allows the person to communicate and control assistive robots by the use of electroencephalogram (EEG) or other brain signals. Though several approaches have been available in the literature for learning EEG signal feature, the deep learning (DL) models need to further explore for generating novel representation of EEG features and accomplish enhanced outcomes for MI classification. With this motivation, this study designs an arithmetic optimization with RetinaNet based deep learning model for MI classification (AORNNDL-MIC) technique on BCIs. The proposed AORNNDL-MIC technique initially exploits Multiscale Principal Component Analysis (MSPCA) approach for the EEG signal denoising and Continuous Wavelet Transform (CWT) is exploited for the transformation of 1D-EEG signal into 2D time-frequency amplitude representation, which enables to utilize the DL model via transfer learning approach. In addition, the DL based RetinaNet is applied for extracting of feature vectors from the EEG signal which are then classified with the help of ID3 classifier. In order to optimize the classification efficiency of the AORNNDL-MIC technique, arithmetical optimization algorithm (AOA) is employed for hyperparameter tuning of the RetinaNet. The experimental analysis of the AORNNDL-MIC algorithm on the benchmark data sets reported its promising performance over the recent state of art methodologies.

1. Introduction

Brain-computer interface (BCI) is a technology that permits us to communicate with the computer, whereby the device forecasts the abstract aspect of cognitive states with brain signals, namely, electroencephalography (EEG). Also, it is named as Brain-computer interface (BCI) that is commonly associated with AI-enabled approach which permits the user

to harness brain, etc [1]. It is a noninvasive approach that gathers brain oscillatory activation patterns from the scalp. The human brain produces electrical signal that is identified by using EEG. Therefore, it is highly reliable and applicable method for receiving the control command for BCI [2]. Studies involving EEG signals when imagining limb or finger movement, widely called motor imagery (MI), to function artificial intelligence (AI) technique has been witnessed in

this study [3]. An effective BCI scheme has two fundamental needs that consist of effective machine learning (ML) method for the classification of feature extraction and an efficient set of EEG feature must be capable of differentiating task induced brain activities. The study aims to identify the MI-task induced EEG patterns [4, 5].

Mostly, BCI system involves filtering or preprocessing to remove this undesirable component that is embedded with the EEG signals which leads to wrong conclusions and bias the analysis of the EEG [6]. Appropriate preprocessing within the BCI scheme results in cleaner EEG signal, thus enhancing the classification outcomes. The study focuses on the quantum mechanics inspired preprocessing phase within the BCI scheme, for extracting further data from the attained noisy EEG signal, and leads to increased classification performance although categorized by using multiple classification methods [7]. Especially, SVM is widely employed for MI classification in BMI. Imagery signal classification is performed by LR method. KNN is utilized in seizure detection, where NB is utilized for detecting the lower limb movement by analyzing EEG signals [8]. At the same time, DT is primarily utilized for hand amplitude modulation and movement interpretation spatial activity. Deep Learning (DL) method could considerably simplify processing channel, allow automated end-to-end training of retrieval, preprocessing, and classification models [9], while guaranteeing better performance in target. Deep neural network (DNN) stimulated by previous methods like multilayer perceptron (MLP).

This study designs an arithmetic optimization with RetinaNet based deep learning model for MI classification (AORNDL-MIC) technique on BCIs. The proposed AORNDL-MIC technique undergoes two stages of reprocessing namely Multiscale Principal Component Analysis (MSPCA) based denoising and Continuous Wavelet Transform (CWT) based decomposition. Besides, the arithmetic optimization algorithm (AOA) based RetinaNet model is as feature extractor which are then classified by the use of ID3 classifier. To ensure the better results of the AORNDL-MIC approach, a number of experiments were carried out and the result is inspected under different aspects.

The rest of the paper is organized as follows. Section 2 offers related works, Section 3 provides proposed model, Section 4 discusses performance validation, and Section 5 draws conclusion.

2. Related Works

Zhang et al. [10] validate and developed a DL-based algorithm for automatically recognizing two distinct MI states by choosing the related EEG channel. It employs an automated channel selection (ACS) approach. Furthermore, we proposed a CNN method for fully exploiting the time-frequency feature, therefore outperforming conventional classification method interms of robustness and accuracy. Kant et al. [11] present an integration of DL-based TL and CWT for solving the problems. CWT transforms 1D-EEG signal into 2D time-frequency-amplitude representation enables users to

make use of deep network via TL method. Corsi et al. [12] adapted a fusion technique that integrates features from instantaneously recorded MEG and EEG signals to enhance classification performance in MI-based BCI. Thomas et al. [13] introduce a discriminatory filter bank (FB) common spatial pattern model for extracting FB for the classification of MI. The presented model improves the classifier performance in BCI datasets.

Dong et al. [14], proposed a hierarchical SVM approach for addressing an EEG-based 4-class MI classification process. Wavelet packet transform is applied for decomposing original EEG signal. EEG feature vector is extracted and a two-layer HSVM approach is developed for classifying this EEG feature vector, whereas ‘‘OVO’’ classifier is utilized in the initial layer as well as ‘‘OVR’’ in the next layer. Zhang et al. [15], proposed a ‘‘brain-ID’’ architecture based hybrid DNN using TL method for handling single difference of 4-class MI tasks. A dedicated HDNN is designed for learning the common feature of MI signals. The suggested algorithm comprises LSTM and CNN models that are employed for decoding the spatiotemporal features of the MI signal. Zhang et al. [16] introduce 5 systems for adoption of a DCNN based EEG-BCI scheme for decoding hand MI. All the systems are widely trained, pretrained method and adapt it to improve the efficiency.

3. The Proposed Model

In our study, an AORNDL-MIC approach was introduced to classify the MI on BCIs. The proposed AORNDL-MIC technique encompasses a series of operations namely MSPCA based denoising, CWT based decomposition, RetinaNet based feature extraction, AOA based hyperparameter tuning, as well as ID3 based classification.

3.1. Data Preprocessing. Initially, the data preprocessing takes place in two stages namely MSPCA based noise removal and CWT based decomposition. Consider a measurement data set with m sensor exists, namely xeR^m . All the sensors in the measurement samples have n sampling data, that is integrated into a data matrix of size $m \times n$. The procedure has been shown as follows [2]:

$$X = [x_1, x_2, x_3, \dots, x_n]. \quad (1)$$

All the columns represent a measurement variable, and all the rows of X denote a sample. The PCA models initiated by normalizing all the samples of X by calculating the covariance matrix of X .

$$cov(x) \approx \frac{X^T \cdot X}{n - 1}. \quad (2)$$

The method of decomposition X in its PCA, in which $PeR^{m \times A}$ has initial A feature vector of $cov(x)$. Once the feature decomposition of X is made, the size of feature value is arranged from larger to smaller. A indicates the amount of PCA, and it is equivalent to the amount of columns in T . $T \in R^{m \times A}$ denotes a matrix, in which all the columns are called as the principal element variable.

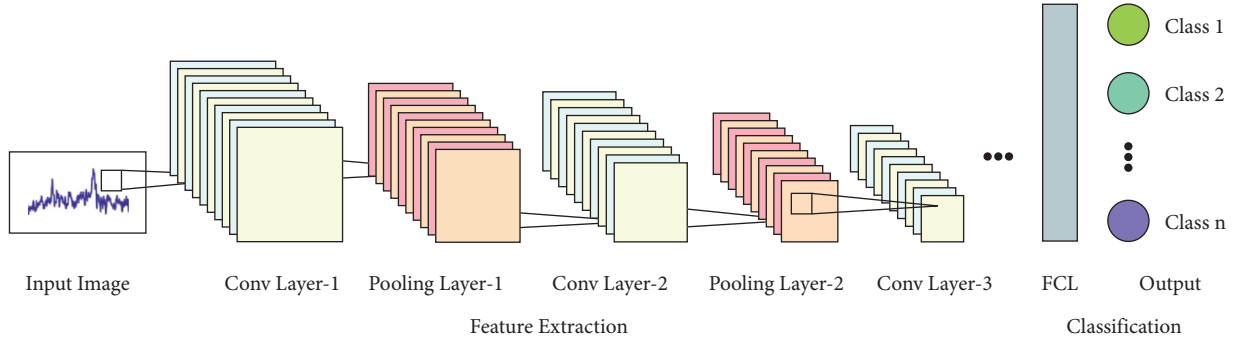


FIGURE 1: Overview of CNN.

$$\begin{cases} X = \hat{X} + Er = T.P^T + Er, \\ T = X.P. \end{cases} \quad (3)$$

In which $\lambda_1, \lambda_2, \dots, \lambda_n$ represent the initial A large eigen values of covariance matrix of X , equation (4) is utilized for determining the principal component covariance,

$$\Lambda = \frac{X^T.X}{n-1} = \begin{bmatrix} \lambda_1 & \dots & \dots & \dots \\ \dots & \lambda_2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \lambda_n \end{bmatrix}. \quad (4)$$

In the study, the wavelet transform is integrated into the PCA model for creating MSPCA to the incoming signal denoising purpose. In MSPCA, the PCA ability for extracting covariance among parameters is integrated to orthonormal wavelets' capability. The capability of PCA is improved by integrating the multi-scale analysis. Simultaneously, it leads to the MSPCA [17]. It finds linearly correlated wavelet coefficient at multilevel sub-bands, attained using wavelet transform. It represents every subband with less features when eliminating the autocorrelated coefficient. The signal is recreated by utilizing the wavelet syntheses. It reduces unnecessary noises from the received signals and generated noise-free and simple signal versions. Also, it can be utilized as a scalogram that is signified by exact value of CWT of the signals. MI signal is gradually changing event peppered by abrupt transient with feature taking place at distinct scales, so lower frequency event, offering maximum time localization to higher frequency, shorter duration event, and higher frequency localization to extended duration, is attained utilizing scalogram.

3.2. RetinaNet Based Feature Extraction. Next to the data preprocessing phase, the AORNDL-MIC technique involves the RetinaNet model as a feature extractor. RetinaNet comprises of two fully convolution networks (FCN), a feature pyramid network (FPN), and residual network (ResNet). ResNet uses distinct network layers. The important role of ResNet is the concept of RL that enables raw input data to be transferred directly to the succeeding layers. The widely employed type of network layer consists of 101-

layer, 152-layer, and 50-layer. The study chooses 101-layers with the optimal training efficiency [18]. Then, extracted the feature of the echocardiography with ResNet and later transmitted to the following subnetworks. FPN is an approach to effectively extract the feature of all the dimensions in a picture with a traditional CNN. Figure 1 illustrates the overview of CNN. Firstly, use single-dimension images as input to ResNet. Next, start from another layer of the convolution network, the feature of each layer was chosen using the FPN and later integrated to generate the last output. The class subnetwork in the FCN implemented the classifier process. Focal loss: it is an amended form of binary cross-entropy expression, as well as the cross-entropy loss:

$$CE(p, y) = \begin{cases} -\log(p), & \text{if } y = 1, \\ -\log(1-p), & \text{otherwise,} \end{cases} \quad (5)$$

whereas $y \in [\pm, 1]$ characterizes the ground truth category and $p \in [0, 1]$ signifies the predicted likelihood of algorithm for $y = 1$.

$$p_t = \begin{cases} p, & \text{if } y = 1, \\ 1-p, & \text{otherwise.} \end{cases} \quad (6)$$

The abovementioned equation is abbreviated as

$$CE(p, y) = CE(p_t) = -\log(p_t). \quad (7)$$

To resolve the problems of the data imbalance among the negative and positive instances, the novel version is changed into the subsequent form:

$$CE(p_t) = -\alpha_t \log(p_t), \quad (8)$$

and amongst them,

$$\alpha_t = \begin{cases} \alpha, & \text{if } y = 1, \\ 1-\alpha, & \text{otherwise,} \end{cases} \quad (9)$$

whereas, $\alpha \in [0, 1]$ characterizes the weight factor. To resolve the problems of complex samples, the variable C is presented for obtaining the last form of focal loss [19]. Figure 2 illustrates the structure of RetinaNet.

$$FL(p_t) = -\alpha_t (1-p_t)^{\gamma} \log(p_t). \quad (10)$$

Since the hyperparameters of the RetinaNet model influence the overall classifier results of the AORNDL-MIC

technique, the AOA is utilized. In general, as other MH approaches, the AOA consists of, exploration, and exploitation phases, stimulated by mathematical operations, like

-, +, *, and /. Firstly, the AOA generates a set of N solutions [20]. Therefore, solution or agent represents X population, as:

$$X = \begin{bmatrix} x_{N-1,1} & x_{N,1} & x_{2,1} & x_{1,1} & x_{N-1,j} & x_{N,j} & x_{2,j} & x_{1,j} & x_{N,n-1} & x_{1,n-1} & x_{N-1,n} & x_{N,n} & x_{2,n} & x_{1,n} \end{bmatrix},$$

$$X = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & x_{1,n-1} & x_{1,n} \\ x_{2,1} & \cdots & x_{2,j} & \cdots & x_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N-1,1} & \cdots & x_{N-1,j} & \cdots & x_{N-1,n} \\ x_{N,1} & \cdots & x_{N,j} & x_{N,n-1} & x_{N,n} \end{bmatrix}. \quad (11)$$

3.3. AOA Based Hyperparameter Tuning. Then, the fitness function of solution is calculated for detecting optimal one X_b . According to the Math Optimizer Accelerated (MOA) values, AOA implements exploitation or exploration methods. Subsequently, MOA is upgraded by

$$MOA(t) = \text{Min} + t \times \left(\frac{\text{Max}_{MOA} - \text{Min}_{MOA}}{M_t} \right), \quad (12)$$

where M_t characterizes the overall amount of iterations. Min_{MOA} and Max_{MOA} signify the minimal and maximal values of the accelerated function, correspondingly, the division (D) and multiplication (M) are applied in the exploration stage of the AOA, as follows:

$$X_{i,j}(t+1) = \begin{cases} X_{bj} \div (M_{OP} + \epsilon) \times ((UB_j - LB_j) \times \mu + LB_j), & r_2 < 0.5, \\ X_{bj} \times M_{OP} \times ((UB_j - LB_j) \times \mu + LB_j), & \text{otherwise.} \end{cases} \quad (13)$$

Next e signifies smaller integer value, LB_j and UB_j shows upper and lower limits of the searching space at j th parameter. $\mu = 0.5$ denotes the control function. Furthermore, Math Optimizer (M_{OP}) is determined by

$$M_{OP}(t) = 1 - \frac{t^{1/\alpha}}{M_t^{1/\alpha}}. \quad (14)$$

$\alpha = 5$ characterizes the dynamic variable which defines the accuracy of the exploitation stage.

Additionally, subtraction (D) and addition operator (A) operators are employed for executing the AOA exploitation phase, as follows.

$$x_{i,j}(t+1) = \begin{cases} X_{bj} - M_{OP} \times ((UB_j - LB_j) \times \mu + LB_j), & r_3 < 0.5, \\ X_{bj} + M_{OP} \times ((UB_j - LB_j) \times \mu + LB_j), & \text{otherwise.} \end{cases} \quad (15)$$

Now r_3 characterizes an arbitrary value in $[0, 1]$. Next, the agent updating procedure is executed by the AOA operator [21]. In summary, Algorithm 1 demonstrates the steps included in AOA.

3.4. ID3 Based Classification. Lastly, the ID3 architecture receives the feature vector as input and carries out the classification process. The ID3 technique selects test elements with computing and relating its information gains (IG). Assume S be the group of data instances. Supposing the class element C has m distinct values that signify m various class labels C_i ($i = 1, 2, \dots, m$). Assume that S_i be the amount of instances from class C_i ($i = 1, 2, \dots, m$). The predictable data amount needed for classifying S was provided in equation (15):

$$I(S_1, S_2, \dots, S_m) = - \sum_{i=1}^m p_i \log_2 p_i, \quad (16)$$

where p_i signifies the probability of samples from S appropriate to class C_i . $I(S_1, S_2, \dots, S_m)$ refers to the average data amount needed for identifying the class label to every instance from S .

Let the element A has v distinct values $\{a_1, a_2, \dots, a_v\}$ from the trained data set S . When A is a nominal element, Afterward, the element separates S as to v subset such that $\{S_1, S_2, \dots, S_v\}$, in that S_j represents the subset of S where sample from S_j has the similar element value a_j on A . But, instance from S_j can have various class labels [22]. Assume S_{ij} be the group of instances that class label is C_i from the subset of $\{S_j | A = a_j, j \in 1, 2, \dots, v, S_j \in S\}$ in which element $A = a_j$. The needed data amount (i.e., entropy) of element A for splitting the trained data set S was measured by (16):

$$E(A) = \sum_{j=1}^v \left(\frac{(s_{1j} + s_{2j} + \dots + s_{mj})}{s} \times I(s_{1j}, s_{2j}, \dots, s_{mj}) \right). \quad (17)$$

The minimum data amount needed, a further purity of sub-dataset is.

$$I(s_{1j}, s_{2j}, \dots, s_{mj}) = - \sum_{i=1}^m p_{ij} \log_2(p_{ij}), \quad (18)$$

where p_{ij} implies the probability of instances from S_j based on class C_i . $I(s_{1j}, s_{2j}, \dots, s_{mj})$ signifies the average data

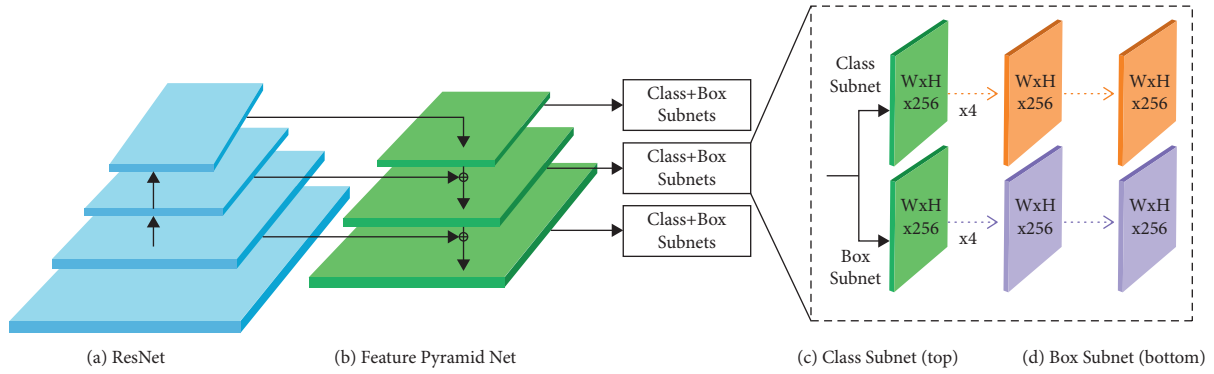


FIGURE 2: RetinaNet network architecture.

```

Input: the parameter of AOA includes overall quantity of iterations  $M_t$ , dynamic exploitation variable ( $\alpha$ ), and number of agent ( $N$ ).
Generate the primary value for the agent  $X_i, i = 1, \dots, N/$ . while ( $t < M_t$ ) do
  Calculate the fitness function for all the agents.
  Define the optimal agent  $X_b$ .
  Upgrade the  $MOA$  and  $M_{OP}$  using equations (11) and (13),
  for  $i = 1$  to  $N$  do
    for  $j = 1$  to  $Di m$  do
      Upgrade the values of  $r_1, r_2$ , and  $r_3$ .
      if  $r_1 > MOA$  then
        Exploration stage
        Employ equation (12) to upgrade the  $X_i$ .
      else
        Exploitation stage
        Employ equations (14) to upgrade the  $X_i$ .
      end if
    end for
  end for
   $t = t + 1$ 
end while
Display the optimal agent ( $X_b$ ).

```

ALGORITHM 1: Pseudocode of AOA.

amount needed for identifying the class labels to every instance from S_j . The IG of A has determined as:

$$\text{Info Gain}(A) = I(S_1, S_2, \dots, S_m) - E(A). \quad (19)$$

Specifically, the count of novel data requirement (only dependent upon class) minus the count of novel data requirements (based the split on element A). Selecting the element with maximal Info Gain (A) as test element that is allocated to internal node from DT. During this process, the required data amount to classify samples is minimal.

4. Results and Discussion

The performance validation of the AORNDL-MIC technique has been validated under two dataset includes BCI competition 2003 dataset III and BCI competition IV data set 2b. The BCI competition 2003, dataset III [23], comprises 3-channel EEG data in healthy females, for the imagination of the right, and left -hand movements. The data from the analysis has of recording in the motor cortex area of brain

utilizing 3 electrodes (C3, Cz, and C4) under the motor imagery of combined right-or-left-hand movement. All individual trail last to 9-second duration of data to all channels C3, Cz, and C4 per trial with every label obtainable. It holds 280 out of which 140 trials were accessible with its labels, and other 140 instances were employed for validation method. The BCI competition IV data set 2b comprises nine subjects all with 5 sessions of motor imagery experimentally, amongst that the initial 2 sessions are verified with no feedback and the remaining 3 sessions are combined online feedback [24].

4.1. Result Analysis on BCI Competition 2003 III Dataset. Figure 3 illustrates the confusion matrices generated by the AORNDL-MIC algorithm under five iterations. At iteration-1, the AORNDL-MIC technique has identified 67 instances in left class and 68 instances in right class. Moreover, at iteration-3, the AORNDL-MIC method has identified 69 instances into left class and 68 instances into right class.

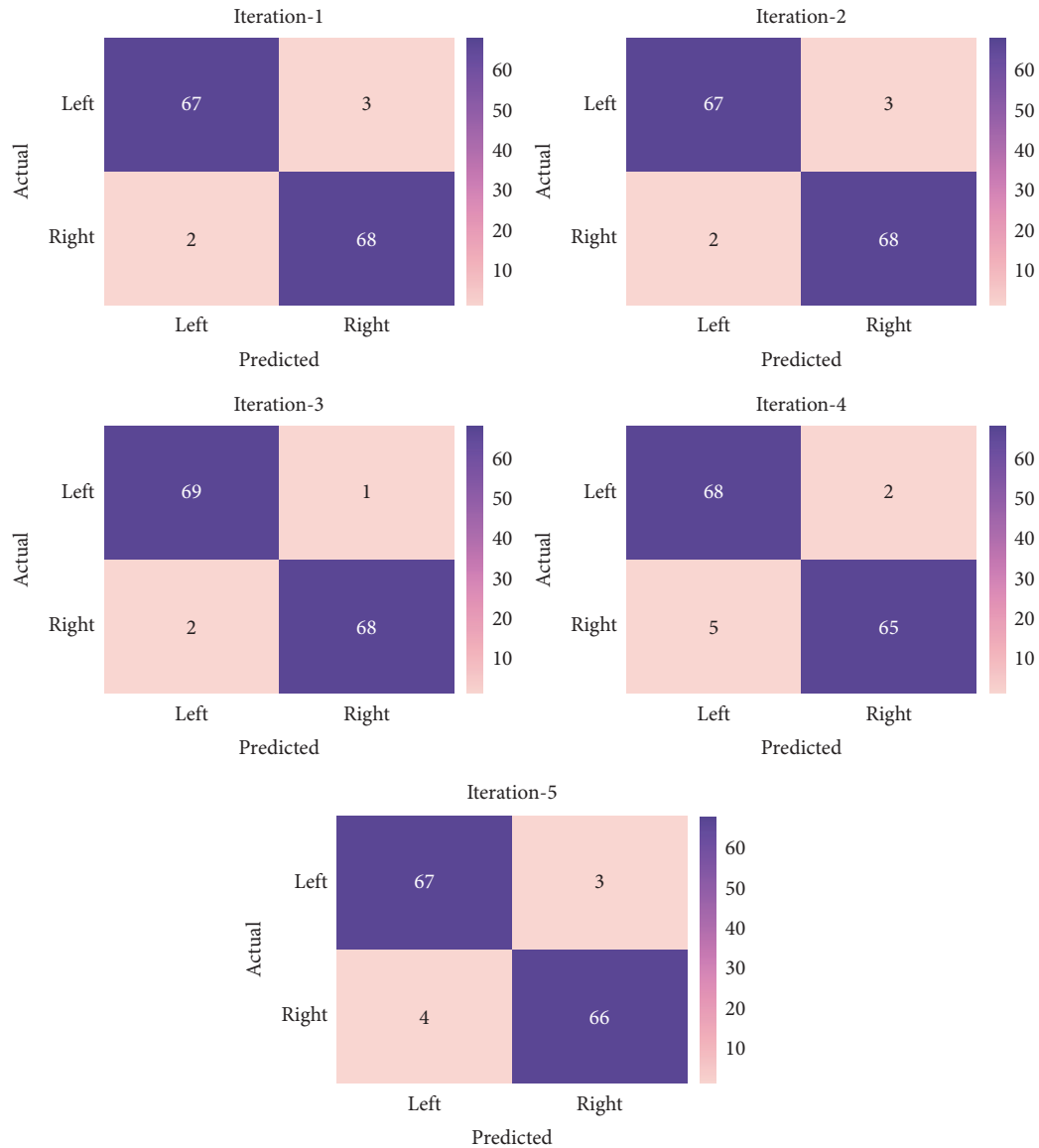


FIGURE 3: Confusion matrix of AORNDL-MIC technique under five iterations.

Furthermore, at iteration-5, the AORNDL-MIC approach has identified 67 samples into left class and 66 samples into right class.

Table 1 and Figures 4, 5 provides a classifier results of the AORNDL-MIC algorithm on BCI competition 2003 III dataset. The experimental result indicates the better outcomes of the AORNDL-MIC technique under each iteration. For example, with iteration-1, the AORNDL-MIC algorithm has gained precision of 97.10%, recall of 95.71%, accuracy of 94.43%, F – score of 96.40%, and kappa of 95.26%. Meanwhile, with iteration-3, the AORNDL-MIC method has reached precision of 97.18%, recall of 98.57%, accuracy of 97.86%, F – score of 97.87%, and kappa of 97.13%. Eventually, with iteration-5, the AORNDL-MIC system has obtained precision of 94.37%, recall of 95.71%, accuracy of 95%, F – score of 95.04%, and kappa of 93.30%.

A comparative analysis of the AORNDL-MIC approach with current methodologies on the test BCI competition

2003, dataset III showed in Figure 6 and Table 2. The result exhibits that the SqueezeNet, ResNet50, GoogleNet, DenseNet201, ResNet18, and ResNet101 techniques have resulted to lower kappa values of 57%, 41%, 44%, 36%, 29%, and 30% correspondingly. Next, the VGG19, AlexNet, and VGG16 models have resulted in slightly increased kappa values of 91%, 87%, and 90%, respectively. However, the proposed AORNDL-MIC technique has accomplished higher kappa value of 94.84%.

A comparative study of the AORNDL-MIC method with recent algorithms on the test BCI competition 2003, dataset III is illustrated in Table 3 and Figure 7. The outcome demonstrates that the CSP-SVM, STFT-KNN, Optimized GA FKNN-LDA, Hybrid KNN, and WTSE-SVM techniques have resulted in minimum accuracy values of 82.86%, 83.57%, 84%, 84.29%, and 86.40%, respectively. Then, the Adaptive PP-Bayesian, STFT-DL, and CWTFB-TL methods have resulted in slightly maximal accuracy values of 90%,

TABLE 1: Classification outcomes of AORNDL-MIC approach on BCI competition 2003 III datasets.

No. of iterations	Precision	Recall	Accuracy	F-score	Kappa
Iteration-1	97.10	95.71	96.43	96.40	95.26
Iteration-2	97.10	95.71	96.43	96.40	95.26
Iteration-3	97.18	98.57	97.86	97.87	97.13
Iteration-4	93.15	97.14	95.00	95.10	93.24
Iteration-5	94.37	95.71	95.00	95.04	93.30
Average	95.78	96.57	96.14	96.16	94.84

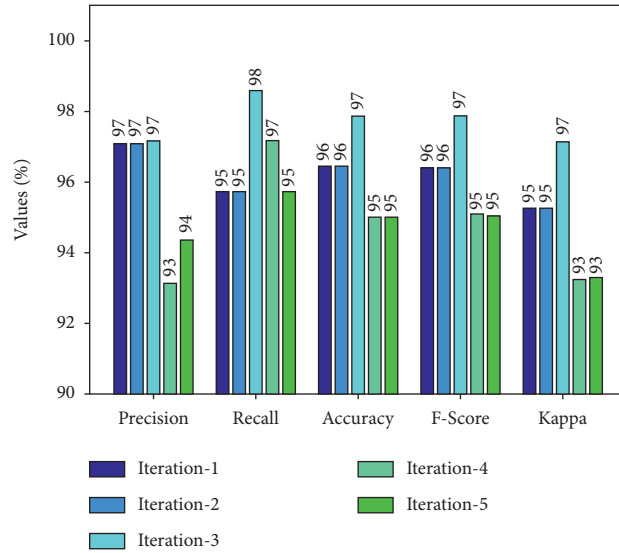


FIGURE 4: Result analysis AORNDL-MIC technique on BCI competition 2003 III datasets.

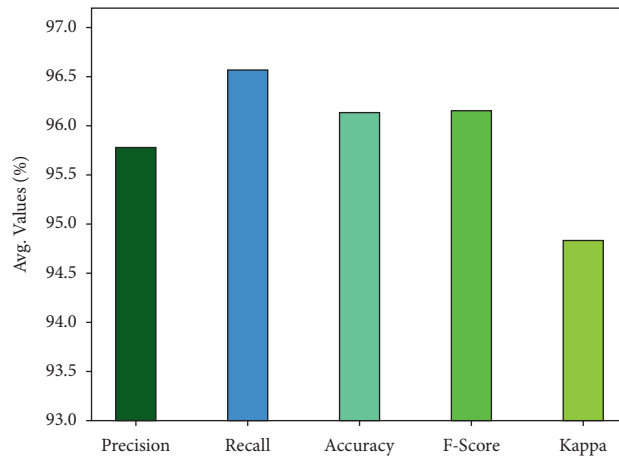


FIGURE 5: Average analysis AORNDL-MIC technique on BCI competition 2003 III dataset.

90%, and 95.71% correspondingly. Lastly, the proposed AORNDL-MIC method has accomplished superior accuracy value of 96.14%.

4.2. Result Analysis on BCI Competition IV Data Set 2b Dataset. A classification results of the AORNDL-MIC method on BCI competition IV data set 2b under several subjects and runs is shown in Table 4 and Figure 8. The experimental value indicates that the AORNDL-MIC

algorithm has demonstrated better performance with an average accuracy of 85.33%, 84.22%, 90.11%, 87.11%, and 85.89% under runs 1–5, respectively.

An average classification results of the AORNDL-MIC method under several subjects are portrayed in Figure 9. The results showed that the AORNDL-MIC system has the ability of accomplishing improved outcomes with the maximum average accuracy of 81.20% under S-1, 87.20% under S-2, 84.60% under S3, 91.60% under S-4, and so on.

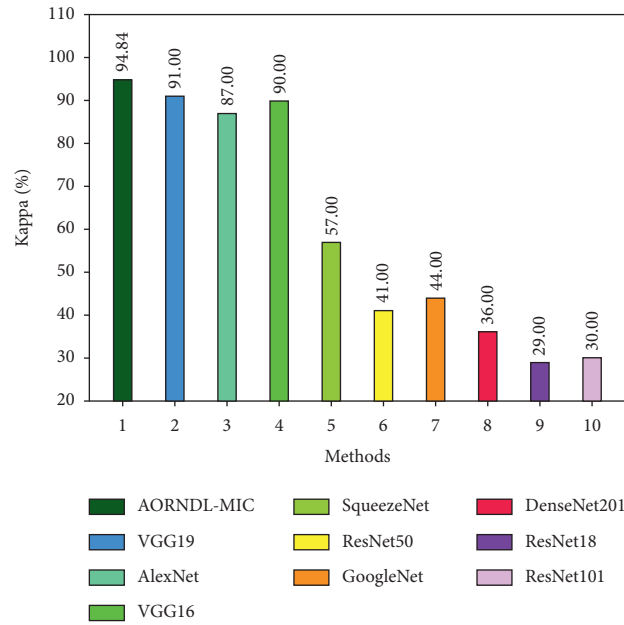


FIGURE 6: Kappa analysis of AORNDL-MIC technique with current approaches.

TABLE 2: Kappa analysis of AORNDL-MIC technique with existing approaches on test BCI competition 2003, dataset III.

Methods	Kappa
AORNDL-MIC	94.84
VGG19	91.00
AlexNet	87.00
VGG16	90.00
SqueezeNet	57.00
ResNet50	41.00
GoogleNet	44.00
DenseNet201	36.00
ResNet18	29.00
ResNet101	30.00

TABLE 3: Accuracy analysis of AORNDL-MIC technique with existing approaches on test BCI competition 2003, dataset III.

Methods	Accuracy
Hybrid KNN	84.29
CSP-SVM	82.86
Adaptive PP-Bayesian	90.00
STFT-KNN	83.57
STFT-DL	90.00
Optimized GA FKNN-LDA	84.00
WTSE-SVM	86.40
CWTFB-TL	95.71
AORNDL-MIC	96.14

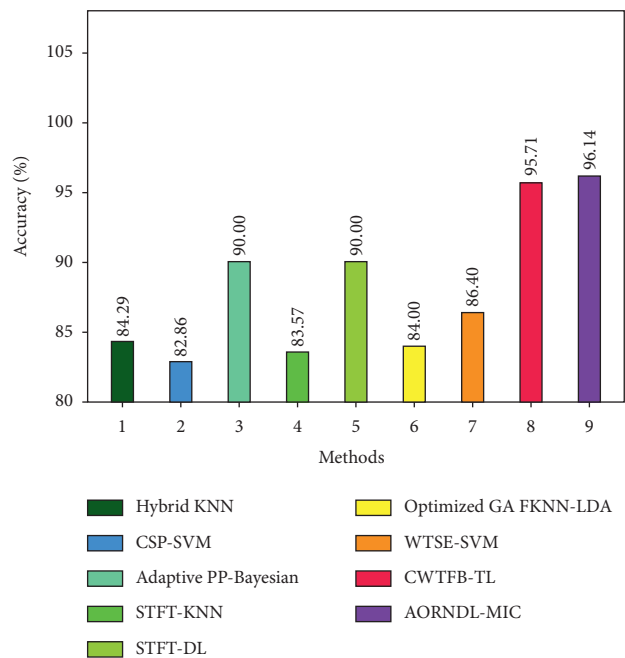


FIGURE 7: Accuracy analysis of AORNDL-MIC approach with current methodologies.

Table 5 and Figure 10 provide a comparative study of the AORNDL-MIC system with current methodologies in terms of accuracy. The experimental results indicated that the AORNDL-MIC technique has resulted in better results over the other methodologies under all subjects. For instance, with S-1, the AORNDL-MIC algorithm has accomplished higher performance of 81.20% whereas the CSP, FBCSP MIBIF, FBCSP

MIRSR, and FDBN techniques have attained lower accuracy of 66%, 68%, 70%, and 81% respectively. Moreover, with S-5, the AORNDL-MIC approach has reached superior accuracy of 85.80% whereas the CSP, FBCSP MIBIF, FBCSP MIRSR, and FDBN methods have attained lesser accuracy of 77%, 93%, 93%, and 93%, respectively. Furthermore, with S-9, the AORNDL-MIC approach has gained superior accuracy of 87.60% whereas the CSP, FBCSP MIBIF, FBCSP MIRSR, and FDBN methods have achieved minimum accuracy of 83%, 88%, 87%, and 91% correspondingly.

TABLE 4: Classification results of the AORNDL-MIC approach under several subjects and runs.

No. of runs	S-1	S-2	S-3	S-4	S-5	S-6	S-7	S-8	S-9	Avg.
R-1	87.00	85.00	88.00	86.00	84.00	76.00	83.00	96.00	83.00	85.33
R-2	71.00	83.00	80.00	94.00	85.00	79.00	91.00	83.00	92.00	84.22
R-3	84.00	96.00	94.00	98.00	88.00	90.00	85.00	89.00	87.00	90.11
R-4	82.00	91.00	75.00	89.00	91.00	92.00	81.00	88.00	95.00	87.11
R-5	82.00	81.00	86.00	91.00	81.00	87.00	95.00	89.00	81.00	85.89
Avg.	81.20	87.20	84.60	91.60	85.80	84.80	87.00	89.00	87.60	86.53

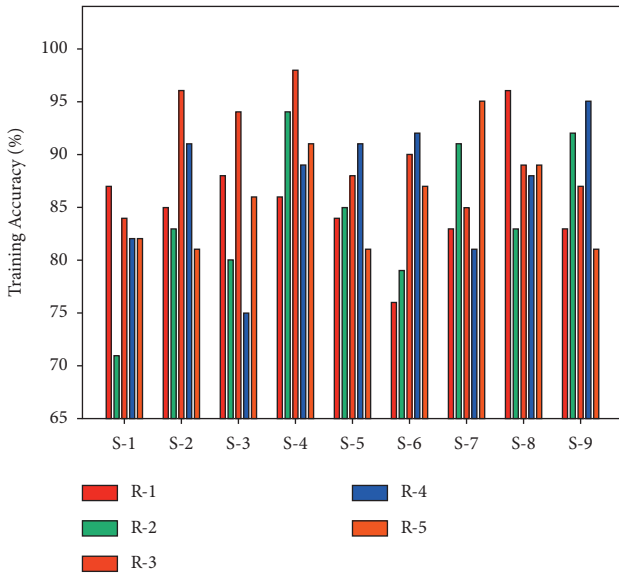


FIGURE 8: Training accuracy analysis of AORNDL-MIC technique.

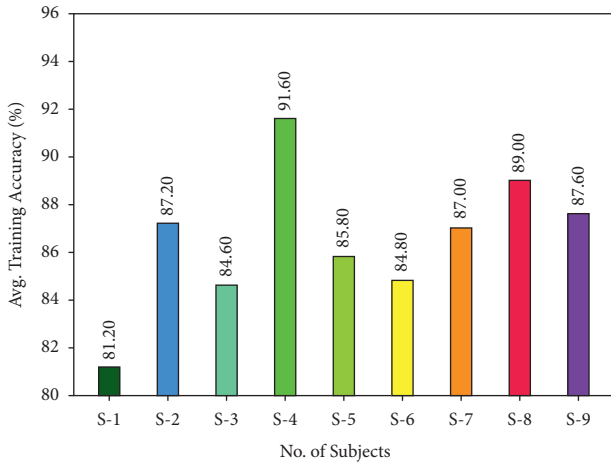


FIGURE 9: Average training accuracy analysis of AORNDL-MIC technique.

For ensuring the improvement of AORNDL-MIC model, an average accuracy analysis is also made in Figure 11. From the figure, it is apparent that the CSP and FBCSP MIBIF techniques have reached lower performance with an average accuracy of 76.33% and 79.56% respectively. In line with, the FBCSP MIRSR and FDBN

TABLE 5: Comparative study of AORNDL-MIC technique with recent methodologies interms of accuracy.

Subject	CSP	FBCSP MIBIF	FBCSP MIRSR	FDBN	AORNDL-MIC
S-1	66.00	68.00	70.00	81.00	81.20
S-2	62.00	59.00	61.00	65.00	87.20
S-3	57.00	59.00	61.00	66.00	84.60
S-4	97.00	98.00	98.00	98.00	91.60
S-5	77.00	93.00	93.00	93.00	85.80
S-6	75.00	80.00	81.00	88.00	84.80
S-7	77.00	78.00	78.00	82.00	87.00
S-8	93.00	93.00	93.00	94.00	89.00
S-9	83.00	88.00	87.00	91.00	87.60
Average	76.33	79.56	80.22	84.22	86.53

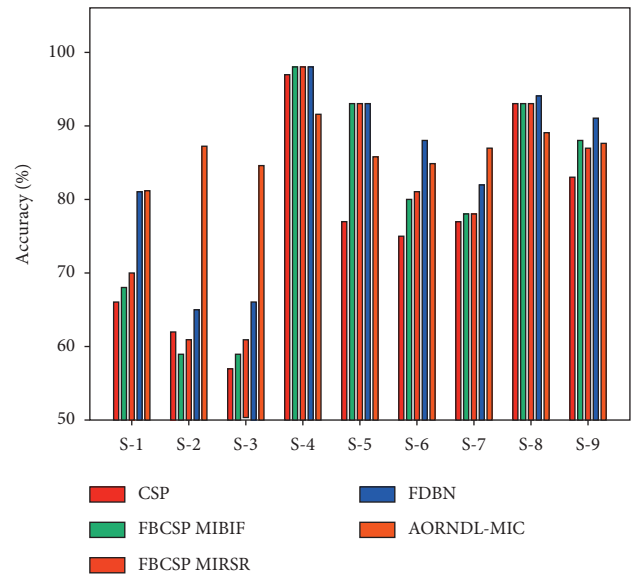


FIGURE 10: Accuracy analysis of AORNDL-MIC technique with recent methods.

systems have resulted in moderately increased average accuracy of 80.22% and 84.22% respectively. However, the AORNDL-MIC approach has gained effective performance over the other methodologies with the maximal average accuracy of 86.53%. By observing the experimental results and discussion, it is confirmed that the AORNDL-MIC approach has shown better results over the other methodologies.

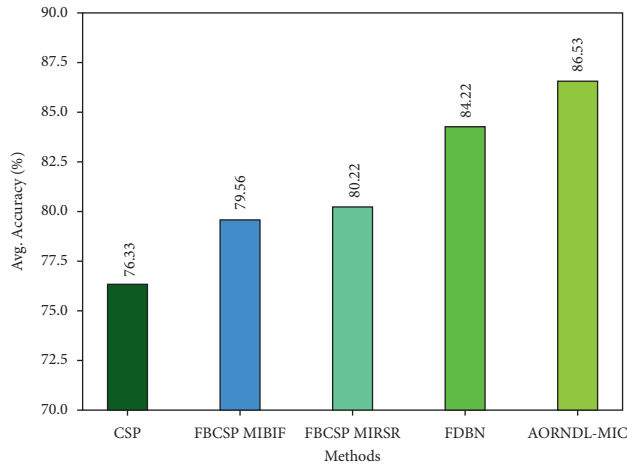


FIGURE 11: Average Accuracy analysis of AORNDL-MIC algorithm with current methodologies.

5. Conclusion

In this study, an AORNDL-MIC system was developed to categorize MI on BCIs. The proposed AORNDL-MIC technique encompasses a series of operations namely MSPCA based denoising, CWT based decomposition, RetinaNet based feature extraction, AOA based hyperparameter, and ID3 based classification. The AOA is employed to tune the hyperparameter of RetinaNet and improves the classification performance of the AORNDL-MIC technique. For ensuring the outcome of the AORNDL-MIC method, a number of experiments were performed and the outcome is examined under different aspects. The experiment results of the AORNDL-MIC algorithm on the benchmark datasets reported its promising outcome over the current state of art approaches. In the future, hybrid DL model can be utilized for boosting the efficacy of the MI classification process.

Data Availability

Data sharing not applicable to this article as no datasets were generated during the current study.

Ethical Approval

This article does not contain any studies with human participants performed by any of the authors.

Consent

Not applicable.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under grant number (RGP 2/18/43). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R151), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to thank Prince Sultan University for its support in paying the Article Processing Charges.

References

- [1] N. Lu, T. Li, X. Ren, and H. Miao, "A deep learning scheme for motor imagery classification based on restricted Boltzmann machines," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 6, pp. 566–576, 2016.
- [2] A. JSubasi and S. Mian Qaisar, "The ensemble machine learning-based classification of motor imagery tasks in brain-computer interface," *Journal of Healthcare Engineering*, vol. 2021, Article ID 1970769, 2021.
- [3] J. Cantillo-Negrete, R. I. Carino-Escobar, P. Carrillo-Mora, D. Elias-Vinas, and J. Gutierrez-Martinez, "Motor imagery-based brain-computer interface coupled to a robotic hand orthosis aimed for neurorehabilitation of stroke patients," *Journal of healthcare engineering*, vol. 2018, Article ID 1624637, 2018.
- [4] Y. Bin He and H. Bin, "Brain-computer interfaces using sensorimotor rhythms: current state and future perspectives," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1425–1435, 2014.
- [5] C. Li, J. Wei, X. Huang, Q. Duan, and T. Zhang, "Effects of a brain-computer interface-operated lower limb rehabilitation robot on motor function recovery in patients with stroke," *Journal of Healthcare Engineering*, vol. 2021, Article ID 4710044, 2021.
- [6] A. Vasilyev, S. Liburkina, L. Yakovlev, O. Perepelkina, and A. Kaplan, "Assessing motor imagery in brain-computer interface training: psychological and neurophysiological correlates," *Neuropsychologia*, vol. 97, pp. 56–65, 2017.
- [7] N. Mora, I. De Munari, and P. Ciampolini, *A Multi-Modal BCI System for Active and Assisted Living*, Springer, New York, NY, USA, 2016.
- [8] N. Lu, T. Li, J. Pan, X. Ren, Z. Feng, and H. Miao, "Structure constrained semi-nonnegative matrix factorization for EEG-based motor imagery classification," *Computers in Biology and Medicine*, vol. 60, pp. 32–39, 2015.
- [9] L. Karl, C. Kaitlin, D. Alexander, S. Kaleb, R. Eitan, and H. Bin, "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface," *Journal of Neural Engineering*, vol. 10, Article ID 046003, 2013.
- [10] H. Zhang, X. Zhao, Z. Wu, B. Sun, and T. Li, "Motor imagery recognition with automatic EEG channel selection and deep learning," *Journal of Neural Engineering*, vol. 18, no. 1, Article ID 016004, 2021.
- [11] P. Kant, S. H. Laskar, J. Hazarika, and R. Mahamune, "CWT Based transfer learning for motor imagery classification for brain computer interfaces," *Journal of Neuroscience Methods*, vol. 345, Article ID 108886, 2020.
- [12] M.-C. Corsi, M. Chavez, D. Schwartz et al., "Integrating EEG and MEG signals to improve motor imagery classification in

- brain-computer interface,” *International Journal of Neural Systems*, vol. 29, no. 1, Article ID 1850014, 2019.
- [13] K. P. Thomas, C. Guan, C. T. Lau, A. P. Vinod, and K. K. Ang, “A new discriminative common spatial pattern method for motor imagery brain-computer interfaces,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 11, pp. 2730–2733, 2009.
- [14] E. Dong, C. Li, L. Li, S. Du, A. N. Belkacem, and C. Chen, “Classification of multi-class motor imagery with a novel hierarchical SVM algorithm for brain-computer interfaces,” *Medical, & Biological Engineering & Computing*, vol. 55, no. 10, pp. 1809–1818, 2017.
- [15] R. Zhang, Q. Zong, L. Dou, X. Zhao, Y. Tang, and Z. Li, “Hybrid deep neural network using transfer learning for EEG motor imagery decoding,” *Biomedical Signal Processing and Control*, vol. 63, Article ID 102144, 2021.
- [16] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, “Adaptive transfer learning for EEG motor imagery classification with deep Convolutional Neural Network,” *Neural Networks*, vol. 136, pp. 1–10, 2021.
- [17] M. T. Sadiq, X. Yu, Z. Yuan et al., “Motor imagery EEG signals decoding by multivariate empirical wavelet transform-based framework for robust brain-computer interfaces,” *IEEE Access*, vol. 7, pp. 171431–171451, 2019.
- [18] Y. Wang, C. Wang, H. Zhang, Y. Dong, and S. Wei, “Automatic ship detection based on RetinaNet using multi-resolution gaofen-3 imagery,” *Remote Sensing*, vol. 11, no. 5, p. 531, 2019.
- [19] M. Yang, X. Xiao, Z. Liu et al., “Deep RetinaNet for Dynamic Left Ventricle Detection in Multiview Echocardiography Classification,” *Scientific Programming*, vol. 2020, Article ID 7025403, 2020.
- [20] L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz, and A. H. Gandomi, “The arithmetic optimization algorithm,” *Computer Methods in Applied Mechanics and Engineering*, vol. 376, Article ID 113609, 2021.
- [21] A. A. Ewees, M. A. A. Al-qaness, L. Abualigah et al., “Boosting arithmetic optimization algorithm with genetic algorithm operators for feature selection: case study on cox proportional hazards model,” *Mathematics*, vol. 9, no. 18, p. 2321, 2021.
- [22] S. Yang, J.-Z. Guo, and J.-W. Jin, “An improved Id3 algorithm for medical data classification,” *Computers & Electrical Engineering*, vol. 65, pp. 474–487, 2018.
- [23] S. Lemm, C. Schafer, and G. Curio, “BCI competition 2003-data set III: probabilistic modeling of sensorimotor/spl mu/rhythms for classification of imaginary hand movements,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1077–1080, 2004.
- [24] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, “Brain-computer communication: motivation, aim, and impact of exploring a virtual apartment,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 4, pp. 473–482, 2007.

Research Article

Optimal Deep Learning Enabled Prostate Cancer Detection Using Microarray Gene Expression

Abdulrhman M. Alshareef ¹, **Raed Alsini** ¹, **Mohammed Alsieni** ², **Fadwa Alrowais**,³
Radwa Marzouk,⁴ **Ibrahim Abunadi**,⁵ and **Nadhem Nemri** ⁶

¹Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

²Department of Pharmacology, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

³Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P. O. Box 84428, Riyadh 11671, Saudi Arabia

⁴Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

⁵Department of Information Systems, College of Computer and Information Sciences, Prince Sultan University, P.O. Box No. 66833, Rafha Street, Riyadh 11586, Saudi Arabia

⁶Department of Information Systems, College of Science and Arts at Muhayel, King Khalid University, Mahayel Aseer, Saudi Arabia

Correspondence should be addressed to Nadhem Nemri; nnemri@kku.edu.sa

Received 15 December 2021; Revised 30 December 2021; Accepted 15 January 2022; Published 10 March 2022

Academic Editor: K. Shankar

Copyright © 2022 Abdulrhman M. Alshareef et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prostate cancer is the main cause of death over the globe. Earlier detection and classification of cancer is highly important to improve patient health. Previous studies utilized statistical and machine learning (ML) techniques for prostate cancer detection. However, several challenges that exist in the investigation process are the existence of high dimensionality data and less number of training samples. Metaheuristic algorithms can be used to resolve the curse of dimensionality and improve the detection rate of artificial intelligence (AI) techniques. With this motivation, this article develops an artificial intelligence based feature selection with deep learning model for prostate cancer detection (AIFSDL-PCD) using microarray gene expression data. The AIFSDL-PCD technique involves preprocessing to enhance the input data quality. In addition, a chaotic invasive weed optimization (CIWO) based feature selection (FS) technique for choosing an optimal subset of features shows the novelty of the work. Moreover, the deep neural network (DNN) model can be applied as a classification model to detect the existence of prostate cancer in the microarray gene expression data. Furthermore, the hyperparameters of the DNN model can be effectively adjusted by the use of RMSprop optimizer. The design of CIWO based FS technique helps for reducing the computational complexity and improve the classification accuracy. The experimental results highlighted the betterment of the AIFSDL-PCD approach on the other techniques with respect to distinct measures.

1. Introduction

In recent times, cancer is the leading cause of death worldwide. Generally, around 1 death from 6 overall deaths is because of cancer [1]. Therefore, in 2030, several new cases predicted annually might increase up to 25 million [2]. But early diagnoses of cancer might save billions of dollars and

countless lives. The earlier prediction and identification of cancer is very crucial for cancer research and patient health. Once cancer is detected at earlier stages, treatment is highly efficient. In the past, classification of cancer is based on clinical and morphological technologies [3]. The innovative technologies have made considerable development in precise observation of hundreds of cancer genes via gene

expression data. This method provides a massive amount of information to the authors for exploring several knowledge; however it has certain problems [4]. The key challenges of microarray data are low sample size and high dimensionality. Additionally, many microarray cancer information is noisy and could not be extremely helpful in the diagnosis of cancer [5]. Nowadays, categorizing cancer type more accurately and precisely and selecting most important genes associated with the cancer is one of the key challenges in the study [6].

Prostate cancer (PCa) is the 3rd one of the general detected cancers around the world, after breast and lung cancer, and the 5th cause of cancer-specific deaths in males [7]. In the past decades, researchers focus more on the prediction, diagnosis, and prognosis of PCa results taking the next step with help of Statistics and Artificial Intelligence (AI) technology. The usage of computer-based learning methods developed a significant research field in PCa. Generally, gene expression data contains large amount of genes; some authors evaluated and analyzed the cancer classification problems by utilizing different machine learning (DL), data mining (DM), and statistical based algorithms [8]. Several ML methods have attained lot of success better and classification performance in the cancer classification [9]. But, still, there are few problems with this approach which makes the cancer classification nontrivial tasks [10]. A disadvantage of conventional ML approach is needing preengineered organization of new input data as to structured data sets. The DL approach is a field of ML that employed layered structure for building sophisticated modules with the capacity to understand complex information [11]. This capability allows DL algorithms to demonstrate conventional ML techniques from multiple domains such as speech recognition computer vision, image classification, and so on.

The gene expression data comprises many redundant, noisy, and irrelevant items. The informative ratio to noisy data is 1 : 10 which degrade the performances of clustering when traditional approaches are employed directly to the comprehensive feature set. Hence, the informative feature selection (FS) technique plays an important role in higher-dimension gene expression data for biological data retrieval [12]. The FS method is separated into two classes. The initial class includes semisupervised, supervised, and unsupervised methods based on availability of historical data. The next class comprises ensemble, filter, wrapper, embedded, and hybrid approaches based on how they concatenate the selection by modeling. Each of these approaches has its disadvantages and advantages. In general, the hybrid approach is superior to the wrapper methods since it is less prone to overfitting. But the ensemble methodology is very flexible and robust [13]. The large dimension of gene expression data includes irrelevant, noise, and redundant items which makes it hard to examine. In this study, the FS methods are employed to lower the dimension of information for analysis of gene expression. Previously, the evolutionary learning method has been used effectively in distinct microarray researches, for example, to select informative subset of genes, for biclustering and sample, and clustering classification.

This article develops an artificial intelligence based feature selection with deep learning model for prostate

cancer detection (AIFSDL-PCD) using microarray gene expression data. The AIFSDL-PCD technique derives a chaotic invasive weed optimization (CIWO) based FS technique for choosing an optimal subset of features. In addition, the deep neural network (DNN) model can be applied to prostate cancer classification utilizing the microarray gene expression data. Besides, the hyperparameters of the DNN model can be effectively adjusted by the use of RMSprop optimizer. For examining the betterment of the AIFSDL-PCD technique, a comprehensive experimental analysis is carried out and the results are examined under several aspects.

The rest of the study is planned as follows. Section 2 offers the literature review, Section 3 presents the proposed model, Section 4 elaborates the performance validation, and Section 5 draws the conclusion.

2. Literature Review

Tavasoli et al. [14] presented a classification technique which employed metaheuristic and SVM algorithms. The optimization of the SVM hyperparameters for the RBF is implemented by utilizing the modified Water Cycle Algorithm (mWCA). The result indicates that the ensemble performance of gene-mWCA SVM (EGmWS) was regarded as effective methodology compared to related methodologies in terms of accuracy and solving the uncertainty problems. Elmarakeby et al. [15] designed a P-NET—a biologically informed DL method—for stratifying patients with PCa by treatment resistance state and gauging molecular driver of treatment resistance to therapeutic target via method interpretability. They demonstrated that P-NET could forecast cancer state by utilizing molecular information with performances, i.e., better than other modeling techniques.

Glaab et al. [16] estimated a rule-based evolutionary ML method, GAssist, and BioHEL, on three public microarray cancer data sets, attaining simple rule-based model for sample classifier. Compared to other standards of microarray, sample classification depends on three different FS methods. Darendeli et al. [17] focused on providing different perspectives of cancer diagnoses with DL method on gene expression data. In this work, RNA-Seq data of around thirty distinct kinds of cancer patients and the normal tissue RNA-Seq data from GTEx and Cancer Genome Atlas (TCGA) have been employed. The input data for the training was converted into RGB formats and the training was performed by a CNN approach.

Nirmalakumari et al. [18] focused on classifying the PCa in an accurate manner. Open-source two-class prostate data which contains 136 samples and 12,600 genes are taken into account. At first, PCA and Kruskal-Wallis test are employed to determine the informative genes. Next, they are categorized by utilizing LDA, SVM, XGB, and KNN classification to classify prostate patients as normal or abnormal. Ahn et al. [19] aimed at addressing how far the DL method could learn for recognizing cancer. They incorporated gene expression data from the GEO, TCGA, TARGET, and GTEx database including 12,842 normal gene expression data and 13,406 cancer from twenty-four distinct tissues. First, a

DNN system is trained for identifying normal and cancer samples with different gene selection approach. Al-Obeidat et al. [20] introduce gene encoder, an unsupervised 2-phase FS method for the classification of cancer sample. Initially, they aggregate three filter methodologies, such as spectral-based FS, PCA, and correlation methods. Then, the GA approach is utilized that estimates the chromosome using the AE-based clustering. The resulting feature subsets are utilized for classifier process.

3. The Proposed Model

In this study, a new AIFSDL-PCD technique has been developed for the detection and classification of PCa. The proposed AIFSDL-PCD technique incorporates different processes, namely, preprocessing, CIWO based FS, DNN based classification, and RMSprop based hyperparameter tuning. The application of CIWO based FS technique helps for reducing the computational complexity and improving the classification accuracy. Figure 1 illustrates the overall working process of AIFSDL-PCD technique.

3.1. Data Preprocessing. The presented work utilizes the preprocessed step as a huge volume of biological information has high level of noise as well as bias. So, the gene term dataset needs the subsequent more than one preprocessed step previously executing design investigation [21].

- (i) The gene expression data demonstrate skewed distributions where lower stated genes were among zero as well as one, but the extreme term genes are among one as well as infinity. Thus, once a parametric statistical test was implemented for such asymmetric data, at the end outcome is from biased result. For overcoming this challenge, the log transformation was utilized for making the data further symmetric that is anticipated for giving an accurate outcome under statistical tests.
- (ii) The replicate of handling look at the repeated gene identify from a dataset that is afterward exchanged by its average value, so extracting the unpredictable repetitions.
- (iii) This design standardized was utilized that removes the scale variance among the features by subtracting the instance average and dividing the value by standard deviation (SD).
- (iv) The occurrence of missing value of gene term has allowed for average form.
- (v) The flat pattern filter was utilized which removes genes for reducing the difficulty of dataset which is employed to biological significant study.

3.2. Design of CIWO-Based Feature Selection Technique. At this stage, the preprocessed data is passed as input to CIWO technique for the optimum selection of feature subsets. The IWO technique is stimulated by the procedure of adaptability, reproduction, and existence [22].

Accordingly, weeds represent unwanted plants which have aggressive behaviour for growth and are threats to another crop and prevent them from growing. This approach is fast, simple, and highly efficient in detecting the optimum point. Indeed, this method is depending on the natural features of weeds like struggle for existence, seed production, and growth. The description of IWO approach is given in the following:

- (i) The evaluation of objective function and the production of arbitrary population initialization (seed distribution) from chosen domain are done, so that an initial population from the problem solving domains are distributed randomly and estimated.
- (ii) Reproduction depends on upgraded SD and competency. All the members of population, based on their capacity, yield seeds according to the maximum and minimum competence among the two pre-determined quantities.

The amount of seeds that every plant could yield linearly differs in the small amount of seeds to the maximal number (S_{\min} ; S_{\max}). The amount of seeds generated near every weed is defined by the following equation:

$$\text{Seed}_i = \text{Round} \left\{ S_{\min} + (S_{\max} - S_{\min}) \times \frac{N_{\text{weed}} - \text{rank}_i}{N_{\text{weed}} - 1} \right\}, \quad (1)$$

where rank_i represents the rank of i seed, Round denotes the function to iteration number, N_{weed} indicates the amount of initial weeds, S_{\max} and S_{\min} signify the least and most seeds which are generated near every weed, correspondingly, and Seed_i implies the amount of seeds generated near i th weed. The seed generated in the searching space is distributed arbitrarily in the problem space with standard distribution (predefined variance and average of zero); the seed is dispersed near to its parent (weeds). The values of SD (r_{iter}) reduce nonlinearity in all iterations in the first value (r_{initial}) to the last values (r_{final}) as follows. For example, the closer we get to the end of the process, the further the seeds are produced near the answer attained and the less distributed they are than at the beginning of the process.

$$\sigma_{\text{iter}_i} = \left(\frac{\text{max_iter} - \text{iter}_i}{\text{max_iter}} \right)^n (\sigma_{\text{initial}} - \sigma_{\text{final}}) + \sigma_{\text{final}}. \quad (2)$$

In equation (2), max_iter denotes the maximal amount of iterations, iter_i indicates i^{th} iteration, n represent the nonlinear coefficient, and σ_{iter_i} indicates the SD of i^{th} iteration. When the weed does not reproduce, it would pass away. Hence, competition among weeds is required for limiting the maximal amount. Assuming that, after many stages of iteration, the amount of seeds owing to reproduction rises, an algorithm must be determined for controlling the entire amount of them. Once the maximal amount of allowed seeds (P_{max}) is attained, the weaker seeds must be removed; thus the seed population remains at the maximal number (P_{max}). This procedure is repeated till the plant reaches the optimal by checking the end condition.

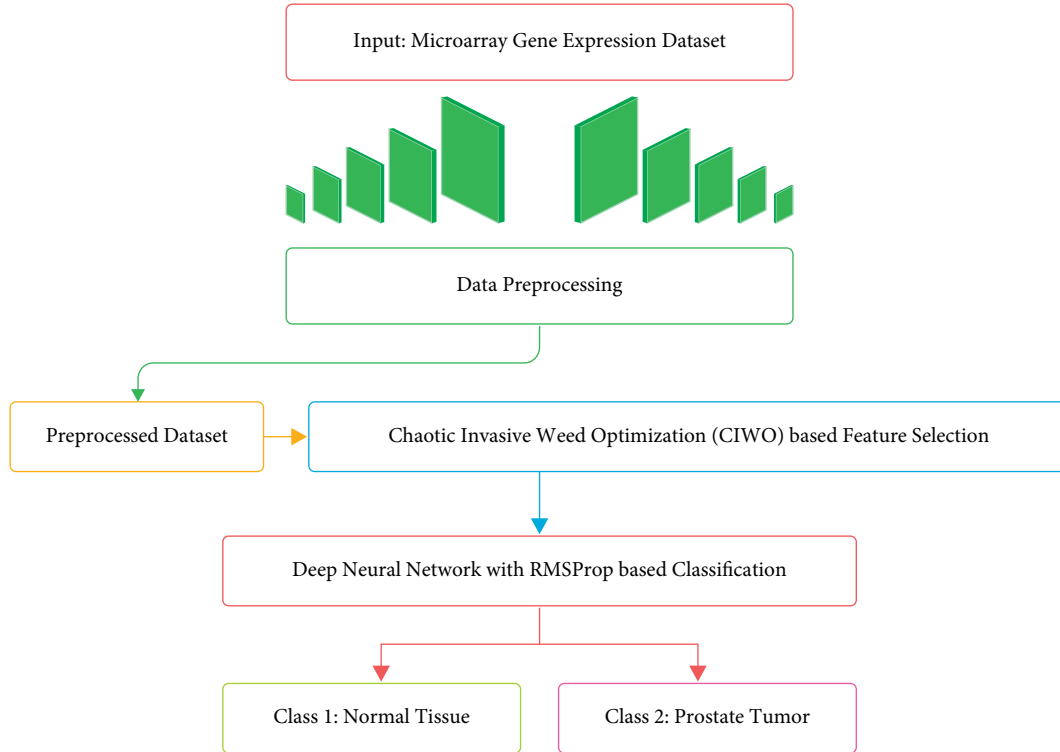


FIGURE 1: Overall process of AIFSDL-PCD technique.

To improve the efficiency of the IWO algorithm, the CIWO algorithm has been derived by the integration of chaos theory. Chaos is a widespread nonlinear phenomenon by its nature and is a feature of randomness, ergodicity, sensitivity to primary states, etc. [23]. Because of the features of ergodicity and randomness, chaotic motion traverses each state from particular range based on its individual law without repetition. So, when it can be utilized with chaos variables for searching optimum, it undoubtedly has further benefits to arbitrary searches. The chaos ergodicity feature was utilized for optimizing the search and avoiding fall as to local minima; so, chaos optimized search technique developed a new optimized approach. The chaotic orders created by distinct mappings are utilized as tent map, sinusoidal map, logistic map, singer map, and sine map. Many chaotic maps are tried and an optimum one is selected for combining with IWO technique. Because of the primary testing, logistic map attained optimum outcomes. Therefore, the chaotic orders were created by utilizing logistic map as

$$x_{i+1} = ux_i(1 - x_i), \quad (3)$$

where u refers to the control parameter and assumes $u = 4$. When $u = 4$, the logistic mapping derives as to detailed chaotic state. Assume $x_i \in (0, 1)$ and $x_i \neq 0.25, 0.5, 0.75$.

The preliminary weed population $Seed_i$ is mapped to chaotic order which is created based on (3), resulting in equivalent chaotic seed population pch .

$$pch = x_i * Seed_i. \quad (4)$$

During the IWO based FS process, when the feature vector size is N , the number of possible feature arrangements is found to be 2^N , which is massive. The IWO algorithm looks for the optimal subset of features in the search space. Algorithm 1 shows the pseudocode of IWO algorithm.

The FS problem can be considered as a multiobjective issue which aims for reducing the number of chosen features and increasing the classification accuracy. Therefore, the fitness function of the IWO algorithm can determine the solutions constructed to maintain a tradeoff among two objectives.

$$\text{fitness} = \alpha \Delta_R(D) + \beta \frac{|Y|}{|T|}, \quad (5)$$

where $\Delta_R(D)$ denotes the error rate of the classification model, $|Y|$ indicates the number of features chosen by the IWO algorithm, and $|T|$ represents the available set of features that exist in the present dataset.

3.3. Design of Optimal DNN-Based Classification Model.

During classification process, the chosen subset of features is passed into the DNN model for PCa detection. The DNN is a version of MLP and that is kind of FFNN with two or more layers with 1 input, 1 output layer, and one or more hidden layers. All layers have many neurons and FC with neurons from forwarding direction [24]. The model is mathematically determined as $O: \mathbb{R}^m \times \mathbb{R}^n$. An input vector $x = x_1, x_2, x_3, \dots, x_m$ and their size is 'm' and resultant

```

Begin {
  (i) Parameter Initiation;
  (ii) Present_iterat = 1;
  (iii) While (Present_iterat < Maxm_iterat) do
  (iv) {
  (v) Determine the optimal and poor fitness in the population
  (vi) Determine the standard deviation  $st$   $d$  based on rounds
  (vii) For every weed  $w$  in the population  $W$ 
  (viii) {
  (ix) Determine the seed count  $w$  based on the fitness
  (x) Elect the seed among the possible solution surrounding the parent weed  $w$  in a neighboring area with uniform
  distribution of mean =  $O$  and standard deviation =  $std$ ;
  (xi) Append seeds created to the population  $W$  If ( $|W|$  Maxm_Size_Population)
  (xii) {
  (xiii) Arrange population  $W$  based on the fitness
  (xiv)  $W = \text{Select\_Better}(\text{weed}, \text{seed}, \text{Maxm\_Size\_Population})$ 
  (xv) } End if
  (xvi) } End for
  (xvii) Present_iterat = Present_iterat + 1;
  (xviii) } End while
  (xix) }
  (xx) End

```

ALGORITHM 1: Pseudocode of IWO algorithm.

vector has $O(x)$ and their size 'n'. The calculation of all hidden layers h_j is determined mathematically as

$$h_j(x_j^{l+1}) = f(Z_{ij} + b_j^{(l+1)}), \quad (6)$$

$$Z_{ij} = x_i^l w_{ij}^{(l,l+1)}. \quad (7)$$

Every lower layer neuron individual is linked to neuron j . In equations (6) and (7), $x_i^{(l)}$ has neuron i activation function at layer l and Z_{ij} refers to the influence of neuron i at layer l to activation of neuron j at layer $l + 1$. The function f refers to the nonlinear activation function, $w_{ij}^{(l,l+1)}$ implies the weight, and b_j^{l+1} represents the bias of neuron j . This technique utilizes softmax function as nonlinear activation function to multiclass classifier. In several stacking hidden layers MLP has been named DNN. Generally, the DNN with several hidden layers is expressed as

$$H_l(x) = H_l(H_{l-1}(H_{l-2}(\dots(H_1(x)))). \quad (8)$$

The DNN framework has 2 hidden layers. It gets inputs $x = x_1, x_2, x_3, \dots, x_m$ and outputs were $o = o_1, o_2, \dots, o_{c-1}, o_c$. Figure 2 showcases the framework of DNN.

Further advanced typical feedforward network DNN can be utilized with all the hidden layers having ReLU nonlinear activation functions. It is used for decreasing the state of vanishing and error gradient problems [24] and is related to another nonlinear activation function ReLU which is quicker and simpler for training the technique with huge hidden layer.

The loss function has optimum parameters that can be vital for achieving higher efficiency. The target and forecast

values variance was computed as utilizing loss function. It could be defined as

$$d(t, p) = \|t - p\|_2^2. \quad (9)$$

It attempts for learning an estimate to identify the function, with the learning procedure explained as minimizing reform error as illustrated in equation (9), where t and p refer to the target as well as forecasted values correspondingly. The loss function is used for identifying that forecasted value diverges in the target value. The target is fed to model along with features for calculating the loss function and classifying the attack. The negative log probability with t and probability distribution $p(pd)$ are utilized to target and forecast classes correspondingly from multiclass classifier. It could be written as

$$d(t, p(p, d)) = -\log p(pd)_t. \quad (10)$$

To effectually tune the hyperparameters of the DNN model, the RMSprop optimizer is utilized. RMSprop is the enhancement form of Adagrad; the upgrade procedure of RMSprop is the same as Adagrad [25]. For RMSprop, an exponentially decaying average of squared gradient is computed initially.

$$\begin{aligned} G_t &= \beta G_{t-1} + (1 - \beta) g_t \odot g_t \\ &= (1 - \beta) \sum_{\tau=1}^t \beta^{t-\tau} g_\tau \odot g_\tau, \end{aligned} \quad (11)$$

where β refers to the decay rate that is generally offered which is fixed to 0.9. And the upgrade value of parameters from RMSprop is similar to Adagrad:

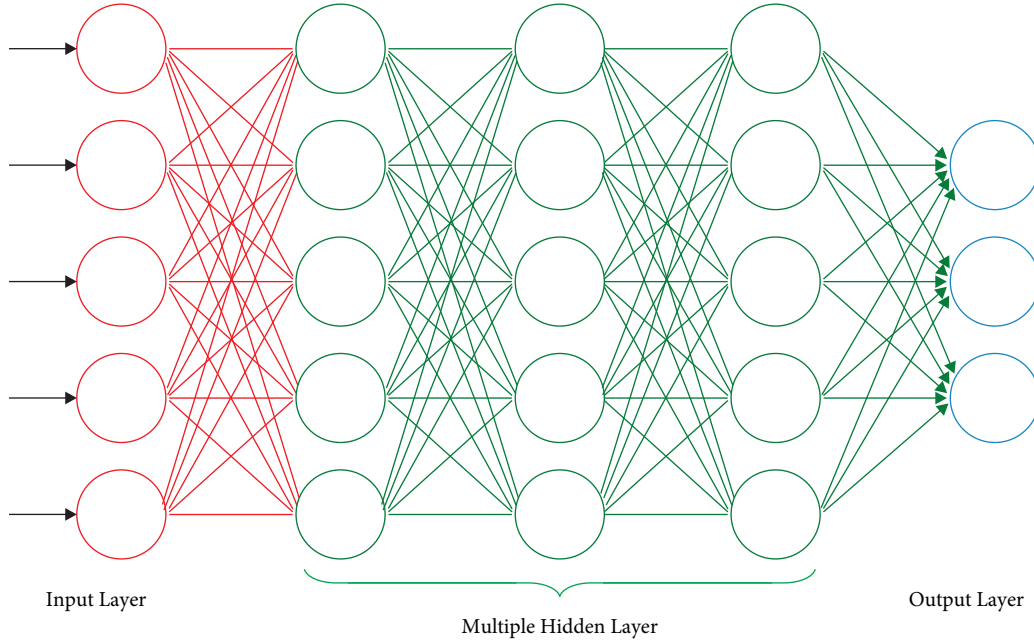


FIGURE 2: DNN structure.

$$\Delta\theta_t = -\frac{\alpha}{\sqrt{G_t + \varepsilon}} \odot g_t. \quad (12)$$

Also, the simplified concept of Adagrad technique is implemented. g'_t is explained as

$$g'_t = \frac{1}{\sqrt{G_t + \varepsilon}} \odot g_t, \quad (13)$$

and the upgrade value of RMSprop has been determined as

$$\Delta\theta_t = -\alpha g'_t. \quad (14)$$

So, the RMSprop is an optimized technique dependent upon gradient actually. To provide analysis, the rate of learning optimized technique was utilized for improving the trained efficiency.

4. Results and Discussion

The performance validation of the AIFSDL-PCD technique takes place using an open access dataset, including 102 tissue instances (52 prostate tumors and 50 normal tissues) with 2135 genes. The proposed model is simulated using Python 3.6.5 tool. Table 1 and Figure 3 illustrate the result analysis of the optimal DNN model under ten iterations. The results exhibited that the optimal DNN algorithm has accomplished satisfactory outcomes. For instance, under iteration 1, the optimal DNN model has provided sens_y , spec_y , prec_n , acc_y , and F_{score} of 96.30%, 95.56%, 96.67%, 96.64%, and 96.32%, respectively.

In line with this, under iteration 4, the optimal DNN method has provided sens_y , spec_y , prec_n , acc_y , and F_{score} of 96.13%, 96.34%, 96.15%, 96.19%, and 96.19% correspondingly. Meanwhile, under iteration 6, the optimal DNN approach has offered sens_y , spec_y , prec_n , acc_y , and F_{score} of

TABLE 1: Result analysis of optimal DNN model

No. of iterations	Sensitivity	Specificity	Precision	Accuracy	F_{score}
Iteration 1	96.30	95.56	96.67	96.64	96.32
Iteration 2	96.20	96.46	96.57	96.50	96.97
Iteration 3	95.82	96.64	96.55	95.99	95.95
Iteration 4	96.13	96.34	96.15	96.19	96.19
Iteration 5	96.25	95.66	96.75	96.51	95.63
Iteration 6	95.59	95.63	96.55	95.86	95.53
Iteration 7	95.92	96.04	96.11	95.99	96.04
Iteration 8	95.56	96.88	96.34	95.72	96.43
Iteration 9	96.17	95.57	96.53	96.31	96.27
Iteration 10	96.44	96.18	96.15	96.38	96.05
Average	96.04	96.10	96.44	96.21	96.14

95.59%, 95.63%, 96.55%, 95.86%, and 95.53%, respectively. Eventually, under iteration 8, the optimal DNN technique has showed sens_y , spec_y , prec_n , acc_y , and F_{score} of 95.56%, 96.88%, 96.34%, 95.72%, and 96.43% correspondingly. At last, under iteration 10, the optimal DNN methodology has provided sens_y , spec_y , prec_n , acc_y , and F_{score} of 96.44%, 96.18%, 96.15%, 96.38%, and 96.05% correspondingly.

The ROC analysis of the optimal DNN approach is implemented in Figure 4. The figure displayed that the optimal DNN approach has accomplished optimum ROC classification performance with the increased ROC of 99.3002.

Table 2 and Figure 5 showcase the result analysis of the AIFSDL-PCD approach under ten iterations. The outcomes showed that the AIFSDL-PCD technique has accomplished satisfactory outcomes. For instance, under iteration 1, the AIFSDL-PCD algorithm has provided sens_y , spec_y , prec_n , acc_y , and F_{score} of 97.75%, 97.26%, 96.87%, 97.47%, and 97.58% correspondingly. Likewise, under iteration 4, the

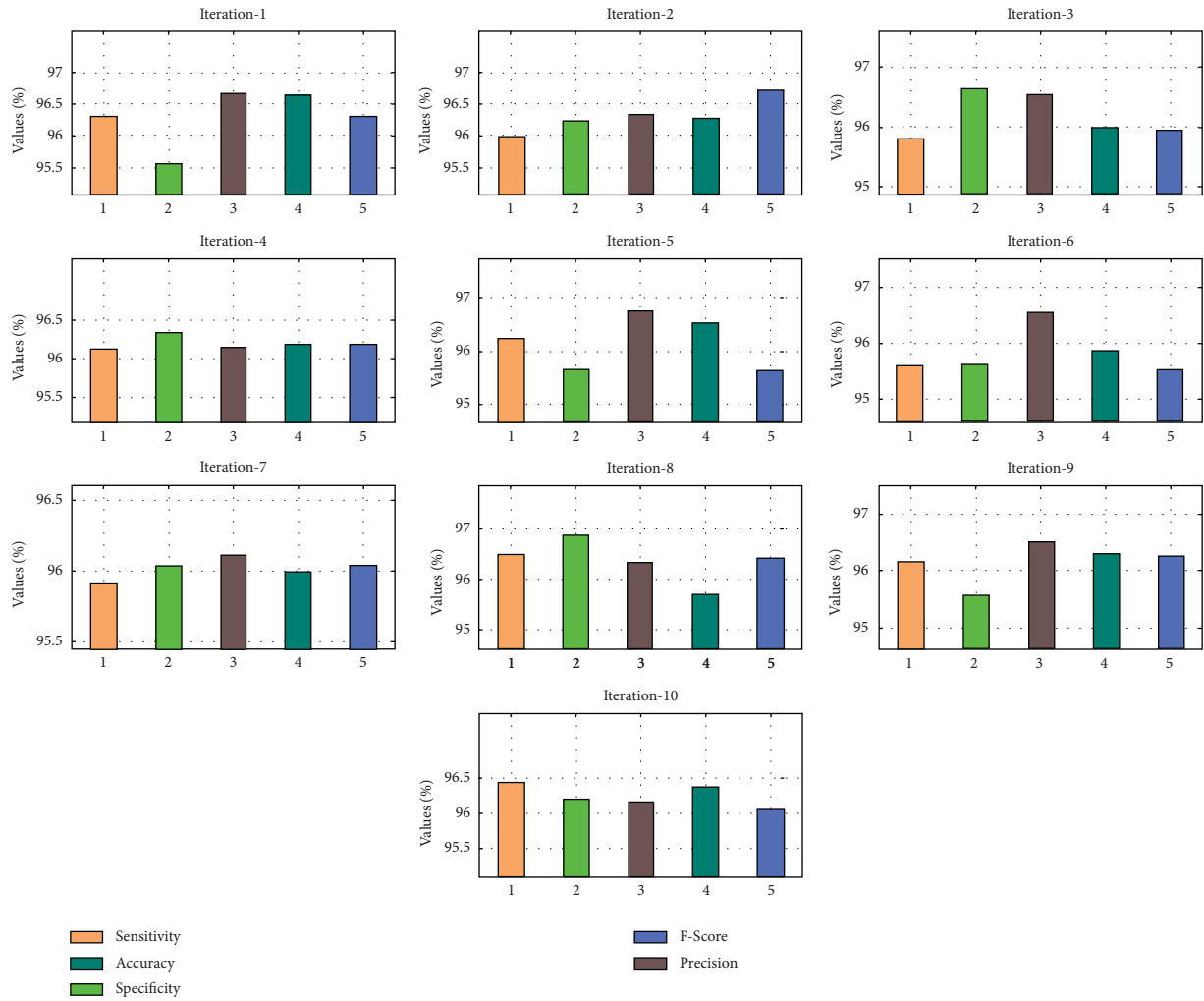


FIGURE 3: Result analysis of optimal DNN technique.

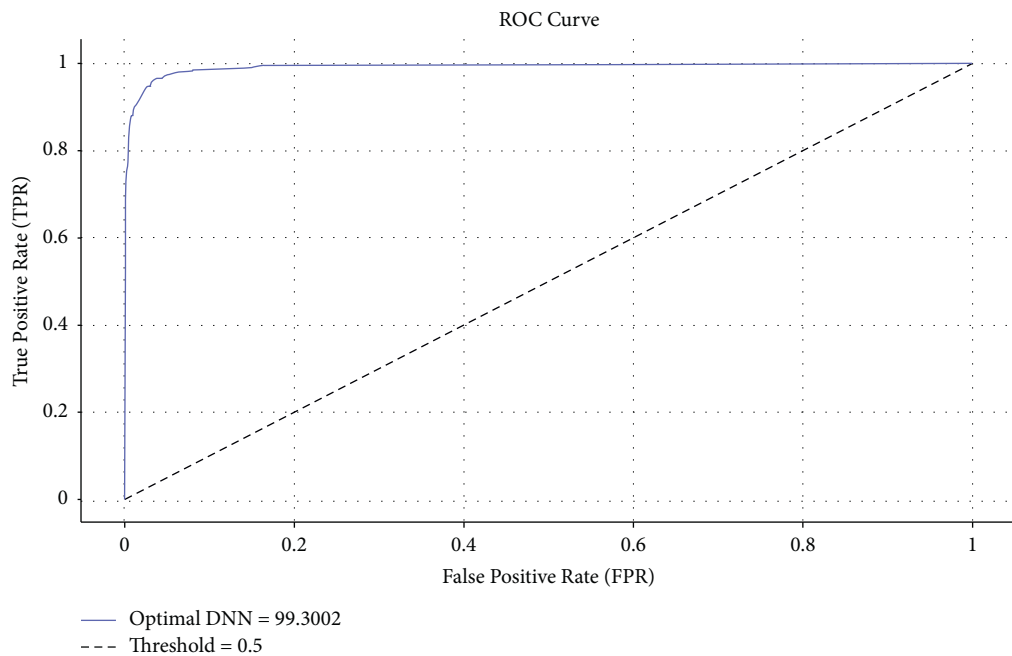


FIGURE 4: ROC analysis of optimal DNN technique.

TABLE 2: Result analysis of proposed AIFSDL-PCD model.

No. of iterations	Sensitivity	Specificity	Precision	Accuracy	F -score
Iteration 1	97.75	97.26	96.87	97.47	97.58
Iteration 2	97.25	97.30	96.69	96.83	97.48
Iteration 3	97.59	97.21	97.34	97.41	97.06
Iteration 4	97.49	97.10	96.92	97.18	97.07
Iteration 5	97.11	96.60	96.90	96.69	97.22
Iteration 6	97.43	96.92	97.39	97.27	96.89
Iteration 7	96.87	97.27	97.34	97.06	97.75
Iteration 8	97.18	97.37	97.43	97.34	97.71
Iteration 9	96.51	97.78	97.25	97.41	97.51
Iteration 10	97.28	97.66	97.23	97.28	96.51
Average	97.25	97.25	97.14	97.19	97.28

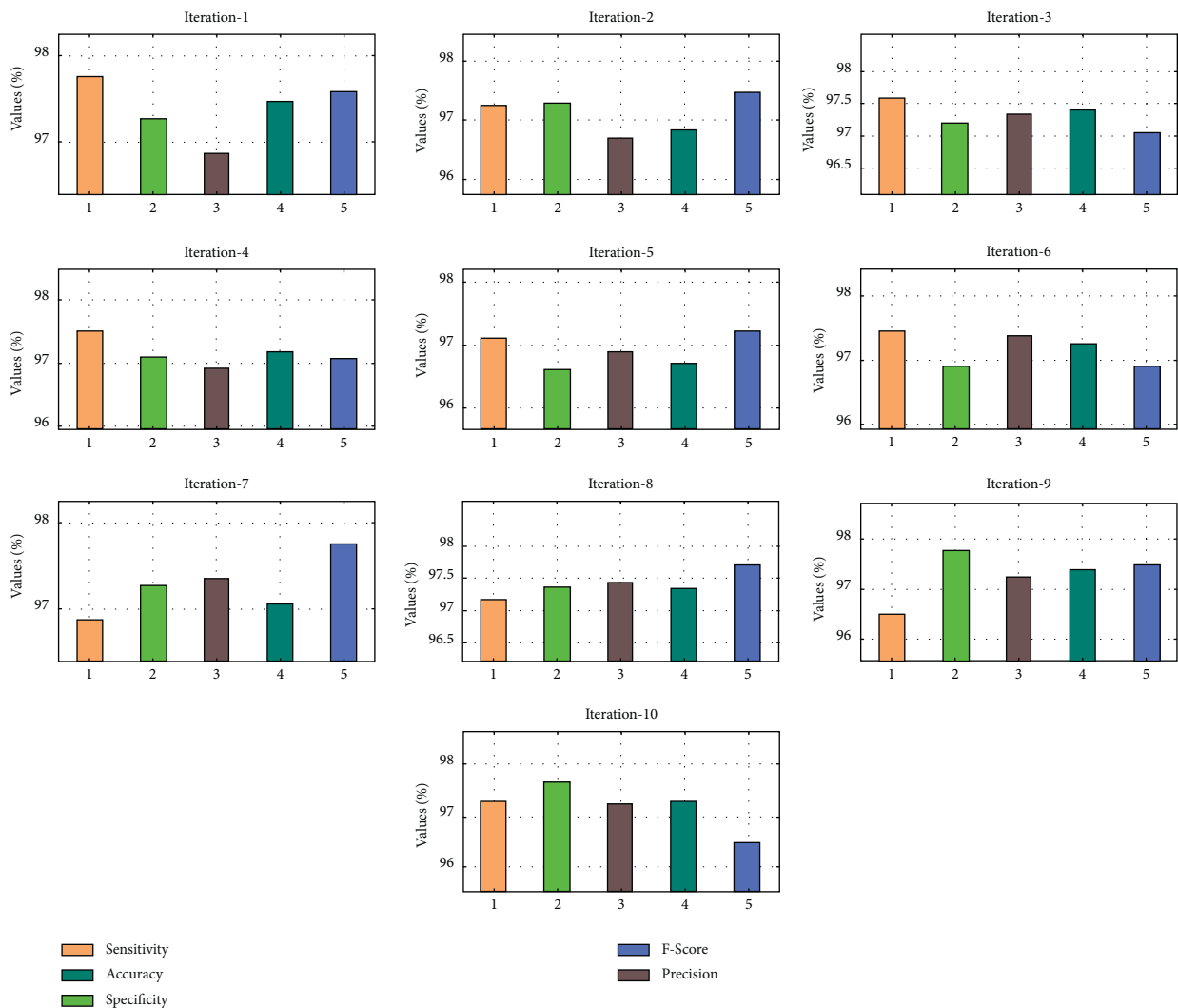


FIGURE 5: Result analysis of AIFSDL-PCD approach.

AIFSDL-PCD technique has given $sens_y$, $spec_y$, $prec_n$, acc_y , and F_{score} of 97.49%, 97.10%, 96.92%, 97.18%, and 97.07% correspondingly. In the meantime, under iteration 6, the AIFSDL-PCD model has provided $sens_y$, $spec_y$, $prec_n$, acc_y , and F_{score} of 97.43%, 96.92%, 97.39%, 97.27%,

and 96.89%, respectively. Eventually, under iteration 8, the AIFSDL-PCD approach has offered $sens_y$, $spec_y$, $prec_n$, acc_y , and F_{score} of 97.18%, 97.37%, 97.43%, 97.06%, and 97.75% correspondingly. At last, under iteration 10, the AIFSDL-PCD model has provided $sens_y$, $spec_y$, $prec_n$,

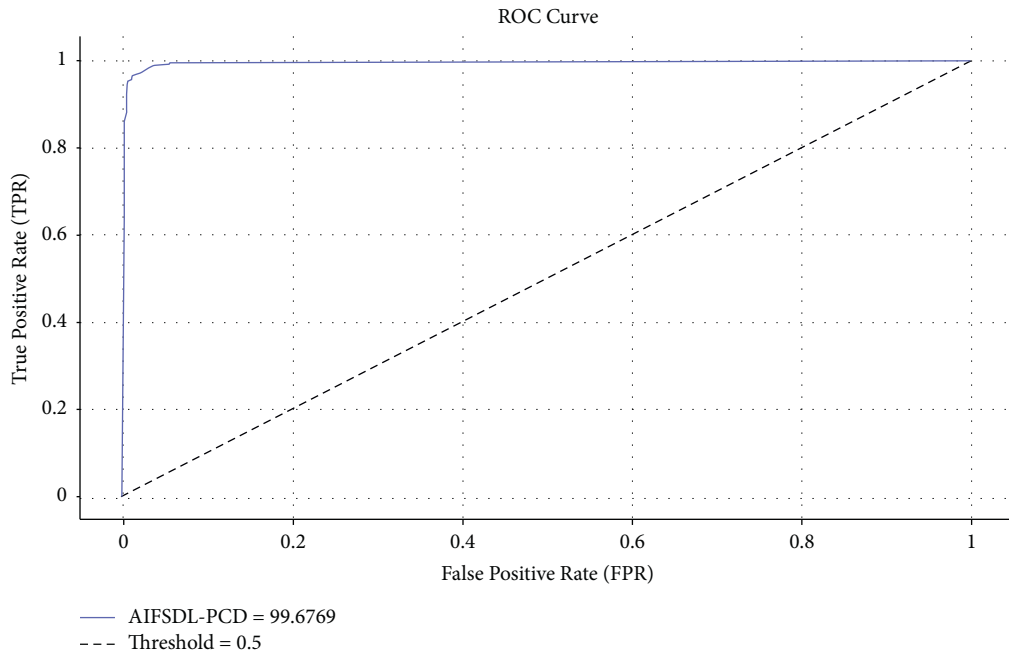


FIGURE 6: ROC analysis of AIFSDL-PCD technique.

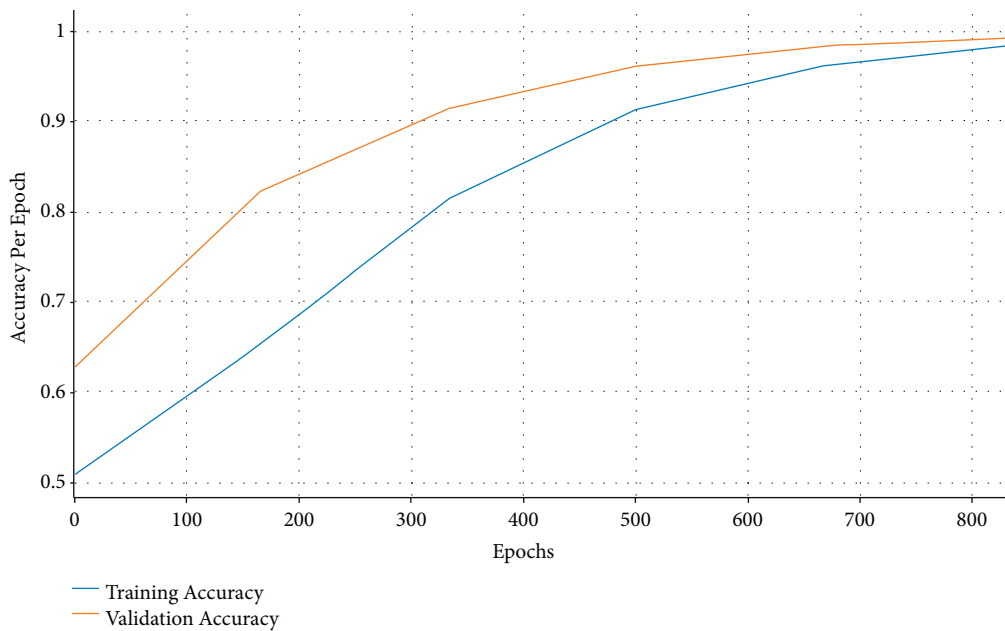


FIGURE 7: Accuracy graph analysis of AIFSDL-PCD technique.

acc_y , and F_{score} of 97.28%, 97.66%, 97.23%, 97.28%, and 96.51% correspondingly.

The ROC analysis of the AIFSDL-PCD technique is performed in Figure 6. The figure exhibited that the AIFSDL-PCD technique has accomplished better ROC classification performance with a maximum ROC of 99.6769.

Figure 7 demonstrates the accuracy analysis of AIFSDL-PCD technique on the test dataset. The outcomes exhibited that the AIFSDL-PCD system has accomplished increased performance with improved training and validation

accuracy. It can be clear that the AIFSDL-PCD methodology has reached enhanced validation accuracy on the training accuracy.

Figure 8 depicts the loss analysis of the AIFSDL-PCD approach on the test dataset. The outcomes recognized that the AIFSDL-PCD methodology has resulted in a proficient outcome with lesser training and validation loss. It can be obvious that the AIFSDL-PCD algorithm has obtainable lesser validation loss on the training loss.

To portray the better classification performance of the AIFSDL-PCD method, a comparative acc_y analysis is

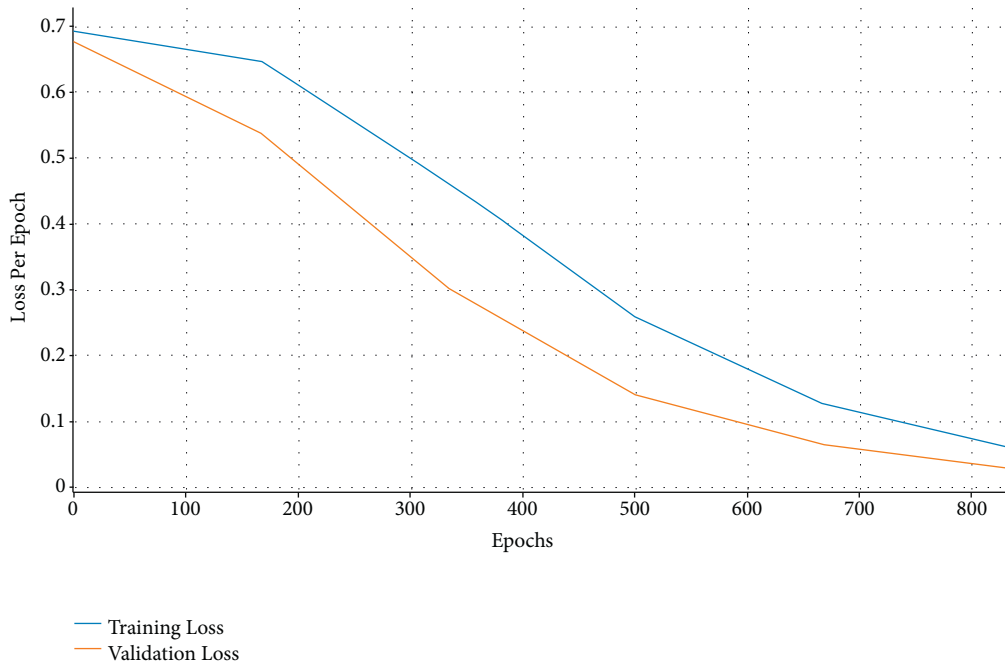


FIGURE 8: Loss graph analysis of AIFSDL-PCD technique.

TABLE 3: Comparative analysis of AIFSDL-PCD approach with existing techniques.

Methods	Accuracy
PLR-MC	0.9460
RFLD-MC	0.9340
Bio-HEL	0.9400
SVM model	0.9120
GA-KNN + SVM	0.8571
CSF-RC	0.9510
Optimal DNN	0.9621
AIFSDL-PCD	0.9719

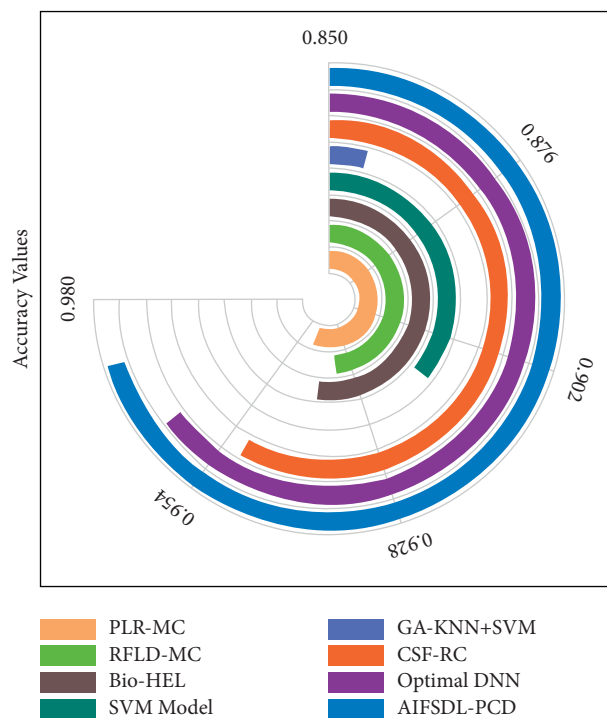


FIGURE 9: Accuracy analysis of AIFSDL-PCD technique with existing manners.

made in Table 3 and Figure 9 [26, 27]. The results show that the GA-KNN+SVM model has failed to achieve proficient classification performance. At the same time, the PLR-MC, RFLD-MC, and Bio-HEL techniques have accomplished moderately closer accuracy values. Along with that, the CSF-RC and optimal DNN techniques have managed to demonstrate reasonable accuracy values. However, the AIFSDL-PCD technique has resulted in superior performance with higher accuracy of 0.9719. From the aforementioned tables and figures, it can be obvious that the AIFSDL-PCD method is found to be an effective tool for PCa detection and classification.

5. Conclusion

In this study, a new AIFSDL-PCD method has been developed for the detection and classification of PCa. The proposed AIFSDL-PCD technique incorporates different processes, namely, preprocessing, CIWO based FS, DNN based classification, and RMSprop based hyperparameter tuning. The application of CIWO based FS technique helps for reducing the computational complexity and improves the classification accuracy. For examining the betterment of the AIFSDL-PCD technique, a comprehensive experimental analysis is carried out and the results are examined under several aspects. The experimental results reported the supremacy of the AIFSDL-PCD technique over the other techniques in terms of different measures. Therefore, the AIFSDL-PCD technique can be applied as a proficient tool for the detection and classification of PCa. As a part of future extension, hybrid DL based classifiers with metaheuristics based hyperparameter optimizers can be developed to boost the PCa detection results.

Data Availability

Data sharing is not applicable to this article as no datasets were generated during the current study.

Ethical Approval

This article does not contain any studies with human participants performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under grant number RGP 2/71/43. Princess Nourah bint Abdulrahman University Researchers

Supporting Project number PNURSP2022R77, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication.

References

- [1] World Health Organization, "Cancer," 2018, <https://www.who.int/news-room/fact-sheets/detail/cancer>.
- [2] M. T. Arslan and A. Kalinli, "A comparative study of statistical and artificial intelligence based classification algorithms on central nervous system cancer microarray gene expression data," *Int J Intell Syst Appl Eng*, vol. 4, 2016.
- [3] H. Dhahri, E. Al Maghayreh, A. Mahmood, W. Elkilani, and M. Faisal Nagi, "Automated breast cancer diagnosis based on machine learning algorithms," *Journal of healthcare engineering*, vol. 2019, Article ID 4253641, 11 pages, 2019.
- [4] A. Khare, M. Jeon, I. K. Sethi, and B. Xu, "Machine learning theory and applications for healthcare," *Journal of healthcare engineering*, vol. 2017, Article ID 5263570, 2 pages, 2017.
- [5] P. Danaee, R. Ghaeini, and D. A. Hendrix, "A deep learning approach for cancer detection and relevant gene identification," in *Pacific symposium on biocomputing* vol. 22, 2017.
- [6] S. M. D. A. C. Jayatilake and G. U. Ganegoda, "Involvement of machine learning tools in healthcare decision making," *Journal of Healthcare Engineering*, vol. 2021, Article ID 6679512, 20 pages, 2021.
- [7] S. Hamena and S. Meshoul, "Multi-class classification of gene expression data using deep learning for cancer prediction," *Int J Mach Learn Comput*, vol. 8, no. 5, pp. 454–459, 2018.
- [8] M. S. Iqbal, I. Ahmad, L. Bin, S. Khan, and J. J. Rodrigues, "Deep learning recognition of diseased and normal cell representation," *Trans Emerg Telecommun Technol*, vol. 32, 2020.
- [9] S. H. Shah, M. J. Iqbal, I. Ahmad, S. Khan, and J. J. Rodrigues, "Optimized gene selection and classification of cancer from microarray gene expression data using deep learning," *Neural Computing & Applications*, pp. 1–12, 2020.
- [10] N. Kumar, N. Narayan Das, D. Gupta, K. Gupta, and J. Bindra, "Efficient automated disease diagnosis using machine learning models," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9983652, 13 pages, 2021.
- [11] M. Jansi Rani and D. Devaraj, "Two-stage hybrid gene selection using mutual information and genetic algorithm for cancer data classification," *Journal of Medical Systems*, vol. 43, no. 8, p. 235, 2019.
- [12] V. Bolo'n-Canedo, N. Sa'nchez-Marono, A. Alonso-Betanzos, J. M. Beni'tez, and F. Herrera, "A review of microarray datasets and applied feature selection methods," *Information Science*, vol. 282, pp. 111–135, 2014.
- [13] M. Ghosh, S. Adhikary, K. K. Ghosh, A. Sardar, S. Begum, and R. Sarkar, "Genetic algorithm based cancerous gene identification from microarray data using ensemble of filter methods," *Medical, & Biological Engineering & Computing*, vol. 57, no. 1, pp. 159–176, 2019.
- [14] N. Tavasoli, K. Rezaee, M. Momenzadeh, and M. Sehati, "An ensemble soft weighted gene selection-based approach and cancer classification using modified metaheuristic learning," *Journal of Computational Design and Engineering*, vol. 8, no. 4, pp. 1172–1189, 2021.
- [15] H. A. Elmarakeby, J. Hwang, R. Arafeh et al., "Biologically informed deep neural network for prostate cancer discovery," *Nature*, vol. 598, no. 7880, pp. 348–352, 2021.

- [16] E. Glaab, J. Bacardit, J. M. Garibaldi, and N. Krasnogor, "Using rule-based machine learning for candidate disease gene prioritization and sample classification of cancer gene expression data," *PLoS One*, vol. 7, no. 7, Article ID e39932, 2012.
- [17] B. N. Darendeli and A. Yilmaz, "Convolutional neural network approach to predict tumor samples using gene expression data," *Journal of Intelligent Systems: Theory and Applications*, vol. 4, no. 2, pp. 136–141, 2021.
- [18] K. Nirmalakumari, H. Rajaguru, and P. Rajkumar, "Microarray prostate cancer classification using eminent genes," in *Proceedings of the 2021 Smart Technologies, Communication and Robotics (STCR)*, pp. 1–5, IEEE, Sathyamangalam, India, October 2021.
- [19] T. Ahn, T. Goo, C. H. Lee et al., "Deep learning-based classification and interpretation of gene expression data from cancer and normal tissues," *International Journal of Data Mining and Bioinformatics*, vol. 24, no. 2, pp. 121–139, 2020.
- [20] F. Al-Obeidat, A. Tubaishat, B. Shah, and Z. Halim, "Gene encoder: a feature selection technique through unsupervised deep learning-based clustering for large gene expression data," *Neural Computing & Applications*, pp. 1–23, 2020.
- [21] T. Muhammad and Z. Halim, "Employing artificial neural networks for constructing metadata-based model to automatically select an appropriate data visualization technique," *Applied Soft Computing*, vol. 49, pp. 365–384, 2016.
- [22] M. Khajenoori, J. Safdari, S. Yousefi-Nasab, M. H. Mallah, J. Karimi-Sabet, and M. H. Askari, "Optimization of a conical cascade using invasive weed optimization (IWO) algorithm for multi-component systems: investigation of the effect of feed flow on the separation factor of centrifuge machines in the cascade by DSMC method," *Annals of Nuclear Energy*, vol. 162, Article ID 108497, 2021.
- [23] Z. Cai, J. Gu, C. Wen et al., "An intelligent Parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach," *Computational and mathematical methods in medicine*, vol. 2018, Article ID 2396952, 24 pages, 2018.
- [24] K. Narayana Rao, K. Venkata Rao, and P. R. P.V.G.D., "A hybrid intrusion detection system based on sparse autoencoder and deep neural network," *Computer Communications*, vol. 180, pp. 77–88, 2021.
- [25] C. Zhang, M. Yao, W. Chen, S. Zhang, D. Chen, and Y. Wu, "Gradient descent optimization in deep learning model training based on multistage and method combination strategy," *Security and Communication Networks*, vol. 2021, Article ID 9956773, 15 pages, 2021.
- [26] D. Singh, P. G. Febbo, K. Ross et al., "Gene expression correlates of clinical prostate cancer behavior," *Cancer Cell*, vol. 1, no. 2, pp. 203–209, 2002.
- [27] A. Gumaiei, R. Sammouda, M. Al-Rakhmi, H. AlSalman, and A. El-Zaart, "Feature selection with ensemble learning for prostate cancer diagnosis from microarray gene expression," *Health Informatics Journal*, vol. 27, no. 1, Article ID 1460458221989402, 2021.

Research Article

Estimation and Prediction of Hospitalization and Medical Care Costs Using Regression in Machine Learning

Ahmed I. Taloba ^{1,2} **Rasha M. Abd El-Aziz** ^{1,3} **Huda M. Alshanbari** ⁴
and **Abdal-Aziz H. El-Bagoury** ⁵

¹Department of Computer Science, College of Science and Arts in Gurayat, Jouf University, Sakakah, Saudi Arabia

²Information System Department, Faculty of Computers and Information, Assiut University, Assiut, Egypt

³Computer Science Department, Faculty of Computers and Information, Assiut University, Assiut, Egypt

⁴Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

⁵Basic Science Department, Higher Institute of Engineering and Technology, El-Mahala El-Kubra, Egypt

Correspondence should be addressed to Abdal-Aziz H. El-Bagoury; azizhel2013@yahoo.com

Received 26 December 2021; Accepted 7 February 2022; Published 2 March 2022

Academic Editor: K. Shankar

Copyright © 2022 Ahmed I. Taloba et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medical costs are one of the most common recurring expenses in a person's life. Based on different research studies, BMI, ageing, smoking, and other factors are all related to greater personal medical care costs. The estimates of the expenditures of health care related to obesity are needed to help create cost-effective obesity prevention strategies. Obesity prevention at a young age is a top concern in global health, clinical practice, and public health. To avoid these restrictions, genetic variants are employed as instrumental variables in this research. Using statistics from public huge datasets, the impact of body mass index (BMI) on overall healthcare expenses is predicted. A multiview learning architecture can be used to leverage BMI information in records, including diagnostic texts, diagnostic IDs, and patient traits. A hierarchy perception structure was suggested to choose significant words, health checks, and diagnoses for training phase informative data representations, because various words, diagnoses, and previous health care have varying significance for expense calculation. In this system model, linear regression analysis, naive Bayes classifier, and random forest algorithms were compared using a business analytic method that applied statistical and machine-learning approaches. According to the results of our forecasting method, linear regression has the maximum accuracy of 97.89 percent in forecasting overall healthcare costs. In terms of financial statistics, our methodology provides a predictive method.

1. Introduction

The incidence of overweight and obesity has increased significantly in most countries in recent decades. Excess weight is associated with an increased incidence of many chronic diseases, including vascular disease, respiratory disease, osteoarthritis, some cancer, type 2 diabetes, and premature death. There is consistent evidence that an increased BMI is associated with higher health costs, and these costs are expected to increase as obesity. Modelling uses machine-learning methods, in which the machine learns from the data and uses it to forecast new data [1, 2]. The most

commonly predictive analytic model used is regression [3–6]. The proposed model for accurate prediction of future outputs has applications in banking, economics, e-commerce, sports, business, entertainment, etc. A method used to forecast healthcare costs for BMI is based on several factors. Multiple linear regression is one of the statistical techniques for estimating the relationship among the dependent (target) and independent variables. The regression method is commonly used to develop a system based on a number of factors to predict the cost [5–11].

The regression analysis is performed to determine the relationship among two or more variables with cause-effect

relationships and to make predictions for the topic using the relationships [12]. If regression used one independent variable, then it is known as univariate regression analysis, or else if it used more than two independent variables then it is known as multivariate regression analysis. Linear regression involves initially uploading the data and then analysing the data. Subsequently, the data are cut, and then, the data are trained and separated to create the model. At last, it will evaluate the accuracy. The main aim of regression is to develop an efficient technique for predicting dependent properties from a set of characteristic variables. A regression problem is the actual or continuous value of the output variables, that is, area, salary, and weight. Regression can be defined as a statistical method used in applications such as predicting the healthcare costs. Regression is used to predict the relationship among the dependent variable and set of independent variables. There are various types of regression techniques available namely simple linear regression, multiple linear regression, polynomial regression, support vector regression, and random forest regression [13].

Fast-growing healthcare costs have become a significant challenge in several developed countries. Existing evidence suggests that healthcare costs have accumulated among a large number of BMI. Even though experiments have attempted to develop accurate models for predicting healthcare costs for BMI, their effectiveness is excellent due to the lack of detailed clinical information in the data used to create complex intervals and prognostic models. Numerous studies on more costs for obesity patient prognostic models have relied on self-report data and electronic health data from claims [14]. Data from laboratory tests are defined—these, more granular and detailed clinical information, lead to improvements in the prognostic model. A recent survey by health research program and claim data shows that there is an improvement in the performance of the machine-learning-based predictive model for health costs for obesity. Still, many insurers and providers worldwide are actively seeking an approach that can accurately predict obesity BMI [15].

However, despite the potential value of advanced machine-learning approaches for risk prediction, payers and providers still rely heavily on linear regression to manage and adapt their patient population [16, 17]. The slow adoption of advanced machine-learning techniques may be partly explained by the lack of familiarity with risk stabilization analysts with such techniques and the combination of complex interpretation and results required in practice. Machine-learning regression models are within the framework of standard linear regression and perform some sophisticated but less explicit machine-learning techniques [18, 19]. This study focused on fine linear regression models, which conducted a complete comparison of penalty regression with linear regression in forecasting overall health costs, which was not reported in the previously published literature. The major focus of this study is to estimate the health costs incurred due to obesity in the population.

The rest of this study is formalized as follows: Section 2 defines the related works on estimating the healthcare costs using various methodology methods. Section 3 designates in detail the workflow of the proposed algorithm. Section 4 represents the experiments with results and comparison graphics with existing works and its discussion. Finally, Section 6 concludes the study.

2. Related Work

Some of the recent literature that describes the various mechanism of estimating the costs of physical healthcare is summarized below. In [20], unplanned 30-day readmissions are a common occurrence among congestive heart failure (CHF) patients, posing major health concerns and increasing healthcare costs. It is critical to implement tailored treatment programs for high hazard patients of readmission in an attempt to prevent readmissions and lower healthcare costs. This necessitates recognizing high individuals at the time of hospital release. They constructed and evaluated a deep learning network to predict 30-day unplanned readmission using actual information from over 7,500 CHF patients hospitalized in Sweden. Using specialist characteristics and situational integration of medical knowledge provides a cost-sensitive implementation of the long short-term memory (LSTM) neural net. Using both machine-derived and professional characteristics, including frequent patterns, and resolving the issue of class imbalances, this research focuses on important parts of an EHR-driven forecasting system in a single framework. We assess each element's impact on forecasting effectiveness (F1 measure, ROC-AUC) and price benefits. In at least 2 evaluating criteria, it shows that the technique with all critical features outperforms the simplified approaches in terms of discriminating capability. Researchers also propose a basic economic assessment to predict annual income if high-risk patients are provided tailored therapies.

Patients with heart failure (HF) require precise hazard classification to implement tailored therapies focused on enhancing their efficiency of living and results [21]. To assess the economic benefit of complementing claim-based forecasting analytics with electronic medical record (EMR)-derived data and to contrast machine-learning techniques to conventional logistic regression in forecasting critical results in patients with HF, healthcare patients with HF from 2 healthcare professional systems in Massachusetts, Boston, were included in predictive research with a one-year follow-up duration. "Providers" comprise therapists, various medical professionals, clinicians, and their organization including the network. Logistic regression, gradient boosted modelling, regression trees, random forests, least absolute shrinkage, classification, and selection operation regression were used to predict all-cause morbidity, top cost decile, HF hospitalization, gradient boosted modelling, and home days loss larger than 25%. Information from network 1 was used to educate all algorithms, which were then evaluated in network 2. The area under high accuracy curves (AUPRCs) and overall value estimations from decision curves were obtained after choosing the best effective modelling strategy

depending on the Brier score, calibration, and discrimination.

The goal of this study was to evaluate the effectiveness of machine-learning methodologies for predicting healthcare expenses connected with spinal fusion in aspects of gains or losses in Taiwan Diagnosis-Related Groups (Tw-DRGs) and to use these techniques to investigate the major features connected with spinal fusion medical costs. Methods: a data collection was gathered from a healthcare facility centre in Taoyuan, Taiwan, containing data on Tw-DRG49702 patients (without problems or comorbidity; posterior and other spinal fusion). Weka 3.8.1 was used to forecast using random forest, support vector machines, Naive Bayesian, C4.5 decision tree, and logistic regression approaches [22, 23]. The research showed that the random forest approach may be used to estimate the healthcare expenditures of Tw-DRG49702 and that it can help institutions improve the financially operational effectiveness of this procedure.

Because of the ageing populations and enhanced therapy of fundamental conditions, cardiac arrest is among the most complicated chronic disorders with a higher incidence. The incidence is projected to gradually climb, reaching 3% of the population in Western countries [24]. It is the leading reason for hospitalizations in people aged 65 and above, leading to substantial expenses and a significant societal effect. In the therapy of HF, the present “one-size-fits-all” strategy does not produce the optimal results for all patients. These facts pose a serious danger to the proper treatment of heart failure patients. It will take an unconventional method from a unique perspective on health care. We offer a unique forecasting, preventive, and personalized healthcare strategy, in which patients are actually in charge of their care, aided by a user-friendly online form that employs artificial intelligence (AI). This technique study outlines the demands in HF care, as well as the necessary paradigm shift and the factors necessary to make it happen. A digital physician is being developed through an exciting combination of medical and high-tech partners from patient coaching, serious gaming, North-West Europe, artificial intelligence, and combining state-of-the-art HF health care. The findings are intended to improve and customize self-care, in which patients conduct routine care chores without the intervention of healthcare experts, allowing them to focus on more difficult problems. This innovative approach to health care will lower prices per patient while increasing results, ensuring the long-term viability of top-tier HF health care.

In [25], DRG codes are useful for price tracking and allocation of resources since healthcare operators obtain predetermined levels of compensation for certain treatments under diagnosis-related group (DRG) payments. Coding, on the other hand, is usually done after the fact, after the patient has been discharged. They want to use normal medical text to forecast DRGs and DRG-based case mix index (CMI) at initial inpatient admission to forecast hospital costs in an acute context. Without manual coding, a deep learning-based natural language processing (NLP) method is tested to forecast cost-reflecting weights and per-episode DRGs on 2 cohorts (paid by All Patient Refined (APR) DRG or Medicare Severity (MS) DRG). In fivefold cross-validation

trials on the first day of ICU admission, it attained macro-averaged area under the receiver operating characteristic curve (AUC) scores of 0.871 (SD 0.011) on MS-DRG and 0.884 (0.003) on APR-DRG. When applied to hypothetical patient populations to predict average cost-reflecting weights, the algorithm improved over time, yielding absolute CMI errors of 12.79 (2.31%) and 2.40 (1.07%) on the first day, correspondingly. Because the system can adjust to changes in admission time and cohort size while requiring no additional manual coding, it has the potential to aid in cost estimation for active patients and enable improved functional outcome in hospitals.

3. The Proposed Method Based on Linear Regression

Linear regression is one of the most common supervisory machine learning statistical analysis techniques [26]. It is commonly used to find linear correlations between two or more responses and predictive variables. The technique is divided into two types depending on the number of variables in the model such as simple linear regression and multiple linear regression. A response variable corresponding to a predictive variable is simple linear regression. Whether more than two response variables correspond to predictive variables is known as multiple linear regression as shown in Figure 1. This work used linear regression to study the relationship among total maintenance and other properties in datasets to obtain the properties most affected by the total cost of maintenance. 75% of the data in the dataset were trained, and 25% of the data were tested. Then, Pearson's correlation coefficient (PCC) for each simple linear regression sample was calculated. The PCC is determined and calculated by the following equation to find the parallel variability and strength of a linear regression relationship between two factors:

$$Y'_i = f_n(X'_i, \beta_p) + e. \quad (1)$$

Here, X'_i and y'_i represent the independent variable and dependent variable; f_n represents the function; β_p represents the unknown parameters; and e represents the error terms. The most commonly used measurements to estimate the performance of a linear regression are the root mean square error (RMSE), the mean absolute error (MAE), and the mean square error (MSE) [26]. The following equations denote the error deviation for regression:

$$\text{RMSE} = \sqrt{\frac{\sum_{j=1}^M (y_i - y'_i)^2}{M}}, \quad (2)$$

$$\text{MAE} = \sum_{i=1}^N \frac{x_i - x}{N}, \quad (3)$$

$$\text{MSE} = \frac{1}{N} \sum_{j=1}^N (y_i - y'_i)^2. \quad (4)$$

These regression measurements are constant variables and standard measurements for determining sample accuracy.

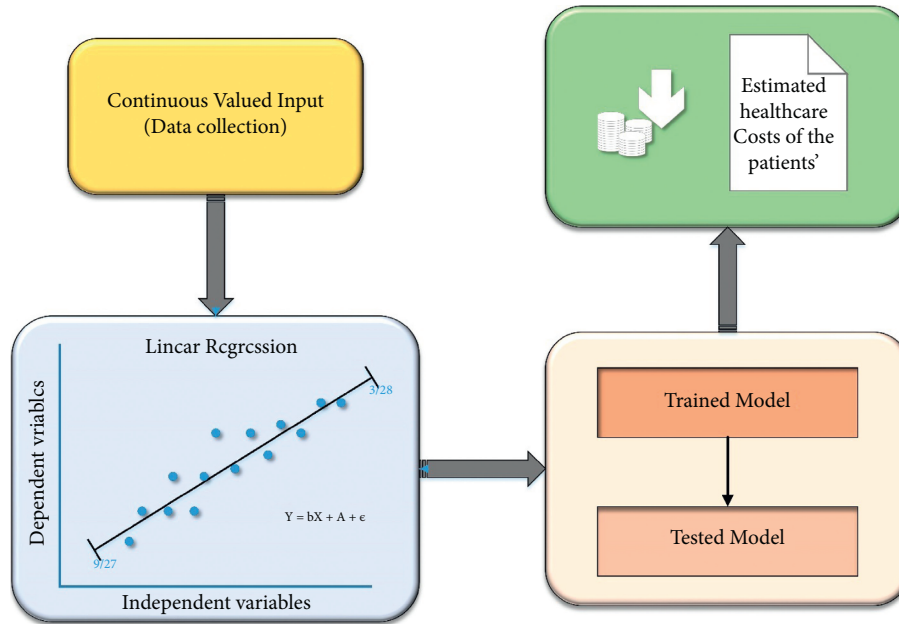


FIGURE 1: Block diagram for the proposed model.

3.1. Regression's Role in Predicting the Costs. Clinics are encouraged to find more meaning in the substantial amount of data they generate and store each day [27]. Regression provides useful predictive accuracy and value for machine-learning clinics' databases with useful methods, features, and structures and contributes to a variety of strategies. The regression method aims to identify the possibility of improving results based on the predictive value of large-scale datasets for annual health costs. This is evidence of effectiveness in dealing with priority tasks, which defines that behaviours have the maximum tendency to cause preferred outcomes.

3.2. Steps for Applying Regression to Datasets. The database used here is a collection of medical expense personal data, which contain anonymous information about people. These data will act as a method learning object to generate functional information. In Table 1, the attributes such as BMI and age are continuous variables, and the attributes such as smoker and sex are categorical variables:

- (i) The next step is data exploration and preparation, and the quality of any machine-learning program is largely based on the quality of the data it uses. This stage requires more human intervention in the machine-learning process. Frequently cited statistics show that 80% of efforts in machine learning are dedicated to data. Most of this time is spent learning more about data and its nuances throughout an exercise known as data analysis.
- (ii) Then, a model on the data is trained. The specific machine-learning task will announce the selection of the suitable method, and the method will denote the data in the form of a model.

TABLE 1: Healthcare attributes and their specifications.

Attributes	Specifications
BMI	Body mass index
Age	Primary beneficiary age
Sex	Gender (male/female)
Smoker	The one who smokes affected by the obesity
Children	Number of children under BMI
Costs	Individual healthcare costs of the respective person

- (iii) Subsequently, the model performance is evaluated. It is important to evaluate how well the method has learned from its past experience as each machine-learning model results in a biased solution to the learning problem. Depending on the type of model used, the accuracy of the sample can be estimated using the experimental database.
- (iv) Finally, the performance of the model is improved. It is necessary to use advanced techniques to increase the performance of the model if better performance is required. Each time, an entirely different type of model may have to be changed. After completing these steps, if the model appears to be operating acceptably, it can be used for its intended purpose. This model can be used to provide score data for forecasting, for financial data forecasting, to generate relevant insights for marketing or research, or to automate tasks.

3.3. Dataset Description. We intended to forecast a patient's healthcare costs for the coming year depending on their insurance payment statistics and previous healthcare data. Tsuyama Chuo Hospital contributed the healthcare record information. These documents come from healthcare insurance applications that the hospital is required to submit

to the administration. Every patient is recognized by an individual identity (ID) in these reports, which include the patient's conditions, medications, operations, and payment details [28]. This claim's comprehensive paperwork can be obtained on the relevant website. We were able to retrieve the following information using this information:

- (i) Patient demographics include age and gender.
- (ii) Patients' characteristics include their body fat percentage, height, weight, and waist circumference.
- (iii) Health care verifies the outcomes of a patient's healthcare check-up tests. Every testing is assigned a code, and the outcome should be provided. Blood pressure (BP) and creatinine levels are two instances. There are 25 various categories of tests, as well as the date that they were gathered.
- (iv) Prognosis: a patient's ailment is diagnosed using ICD-10 codes and is tracked by date.
- (v) Payment details: for every session or hospital stay, every patient was assigned a score. This result effectively corresponds to the expense of a patient payment, which is the figure we needed to forecast for the following years.

It has been demonstrated that predicting patients' healthcare costs solely based on medical data is difficult. Preceding healthcare expenses are the strongest predictor of future expenditures: a longer history of healthcare expenditures is considered to increase forecasting. Depending on this fact, it is easier to anticipate future healthcare expenses when patients' information is available for multiple periods. When attempting to forecast expenditures for a single year, at least a two-year history is required [29].

Patients' monthly histories were included in our database. Furthermore, since many patients only had limited claims per year, there are several missing data. As a result, we decided to arrange claims by year to reduce the number of missing information. This technique did not work out as planned because many patients only had data [30]. We next screened out these patients, leaving only those with clinical history. The fundamental characteristics of these patients are shown in Table 2.

Figure 2 forecasts every patient's scores for the following year. These scores are directly proportional to the amount of cost a patient spent on health care. The range of patient values is depicted in the graph. As anticipated for healthcare expenses, the scores exhibit all similar patterns as indicated previously, with a spike at zero and a lengthy right-hand tail.

It has been claimed that using medical characteristics produces similar results as using solely expense predictions. Although the fact that medical record appears to have little effect on forecast accuracy, we choose to maintain it because it can enhance the range of variables in the algorithm, which might enhance vector differentiation. Every resource accessible as characteristics was used to encode a patient's history. Demographics, health check-up results, ICD-10 diagnostic groupings, real score, and preceding score are the inputs [31, 32]. The patient's vector is described in full in the

TABLE 2: Patients' characteristics and their predicted value.

Statistics	Predicted value
Total no. of patients	24,353
Mean value for expenses	10,538
Mean (age)	46.08
Male (%)	47.48
Female (%)	50.30

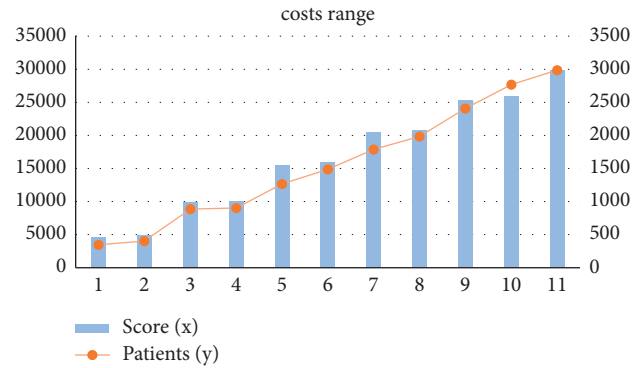


FIGURE 2: Graphic representation of cost range for patients' score.

table. We employed all of the parameters listed in the table as input vectors, with the exception of the real score, which was used as our target attribute.

3.4. Training Phase. We must determine the ideal hyper-parameters of our system for a forecast to adapt as closely as feasible to its true value. The weights of every dimension used in the distance function and g in the discount function are these parameters. For the training process, we used the gradient-based methods since they have a strong mathematical foundation for achieving optimal results.

The gradient descent technique is an automated approach for minimizing or maximizing a target function by optimizing variable values. As our objective parameter, we used the mean absolute error (MAE), which is calculated as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n \|y_i - \hat{y}_i\|. \quad (5)$$

The target is to minimize the values of the MAE equation, which is dependent on the variable v_t that could be either γ or ω . The following equation gives the updated value of v_t , termed v_{t+1} as follows:

$$v_{t+1} = v_t - \alpha \frac{\partial \text{MAE}}{\partial v}. \quad (6)$$

This technique offers us a series of numbers for v_0, \dots, v_k that minimizes the MAE, with the first value for v (i.e., v_0) generally chosen at random. During the training phase, we use all of the remaining $N - 1$ patients in L as evidence to try to forecast the expense of every patient p_i in the training set L . An epoch is an execution that computes forecasts for every N patient; the gradient descent approach accumulates by completing repeated epochs.

3.5. Time Optimizing in Computing. A prediction's computing duration scales linearly with the size of the training phase. To find the mass of vectors of dimension m in a database with a training dataset of size n , we must firstly use the discounting function, which has a $O(m)$ complexity. With the training set, we can estimate any discounting functions of the input vector in $O(mn)$. Then, we can estimate K (9), which requires $O(mn)$. for every output series and $O(mn)$ for the accumulation; thus, we can estimate K in $O(mn)$. time. Lastly, we require the discounting function, K , and a product series to get the mass, so we estimate the weights of the input vector (8) while keeping $O(mn)$. Therefore, given $O(mn)$. complexity, we could obtain the forecast.

According to reference, a K -nearest neighbour technique could be used to accelerate up calculation without sacrificing efficiency. For the actual closest neighbour's searches depending on product quantization, we used [33] methodology. Using this technique, we can generate indices for the K -nearest searches in time $O(mn + Kn)$ within the training step. The weights of the K -nearest neighbours, which will be estimated in $O(Km)$, are thus all that is required for a fresh forecast; the other weights are presumed to be null. Whenever the algorithm has been trained, the prediction's complexity is $O(Km)$.

3.6. Interpretability. The IEVREG is a framework that is accessible. For every forecasting we generate, we could calculate the proportion (mass) of every element of information in the testing phase L. As a result, we have a complete understanding of how the anticipated quantity is calculated. This prototype is already interpretable, but to make it completely understandable, we will write a system of regulations for every forecast using the weights from the training dataset and the masses of every dimension gained all through the training step [34]. The idea is to calculate how much every piece of proof adds to the forecast. Firstly, using the weights of the existing N1 patients in the training dataset, we establish a system of regulations for each of the patients in the training phase for forecasting. Using the weights of the remaining N1 patients in the phases and the weights of the dimensions, we firstly build a system of regulations for each of the patients in the training phase [35, 36]. The limits of the measurements for each of the input characteristics, as well as their weights, are encoded by these principles. The algorithm then chooses the patients in the training phase who are the most identical and combines their principles to generate a new collection of criteria for that forecast.

We use a tiny healthcare coverage database only with 5 characteristics as input to demonstrate how we get the regulations with the IEVREG framework. Table 3 shows the 5 data inputs (measurements) and the anticipated result for the healthcare expenses.

We used only the 60 closest neighbours to forecast this patient's result. The most significant principles (greater values) for expense forecasting are then obtained, as illustrated in Table 4. These are the limits and parameters that the

patients have in common with the patients in the training phase.

We could see how a patient's expense projection is interpreted in Table 4. Low weight is associated with age in the IEVREG framework, while higher weight is associated with others. As a result, the method seeks out individuals with identical genders, BMIs, children, and smoking statuses, while ignoring age.

Algorithm 1 represents the steps of the linear regression model.

The flowchart for the proposed linear regression model is shown in Figure 3.

4. Results and Analysis

The average annual rates and costs of consultations, tests, and prescription items were estimated by BMI category at the time of recruitment as shown in Figure 4. Percentage differences in rates and average annual costs were calculated for women with a BMI greater than 2 kg/m^2 and a BMI greater than 20 kg/m^2 , both overall and according to the type of drug use. All models were evaluated using semi-possible generalized linear models with variations such as record link and Poisson. At the beginning of each year, annual expenses are estimated in subgroups defined by alcohol consumption, socioeconomic status, smoking level, educational qualifications, and strenuous exercise in recruitment [37]. The diversity of the proportional increases in annual costs among the types of each subgroup was estimated using the chi-square test.

The mean absolute error, moreover, is ineffective for comparing outcomes with costs stated in various dollars, so we will use the mean absolute percentage error (MAPE), a customized absolute error in which the MAE is reduced by the mean cost and calculated as follows:

$$\text{MAPE} = \frac{(1/n) \sum_{i=1}^n ||y_i - \hat{y}_i||}{m} \quad (7)$$

Here, \hat{y}_i is the estimated output for parameter y_i and m is the mean of variable y , denoted as follows:

$$m = \frac{1}{n} \sum_{i=1}^n y_i. \quad (8)$$

We will also use additional metric, the R2, which reflects how closely we are to the true cost curve and is defined as the Pearson correlation among projected and real healthcare costs. The following formula is used to determine this significance:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - m)^2} \quad (9)$$

5. Discussion

We provided a novel linear regression technique that can simply demonstrate the purposes for producing a certain forecast regarding potential healthcare expenses, which is a useful capacity in the medical field. We evaluated its

TABLE 3: Details of the patients.

Gender	BMI	Smoker	Age	Children	Actual value	Forecasted value
Female	29.98	No	37	1	6245	7154
Male	32.12	No	40	2	6725	7540

TABLE 4: Estimated values.

Gender	Estimated values	Weights
Male	30.6530 < BMI < 31.8560	0.45
	Gender = 0.0	0.45
	Children = 0.0	0.45
	Smoker = 0.0	0.45
	39.2016 < age < 40.2451	0.22
Female	28.5421 < BMI < 29.7451	0.39
	Gender = 0.0	0.39
	Children = 0.0	0.39
	Smoker = 0.0	0.39
	36.2016 < age < 37.2452	0.19

Require: Training data D , number of epochs e , learning rate η , and standard deviation σ .

Ensure: Weights. $\omega_0, \omega_1, \dots, \omega_k$

- (1) Initialize weights $\omega_0, \omega_1, \dots, \omega_k$ from standard normal distribution with zero mean and standard deviation σ .
 - for** epoch in $1, \dots, e$ **do**
 - for** each $(x, y) \in D$ in random order **do**
 - $\hat{y} \leftarrow \omega_0 + \sum_{i=1}^k \omega_i x_i$
 - if** $(\hat{y} > 1 \text{ and } y = 1)$ or $(\hat{y} < -1 \text{ and } y = -1)$ **then**
 - Continue**
 - $\omega_0 \leftarrow \omega_0 - \eta 2 (\hat{y} - y)$
 - for** i in $1, \dots, k$ **do**
 - $\omega_i \leftarrow \omega_i - \eta 2 (\hat{y} - y) x_i$
 - end for**
 - end for**
 - return** $\omega_0, \omega_1, \dots, \omega_k$

ALGORITHM 1: Linear regression (LR).

outcomes to the forecasting produced by the finest algorithms from the analysed research and reported to see how well it predicted. The linear regression is what we are talking about here. When we compare the outcomes of previous designs for the cost of healthcare forecasting approach, we can see that our system is more efficient, demonstrating that a more explicit approach for an issue such as healthcare cost forecasting is conceivable [38–40]. Our research, on the other hand, clearly reveals that healthcare spending is highly connected inside the Medicare program. There are approximately numerous people enrolled in the program. This finding could lead to preventive measures. Autocorrelation shows an inherent methodology that could be influenced by variables that can be changed. As a result, clinicians can use more accurate machine-learning algorithms to target these therapies to the proper HCHN group. There are a few flaws in this research. Initially, we performed the research within the context of a single state’s Medicare system. The outcomes

might differ depending on the state or kind of payer. Secondly, only general-purpose machine-learning algorithms were used. Certain customized versions might function optimally. Thirdly, the prediction algorithms offer no direction on the preventive characteristics that should be considered when developing treatments. Lastly, determining overall health solely based on claim statistics is restricted. Further input resources, such as descriptive elements of electronic health records (EHRs), illness intensity assessments, and socioeconomic determinants of health care, might well be required. A few of these restrictions will be addressed in the forthcoming research. We intend to broaden the scope of the study to include various sorts of healthcare initiatives. We will additionally collect the abovementioned extra data to assess predicted effectiveness [40]. We will also work with physicians and policymakers to make the algorithms more medically applicable using domain expertise to effectively target risk reduction actions.

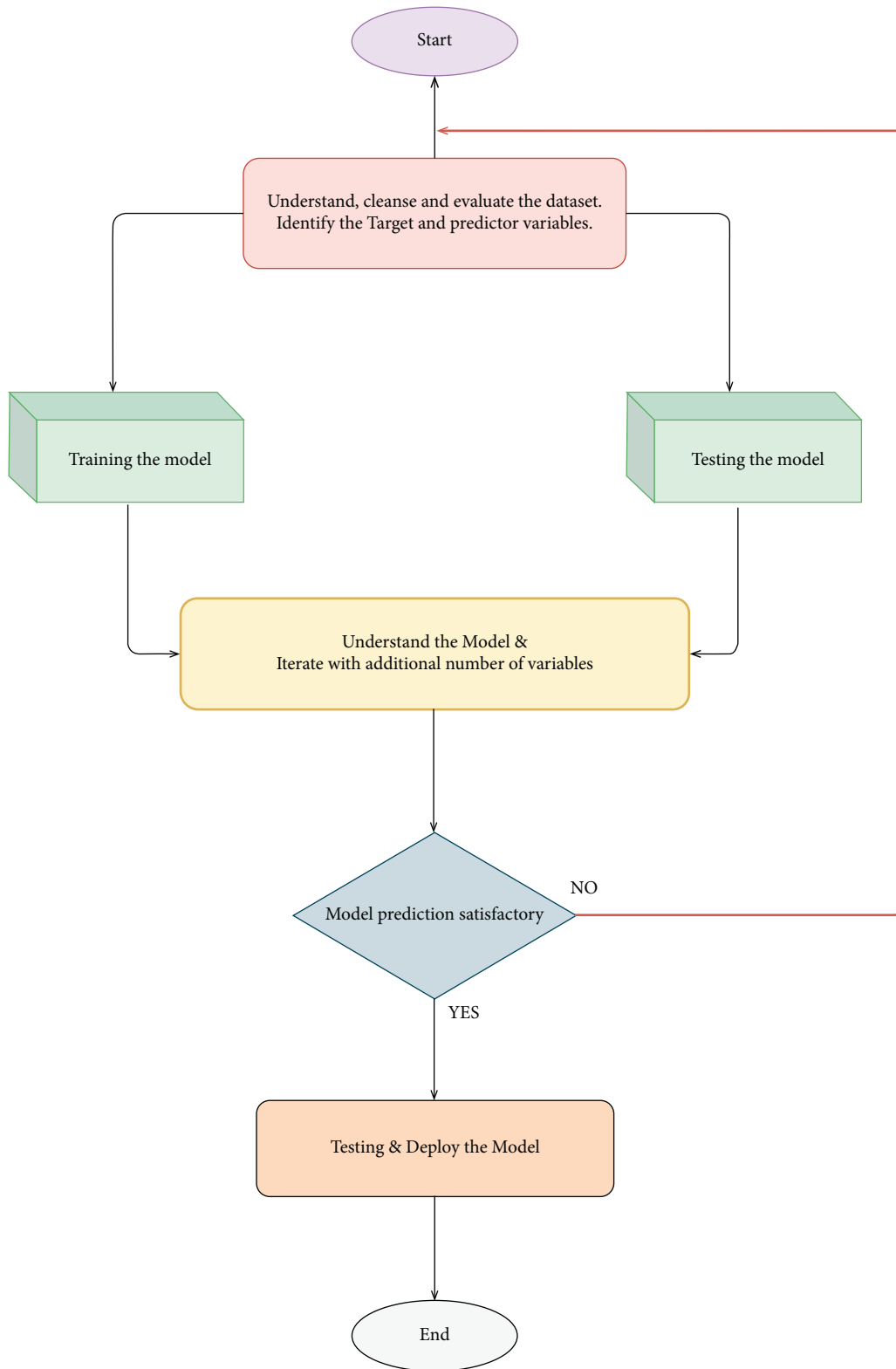


FIGURE 3: Flowchart for estimating the healthcare costs.

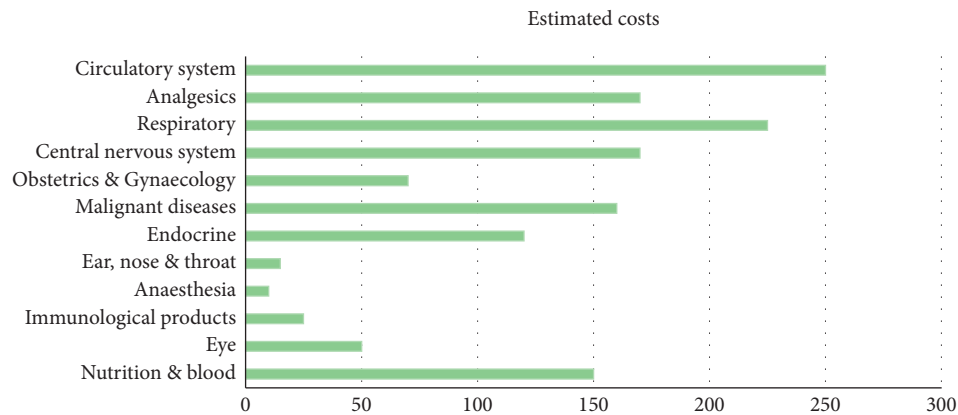


FIGURE 4: Healthcare expenses attributable to obesity and overweight between people on a yearly basis.

6. Conclusion

We provided a new linear regression that can easily demonstrate the reasons for producing a certain forecast regarding potential healthcare expenses, which is a useful capacity in the healthcare area. The linear regression algorithm is used to estimate the healthcare costs of the patients such as obesity (BMI) using certain devices such as smartphones and smart devices. For estimation, by the use of linear regression, supervised learning performs more accurately. By providing comprehensive evidence, regression methodology can be effectively used for prognosis in conjunction with the dataset. The domain and time accuracy will determine the prediction model and the estimation of healthcare expenses. The proposed method reduces the risk of overfitting, and also, training time is less. This method is effective in estimating the healthcare costs of patients with an accuracy rate of 97.89%. The extensive tests on a real-time world database have confirmed the efficiency of our method.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

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

Acknowledgments

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R299), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

- [1] A. K. Dutta, N. Ali Aljarallah, T. Abirami et al., "Optimal deep-learning-enabled intelligent decision support system for SARS-CoV-2 classification," *Journal of Healthcare Engineering*, vol. 2022, Article ID 4130674, 14 pages, 2022.
- [2] J. Mohana, B. Yakkala, S. Vimalnath et al., "Application of internet of things on the healthcare field using convolutional neural network processing," *Journal of Healthcare Engineering*, vol. 2022, Article ID 1892123, 2022.
- [3] L. Hu, L. Li, J. Ji, and M. Sanderson, "Identifying and understanding determinants of high healthcare costs for breast cancer: a quantile regression machine learning approach," *BMC Health Services Research*, vol. 20, no. 1, pp. 1066–1110, 2020.
- [4] M. A. Aefa, M. Mahmoud, and M. M. Nassar, "Parameter estimation for a mixture of inverse chen and inverse compound Rayleigh distribution based on type-I hybrid censoring scheme," *Journal of Statistics Applications & Probability*, vol. 10, no. 3, pp. 647–663, 2021.
- [5] W. A. Afifi and A. H. El-Bagoury, "Optimal multiplicative generalized linear search plan for a discrete randomly located target," *Information Sciences Letters*, vol. 10, no. 1, pp. 153–158, 2021.
- [6] R. A. Ganaie, V. Rajagopalan, and S. Aldulaimi, "The weighted power shanker distribution with characterizations and applications of real life time data," *Journal of Statistics Applications & Probability*, vol. 10, no. 1, pp. 245–265, 2021.
- [7] M. H. Abu-Moussa, A. M. Abd-Elfattah, and E. H. Hafez, "Estimation of stress-strength parameter for Rayleigh distribution based on progressive type-II censoring," *Information Sciences Letters*, vol. 10, no. 1, pp. 101–110, 2021.
- [8] S. Sana and M. Faizan, "Bayesian estimation using lindley's approximation and prediction of generalized exponential distribution based on lower record values," *Journal of Statistics Applications & Probability*, vol. 10, no. 1, pp. 61–75, 2021.
- [9] K. Sahu and R. K. Srivastava, "Needs and importance of reliability prediction: an industrial perspective," *Information Sciences Letters*, vol. 9, no. 1, pp. 33–37, 2020.
- [10] A. A. Soliman, Al-W. A. Farghal, and G..A. Abd-Elmougod, "Statistical inference under copula approach of accelerated dependent generalized inverted exponential failure time with progressive hybrid censoring scheme," *Applied Mathematics & Information Sciences*, vol. 15, no. 6, pp. 687–699, 2021.
- [11] S. Kent, J Green, G Reeves et al., "Hospital costs in relation to body-mass index in 1.1 million women in England: a prospective cohort study," *The Lancet Public Health*, vol. 2, no. 5, pp. e214–e222, 2017.
- [12] V. S. Kadam, S. Kanhere, and S. Mahindrakar, "Regression techniques in machine learning & applications: a review,"

- International Journal for Research in Applied Science and Engineering Technology*, vol. 8, no. 10, pp. 826–830, 2020.
- [13] B. Panay, N. Baloian, J. Pino, S. Peñafiel, H. Sanson, and N. Bersano, “Predicting health care costs using evidence regression,” *Proceedings*, vol. 31, p. 74, 2019, <https://www.mdpi.com/2504-3900/31/1/74>.
- [14] B. J. Moore, S. White, R. Washington, N. Coenen, and A. Elixhauser, “Identifying increased risk of readmission and in-hospital mortality using hospital administrative data,” *Medical Care*, vol. 55, no. 7, pp. 698–705, 2017.
- [15] R. S. Suidan, W. He, C. C. Sun et al., “Impact of body mass index and operative approach on surgical morbidity and costs in women with endometrial carcinoma and hyperplasia,” *Gynecologic Oncology*, vol. 145, no. 1, pp. 55–60, 2017.
- [16] H. J. Kan, H. Kharrazi, H.-Y. Chang, D. Bodycombe, K. Lemke, and J. P. Weiner, “Exploring the use of machine learning for risk adjustment: a comparison of standard and penalized linear regression models in predicting health care costs in older adults,” *PloS one*, vol. 14, no. 3, Article ID e0213258, 2019.
- [17] S. Kent, S. A. Jebb, A. Gray et al., “Body mass index and use and costs of primary care services among women aged 55-79 years in England: a cohort and linked data study,” *International Journal of Obesity*, vol. 43, no. 9, pp. 1839–1848, 2019.
- [18] J. A. Irvin, A. A. Kondrich, M. Ko et al., “Incorporating machine learning and social determinants of health indicators into prospective risk adjustment for health plan payments,” *BMC Public Health*, vol. 20, no. 1, pp. 608–610, 2020.
- [19] S. Kent, F. Fusco, A. Gray, S. A. Jebb, B. J. Cairns, and B. Mihaylova, “Body mass index and healthcare costs: a systematic literature review of individual participant data studies,” *Obesity Reviews*, vol. 18, no. 8, pp. 869–879, 2017.
- [20] A. Ashfaq, A. Sant’Anna, M. Lingman, and S. Nowaczyk, “Readmission prediction using deep learning on electronic health records,” *Journal of Biomedical Informatics*, vol. 97, Article ID 103256, 2019.
- [21] R. J. Desai, S. V. Wang, M. Vaduganathan, T. Evers, and S. Schneeweiss, “Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes,” *JAMA Network Open*, vol. 3, no. 1, Article ID e1918962, 2020.
- [22] C.-Y. Kuo, L.-C. Yu, H.-C. Chen, and C.-L. Chan, “Comparison of models for the prediction of medical costs of spinal fusion in Taiwan diagnosis-related groups by machine learning algorithms,” *Healthcare informatics research*, vol. 24, no. 1, pp. 29–37, 2018.
- [23] E. Frank, M. A. Hall, and I. H. Witten, *The WEKA Workbench. Online Appendix for “Data Mining: Practical Machine Learning Tools and Techniques”*, Morgan Kaufmann, Massachusetts, MA, USA, Fourth edition, 2016.
- [24] M. Barrett, J. Boyne, J. Brandts et al., “Artificial intelligence supported patient self-care in chronic heart failure: a paradigm shift from reactive to predictive, preventive and personalised care,” *The EPMA Journal*, vol. 10, no. 4, pp. 445–464, 2019.
- [25] J. Liu, D. Capurro, A. Nguyen, and K. Verspoor, “Early prediction of diagnostic-related groups and estimation of hospital cost by processing clinical notes,” *NPJ Digital Medicine*, vol. 4, no. 1, pp. 1–8, 2021.
- [26] H. N. Alhazmi, A. Alghamdi, F. Alajlani, S. Abuayied, and F. M. Aldosari, “Care cost prediction model for orphanage organizations in Saudi Arabia,” *IJCSNS*, vol. 21, no. 4, p. 84, 2021.
- [27] B. Nithya and V. Ilango, “Predictive analytics in health care using machine learning tools and techniques,” in *Proceedings of the 2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 492–499, Madurai, India, June 2017.
- [28] Z. J. Ward, S. N. Bleich, M. W. Long, and S. L. Gortmaker, “Association of body mass index with health care expenditures in the United States by age and sex,” *PloS one*, vol. 16, no. 3, Article ID e0247307, 2021.
- [29] I. Osawa, T. Goto, Y. Yamamoto, and Y. Tsugawa, “Machine-learning-based prediction models for high-need high-cost patients using nationwide clinical and claims data,” *NPJ digital medicine*, vol. 3, no. 1, pp. 148–149, 2020.
- [30] C. Yang, C. Delcher, E. Shenkman, and S. Ranka, “Machine learning approaches for predicting high cost high need patient expenditures in health care,” *BioMedical Engineering Online*, vol. 17, no. 1, pp. 131–220, 2018.
- [31] H. Kharrazi, H.-Y. Chang, S. E. Heins, J. P. Weiner, and K. A. Gudzone, “Assessing the impact of body mass index information on the performance of risk adjustment models in predicting health care costs and utilization,” *Medical Care*, vol. 56, no. 12, pp. 1042–1050, 2018.
- [32] F. Wang, T. McDonald, J. Bender, B. Reffitt, A. Miller, and D. W. Edington, “Association of healthcare costs with per unit body mass index increase,” *Journal of Occupational and Environmental Medicine*, vol. 48, no. 7, pp. 668–674, 2006.
- [33] C. S. Florence, G. Bergen, A. Atherly, E. Burns, J. Stevens, and C. Drake, “Medical costs of fatal and nonfatal falls in older adults,” *Journal of the American Geriatrics Society*, vol. 66, no. 4, pp. 693–698, 2018.
- [34] M. Ravaut, H. Sadeghi, K. K. Leung et al., “Predicting adverse outcomes due to diabetes complications with machine learning using administrative health data,” *NPJ digital medicine*, vol. 4, no. 1, pp. 1–12, 2021.
- [35] C. Wu, F. Wu, Y. Huang, and X. Xie, “NICE: neural in-hospital cost estimation from medical records,” in *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pp. 2409–2412, Beijing, China, 2019.
- [36] N. I. Jha, I. Ghergulescu, and A.-N. Moldovan, “OULAD MOOC dropout and result prediction using ensemble, deep learning and regression techniques,” in *Proceedings of the 11th International Conference on Computer Supported Education CSEU*, no. 2, pp. 154–164, Heraklion, Crete, Greece, MAY 2019.
- [37] M. P. Shyamala Devi, M. Swathi, V. Purushotham Reddy et al., “Linear and ensembling regression based health cost insurance prediction using machine learning,” in *In: Smart Computing Techniques and Applications. Smart Innovation, Systems and Technologies*, S. C. Satapathy, V. Bhateja, M. N. Favorskaya, and T. Adilakshmi, Eds., vol. 224, Singapore, Springer, 2019.
- [38] S. Baharvand, A. Jozaghi, R. Fatahi-Alkouhi, S. Karimzadeh, R. Nasiri, and B. Lashkar-Ara, “Comparative study on the machine learning and regression-based approaches to predict the hydraulic jump sequent depth ratio,” *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, vol. 45, pp. 2719–2732, 2021.
- [39] N. M. Mohamed, “Estimation on kumaraswamy-inverse weibull distribution with constant stress partially accelerated life tests,” *Applied Mathematics & Information Sciences*, vol. 15, no. 4, pp. 503–510, 2021.
- [40] G. Manogaran and D. Lopez, “Health data analytics using scalable logistic regression with stochastic gradient descent,” *International Journal of Advanced Intelligence Paradigms*, vol. 10, no. 1–2, pp. 118–132, 2018.

Research Article

Optimal Deep-Learning-Enabled Intelligent Decision Support System for SARS-CoV-2 Classification

Ashit Kumar Dutta,¹ Nasser Ali Aljarallah,^{2,3} T. Abirami,⁴ M. Sundarrajan,⁵ Seifedine Kadry,⁶ Yunyoung Nam ,⁷ and Chang-Won Jeong⁸

¹Department of Computer Science and Information Systems, College of Applied Sciences, AlMaarefa University, Ad Diriyah Riyadh 13713, Saudi Arabia

²Department of Business Administration, AlMajmaah University, Saudi Arabia

³AlMaarefa University, Riyadh, Saudi Arabia

⁴Department of Information Technology, Kongu Engineering College, Erode 638060, India

⁵Department of Computer Science and Engineering, K. Ramakrishnan College of Engineering, Tiruchirappalli 621112, India

⁶Department of Applied Data Science, Noroff University College, Kristiansand, Norway

⁷Department of Computer Science and Engineering, Soonchunhyang University, Asan, Republic of Korea

⁸Medical Convergence Research Center, Wonkwang University, Iksan, Republic of Korea

Correspondence should be addressed to Yunyoung Nam; ynam@sch.ac.kr

Received 15 November 2021; Revised 9 December 2021; Accepted 15 December 2021; Published 25 January 2022

Academic Editor: K. Shankar

Copyright © 2022 Ashit Kumar Dutta et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Intelligent decision support systems (IDSS) for complex healthcare applications aim to examine a large quantity of complex healthcare data to assist doctors, researchers, pathologists, and other healthcare professionals. A decision support system (DSS) is an intelligent system that provides improved assistance in various stages of health-related disease diagnosis. At the same time, the SARS-CoV-2 infection that causes COVID-19 disease has spread globally from the beginning of 2020. Several research works reported that the imaging pattern based on computed tomography (CT) can be utilized to detect SARS-CoV-2. Earlier identification and detection of the diseases is essential to offer adequate treatment and avoid the severity of the disease. With this motivation, this study develops an efficient deep-learning-based fusion model with swarm intelligence (EDLFM-SI) for SARS-CoV-2 identification. The proposed EDLFM-SI technique aims to detect and classify the SARS-CoV-2 infection or not. Also, the EDLFM-SI technique comprises various processes, namely, data augmentation, preprocessing, feature extraction, and classification. Moreover, a fusion of capsule network (CapsNet) and MobileNet based feature extractors are employed. Besides, a water strider algorithm (WSA) is applied to fine-tune the hyperparameters involved in the DL models. Finally, a cascaded neural network (CNN) classifier is applied for detecting the existence of SARS-CoV-2. In order to showcase the improved performance of the EDLFM-SI technique, a wide range of simulations take place on the COVID-19 CT data set and the SARS-CoV-2 CT scan data set. The simulation outcomes highlighted the supremacy of the EDLFM-SI technique over the recent approaches.

1. Introduction

Intelligent decision support systems (IDSS) has become widely used in several applications of healthcare. Internet of things (IoT), wearables, manual data entry, and online sources are some of the instances of complex data sources for IDSS. The data sustained by IDSS significantly helps in the earlier identification of diseases and equivalent treatments.

The coronavirus disease 2019 (COVID-19) epidemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), began in Wuhan city, Hubei province, in December, 2019, and has spread throughout China. COVID-19 is an infectious disease caused by the novel coronavirus named SARS-CoV-2. The virus is extremely infectious, and can be transmitted by indirect or direct contact with diseased persons with respiratory droplets while they cough, sneeze,

or even talk [1]. Now, the real-time polymerase chain reaction (RT-PCR) test is the common method used to confirm COVID-19 infection, and with the rapid rise in the number of diseased persons, almost all countries are confronting a shortage of testing kit. Furthermore, RT-PCR testing can have a higher false-negative rate and turnaround times [2]. Therefore, it is appropriate for considering other testing tools for detecting COVID-19 infected people to isolate them and alleviate the pandemic impact on the lives of several people. The chest computed tomography (CT) is an appropriate supplement to RT-PCR testing and plays a role in diagnosing and screening COVID-19 infection. In current works [3], the researchers manually investigated chest CT scans of over thousands of patients and confirmed the helpfulness of chest CT scan in COVID-19 detection with a higher sensitivity rates.

In certain cases, the patient had a negative PCR test at first, but confirmation was depending on their CT results. Additionally, chest CT screening was suggested, while the patient shows compatible symptoms with COVID-19; however, the outcome of its PCR tests is negative [4]. Hence, it is necessary for an automatic detection tool that exploits the current developments in deep learning (DL) and artificial intelligence (AI), as well as the accessibility of CT images to construct AI-based tools to prevent further spreading and expedite the diagnoses method [5]. In order to mitigate the shortage and inefficiency of current tests for COVID-19 infection, various attempts have been dedicated to seeking alternate testing tools [6]. Various researches have exposed that CT scans manifest strong radiological results of COVID-19 and are promising in serving as an accessible and more efficient testing tool because of the wider accessibility of CT devices, which could achieve results at the highest rate. Furthermore, to mitigate the burden of medical specialists from reading CT scans, numerous studies have designed DL algorithms that could automatically interpret CT images and forecast whether the CT is positive for COVID-19 infection. When this work has demonstrated effective outcomes, they have two limitations [7]. Initially, the CT scans data set utilized in this study are not accessible to the public because of security concerns.

Accordingly, their results could not be reproduced, and the trained methods could not be utilized in other hospitals. In addition, the lack of open-sourced annotated COVID-19 CT data sets seriously hinders the development and research of innovative AI tools for precise CT-based testing of COVID-19 infection [8]. Next, this study requires a wide range of CTs at the time of model training to accomplish performances that meet the medical standards. These requirements are practically stringent, and it could not met by several hospitals, particularly under the circumstance that medical experts are very occupied in handling COVID-19 infected patients and do not have time to annotate and collect a huge amount of COVID-19 CT scans.

This study develops an efficient deep-learning-based fusion model with swarm intelligence (EDLFM-SI) for SARS-CoV-2 identification for complex healthcare applications. Moreover, the EDLFM-SI technique comprises a

fusion of capsule network (CapsNet) and MobileNet based feature extractors are employed. Furthermore, a water strider algorithm (WSA) is applied to fine-tune the hyperparameters involved in the DL models. Lastly, a cascaded neural network (CNN) classifier is applied to detect the existence of SARS-CoV-2. For examining the enhanced outcomes of the EDLFM-SI technique, a comprehensive experimental analysis is carried out on the COVID-19 CT data set and the SARS-CoV-2 CT scan data set.

The rest of the paper is organized as follows. Section 2 offers the related works; Section 3 elaborates the proposed model; Section 4 provides the result analysis; and Section 5 draws the conclusions.

2. Related Works

This section provides a comprehensive review of existing COVID-19 detection models. Biswas et al. [9] aimed to determine a strong COVID-19 predictive method via chest CT images through effective TL methods. At first, they utilized three typical DL algorithms, such as Xception, VGG-16, and ResNet50, for COVID-19 prediction. Next, they presented a method to integrate the abovementioned pre-trained method for the general enhancement of the predictive capacity of the model. Ibrahim et al. [10] proposed a new computer-aided framework (COV-CAF) to categorize the severity level of the disease from three-dimensional CT Volumes. COV-CAF integrates conventional and DL methods. The presented COV-CAF method contains two stages: the preparatory stage and the feature analysis and classification stage. The feature analysis and classification stage integrates fuzzy clustering for feature fusion and automated RoI segmentation.

In Dansana et al. [11], the CNN approach is utilized to binary classification pneumonia-based transformation of Inception_V2, DT, and VGG-19 methods on CT scan and X-ray image data sets that have 360 images. It could gather that fine-tuned VGG-19, Inception_V2, and DT methods show outstanding performances with an increased rate of validation and training accuracy. Wang et al. [12] hypothesized that AI method that could extract certain graphical features of COVID-19 and offer medical diagnoses in advance of the pathogenic test, therefore saving critical time for controlling the disease. They gathered 1,065 CT images of pathogen-confirmed COVID-19 cases and persons who were diagnosed previously with standard COVID-19. They adapted the inception TL method for establishing the model, followed by external and internal validations.

Mei et al. [13] employed AI methods for integrating chest CT results with laboratory testing, medical symptoms, and exposure history to quickly analyze persons with positive for COVID-19. Goel et al. [14] presented a novel architecture for exploiting effective features extracted from the AE and GLCM, integrated with the RF model for the effective and faster diagnosis of COVID-19 with CT images. Mohammed et al. [15] presented an automatic CAD system for COVID-19-based chest X-ray image analyses. It is developed for COVID-19 diagnosis from another ARDS, MERS, and SARS infection. The optimum threshold values for chest images

segmentation are deduced by using Li's model and PSI method. Then, Laws' mask is employed in the chest image segmentation for highlighting secondary characteristics. Next, nine distinct vectors of features are extracted from the GLCM of every Laws' mask finding. The ensemble SVM methods are constructed according to the extracted feature vector. Munir et al. [16] presented a DNN method that is trained on the X-ray image of the COVID-19 and standard X-ray images to the COVID-19 diagnosis. Alquzi et al. [17] developed a result to detect persons with COVID-19 from CT images and ML models. This method is depending on a CNN model named EfficientNet.

3. The Proposed EDLFM-SI Technique

In this study, an effective EDLFM-SI technique is designed to detect and classify the SARS-CoV-2 infection or not. Also, the EDLFM-SI technique comprises various processes namely data augmentation, preprocessing, fusion-based feature extraction, WSA-based hyperparameter optimization, and CNN-based classification. At the same time, a fusion of CapsNet and MobileNet based feature extractors are employed. Figure 1 illustrates the overall process of the EDLFM-SI model. The working principle of every process is elaborated in the succeeding sections.

3.1. Preprocessing and Data Augmentation. Primarily, median filtering is applied for removing the noise present in the test CT images. Next, data augmentation comprises raising the number of training instances by the transformation of the images with no loss of semantic details. In this study, data augmentation takes place in three ways such as rotation, horizontal flip, and scaling.

3.2. Fusion-Based Feature Extraction. At this stage, the fusion-based feature extraction process is employed in which the fusion of MobileNet and CapsNet features is extracted.

3.2.1. MobileNet Model. The MobileNet V2 enhances efficiency of mobile techniques on several tasks and benchmarks and through the spectrum of various technique sizes. The basic principle behind MobileNet technique is the replacement of convolutional layers with depthwise separable convolution blocks where the depthwise convolution layer is trailed by the pointwise convolution layer to create effective feature vectors. It can be much greater than the regular convolutional with around similar outcomes. In MobileNet V2, all the blocks include 1×1 development layer from more depthwise and pointwise convolution layers. Different V1, the pointwise convolution layer of V2 recognized as the prediction layer projects information with the maximum amount of channels as to tensor with a considerably minimum amount of channels. MobileNetv2 is based on an inverted residual structure where the residual connections exist among the bottleneck layers. A 1×1 expansion convolution layer has increased the amount of channels

dependent upon expansion issue from the data previously as it goes to depthwise convolutions. The second novel thing from MobileNet V2's structure block has remaining linking [18]. The remaining linking uses the flow of gradient with networks.

Computation cost is considerably lower than the typical convolution with a compromise in slightly reduced accuracy.

3.2.2. CapsNet Model. For overcoming the limitations of CNN and generating it nearby the cerebral cortex activity framework, Hinton [19] presented a maximum dimension vector named as "capsule" for representing an entity (an object or part of object) with a set of neurons before a single neuron. All the capsules learn an implicit explanation of visual entity that output the probabilities of the entity and the group of "instantiated parameter containing the precise pose (place, size, and orientation), deforming, velocity, albedo, hue, texture, and so on.

The structure of CapsNet has been distinct in other DL techniques. The outcomes of input and output of CapsNets have been vectors whose norm and way demonstrate the existence probabilities and several attributes of entity correspondingly. If the several forecasts have been consistent, the higher level of one capsule is developed actively. Figure 2 depicts the framework of the CapsNet model. The structure has been shallow with only two convolution layers (Conv1, and PrimaryCaps) and one fully connected (FC) layer (EntityCaps). In detail, Conv1 has the typical convolution layer that alters images to initial features and outcomes to PrimaryCaps with a convolutional filter with a size of $13 \times 13 \times 256$. During the case where the original image is not appropriate to the input of the primary layer of the CapsNet, the rule feature then convolution was implemented.

The second convolution layer generates the equivalent vector framework as input of the capsule layer [20]. The typical convolutional of all output is scalars; however, the convolutional of PrimaryCaps has distinct from the classical one. It is considered 2-D convolutional of eight distinct weights to the input of $15 \times 15 \times 256$. The third layer (EntityCaps) has been the resultant layer that has nine typical capsules equivalent to nine distinct classes.

A layer of CapsNet has been separated into several computational units called capsules. Consider a capsule i with activity neuron i , it can be given as capsule j for generating activity level v_j of EntityCaps. The propagating and upgrading have been conducted utilizing vectors among PrimaryCaps and EntityCaps. The matrix model was employed to scalar input from all the layers of typical NN that is basically a linear combination of outcomes. The capsule modeling input has been separated into two phases: linear combination as well as routing. The linear combination represents an idea of modeling scalar input with NN that implies processing the connection among two objects from the scene with a visual alteration matrix but maintaining its relative relation. In detail, the linear combination was expressed as follows:

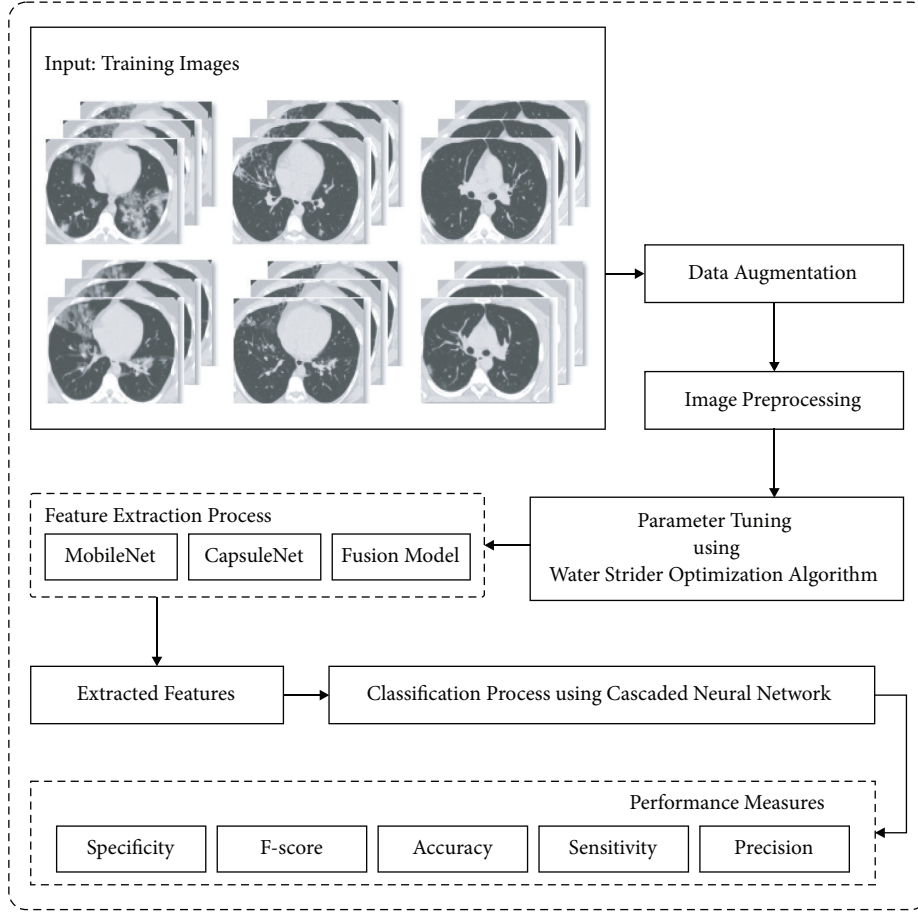


FIGURE 1: Overall block diagram of EDLFM-SI model.

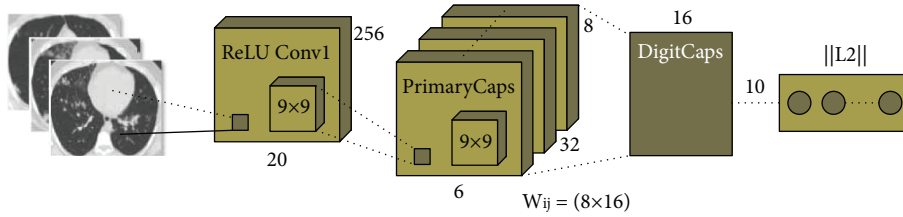


FIGURE 2: Structure of the CapsNet model.

$$\hat{u}_{j|i} = u_i W_{ij}, \quad (1)$$

where \hat{u} refers to the forecast vector created by changing the outcome u_i of the capsule from the layer under by weight W_{ij} . Afterward, during the routing phase, the input vector s_j of the capsule, j is determined as follows:

$$s_j = \sum_i c_{ij} \hat{u}_{j|i}, \quad (2)$$

where c_{ij} implies the coupling coefficient defined as the iterative dynamic routing procedure. The routing part comprises a weighted sum of \hat{u} coupling coefficients. The vector output of capsule j has computed by implementing a non-linear squashing function produces

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|}. \quad (3)$$

Noticeably, the capsule activation function essentially suppresses as well as redistributes vector length. Its output has been utilized as probabilities of entity signified as the capsule from the present type. The entire loss function of novel CapsNet has a weighted summation of marginal loss and reconstructing loss. The MSE has utilized from the novel reconstructing loss function that degrades this technique considerably if modeling noisy data.

3.2.3. Fusion Process. Data fusion is employed in many applications of ML and CV methods. Feature fusion is an

important task that integrates one or more feature vectors. The proposed method is dependent on feature

fusion through entropy. The two vectors are described as follows:

$$\begin{aligned} f_{\text{CapsNet}_{1 \times n}} &= \{\text{CapsNet}_{1 \times 1}, \text{CapsNet}_{1 \times 2}, \text{CapsNet}_{1 \times 3}, \dots, \text{CapsNet}_{1 \times n}\}, \\ f_{\text{EfficientNet}_{191 \times m}} &= \{\text{MobileNet}_{1 \times 1}, \text{MobileNet}_{1 \times 2}, \text{MobileNet}_{1 \times 3}, \dots, \text{MobileNet}_{1 \times m}\}. \end{aligned} \quad (4)$$

Besides, the extracted features are integrated into a single vector using the following equation:

$$\text{Fused (features vector)}_{1 \times q} = \sum_{i=1}^2 f_{\text{CapsNet}_{1 \times n}}, f_{\text{MobileNet}_{1 \times m}}, \quad (5)$$

where f represents fused vectors (1×1186). The entropy is utilized on features vectors for the selection of optimum features according to the score.

3.3. Hyperparameter Optimization. In order to optimally adjust the hyperparameters involved in the fusion model, the WSA is applied to it. The WSA is a population-based method that stimulates succession of water strider bugs, territorial behavior, feeding mechanism, mating style, and intelligent ripple communication. This method is described briefly in the following steps.

3.3.1. Initial Birth. The candidate solution/water strider (WS) is arbitrarily caused in the searching space as follows:

$$\begin{aligned} WS_i^0 &= Lb + \text{rand}, \\ (Ub - Lb); \quad & i = 1, 2, \dots, nws, \end{aligned} \quad (6)$$

where WS_i^0 represents the first position of i -th WS in the lake (search space). Lb and Ub represent lower and upper bounds, respectively. rand denotes an arbitrary value in the range of zero and one, and nws indicates the amount of WS_s (population size). The first position of WS_s is estimated by an objective function to evaluate the fitness.

3.3.2. Territory Establishment. To determine nt amount of territories, WS_s is arranged based on their fitness, and nws/nt amount of groups are generated orderly. The j -th member of all the groups is allocated to the j -th territory, where $j = 1, 2, \dots, nt$. Thus, the amount of WS_s lives in all the territories are equivalent to nws/nt . The position in all the territories with the best and worst fatness is considered female and male (keystone), respectively.

3.3.3. Mating. The male WS transmits ripple to female WS for mating. As the response of females is unknown, a probability p is determined for attraction or else repulsion [21]. The p is fixed to 0.5. The location of the male WS is upgraded as follows:

$$\begin{cases} WS_i^{t+1} = WS_i^t + R \cdot \text{rand}; & \text{if mating happens (with probability of } p), \\ WS_i^{t+1} = WS_i^t + R \cdot (1 + \text{rand}); & \text{otherwise.} \end{cases} \quad (7)$$

The length of R is estimated as follows:

$$R = WS_F^{t-1} - WS_M^{t-1}, \quad (8)$$

where WS_M^{t-1} and WS_F^{t-1} denotes the male and female WS in the $(t-1)^{th}$ cycle, respectively.

3.3.4. Feeding. Mating expends numerous energies for water strider, and the male WS forages to food afterward mating. During the latter scenario, the male WS move towards the

optimal WS of lake (WS) for finding foods based on the following equation:

$$WS_i^{t+1} = WS_i^t + 2\text{rand}^* (WS_{BL}^t - WS_i^t). \quad (9)$$

3.3.5. Death and Succession. In the novel location, the male WS could not find food; it would pass away; and a novel WS would replace it as follows:

$$WS_i^{t+1} = Lb_j^t + \text{rand}^*(Ub_j^t - Lb_j^t), \quad (10)$$

where Ub_j^t and Lb_j^t are the maximal and minimal values of WS^t s located inside the j -th territory.

3.3.6. WSA Termination. When the end criteria are met, the process would return to the mating step for a novel loop. Now, the maximum amount of function evaluation (MaxNFEs) is considered an end criterion.

3.4. CNN-Based Classification. Finally, the features are fed into the CNN model to allot the classes that exist in it. The perceptron linking that has been designed among the input and output has a procedure of direct relation, but FFNN linked generated among input and output was an indirect connection. The link was non-linear from shape with activation function from the hidden layer. When the link generated on perceptron and multilayer network has been joined, afterward, the network with direct link among the input and output layers is created. The network generated in this linking design was named CNN. The formulas are created in the CNN technique that is expressed as follows:

$$y = \sum_{i=1}^n f^i \omega_i^i x_i + f^o \left(\sum_{j=1}^k \omega_j^o f_j^h \left(\sum_{i=1}^n \omega_{ji}^h x_i \right) \right), \quad (11)$$

where f refers to the activation function in the input-output layers and ω_i^i implies the weight in the input-output layers [22]. When the bias has more than the input layers and activation function of all the neurons from the hidden layer is f^h , then

$$y = \sum_{i=1}^n f^i \omega_i^i x_i f^o \left(\omega^b + \sum_{j=1}^k \omega_j^o f \left(\omega_j^b + \sum_{j=1}^k \omega_j^o f_j^h \right) \right). \quad (12)$$

During this case, the CFNN technique was executed from the time sequences data. So the neurons from the input layer are the delays of time sequences data $X_{t-1}, X_{t-2}, \dots, X_{t-p}$, but the output has the present data X_t .

4. Performance Validation

4.1. Implementation Data. The experimental validation of the EDLFM-SI technique takes place using two benchmark data set, namely, SARS-CoV-2 CT scan [23] and COVID-19 CT [24, 25] data sets. The former contains a set of 2,482 CT scans with 1,252 scans under SARS-CoV-2 and 1,230 scans under other lung diseases. The next data set includes 746 CT images, with 349 CT images under COVID-19 and 397 CT images under non-COVID-19. Few sample images are demonstrated in Figure 3.

4.2. Result Analysis on SARS-CoV-2 CT Scan Data Set. Figure 4 demonstrates the confusion matrices produced by the EDLFM-SI technique on test data set-1. The results exhibited that the EDLFM-SI technique has identified the COVID-19 and non-COVID-19 images correctly under all



FIGURE 3: Sample images.

runs. For instance, with run-1, the EDLFM-SI technique has classified 1,234 images into COIVID-19 and 1,214 images into non-COVID-19. At the same time, with run-4, the EDLFM-SI approach has classified 1,241 images into COIVID-19 and 1,214 images into non-COVID-19. Followed by, with run-6, the EDLFM-SI method has classified 1,237 images into COIVID-19 and 1,216 images into non-COVID-19. Moreover, with run-8, the EDLFM-SI system has classified 1,236 images into COIVID-19 and 1,215 images into non-COVID-19. Furthermore, with run-10, the EDLFM-SI methodology has classified 1,238 images into COIVID-19 and 1,218 images into non-COVID-19.

Table 1 and Figure 5 provide the overall COVID-19 classification outcomes analysis of the EDLFM-SI technique on data set-1. The table depicted that the EDLFM-SI technique has the ability to classify images under all runs. For instance, with run-1, the EDLFM-SI technique has gained increased pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9872, 0.9856, 0.9870, 0.9863, and 0.9864, respectively. Along with that, with run-2, the EDLFM-SI system has reached enhanced pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9888, 0.9904, 0.9886, 0.9895, and 0.9896, respectively. In line with that, with run-6, the EDLFM-SI methodology has attained improved pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9888, 0.9880, 0.9886, 0.9883, and 0.9884, respectively. Followed by that, with run-8, the EDLFM-SI technique has gained increased pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9880, 0.9872, 0.9878, 0.9875, and 0.9876, respectively. Lastly, with run-10, the EDLFM-SI approach has achieved higher pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9904, 0.9888, 0.9902, 0.9895, and 0.9896, respectively.

Figure 6 showcases the accuracy graph analysis of the EDLFM-SI technique on the test data set 1. The figure revealed that the EDLFM-SI technique has resulted in maximum training and validation accuracies. It is observed that the EDLFM-SI technique has accomplished increased validation accuracy compared to training accuracy.

Next, the loss graph analysis of the EDLFM-SI technique under data set-1 takes place in Figure 7. The figure reported that the EDLFM-SI technique has attained minimal training and validation losses. It is also noticeable that the EDLFM-SI

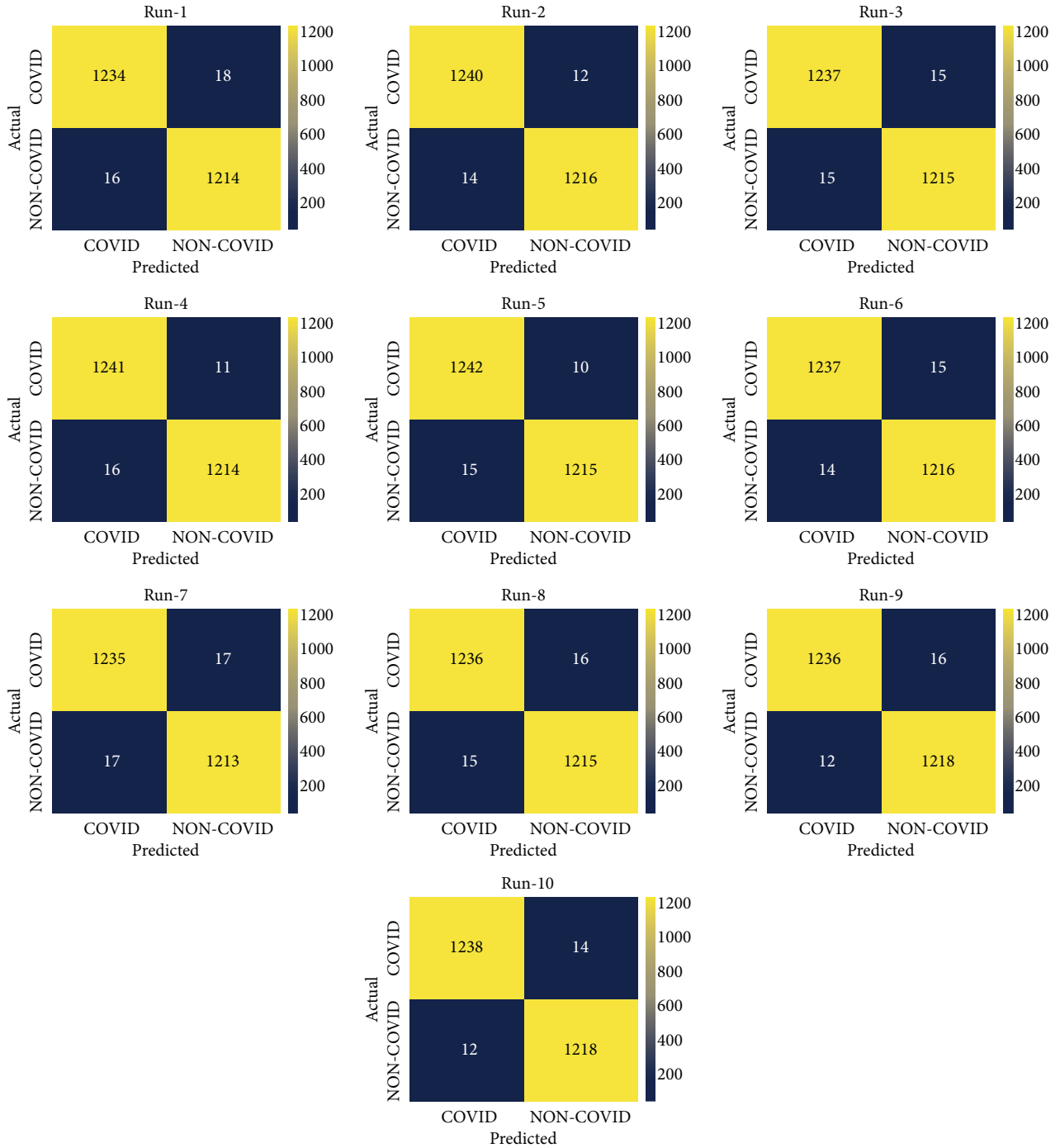


FIGURE 4: Confusion matrix of EDLFM-SI model under data set-1: (a) run-1, (b) run-2, (c) run-3, (d) run-4, (e) run-5, (f) run-6, (g) run-7, (h) run-8, (i) run-9, and (j) run-10.

TABLE 1: Result analysis of EDLFM-SI model with distinct measures under data set-1.

No. of runs	Precision	Sensitivity	Specificity	Accuracy	F-score
Run-1	0.9872	0.9856	0.9870	0.9863	0.9864
Run-2	0.9888	0.9904	0.9886	0.9895	0.9896
Run-3	0.9880	0.9880	0.9878	0.9879	0.9880
Run-4	0.9873	0.9912	0.9870	0.9891	0.9892
Run-5	0.9881	0.9920	0.9878	0.9899	0.9900
Run-6	0.9888	0.9880	0.9886	0.9883	0.9884
Run-7	0.9864	0.9864	0.9862	0.9863	0.9864
Run-8	0.9880	0.9872	0.9878	0.9875	0.9876
Run-9	0.9904	0.9872	0.9902	0.9887	0.9888
Run-10	0.9904	0.9888	0.9902	0.9895	0.9896
Average	0.9883	0.9885	0.9881	0.9883	0.9884

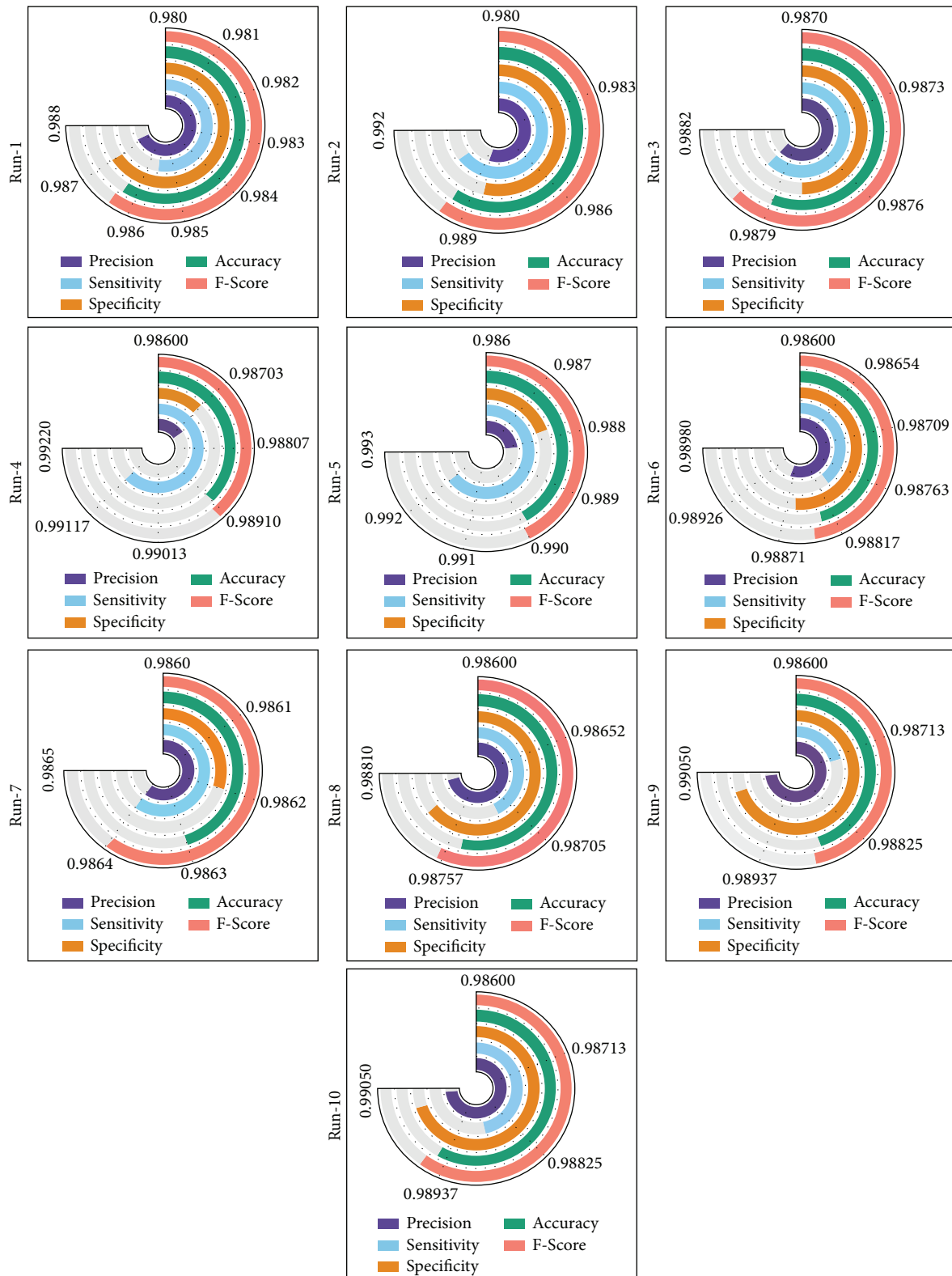


FIGURE 5: Result analysis of EDLFM-SI model under data set-1.

technique has resulted in reduced validation loss over the training loss.

A detailed comparative results analysis of the EDLFM-SI technique with recent techniques takes place on data set-1 in

Table 2 and Figure 8. The figure shows that the DT model has gained poor outcomes with the lower classification. At the same time, the GN, VGG-16, RN, and AN models have reached moderately closer classification performance. Along

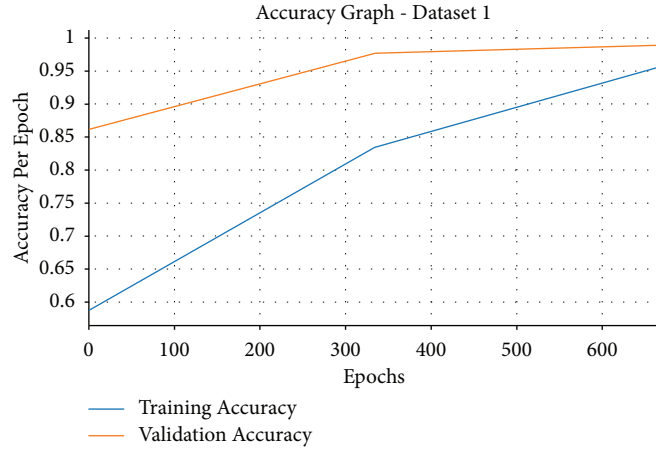


FIGURE 6: Accuracy analysis of EDLFM-SI model under data set-1.

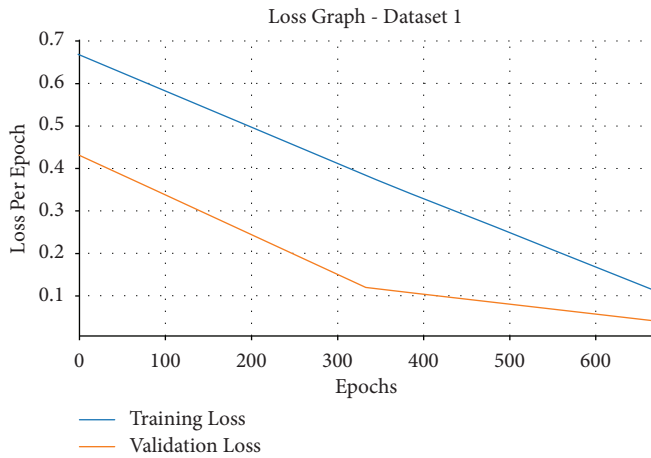


FIGURE 7: Loss analysis of EDLFM-SI model under data set-1.

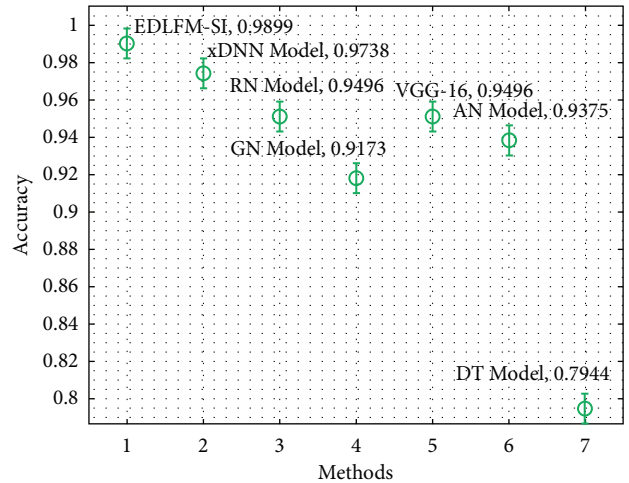


FIGURE 8: Comparative analysis of EDLFM-SI model under data set-1.

TABLE 2: Comparative analysis of EDLFM-SI model with existing approaches under data set-1.

Methods	Precision	Sensitivity	Accuracy	F-score
EDLFM-SI	0.9904	0.9920	0.9899	0.9900
xDNN model	0.0916	0.9553	0.9738	0.9731
RN model	0.9300	0.9715	0.9496	0.9503
GN model	0.9020	0.9350	0.9173	0.9182
VGG-16	0.9402	0.9543	0.9496	0.9497
AN model	0.9498	0.9228	0.9375	0.9361
DT model	0.7681	0.8313	0.7944	0.7984

with that, the xDNN model has accomplished reasonable classification performance over the other techniques. At last, the proposed EDLFM-SI technique has outperformed the other methods with the maximum pre_n , sen_y , acc_y , and F_{score} of 0.9904, 0.9920, 0.9899, and 0.9900, respectively.

4.3. Results Analysis on COVID-19 CT Data Set. Figure 9 exhibits the confusion matrices formed by the EDLFM-SI system on the test data set-2. The outcomes showcased that the EDLFM-SI manner has identified the COVID-19 and non-COVID-19 images correctly under all runs.

For sample, with run-1, the EDLFM-SI scheme has classified 331 images into COIVD-19 and 381 images into non-COVID-19. Likewise, with run-4, the EDLFM-SI algorithm has classified 335 images into COIVD-19 and 382 images into non-COVID-19. Similarly, with run-6, the EDLFM-SI technique has classified 333 images into COIVD-19 and 378 images into non-COVID-19. In addition, with run-8, the EDLFM-SI method has classified 332 images into COIVD-19 and 377 images into non-COVID-19. At last, with run-10, the EDLFM-SI approach has classified 331 images into COIVD-19 and 383 images into non-COVID-19.

Table 3 and Figure 10 offer the overall COVID-19 classification outcomes analysis of the EDLFM-SI approach on data set-2. The table outperformed that the EDLFM-SI system has the ability to classify images in all runs. For instance, with run-1, the EDLFM-SI approach has attained maximal pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9539, 0.9484, 0.9597, 0.9544, and 0.9511, respectively. At the same time, with run-4, the EDLFM-SI methodology has attained superior pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9571, 0.9599,

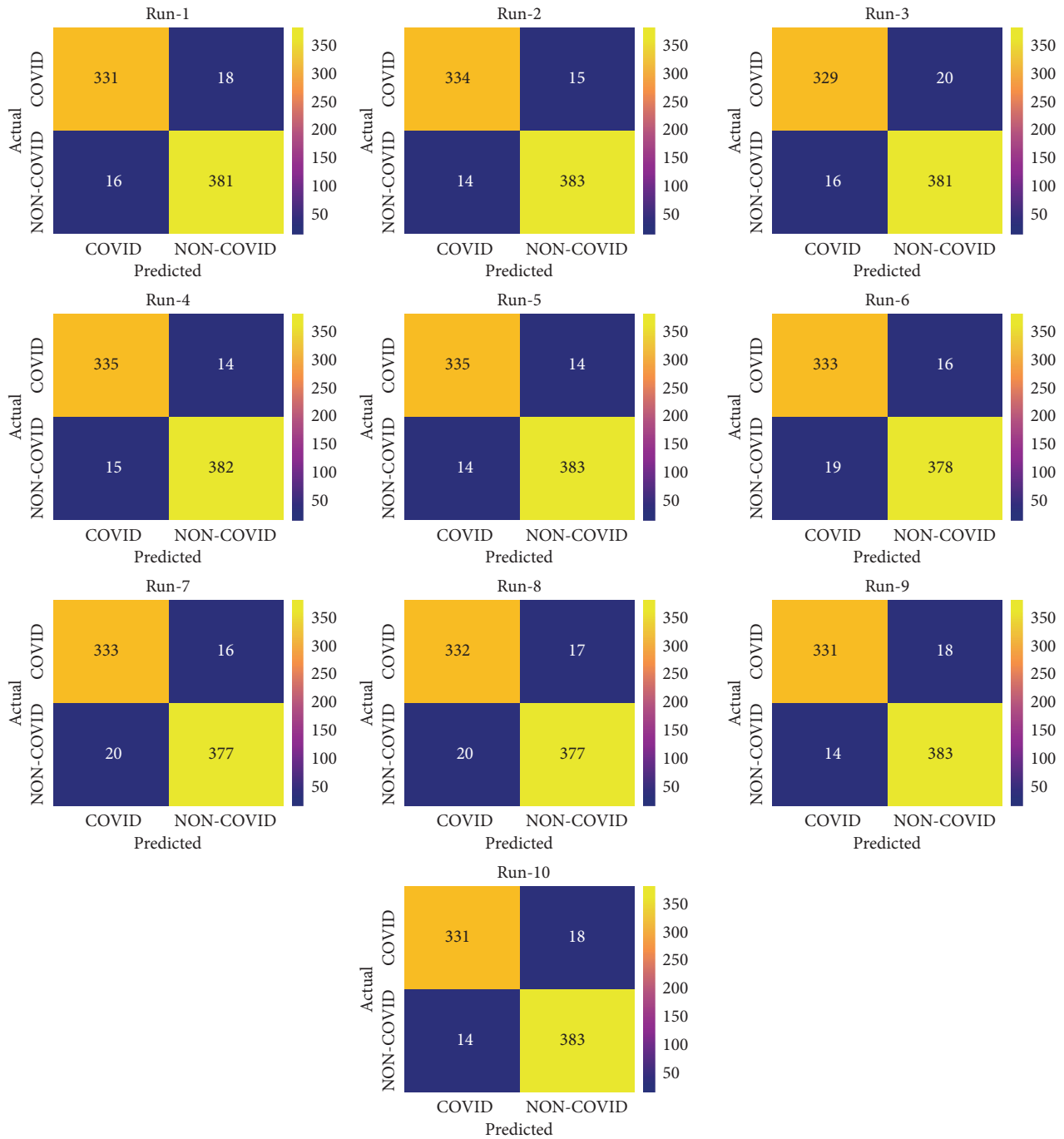


FIGURE 9: Confusion matrix analysis of EDLFM-SI model under data set-2: (a) run-1, (b) run-2, (c) run-3, (d) run-4, (e) run-5, (f) run-6, (g) run-7, (h) run-8, (i) run-9, and (j) run-10.

TABLE 3: Result analysis of EDLFM-SI model under data set-2.

No. of runs	Precision	Sensitivity	Specificity	Accuracy	F-score
Run-1	0.9539	0.9484	0.9597	0.9544	0.9511
Run-2	0.9598	0.9570	0.9647	0.9611	0.9584
Run-3	0.9536	0.9427	0.9597	0.9517	0.9481
Run-4	0.9571	0.9599	0.9622	0.9611	0.9585
Run-5	0.9599	0.9599	0.9647	0.9625	0.9599
Run-6	0.9460	0.9542	0.9521	0.9531	0.9501
Run-7	0.9433	0.9542	0.9496	0.9517	0.9487
Run-8	0.9432	0.9513	0.9496	0.9504	0.9472
Run-9	0.9594	0.9484	0.9647	0.9571	0.9539
Run-10	0.9594	0.9484	0.9647	0.9571	0.9539
Average	0.9536	0.9524	0.9592	0.9560	0.9530

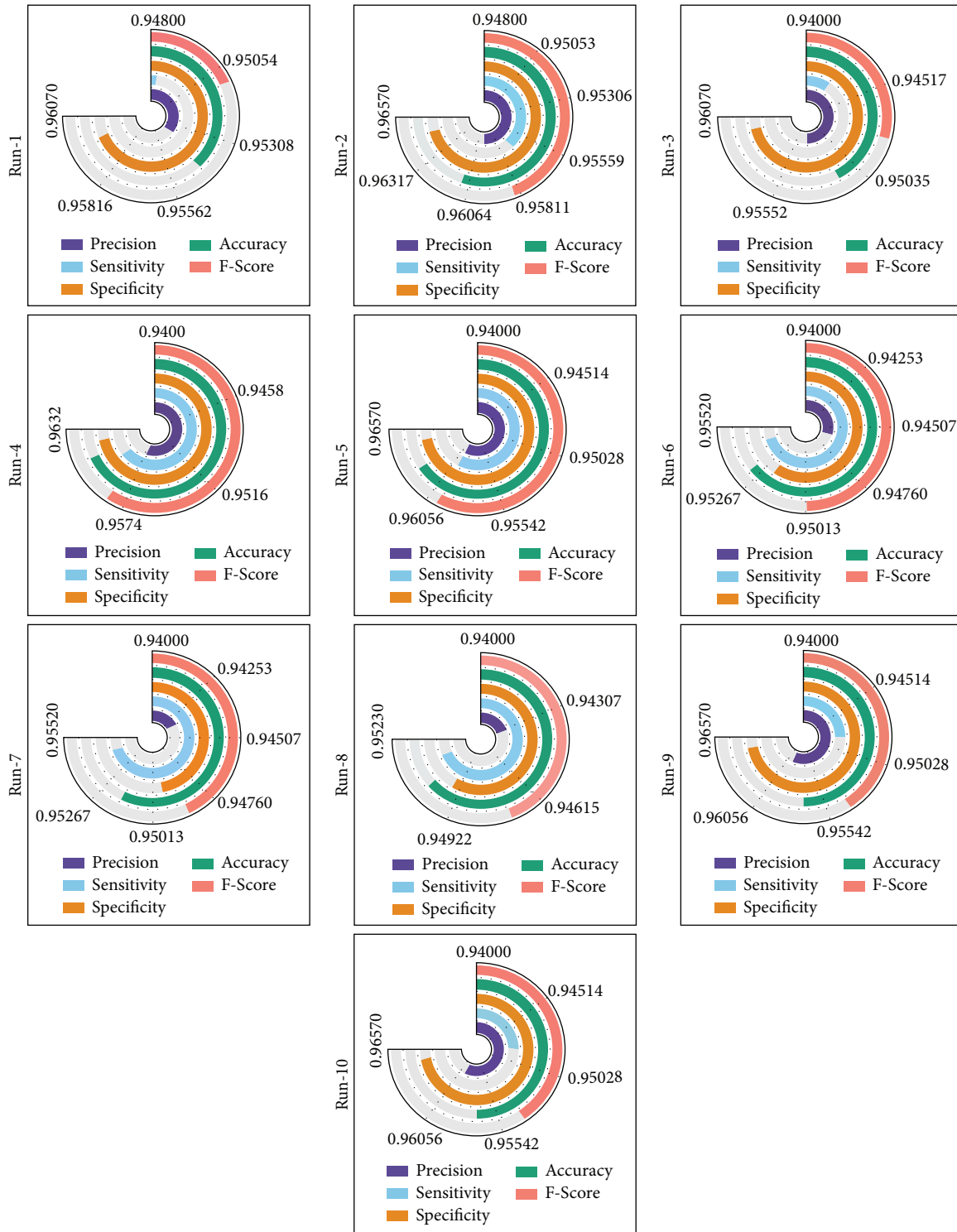


FIGURE 10: Result analysis of EDLFM-SI model under data set-2.

0.9622, 0.9611, and 0.9585, respectively. Besides, with run-6, the EDLFM-SI system has gained maximum pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9460, 0.9542, 0.9521, 0.9531, and 0.9501, respectively. Moreover, with run-8, the EDLFM-SI technique has gained higher pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9432, 0.9513, 0.9496, 0.9517, and 0.9487, respectively. Eventually, with run-10, the EDLFM-SI methodology has

gained improved pre_n , sen_y , spe_y , acc_y , and F_{score} of 0.9594, 0.9484, 0.9647, 0.9571, and 0.9539, respectively.

Figure 11 illustrates the accuracy graph analysis of the EDLFM-SI approach on the test data set 2. From the figure, it is obvious that the EDLFM-SI technique has resulted in maximal training and validation accuracies. It can be clear that the EDLFM-SI technique has

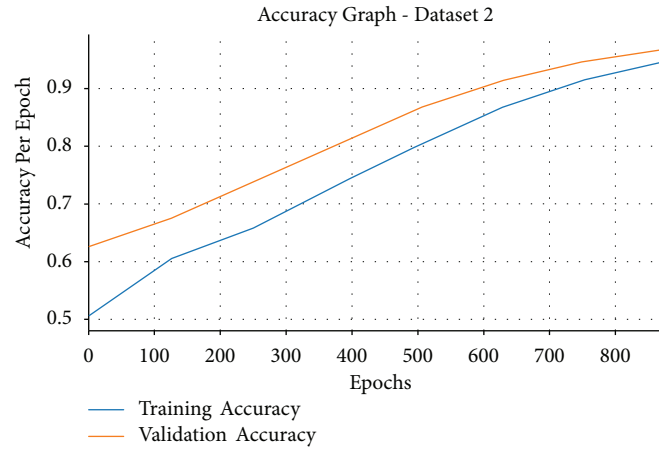


FIGURE 11: Accuracy analysis of EDLFM-SI model under data set-2.

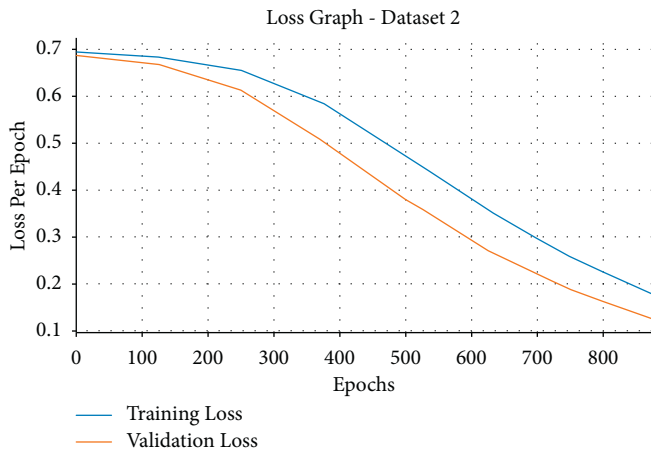


FIGURE 12: Loss analysis of EDLFM-SI model under data set-2.

accomplished increased validation accuracy related to training accuracy.

Then, the loss graph analysis of the EDLFM-SI approach on the test data set-2 takes place in Figure 12. The figure stated that the EDLFM-SI system has reached lesser training and validation losses. It can be also obvious that the EDLFM-SI methodology has resulted in decreased validation loss over the training loss.

A brief comparative outcomes analysis of the EDLFM-SI approach with recent systems takes place on data set-2 in Table 4 and Figure 13. The figure demonstrated that the Xception manner has attained worse results with minimum classification. Simultaneously, the DN-121, InceptionV3, RN-101, and DN-169 methods have obtained moderately closer classification performance. Also, the DN-201 model has accomplished reasonable classification performance over the other techniques. At last, the presented EDLFM-SI algorithm has outperformed the other methodologies with the maximal pre_n , sen_y , acc_y , and F_{score} of 0.9599, 0.9599, 0.9625, and 0.9060, respectively.

By looking into the detailed tables and figures, it is obvious that the EDLFM-SI technique has resulted in improved COVID-19 detection and classification performance over the recent methods.

TABLE 4: Comparative analysis of EDLFM-SI model with existing approaches under data set-2.

Methods	Precision	Sensitivity	Accuracy	F -score
EDLFM-SI	0.9599	0.9599	0.9625	0.9599
RN-101 model	0.8810	0.9310	0.9090	0.9060
InceptionV3	0.8770	0.9000	0.8940	0.8880
Xception	0.8730	0.8830	0.8850	0.8770
DN-121 model	0.8760	0.8880	0.8890	0.8820
DN-169 model	0.8810	0.9370	0.9120	0.9080
DN-201 model	0.9130	0.9370	0.9290	0.9250

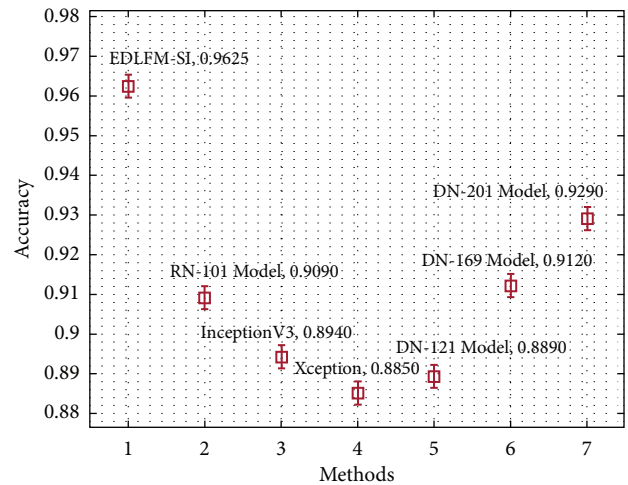


FIGURE 13: Comparative analysis of EDLFM-SI model under data set-2.

5. Conclusion

In this study, an effective EDLFM-SI technique is designed to detect and classify the SARS-CoV-2 infection for complex healthcare applications. Also, the EDLFM-SI technique comprises various processes, namely, data augmentation, preprocessing, fusion-based feature extraction, WSA-based hyperparameter optimization, and CNN-based classification. The fusion-based feature extraction process is

employed in which the fusion of MobileNet and CapsNet features is extracted. To optimally adjust the hyperparameters involved in the fusion model, the WSA was executed to it. Finally, the features are fed into the CNN model to allot the classes that exist in it. For examining the enhanced outcomes of the EDLFM-SI technique, a comprehensive experimental analysis is carried out on the COVID-19 CT data set and the SARS-CoV-2 CT scan data set. The simulation outcomes highlighted the supremacy of the EDLFM-SI technique over the recent approaches. As a part of the future scope, the classification performance of the proposed EDLFM-SI technique can be employed for SARS-CoV-2 detection by the use of hybrid metaheuristic-based optimization algorithms.

Data Availability

The data set used in this paper is publicly available at the following link: <https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset> and <https://github.com/UCSD-A14H/COVID-CT>.

Consent

Not applicable.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

All authors contributed equally to this work. All authors have given approval to the final version of the manuscript.

Acknowledgments

This study was supported by the grants of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health and Welfare (HI18C1216), the grant of the National Research Foundation of Korea (NRF-2020R1I1A1A01074256), and the Soonchunhyang University Research Fund.

References

- [1] H. Alshazly, C. Linse, E. Barth, and T. Martinetz, "Explainable covid-19 detection using chest ct scans and deep learning," *Sensors*, vol. 21, no. 2, p. 455, 2021.
- [2] F. S. Yildirim, M. Sayan, T. Sanlidag, B. Uzun, D. U. Ozsahin, and I. Ozsahin, "Comparative evaluation of the treatment of COVID-19 with multicriteria decision-making techniques," *Journal of Healthcare Engineering*, vol. 2021, pp. 1–11, Article ID 8864522, 2021.
- [3] S. Tello-Mijares and L. Woo, "Computed tomography image processing analysis in COVID-19 patient follow-up assessment," *Journal of Healthcare Engineering*, vol. 2021, Article ID 8869372, 12 pages, 2021.
- [4] M. Kaur, V. Kumar, V. Yadav, D. Singh, N. Kumar, and N. N. Das, "Metaheuristic-based deep COVID-19 screening model from chest X-ray images," *Journal of Healthcare Engineering*, vol. 2021, Article ID 8829829, 2021.
- [5] I. Ghinai, T. D. McPherson, J. C. Hunter et al., "First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in the USA," *The Lancet*, vol. 395, no. 10230, pp. 1137–1144, 2020.
- [6] N. Chen, M. Zhou, X. Dong et al., "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study," *The Lancet*, vol. 395, no. 10223, pp. 507–513, 2020.
- [7] T. Ai, Z. Yang, H. Hou et al., "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, 2020.
- [8] F. Shi, L. Xia, F. Shan et al., "Large-scale screening of covid-19 from community acquired pneumonia using infection size-aware classification," 2020, <https://arxiv.org/abs/2003.09860>.
- [9] S. Biswas, S. Chatterjee, A. Majee, S. Sen, F. Schwenker, and R. Sarkar, "Prediction of covid-19 from chest ct images using an ensemble of deep learning models," *Applied Sciences*, vol. 11, no. 15, p. 7004, 2021.
- [10] M. R. Ibrahim, S. M. Youssef, and K. M. Fathalla, "Abnormality detection and intelligent severity assessment of human chest CT scans using deep learning: a case study on SARS-COV-2 assessment," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–24, 2021.
- [11] D. Dansana, R. Kumar, A. Bhattacharjee et al., "Early diagnosis of COVID-19-affected patients based on X-ray and computed tomography images using deep learning algorithm," *Soft Computing*, pp. 1–9, 2020.
- [12] S. Wang, B. Kang, X. Zeng et al., "A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)," *European Radiology*, pp. 1–9, 2021.
- [13] X. Mei, H.-C. Lee, K.-y. Diao et al., "Artificial intelligence-enabled rapid diagnosis of patients with COVID-19," *Nature Medicine*, vol. 26, no. 8, pp. 1224–1228, 2020.
- [14] C. Goel, A. Kumar, S. K. Dubey, and V. Srivastava, *Efficient Deep Network Architecture for Covid-19 Detection Using Computed Tomography Images*, medRxiv, 2020.
- [15] S. Mohammed, F. Alkinani, F. Alkinani, and Y. Hassan, "Automatic computer aided diagnostic for COVID-19 based on chest X-Ray image and particle swarm intelligence," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 5, pp. 63–73, 2020.
- [16] K. Munir, H. Elahi, M. U. Farooq, S. Ahmed, F. Frezza, and A. Rizzi, "Detection and screening of COVID-19 through chest computed tomography radiographs using deep neural networks," *Data Science for COVID-19*, vol. 19, pp. 63–73, 2021.
- [17] S. Alquzi, H. Alhichri, and Y. Bazi, "Detection of COVID-19 using EfficientNet-B3 CNN and chest computed tomography images," in *Proceedings of the International Conference on Innovative Computing and Communications*, pp. 365–373, Springer, Singapore, August 2021.
- [18] A. Michele, V. Colin, and D. D. Santika, "Mobilenet convolutional neural networks and support vector machines for palmprint recognition," *Procedia Computer Science*, vol. 157, pp. 110–117, 2019.
- [19] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in *Proceedings of the 31st Conference Neural Information Process System*, pp. 1–11, Long Beach, CA, USA, December 2017.
- [20] H.-C. Li, W.-Y. Wang, L. Pan, W. Li, Q. Du, and R. Tao, "Robust capsule network based on maximum correntropy

- criterion for hyperspectral image classification,” *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 738–751, 2020.
- [21] A. Kaveh, M. I. Ghazaan, and A. Asadi, “An improved water strider algorithm for optimal design of skeletal structures,” *Periodica Polytechnica: Civil Engineering*, vol. 64, no. 4, pp. 1284–1305, 2020.
- [22] B. Warsito, R. Santoso, H. Suparti, and H. Yasin, “Cascade forward neural network for time series prediction,” *Journal of Physics: Conference Series*, vol. 1025, no. 1, Article ID 012097, 2018.
- [23] <https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset>.
- [24] <https://github.com/UCSD-AI4H/COVID-CT>.
- [25] X. He, X. Yang, S. Zhang et al., “Sample-Efficient Deep Learning for COVID-19 Diagnosis Based on CT Scans,” *MedRxiv*, pp. 1–10, 2020.

Research Article

Fusion-Based Deep Learning with Nature-Inspired Algorithm for Intracerebral Haemorrhage Diagnosis

Nada M. Alfaer ¹, Hassan M. Aljohani ¹, Sayed. Abdel-Khalek ^{1,2},
Abdulaziz S. Alghamdi ³ and Romany F. Mansour ⁴

¹Department of Mathematics and Statistics, College of Science, Taif University, Taif 21944, Saudi Arabia

²Department of Mathematics, Sohag University, Sohag 82524, Egypt

³Department of Mathematics, College of Science and Arts, King Abdulaziz University, P. O. Box 344, Rabigh 21911, Saudi Arabia

⁴Department of Mathematics, New Valley University, El-Kharga 72511, Egypt

Correspondence should be addressed to Abdulaziz S. Alghamdi; ashalghamedi@kau.edu.sa

Received 10 November 2021; Revised 10 December 2021; Accepted 14 December 2021; Published 18 January 2022

Academic Editor: K. Shankar

Copyright © 2022 Nada M. Alfaer et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Natural computing refers to computational processes observed in nature and human-designed computing inspired by nature. In recent times, data fusion in the healthcare sector becomes a challenging issue, and it needs to be resolved. At the same time, intracerebral haemorrhage (ICH) is the injury of blood vessels on the brain cells, which is mainly liable for stroke. X-rays and computed tomography (CT) scans are widely applied for locating the haemorrhage position and size. Since manual segmentation of the CT scans by planimetry by the use of radiologists is a time-consuming process, deep learning (DL) is used to attain effective ICH diagnosis performance. This paper presents an automated intracerebral haemorrhage diagnosis using fusion-based deep learning with swarm intelligence (AICH-FDLSI) algorithm. The AICH-FDLSI model operates on four major stages namely preprocessing, image segmentation, feature extraction, and classification. To begin with, the input image is preprocessed using the median filtering (MF) technique to remove the noise present in the image. Next, the seagull optimization algorithm (SOA) with Otsu multilevel thresholding is employed for image segmentation. In addition, the fusion-based feature extraction model using the Capsule Network (CapsNet) and EfficientNet is applied to extract a useful set of features. Moreover, deer hunting optimization (DHO) algorithm is utilized for the hyperparameter optimization of the CapsNet and DenseNet models. Finally, a fuzzy support vector machine (FSVM) is applied as a classification technique to identify the different classes of ICH. A set of simulations takes place to determine the diagnostic performance of the AICH-FDLSI model using the benchmark intracranial haemorrhage data set. The experimental outcome stated that the AICH-FDLSI model has reached a proficient performance over the compared methods in a significant way.

1. Introduction

In the last few years, traumatic brain injury (TBI) is the primary cause of growing death rates and disability in the USA. Nearly 30% of injury deaths have been reported [1]. After that, TBI, extra-axial intracranial tumours such as intracranial hemorrhages (ICH), may take place. The ICH disease is the main reason for death worldwide that happens for all ages. At first, the disease is initiated in the brain due to the leakage in the blood vessel and removes the path of interactions (follows the brain function and instruction consequently) and internal organ that results in inactive

body functions such as memory loss, loss of eyesight, speech, and so on [2–4]. The most important risk factors such as high blood pressure (BP), head trauma, leakage in veins, and infected blood vessel walls are related to the ICH. To inspect this disorder, the screening modalities such as single-photon emission computed tomography (SPECT), X-ray, positron emission tomography (PET), and computed tomography (CT) are accessed via brain haemorrhage imaging. In comparison to other methods, a CT scan is widely employed in haemorrhage diagnosis as it is widely available, limited duration, and inexpensive for imaging. Therefore, CT scans are highly desired for ICH detection. The manifestation of

ICH clots on CT scans depends on external factors such as volume, density, location, and slice intensity.

The early prediction of ICH is indispensable for sufficient scheduling of scanning and providing better treatment. Therefore, enormous designers have used the computer-based detection (CAD) method for ICH segmentation. The recently proposed computer-based CAD method of ICH is based on the aspects such as automatic segmentation of haemorrhage that can be forecasted without manual segmentation, professional contribution, in which the human experts have to offer a suitable input for segmentation. The current deployment in convolutional neural network (CNN) and deep learning (DL) served remarkable performances in automatic image segmentation and classification processes [5]. Thus, the DL technique can able to make automated ICH segmentation and prediction.

In recent times, researchers have attempted to employ the DL technique for the diagnosis of ICH on CT scans [6]. This DL technique is a kind of machine learning (ML) that employs various processing layers to learn a representation of data with many levels of abstraction. Earlier researchers utilizing this technique presented tremendous diagnostic performances to detect ICH in every single CT scan, same as that of expert radiotherapists. Additionally, the fully 3D DL method (not on single CT scans) for diagnosing ICH has been stated. Few researchers utilized the back-propagation (BP) model for the learning approach and the CNN that has pattern recognition and self-organization capacities without human programming. Consequently, this method is a problem agnostic and generic technique, not a problem-specific and rule-based model [7]. But it remains challenging to explicate how this technique generates the outcomes from the input data.

This paper presents an automated intracerebral haemorrhage diagnosis using fusion-based deep learning with swarm intelligence (AICH-FDLSI) algorithm. The AICH-FDLSI model employs a seagull optimization algorithm (SOA) with Otsu multilevel thresholding is employed for image segmentation. Besides, the fusion-based feature extraction using the Capsule Network (CapsNet) and EfficientNet is applied to extract a useful set of features. At the same time, deer hunting optimization (DHO) algorithm is utilized for the hyperparameter optimization of the CapsNet and DenseNet models. Lastly, a fuzzy support vector machine (FSVM) is employed as a classifier to determine various classes of ICH. To showcase the improved classifier results of the proposed model, a wide range of experiments is performed using the test benchmark intracranial haemorrhage data set.

The rest of the study is planned as follows. Section 2 provides the related works; Section 3 offers the proposed model; Section 4 discusses the performance validation; and Section 5 concludes the study.

2. Literature Review

Mansour et al. [8] proposed an innovative DL-based ICH diagnoses and classification (DL-ICH) method with the help of optimum image segmentation using inception network.

The presented method includes segmentation, preprocessing, classification, and feature extraction. First, the input data undergoes conversion format in which the NIfTI files are transformed into JPEG form. Anupama et al. [9] presented DL-based ICH diagnoses with GrabCut-based segmentation using SDL, called GC-SDL algorithm. Furthermore, GrabCut-based segmentation is utilized to identify the infected portion efficiently in an image. To execute the process of feature extraction, the SDL method is employed, and lastly, the SM layer is applied as a classifier.

Venugopal et al. [10] proposed a unique multimodal data fusion-based feature extraction method using a DL algorithm, called FFE-DL for ICH Classification and Detection, named as FFEDL-ICH. The presented method consists of classification, preprocessing, image segmentation, and feature extraction. First, the input images are preprocessed by the GF method for removing noise. Next, the DFCM method is employed for segmenting the image. Moreover, the fusion-based feature extraction method is performed by deep features (residual network 152) and handcrafted features (local binary patterns) for extracting appropriate features. Lastly, the DNN method is performed as a classification method to distinguish different types of ICH. A new DL method for ANN, totally distinct from the BP algorithm, was proposed in earlier research [11]. The objective is to measure the possibility of utilizing the model for ICH classification and detection of its subclasses, without applying the CNN method.

Wang et al. [12] focused on evaluating the accuracy and performance of a DL-based automatic segmentation method in segmenting spontaneous ICH volume either with/without IVH extensions. They related this automatic method with two manual segmentation methods. Ginat [7] examines the execution of DL for the work list prioritization and detection of acute ICH on NCCT in different medical sceneries at an academic medical centre. The images were categorized based on the type and presence of haemorrhage, whether this is follow-up/initial images, and patient visit location, involving outpatient, emergency or trauma, and inpatient sections. Yu et al. [13] intended to improve a strong DL segmentation technique for accurate and fast HV analyses via CT. Luong et al. [14] presented a CAD that integrates a DL method and image processing methods for determining patient who suffers from ICH because of their CT scans. The DL method-based MobileNetV2 framework was trained.

Ngo et al. [15] developed a newfangled method for training slice-level classifier on CT-based descriptor of the nearby slices alongside the axis; all of them are extracted by the CNN method. This technique focuses on predicting the existence of ICH and categorizes it into five distinct subclasses. They examine a two-phase training system. Initially, CT images are processed simply as a group of two-dimensional images, and an advanced CNN classifier is trained that is pretrained on ImageNet. In the training phase, all the slices are tested together with the three slices beforehand and the three slices afterward, which makes the batch size a multiple of 7. Next, the output descriptor of all the blocks of seven successive slices attained from phase 1 are stacked into images and fed into other CNNs for the last predictions of

middle slices. Hssayeni et al. [16] developed a method for collecting and eighty-two CT scan data sets of subjects with a traumatic brain injury. Then, the ICH regions were manually delineated in every slice by a consensus decision of two radiologists. The data set is an open-source platform at the PhysioNet repository for upcoming comparisons and analyses. Besides publishing the data set, that is, the major objective of this manuscript, they executed a deep FCN model called as UNet, for segmenting the ICH region from the CT images in a fully automatic methodology.

3. The Proposed Model

This paper has developed a novel AICH-FDLSI technique for ICH detection and classification. The proposed AICH-FDLSI technique encompasses MF-based preprocessing, SOA with Otsu multilevel thresholding-based segmentation, DHO-based feature extraction, and FSVM-based classification. The detailed working of these processes is offered in the succeeding sections.

3.1. Image Preprocessing. Primarily, the MF technique is applied as a preprocessing tool to eliminate the presence of noise involved in it. The MF is nonlinear statistical filtering that changes the existing pixel values with the median value of pixels under the adjacent area. A naive execution primary makes a cumulative histogram to the neighbor area and afterward defines the primary index elsewhere half the amount of pixels from the histograms. An essential issue of this manner on GPU is all the threads required for computing whole histograms. For 8-bit images, a histogram made of 256 bins is generated. It can be useless on present GP as there are not sufficient hardware registers obtainable to all the threads, and utilizing global memory to histogram calculation was too slow. For resolving this issue, the presented model depends upon a bisection search on histogram ranges. This technique does not calculate the actual histogram then iteratively improves the histogram range that contains the median value. In all rounds, the existing valid range was separated into two halves, and the half that is the huge amount of pixels is elected to the next iteration. This procedure was repeated till the range converged to a single bin.

3.2. Image Segmentation. During the image segmentation process, the SOA with Otsu multilevel thresholding is applied to determine the affected regions. The Otsu is also named as maximal difference between clusters [17]. An image histogram as fundamental and maximal difference between target as well as background as the selective condition, this technique obtained an optimum threshold from several cases. An image whose gray-scale range has $\{0, 1, \dots, L-1\}$ was separated as to destination and background by thresholds t . The possibility of gray i is p_i . The likelihood of objective has $\omega_0(t) = \sum_{i=0}^t p_i$. The possibility of background is $\omega_1(t) = \sum_{i=t+1}^{L-1} p_i$. The mean of objective is $u_0(t) = \sum_{i=0}^t i p_i / \omega_0$. The mean background is $u_1(t) = \sum_{i=t+1}^{L-1} i p_i / \omega_1$. The formulation of difference between two parts is $d(t) = \omega_0(t) \omega_1(t) (u_0(t) - u_1(t))^2$. An optimum

threshold t^* generates the difference maximal. Therefore, the process of multithreshold segmentation is as follows:

$$\begin{aligned} d(t_1, t_2, \dots, t_k) &= \omega_0 \omega_1 (u_0 - u_1)^2 + \omega_0 \omega_2 (u_0 - u_2)^2 \\ &+ \omega_0 \omega_3 (u_0 - u_3)^2 \dots + \omega_0 \omega_k (u_0 - u_k)^2 \\ &+ \omega_1 \omega_2 (u_1 - u_2)^2 + \omega_1 \omega_3 (u_1 - u_3)^2 \\ &+ \dots + \omega_1 \omega_k (u_1 - u_k)^2 \\ &+ \dots + \omega_{k-1} \omega_k (u_{k-1} - u_k)^2 \end{aligned}$$

$$\omega_{n-1}(t) = \sum_{i=t_{n-1}+1}^{t_n} p_i, u_{n-1}(t) = \frac{\sum_{i=t_{n-1}+1}^{t_n} i p_i}{\omega_{n-1}}, \quad 1 \leq n \leq (k+1). \quad (1)$$

Optimum thresholds $t_1^*, t_2^*, \dots, t_k^*$ create the entire difference maximal as defined below:

$$t_1^*, t_2^*, \dots, t_k^* = \text{Arg max}_{0 < t_1 < t_2 < \dots < t_k} d(t_1, t_2, \dots, t_k). \quad (2)$$

In this study, the optimal threshold values of the Otsu method are decided by the SOA. The SOA is based on the migration and attacking behavior of the seagulls in nature [18]. The mathematical model of attacking and migrating the prey is described below. The migration (exploration) method inspires how the group of seagulls moves everywhere. In this stage, the seagulls need to fulfill three criteria:

To prevent collision between neighbors (i.e., other seagulls), a further parameter A is applied for the assessment of the new search location as follows:

$$\vec{C}_s = A \times \vec{P}_s(x), \quad (3)$$

where \vec{C}_s signifies the location of search agent that does not collide with other searching agents, \vec{P}_s implies the existing location of the search agent, and x means the existing iteration as follows:

$$A = f_c - \left(x \times \left(\frac{f_c}{\text{Max}_{\text{iteration}}} \right) \right), \quad (4)$$

where $x = 0, 1, 2, \dots, \text{Max}_{\text{iteration}}$. f_c controls the frequency of A that is decreased gradually from f_c to 0. In this study, the value of f_c is set to 2. After evading the collision between neighbors, the searching agent is moving towards the direction of the optimal neighbor.

$$\vec{M}_s = B \times \left(\vec{P}_{bs}(x) - \vec{P}_s(x) \right). \quad (5)$$

Let \vec{M}_s be the position of searching agent \vec{P}_s towards the optimal fit searching agent \vec{P}_{bs} (viz., appropriate seagull). The behavior of B is randomly assigned, that is, accountable for appropriate balancing between exploitation and exploration. B is evaluated by

$$B = 2 \times A^2 \times r d, \quad (6)$$

where $r d$ represents an arbitrary value within [0.1]. Finally, the searching agent could upgrade its location regarding optimal search agent as follows:

$$\vec{D}_s = |\vec{C}_s + \vec{M}_s|. \quad (7)$$

Let \vec{D}_s be the distance between the optimal fit search agent and search agent (viz., optimal seagulls that fitness value is lesser). The exploitation focuses on exploiting the history and experience of the searching method. Seagulls are capable of changing the speed and angle of attack continuously in migration. They retain their altitude with their weight and wings. During prey attacking, the spiral movement behavior takes place in the air. The x , y , and z planes are shown as follows:

$$\begin{aligned} x' &= r \times \cos(k), \\ y' &= r \times \sin(k), \\ z' &= r \times k, \\ r &= u \times e^{kv}, \end{aligned} \quad (8)$$

where r indicates the radius of every turn of the spiral, k represents an arbitrary value within $[0 \leq k \leq 2A]$, u and v

denote constant for determining the spiral shape, and e represents the base of the natural logarithm. It can be evaluated by

$$\vec{P}_s(x) = (\vec{D}_s \times x' \times y' \times z') + \vec{P}_{bs}(x), \quad (9)$$

where $\vec{P}_s(x)$ saves the optimal solutions and upgrades the position of another search agent. The presented SOA initiates by an arbitrarily made population. The search agent might update their location regarding the optimum search agent in the iteration method. For smooth transition between exploitation and exploration, B is in charge. Therefore, the SOA is regarded as a global optimizer as a result of its good exploitation and exploration capacity.

3.3. Feature Extraction. Once the images are segmented, the next stage is to derive a fusion of feature vectors using the CapsNet and EfficientNet models. The two vectors can be defined as follows:

$$\begin{aligned} f_{\text{CapsNet}_{1 \times n}} &= \{\text{CapsNet}_{1 \times 1}, \text{CapsNet}_{1 \times 1} \times 2, \text{CapsNet}_{1 \times 1 \times 3}, \dots, \text{CapsNet}_{1 \times 1 \times n}\}, \\ f_{\text{EfficientNet}_{m \times m}} &= \{\text{EfficientNet}_{1 \times 1}, \text{EfficientNet}_{1 \times 2}, \text{EfficientNet}_{1 \times 3}, \dots, \text{EfficientNet}_{1 \times n}\}. \end{aligned} \quad (10)$$

In addition, the derived individual features are combined into a single vector, using the following equation:

$$\text{Fused (features vector)}_{1 \times q} = \sum_{i=1}^2 f_{\text{CapsNet}_{1 \times n}}, f_{\text{EfficientNet}_{1 \times m}}, \quad (11)$$

where f represents fused vectors (1×1186). The entropy is applied on the features vectors to choose optimal features based on the score to the classifier for differentiating the healthier and glioma images.

3.3.1. CapsNet Model. To address the limitations of CNN, Hinton [19] presented a higher dimension vector named ‘‘capsule’’ for representing an entity (object or a portion of object) by a set of neurons instead of an individual neuron. The activity of the neuron in the active capsule signifies different features of a certain entity, that is, existing in an image. Every capsule learns an implicit description of a visual entity that outputs the likelihood and a group of instantiated parameters that includes the accurate posture (orientation, position, and size), albedo, hue, texture, deformation, and so on. The framework of CapsNet is dissimilar to other DL methods. The outcomes of input and output of CapsNet are vector that direction and norm represent the various attributes and existence probability of the entity, correspondingly. The similar levels of capsule assist to forecast the instantiation parameter of a high-level capsule over a conversion matrix, and then dynamic routing is adapted for making the predictions reliable. Once the various predictions are

reliable, the high-level of the single capsule would turn out to be active.

A simple CapsNet framework has been demonstrated in Figure 1, where the framework is shallow by only one fully connected layer (EntityCaps) and two convolution layers (PrimaryCaps and Convl). Especially, Convl is the typical convolution layer that converts the output to PrimaryCaps and images to primary features via a convolutional filter with $13 \times 13 \times 256$ size. In case, the original images are not appropriate for the input of the primary layer of the CapsNet, and the primary feature afterward convolution is adapted. The next convolution layer creates the respective vector as input of the capsule layer. The standard convolutions of all the outputs are a scalar; however, the convolution of PrimaryCaps is dissimilar to the standard one. It is considered as a two-dimensional convolution of eight distinct weights for the input of $15 \times 15 \times 256$. The PrimaryCaps generate a thrity-two size of 11×11 steps to 2 convolutions and output. The third layer (EntityCaps) is the output layer, which has nine traditional capsules respective to nine distinct categories.

3.3.2. EfficientNet Model. The EfficientNet technique was utilized as a feature extraction component for generating a helpful group of feature vectors of the input satellite image [20]. The DL is the most well-known framework as DL approaches have been learned significant features in an input image at a different convolutional level similar to the purpose of the human brain. The DL was solving complex problems usually well as well as quickly with high classifier accuracy and lower error rate. The DL approach was

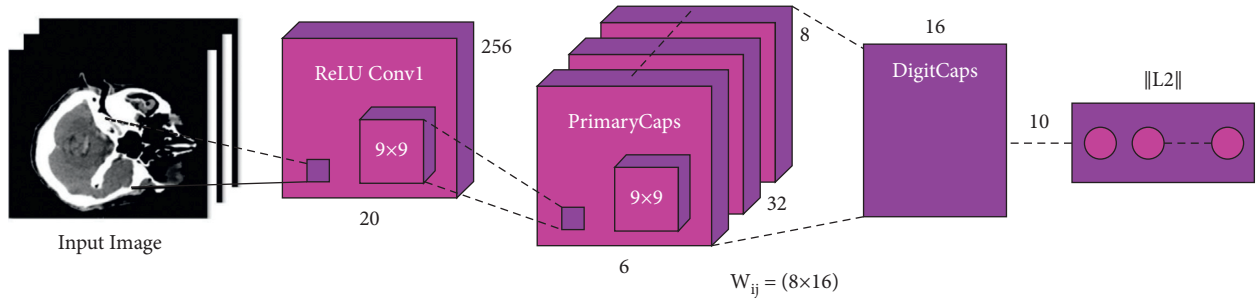


FIGURE 1: Process of CapsNet [19].

contained different modules (convolutional, pooling layer, and fully connected (FC) layers, and activation function). The DL models have the capability of attaining optimal performance over the machine learning models with high computational complexity. Distinct from other existing DL approaches, the EfficientNet structure was a compound scaling manner that employs the compound coefficients to uniformly scale network width, depth, and resolution. An EfficientNet has eight different methods from B0 to B7. The EfficientNet employs inverted bottleneck convolution which is primarily well-known from the MobileNetV2 approach that is a layer that primarily expands the network and next compresses the channel. This structure reduced computation with the factor of 2 as compared with normal convolution, where f signifies the filter size. It is depicted that EfficientNetB0 was the easiest of all eight approaches as well as employs minimal parameters. So it can be directly employed EfficientNetB0 to evaluate performance.

3.3.3. DHO-Based Hyperparameter Tuning. In this work, a new metaheuristic DHO method has been developed for the hyperparameter tuning process, stimulated from deer hunting by a group of hunters [21]. For deer hunting, the hunter encircles it as well as gets closer to them by using some strategies. This strategy includes the deliberation of several parameters, such as the deer position, wind angle, and so on. Cooperation between the hunters is another relevant standard that makes hunting very efficient. Lastly, they attain the target as per the location of the successor and leader. The objective function of this presented model is shown below:

$$f(x) = \max(\text{accuracy}). \quad (12)$$

The weight optimization with the DHO method is described as follows: because of the unique capabilities of deer, it could escape easily from hunting. The process initiates by a vector of an arbitrary population named hunter. It is described by the following equation:

$$X = \{X_1, X_2, \dots, X_m\} | 1 < j < m, \quad (13)$$

where m means the amount of hunter's population (weight), and the overall amount of weight employed to the optimization is denoted as follows. Next, the key parameters such as position angle (weight) and wind angle are employed. The whole searching space is deliberated as a

circle; hence, the wind angle can be defined as the circumference of the circle.

$$\theta_j = 2\pi a, \quad (14)$$

where the arbitrary value within $[0, 1]$ is denoted as a , and the existing iteration is signified as J . Now, θ implies the wind angle. Subsequently, the location propagation with the leader position (X_l) and successor location (X_s) for optimization is presented. The successor location defines the location of subsequent weights, while the leader location defines the primary location of the hunter.

Propagation via (X_l). Afterward initiating the optimal location, all the weights in the population try to attain the optimal location. Then, the location updating algorithm starts by modeling the encircling behavior as follows:

$$X_{j+1} = X_l - Y \cdot p \cdot |L \times X_l - X_j|. \quad (15)$$

Let X_j be the location at the existing iteration and the succeeding iteration location is denoted as X_{j+1} . The Z and K coefficient vectors are involved in this process. The arbitrary value, that is, presented by considering the wind speed is denoted as p , and it comprises values $f_{i,om0}$ to 2. The expression to estimate the Z and K coefficient vectors are given below:

$$Z = \frac{1}{4} \log \left(j + \frac{1}{j_{\max}} \right) b, \quad (16)$$

$$K = 2 \cdot c,$$

where the maximal iteration is denoted as j_{\max} . The b variable has values ranging from -1 to 1 , besides the value of other variables lies within $[0, 1]$. The first location of the hunter is signified as (X, Y) that gets upgraded according to the location of prey. Both Z and K coefficient vectors are modified to reach the optimal location (X_b, Y_b) . When the value of $p < 1$, the location updation algorithm takes place that implies the hunter could arbitrarily move in a different direction without considering the angle location. *Propagation through Angle Location.* The angle location updating is considered to rise the searching space. For making the hunting method more efficient, it is crucial to describe the angle location of the hunter. It can be implemented by

$$X_{j+1} = X_l - p \cdot \left| \cos(v) \times X_l - X_j \right|, \quad (17)$$

where p denotes the arbitrary values and the optimal location can be depicted as $B = \varphi_{j+1}, X_{b_j}$ and p . The individual location is found the opposite to the angle location; hence, the prey does not have any alertness of the hunter. *Propagation via Successor Location*. In the exploration, the vector K is presented within the encircle behavior. At first, the arbitrary searching method is performed by considering the K values as less than 1. Lastly, the location updating algorithm takes place based on a successor location instead of considering the optimal location. Next, the global searching is carried out by

$$X_{j+1} = X_s - Z \cdot p \cdot |K \times X_s - X_j|. \quad (18)$$

The location updating method is performed for identifying the optimal location (viz., termination condition).

3.4. Image Classification. At the final stage, the FSVM model is applied to determine the suitable class labels for the test images. In conventional SVM, each data point is regarded as equally significant and allotted a similar penal variable. Though in several real-time classification applications, few sample points, such as noises/outliers, may not be accurately allotted to one of these two classes, and all the sample points do not have a similar meaning to the decision surface. The hyperplanes in the SVM model are shown in Figure 2. To resolve this issue, the FSVM concept was initially presented [22]. Fuzzy membership to all the sample points is proposed so that discrete sample points might generate distinct contributions to the generation of decision surfaces. The trained sample is considered as follows:

$$S = \{(x_i, y_i, s_i), \quad i = 1, \dots, N\}. \quad (19)$$

Let $x_i \in R^n$ be the n -dimensional sample point, $y_i \in \{-1, +1\}$ signifies its class label, and s_i ($i = 1, \dots, N$) implies a fuzzy membership that fulfills $\sigma \leq s_i \leq 1$ with small constant $\sigma > 0$. The quadratic optimization problems for classification can be represented by

$$\begin{aligned} \min_{w, s, \xi} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^l s_i \xi_i, \\ \text{s.t.} \quad & y_i (w^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i = 1, \dots, l, \end{aligned} \quad (20)$$

where w indicates a standard vector of the separating hyperplane, b denotes a bias, and C represents a parameter that needs to be determined earlier to control the trade-offs amongst the cost of misclassification error and the classification margin. As s_i represent the attitude of the respective point x_i towards one class and the slack parameter ξ_i is a measure of error, the $s_i \xi_i$ term could consider the measure of error with discrete weights. It is considered that the larger the s_i is, the more significantly the respective point is treated; the lower the s_i is, the less outstandingly the respective point is treated. Hereafter, FSVM could discover a strong hyperplane by maximalizing the margin by letting some misclassification of lesser significant points.

For resolving the FSM problems, (2) is transformed into the subsequent dual problem by introducing Lagrangian multiplier α_i as follows:

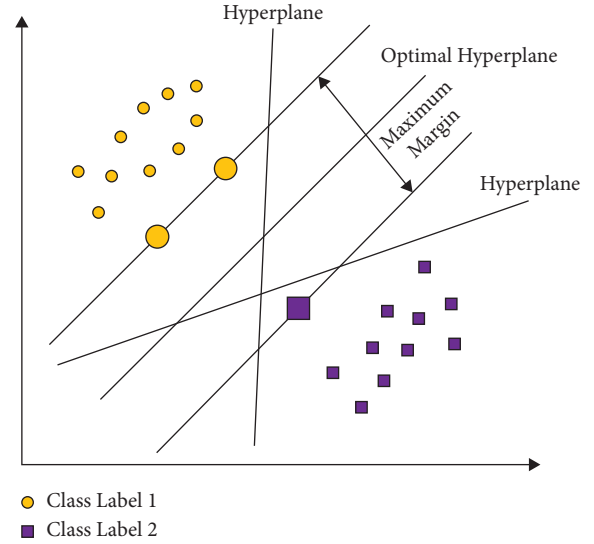


FIGURE 2: SVM hyperplanes.

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j x_i x_j, \\ \text{s.t.} \quad & \sum_{i=1}^N y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq s_i C, \quad i = 1, \dots, N. \end{aligned} \quad (21)$$

When compared to the typical SVM, the above stated has only a small difference, that is, the upper bounds of the value of α_i . By solving this dual problem in (3) for optimal α_i , w , and b could be recovered in the same way as in the typical SVM.

4. Performance Validation

The performance validation of the proposed model takes place using a benchmark CT ICH data set, including 341 images [23]. It comprises 171 images under epidural (EPI) class, 24 images under intraventricular (IVT), 72 images under intraparenchymal (IPC), 56 images under subdural (SBD), and 18 images under subarachnoid (SAD) class. The size of the image is 512 * 512 pixels. Figure 3 shows the sample test images. The data sets include ICH masks and CT scans, in JPG and NifTI format at PhysioNet repository. NifTI is a type of file format for neuroimaging, which is used very commonly in imaging informatics for neuroscience and even neuroradiology research.

Figure 4 showcases the confusion matrix of the AICH-FDLSI technique on the test images under run-1. The figure reported that the AICH-FDLSI technique has classified 19 images under IVT, 64 images under IPC, 12 images under SAD, 170 images under EPI, and 54 images under SBD.

Table 1 reports the ICH classification results analysis of the AICH-FDLSI technique under run-1. The results demonstrated that the AICH-FDLSI technique has classified the IVT class with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.7917, 0.9811, 0.7600, and 0.9677, respectively. In line with, the AICH-FDLSI technique has identified the IPC class with

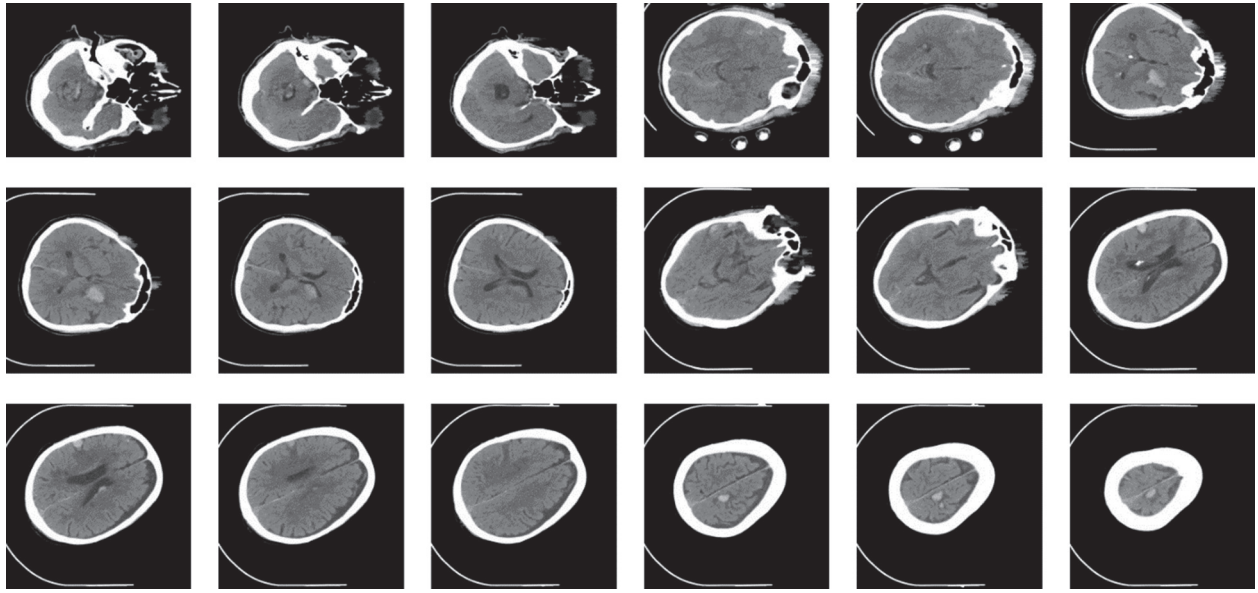


FIGURE 3: Sample images.

		Run-1					
		Intraventricular	Intraparenchymal	Subarachnoid	Epidural	Subdural	
Actual	Intraventricular	19	2	0	0	3	150 100 50 0
	Intraparenchymal	6	64	0	2	0	
	Subarachnoid	0	0	12	6	0	
	Epidural	0	1	0	170	0	
	Subdural	0	0	2	0	54	
			Intraventricular	Intraparenchymal	Subarachnoid	Epidural	
		Predicted					

FIGURE 4: Confusion matrix of AICH-FDLSI technique under run-1.

TABLE 1: ICH classification results analysis of AICH-FDLSI technique on test run-1.

Classes	Sensitivity	Specificity	Precision	Accuracy
IVT	0.7917	0.9811	0.7600	0.9677
IPC	0.8889	0.9888	0.9552	0.9677
SAD	0.6667	0.9938	0.8571	0.9765
EPI	0.9942	0.9529	0.9551	0.9736
SBD	0.9643	0.9895	0.9474	0.9853
Average	0.8611	0.9812	0.8950	0.9742

the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.8889, 0.9888, 0.9552, and 0.9677, respectively. Moreover, the AICH-FDLSI technique has identified the instances under SBD with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.9643, 0.9895, 0.9474, and 0.9853, respectively.

Figure 5 displays the confusion matrix of the AICH-FDLSI technique on the test images under run-2. The figure revealed that the AICH-FDLSI technique has identified 20

		Run-2					
		Intraventricular	Intraparenchymal	Subarachnoid	Epidural	Subdural	
Actual	Intraventricular	20	2	0	2	0	150 100 50 0
	Intraparenchymal	6	64	0	2	0	
	Subarachnoid	0	0	12	6	0	
	Epidural	0	1	0	170	0	
	Subdural	0	0	2	0	54	
			Intraventricular	Intraparenchymal	Subarachnoid	Epidural	
		Predicted					

FIGURE 5: Confusion matrix of AICH-FDLSI technique under run-2.

images under IVT, 64 images under IPC, 12 images under SAD, 170 images under EPI, and 54 images under SBD.

Table 2 offers the ICH classification results analysis of the AICH-FDLSI technique under run-2. The experimental values stated that the AICH-FDLSI technique has classified the IVT class with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.8333, 0.9811, 0.7692, and 0.9707, respectively. Moreover, the AICH-FDLSI technique has categorized the IPC class with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.8889, 0.9888, 0.9552, and 0.9677, respectively. Eventually, the AICH-FDLSI technique has determined the images under SBD with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.9643, 1.0000, 1.0000, and 0.9941, respectively.

Figure 6 demonstrates the confusion matrix of the AICH-FDLSI technique on the test images under run-3. The figure shows that the AICH-FDLSI technique has identified 22 images under IVT, 64 images under IPC, 12 images under SAD, 170 images under EPI, and 54 images under SBD.

TABLE 2: ICH classification results analysis of AICH-FDLSI technique on test run-2.

Classes	Sensitivity	Specificity	Precision	Accuracy
IVT	0.8333	0.9811	0.7692	0.9707
IPC	0.8889	0.9888	0.9552	0.9677
SAD	0.6667	0.9938	0.8571	0.9765
EPI	0.9942	0.9412	0.9444	0.9677
SBD	0.9643	1.0000	1.0000	0.9941
Average	0.8695	0.9810	0.9052	0.9754

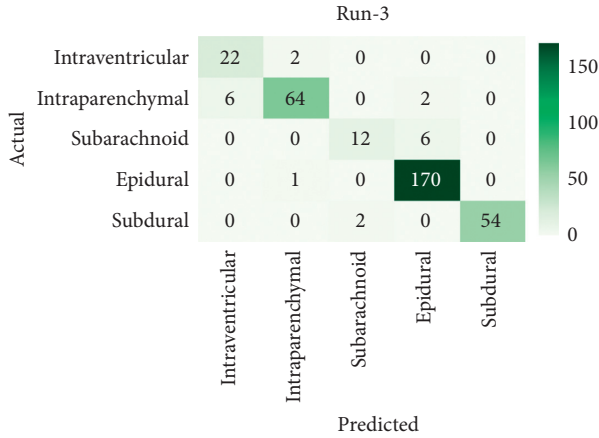


FIGURE 6: Confusion matrix of AICH-FDLSI technique under run-3.

Table 3 depicts the ICH detection performance analysis of the AICH-FDLSI technique under run-1. The results show that the AICH-FDLSI technique has effectively identified the IVT class with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.9167, 0.9811, 0.7857, and 0.9765, respectively. Meanwhile, the AICH-FDLSI technique has identified the IPC class with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.8889, 0.9888, 0.9552, and 0.9677, respectively. Lastly, the AICH-FDLSI technique has identified the instances under SBD with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.9643, 1.0000, 1.0000, and 0.9941, respectively.

Table 4 and Figure 7 offer an overall result analysis of the AICH-FDLSI technique under three different runs. The results show that the AICH-FDLSI technique has accomplished maximum classification performance under three test runs. For instance, under run-1, the AICH-FDLSI technique has classified the ICH with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.8611, 0.9812, 0.8950, and 0.9742, respectively. Likewise, under run-2, the AICH-FDLSI technique has classified the ICH with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.8695, 0.9810, 0.9052, and 0.9754, respectively. Similarly, under run-3, the AICH-FDLSI technique has classified the ICH with the $sens_y$, $spec_y$, $prec_n$, and $accu_y$ of 0.8861, 0.9833, 0.9106, and 0.9777, respectively.

Figure 8 investigates the accuracy graph of the AICH-FDLSI technique on the test data set. The figure demonstrated that the AICH-FDLSI technique has resulted in improved training and validation accuracies.

TABLE 3: ICH classification results analysis of AICH-FDLSI technique on test run-3.

Classes	Sensitivity	Specificity	Precision	Accuracy
IVT	0.9167	0.9811	0.7857	0.9765
IPC	0.8889	0.9888	0.9552	0.9677
SAD	0.6667	0.9938	0.8571	0.9765
EPI	0.9942	0.9529	0.9551	0.9736
SBD	0.9643	1.0000	1.0000	0.9941
Average	0.8861	0.9833	0.9106	0.9777

TABLE 4: Overall ICH results analysis of AICH-FDLSI technique.

No. of runs	Sensitivity	Specificity	Precision	Accuracy
Run-1	0.8611	0.9812	0.8950	0.9742
Run-2	0.8695	0.9810	0.9052	0.9754
Run-3	0.8861	0.9833	0.9106	0.9777

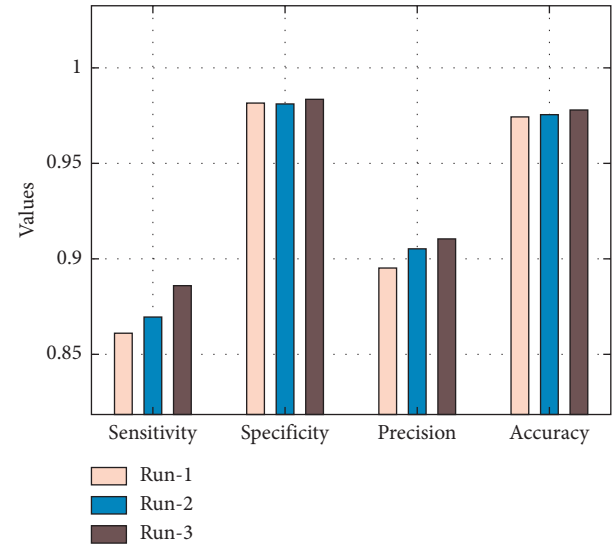


FIGURE 7: Results analysis of AICH-FDLSI technique under three different runs.

The loss graph analysis of the AICH-FDLSI technique takes place on the test data set in Figure 9. The results highlighted that the loss values tend to decrease with the increased epoch count, and it is observable that the validation loss seems to be lower than the training loss.

Table 5 provides a brief result analysis of the AICH-FDLSI with recent techniques. A brief $sens_y$ analysis of the AICH-FDLSI technique with existing approaches [16,24–27] is provided in Figure 10. The figure shows that the UNet, WANN, and SVM techniques have attained lower $sens_y$ values of 63.10%, 60.18%, and 76.38%, respectively. Eventually, the WEM-DCNN and convolutional NN techniques have resulted in reasonable $sens_y$ of 83.33%, and 87.06%, respectively. But the AICH-FDLSI technique has surpassed the other ones with the increased $sens_y$ of 88.61%.

A comparative $prec_n$ analysis of the AICH-FDLSI technique with other techniques is shown in Figure 11. The figure reported that the WANN and SVM techniques have

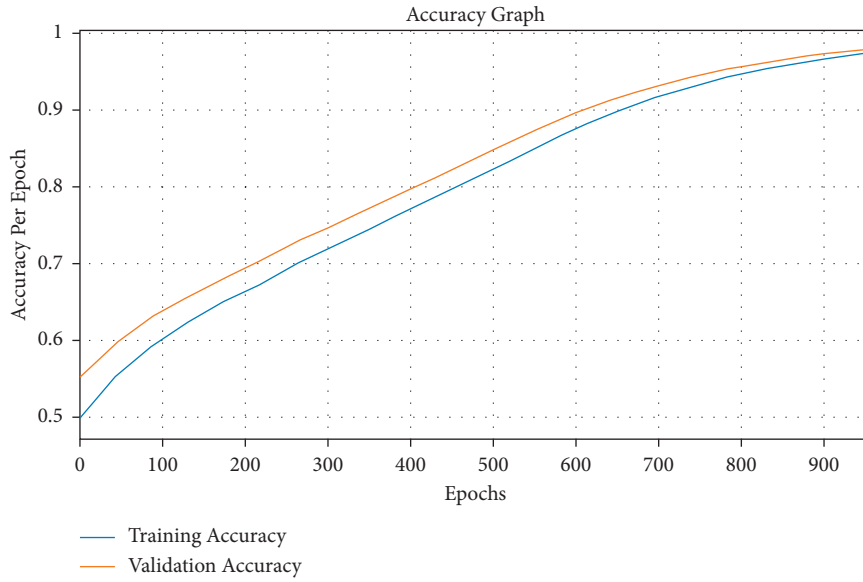


FIGURE 8: Accuracy graph analysis of AICH-FDLSI technique.

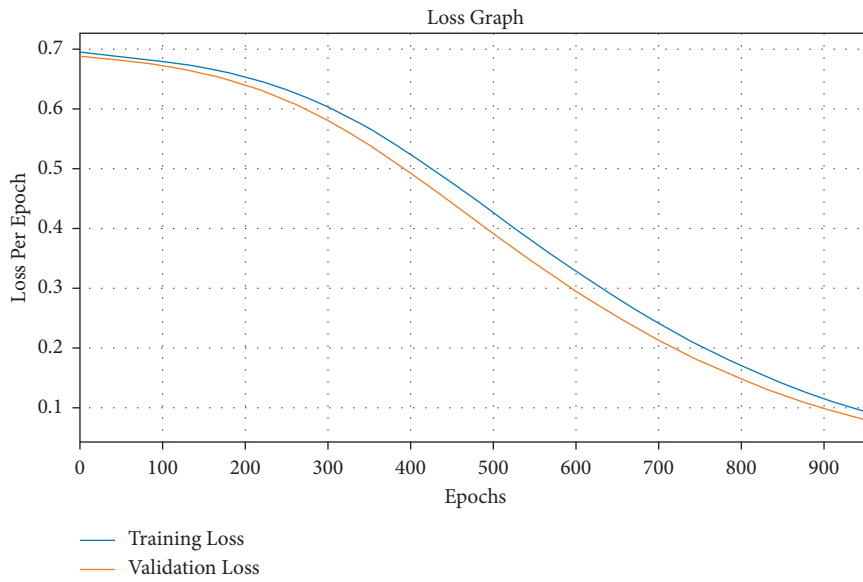


FIGURE 9: Loss graph analysis of AICH-FDLSI technique.

TABLE 5: Comparative results analysis of AICH-FDLSI with recent methods-I.

Methods	Sensitivity	Precision
AICH-FDLSI	88.61	91.06
UNet model	63.10	88.19
WANN	60.18	70.08
WEM-DCNN	83.33	89.90
Convolutional NN	87.06	87.98
SVM	76.38	77.53

attained lower $prec_n$ values of 70.08% and 77.53%, respectively. Along with that, the UNet, WEM-DCNN, and convolutional NN techniques have obtained moderate $prec_n$ of 88.19%, 89.90%, and 87.98%, respectively. However, the

AICH-FDLSI technique has outperformed the other ones with the higher $prec_n$ of 91.06%.

Table 6 offers a comparative results analysis of the AICH-FDLSI with recent techniques in terms of $spec_y$ and $accu_y$. Figure 12 depicts the comparative $spec_y$ analysis of the AICH-FDLSI system with other techniques. From the figure, it is notable that the UNet, WANN, Res-NexT, convolutional NN, and SVM techniques have accomplished minimal classification performance with the $spec_y$ values of 88.60%, 70.13%, 90.42%, 88.18%, and 77.53%, respectively. Next to that, the DN-ELM and WEM-DCNN techniques have resulted to reasonable $spec_y$ of 97.70%, and 97.48%, respectively. However, the AICH-FDLSI technique has gained improved performance with the superior $spec_y$ of 98.33%.

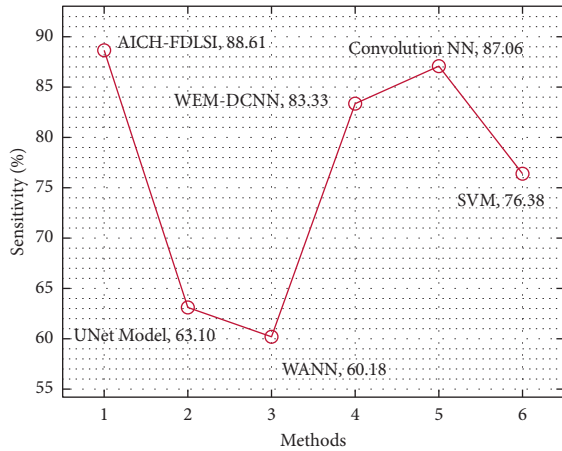


FIGURE 10: Comparative sens_y analysis of AICH-FDLSI technique.

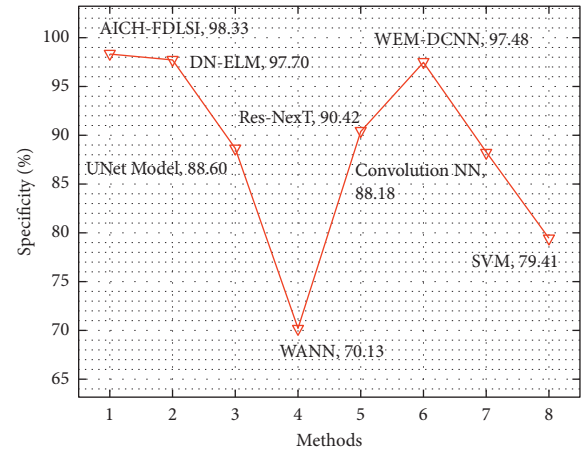


FIGURE 12: Comparative spec_y analysis of AICH-FDLSI technique.

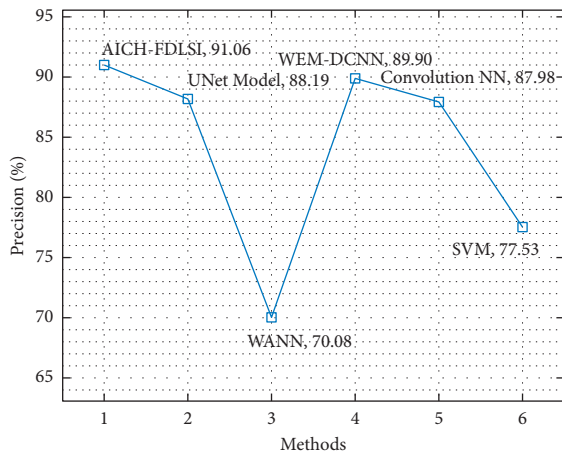


FIGURE 11: Comparative prec_n analysis of AICH-FDLSI technique.

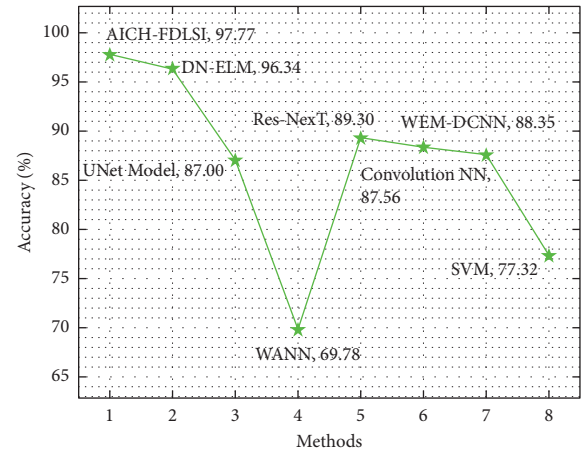


FIGURE 13: Comparative accu_y analysis of AICH-FDLSI technique.

TABLE 6: Comparative results analysis of AICH-FDLSI with recent methods-II.

Methods	Specificity	Accuracy
AICH-FDLSI	98.33	97.77
DN-ELM	97.70	96.34
UNet model	88.60	87.00
WANN	70.13	69.78
Res-NexT	90.42	89.30
WEM-DCNN	97.48	88.35
Convolutional NN	88.18	87.56
SVM	79.41	77.32

Figure 13 portrays the comparative accu_y analysis of the AICH-FDLSI system with other techniques. From the figure, it is notable that the UNet, WANN, Res-NexT, convolutional NN, WEM-DCNN, and SVM techniques have accomplished minimal classification performance with the accu_y values of 87%, 69.78%, 89.30%, 87.56%, 88.35%, and 77.32%, respectively. Following that, the DN-ELM model has offered competitive accu_y of 96.34%. But the AICH-FDLSI technique has surpassed the other ones with the maximum accu_y of 97.77%.

Finally, the CT analysis of the AICH-FDLSI methodology with recent approaches is shown in Figure 14. The

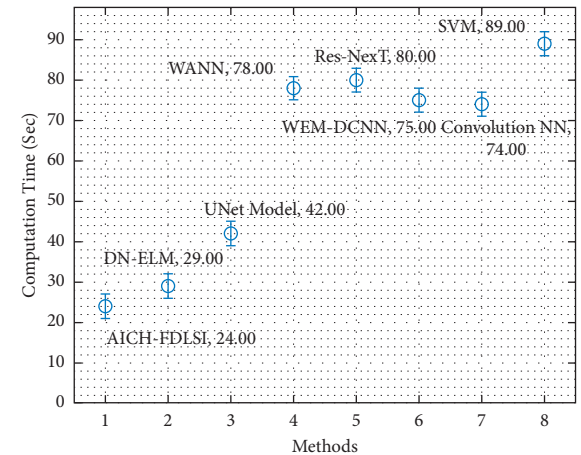


FIGURE 14: Comparative CT analysis of AICH-FDLSI technique.

results portrayed that the WANN, Res-NexT, and SVM models have obtained worse outcomes with maximum CT of 78 s, 80 s, and 89 s, respectively. Following that, the WEM-DCNN and convolutional NN techniques have attained moderately closer CT of 75% and 74%, respectively. Along with that, the DN-ELM and UNet models have obtained

reasonable CT of 29 s and 42 s, respectively. However, the AICH-FDLSI technique has accomplished improved performance with the CT of 24 s. From the above-mentioned results, it is evident that the AICH-FDLSI process is found to be an efficient tool for ICH detection and classification.

5. Conclusion

This paper has developed a novel AICH-FDLSI technique for ICH detection and classification. The proposed AICH-FDLSI technique encompasses MF-based preprocessing, SOA with Otsu multilevel thresholding-based segmentation, fusion-based feature extraction, DHO-based feature extraction, and FSVM-based classification. The application of SOA and DHO algorithms helps improvise the overall ICH classification performance. To showcase the improved classifier results of the proposed model, a wide range of experiments is performed using the test benchmark intracranial haemorrhage data set. The experimental outcome stated that the AICH-FDLSI model has reached a proficient performance. Therefore, the proposed AICH-FDLSI technique can be applied as a proficient tool for ICH diagnosis and classification. In the future, the ICH classification performance of the AICH-FDLSI technique can be improvised by the use of hybrid DL models.

Data Availability

The data set used in this paper is publicly available at <https://physionet.org/content/ct-ich/1.3.1/>

Ethical Approval

This article does not contain any studies with human participants performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgments

This paper was supported by Taif University Researchers Supporting Project (number: TURSP-2020/316), Taif University, Taif, Saudi Arabia.

References

- [1] A. Majumdar, L. Brattain, B. Telfer, C. Farris, and J. Scalera, "Detecting intracranial hemorrhage with deep learning," in *Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 583–587, IEEE, Honolulu, HI, USA, 2018, July.
- [2] M. R. Arbabshirani, B. K. Fornwalt, G. J. Mongelluzzo et al., "Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration," *NPJ digital medicine*, vol. 1, no. 1, pp. 9–7, 2018.
- [3] X. Gou and X. He, "Deep learning-based detection and diagnosis of subarachnoid hemorrhage," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9639419, 10 pages, 2021.
- [4] N. Kumar, N. Narayan Das, D. Gupta, K. Gupta, and J. Bindra, "Efficient automated disease diagnosis using machine learning models," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9983652, 13 pages, 2021.
- [5] L. Luo, X. Xu, Y. Jiang, and W. Zhu, "Predicting intracerebral hemorrhage patients' length-of-stay probability distribution based on demographic, clinical, admission diagnosis, and surgery information," *Journal of healthcare engineering*, vol. 2019, Article ID 4571636, 12 pages, 2019.
- [6] H. Chen, S. Khan, B. Kou, S. Nazir, W. Liu, and A. Hussain, "A Smart Machine Learning Model for the Detection of Brain Hemorrhage Diagnosis Based Internet of Things in Smart Cities," *Complexity*, vol. 2020, Article ID 3047869, 10 pages, 2020.
- [7] D. T. Ginat, "Analysis of head CT scans flagged by deep learning software for acute intracranial hemorrhage," *Neuroradiology*, vol. 62, no. 3, pp. 335–340, 2020.
- [8] R. F. Mansour and N. O. Aljehane, "An optimal segmentation with deep learning based inception network model for intracranial hemorrhage diagnosis," *Neural Computing & Applications*, vol. 33, no. 20, pp. 13831–13843, 2021.
- [9] C. S. S. Anupama, M. Sivaram, E. L. Lydia, D. Gupta, and K. Shankar, "Synergic Deep Learning Model-Based Automated Detection and Classification of Brain Intracranial Hemorrhage Images in Wearable Networks," *Personal and Ubiquitous Computing*, pp. 1–10, 2020.
- [10] D. Venugopal, T. Jayasankar, M. Yacin Sikkandar et al., "A novel deep neural network for intracranial haemorrhage detection and classification," *Computers, Materials & Continua*, vol. 68, no. 3, pp. 2877–2893, 2021.
- [11] J. Y. Lee, J. S. Kim, T. Y. Kim, and Y. S. Kim, "Detection and classification of intracranial haemorrhage on CT images using a novel deep-learning algorithm," *Scientific Reports*, vol. 10, no. 1, Article ID 20546, 2020.
- [12] T. Wang, N. Song, L. Liu et al., "Efficiency of a deep learning-based artificial intelligence diagnostic system in spontaneous intracerebral hemorrhage volume measurement," *BMC Medical Imaging*, vol. 21, no. 1, pp. 1–9, 2021.
- [13] N. Yu, H. Yu, H. Li, N. Ma, C. Hu, and J. Wang, "A Robust Deep Learning Segmentation Method for Hematoma Volumetric Detection in Intracerebral Hemorrhage," *Stroke*, p. 120, 2021.
- [14] K. G. Luong, H. N. Duong, C. M. Van et al., "A computer-aided detection to intracranial hemorrhage by using deep learning: a case study," in *Proceedings of the International Conference on Green Technology and Sustainable Development*, pp. 27–38, Springer, Cham, Da Nang City, Vietnam, November, 2020.
- [15] D. T. Ngo, D. B. Nguyen, H. T. Nguyen, H. H. Pham, and H. Q. Nguyen, "Slice-level Detection of Intracranial Hemorrhage on CT Using Deep Descriptors of Adjacent Slices, 2020.
- [16] M. D. Hssayeni, M. S. Croock, A. D. Salman, H. F. Al-khafaji, Z. A. Yahya, and B. Ghoraani, "Intracranial hemorrhage segmentation using a deep convolutional model," *Data*, vol. 5, no. 1, p. 14, 2020.
- [17] A. Akagic, E. Buza, S. Omanovic, and A. Karabegovic, "May. Pavement crack detection using Otsu thresholding for image segmentation," in *Proceedings of the 2018 41st International*

- Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 1092–1097, IEEE, Opatija, Croatia, May, 2018.
- [18] G. Dhiman and V. Kumar, “Seagull optimization algorithm: theory and its applications for large-scale industrial engineering problems,” *Knowledge-Based Systems*, vol. 165, pp. 169–196, 2019.
- [19] S. Sabour, N. Frosst, and G. E. Hinton, “Dynamic Routing between Capsules,” 2017, <https://arxiv.org/abs/1710.09829>.
- [20] G. Marques, D. Agarwal, and I. de la Torre Díez, “Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network,” *Applied Soft Computing*, vol. 96, Article ID 106691, 2020.
- [21] G. Brammya, S. Praveena, N. S. Ninu Preetha, R. Ramya, B. R. Rajakumar, and D. Binu, “Deer hunting optimization algorithm: a new nature-inspired meta-heuristic paradigm,” *The Computer Journal*, 2019.
- [22] Q. Fan, Z. Wang, D. Li, D. Gao, and H. Zha, “Entropy-based fuzzy support vector machine for imbalanced datasets,” *Knowledge-Based Systems*, vol. 115, pp. 87–99, 2017.
- [23] M. D. Hssayeni, M. S. Croock, A. D. Salman, H. F. Al-khafaji, Z. A. Yahya, and B. Ghoraani, “Computed Tomography Images for Intracranial Hemorrhage Detection and Segmentation,” *Intracranial Hemorrhage Segmentation Using A Deep Convolutional Model. Data*, vol. 5, no. 1, 2020.
- [24] S. Santhoshkumar, V. Varadarajan, S. Gavaskar, J. J. Amalraj, and A. Sumathi, “Machine learning model for intracranial hemorrhage diagnosis and classification,” *Electronics*, vol. 10, no. 21, p. 2574, 2021.
- [25] V. Davis and S. Devane, “Diagnosis & classification of brain hemorrhage,” in *Proceedings of the 2017 International Conference on Advances in Computing, Communication and Control (ICAC3)*, pp. 1–6, IEEE, 2017, December.
- [26] G. Danilov, K. Kotik, A. Negreeva et al., “Classification of intracranial hemorrhage subtypes using deep learning on CT scans,” *Studies in Health Technology and Informatics*, vol. 272, pp. 370–373, 2020.
- [27] M. Karki, J. Cho, E. Lee et al., “CT window trainable neural network for improving intracranial hemorrhage detection by combining multiple settings,” *Artificial Intelligence in Medicine*, vol. 106, Article ID 101850, 2020.

Research Article

Metaheuristics with Deep Learning-Enabled Parkinson's Disease Diagnosis and Classification Model

Adel A. Bahaddad ¹, Mahmoud Ragab ^{2,3,4}, Ehab Bahaudien Ashary ⁵
and Eied M. Khalil ^{4,6}

¹Information Systems Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

²Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

³Centre of Artificial Intelligence for Precision Medicines, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁴Mathematics Department, Faculty of Science, Al-Azhar University, Nasr City 11884, Cairo, Egypt

⁵Electrical and Computer Engineering Department, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁶Mathematics Department, College of Science, Taif University, Taif 21944, Saudi Arabia

Correspondence should be addressed to Mahmoud Ragab; mragab@kau.edu.sa

Received 11 November 2021; Revised 13 December 2021; Accepted 17 December 2021; Published 10 January 2022

Academic Editor: K. Shankar

Copyright © 2022 Adel A. Bahaddad et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Parkinson's disease (PD) affects the movement of people, including the differences in writing skill, speech, tremor, and stiffness in muscles. It is significant to detect the PD at the initial stages so that the person can live a peaceful life for a longer time period. The serious levels of PD are highly risky as the patients get progressive stiffness, which results in the inability of standing or walking. Earlier studies have focused on the detection of PD effectively using voice and speech exams and writing exams. In this aspect, this study presents an improved sailfish optimization algorithm with deep learning (ISFO-DL) model for PD diagnosis and classification. The presented ISFO-DL technique uses the ISFO algorithm and DL model to determine PD and thereby enhances the survival rate of the person. The presented ISFO is a metaheuristic algorithm, which is inspired by a group of hunting sailfish to determine the optimum solution to the problem. Primarily, the ISFO algorithm is applied to derive an optimal subset of features with a fitness function of maximum classification accuracy. At the same time, the rat swarm optimizer (RSO) with the bidirectional gated recurrent unit (BiGRU) is employed as a classifier to determine the existence of PD. The performance validation of the ISFO-DL model takes place using a benchmark Parkinson's dataset, and the results are inspected under several dimensions. The experimental results highlighted the enhanced classification performance of the ISFO-DL technique, and therefore, the proposed model can be employed for the earlier identification of PD.

1. Introduction

Parkinson's disease (PD) is a brain disorder that occurs as a consequence of the loss of brain cells. It mainly affects body mobility. Its symptom gradually becomes evident. Some of these symptoms that perform at early stages are tremors, slowness in movement, poor body posture, rigidity in muscles, deviation in speech, handwriting strokes, and imbalance [1]. In this disorder, a person's nerve cell

gradually loses their ability to communicate between them, which results in nervous system disorders such as depression. This disease must be diagnosed at earlier stages because it is incurable. When the accurate symptom of PD is recognized with their relative weightage, then doctors can suggest a pathology lab test for this feature and diagnosis might take place at an initial consultation itself. It will result in an earlier diagnosis of Parkinson's disease. The symptoms such as changes in speaking patterns and handwriting

strokes might assist in an earlier diagnosis of this disorder [2]. Erdogu Sakar and team lately received a speech dataset by examining the pronunciation of vowels “a” and “o” of disease-affected persons. Except speaking patterns, handwriting stroke patterns might help in detecting the disorder [3]. Factors studied for distinguishing a person from a healthier patient are individual age, fare handedness (right/left), maximum and mean distance among given summary in test, handwriting strokes noted in the drawing, and test time duration.

Recently, data have been improved by number of instances and numbers of features that make data noisier [4]. The noisier datasets could create the model to decrease the predicted accuracy, increase the computation cost, increase the complexity, and train the data slower. Therefore, feature selection developed an essential task for machine learning (ML) beforehand training the models [5]. The feature selection (FS), also known as attribute selection, is a method that focuses on finding a subset from the provided comprehensive set of features and fewer downgrades of the system performance; thus, the subsets of feature forecast the target with accuracy analogous to the performances of the original set of features and with the reducing computation costs. The FS method is categorized into wrapper-based and filter-based algorithms. The filter-based method utilizes a statistical method for finding the vital of all features (attributes). The wrapper-based method utilizes the machine-learning (ML) method. The wrapper-based method is computationally costly when compared to the filter-based method [6]. The wrapper method is additionally classified as heuristic search algorithm and sequential search algorithm.

An evolutionary algorithm is a part of artificial intelligence (AI) system that primarily focused on biological evolution. Biological evolution includes 4 major procedures such as selection, reproduction, mutation, and recombination [7]. Different from conventional optimization models, evolutionary algorithms depend on random sampling. This process is continuously employed on the solution officially reported as population, and the FF was employed for determining the quality of solutions. This solution changes based on the evolutionary procedure that finally assists to discover the global solution to the problems [8]. The evolutionary method has been recognized for performing well under distinct scenarios since it does not consider the fundamental fitness landscape. Even an easy evolutionary algorithm could easily resolve difficult challenges [9]. The only drawbacks in the evolutionary algorithm are the computational cost factor that is decreased by the fitness function calculation.

This study presents an improved sailfish optimization algorithm with deep learning (ISFO-DL) model for PD diagnosis and classification. The presented ISFO-DL technique designs an ISFO-based feature selection technique to derive an optimal subset of features with a fitness function of maximum classification accuracy. At the same time, the rat swarm optimizer (RSO) with the bidirectional gated recurrent unit (BiGRU) is employed as a classifier to determine the existence of PD. The experimental validation of the ISFO-DL model is carried out using a benchmark

Parkinson’s dataset, and the results are inspected under several dimensions.

The rest of the paper is arranged as follows. Section 2 offers the related works, Section 3 provides the proposed model, Section 4 inspects the performance validation, and Section 5 draws the conclusion.

2. Related Works

Huseyn [10] presented the DL methodology for realizing healthy people, analysis of PD, and multiple system atrophy. Oh et al. [11] employed the EEG signal of 20 PD and 20 standard subjects in this work. A 13-layer CNN framework could conquer the requirement for the traditional feature representation phases that are carried out. Wang et al. [12] introduced a novel deep-learning model for the earlier detection and classification of PD using the premotor features. In particular, to diagnose PD at earlier stages, various symptoms have been taken into account. Shahid and Singh [13] developed a DNN method with the decreased input feature space of Parkinson’s telemonitoring datasets for predicting PD evolution. PD is a progressive and chronic nervous system disorder, which impacts the motion of body. PD is measured by utilizing the unified PD rating scale (UPDRS).

Kaur et al. [14] surged a feasible medical decision-making method, which assists the medical professionals in detecting the PD-affected person. In this study, a certain architecture-based grid searching optimization method is presented for developing an enhanced DL algorithm to forecast the earlier diagnosis of PD; therefore, various hyperparameters are to be tuned and set for the assessment of DL algorithm. The grid searching optimization method includes its performance, the optimization of DL method, and the hyperparameters. In the study by Sivaranjini S. and Sujatha [15], an effort has been made for classifying the MR images of healthier control and PD subjects with the DL-NN model. The CNN framework AlexNet is utilized for refining the detection of PD. The MR image is tested to provide the accuracy measures and trained with the transfer learned network.

Quan et al. [16] presented a Bi-LSTM method for capturing the time-series dynamic feature of a speech signal to PD diagnosis. The dynamic speech feature is evaluated on the basis of energy content evaluation from the transition under voiced to unvoiced segments (offset) and the transition from unvoiced to voiced segments (onset). Sigcha et al. [17] proposed a novel methodology-based RNN and a single waist-worn triaxial accelerometer for enhancing the FOG recognition accuracy to be utilized in real home environment.

Leung et al. [18] focused on developing DL, an ensemble method for the prediction in person with PD. The initial and next phases of the method extracted features from DaTscan and medical measures of motor symptoms, respectively. Then, an ensemble of DNN model was trained on distinct subsets of the extracted feature for predicting the person results from 4 years afterward early baseline screening. Masud et al. [19] introduced an ACSA- and DL-based optimal FS technique. The presented method is the

integration of CROW Search and DL (CROWD) SSAE-NN. PD dataset has been taken for experimental purposes.

3. The Proposed ISFO-DL Model

In this study, the ISFO-DL technique has been developed for PD detection and classification. The proposed ISFO-DL technique is mainly intended to determine PD and thereby enhance the survival rate of the person. The presented ISF-DL technique involves three major processes namely ISFO-based feature selection, BiGRU-based classification, and RSO-based hyperparameter optimization. These three processes are elaborated in the succeeding sections.

3.1. Design of ISFO-Based Feature Selection Technique. At this stage, the ISFO algorithm is employed to choose an optimal subset of features and thereby boost the classifier results. Research has established that group hunting is the major social behavior in groups of fish, birds, mammals, and arthropods. In comparison with individual hunting, group hunting could save the energy utilization of the hunter to attain the aim of catching prey. Sailfish is employed for saving the present optimum solution, although sardines are applied in the searching space for finding an optimal solution. The arithmetical expression of the model is given as follows.

The population locations of sardines and sailfish are arbitrarily initiated, and every sardine and sailfish are allocated a randomized location $X_{SF(i)}^k$ and $X_{SD(j)}^k$, successively, where $i \in \{\text{sail fish}\}$, $j \in \{\text{sardines}\}$, and k represent the iteration count. The upgraded location of sailfish has been arithmetically given as follows:

$$X_{SF(i)}^{k+1} = X_{elite}^k - \mu_k \times \left(\text{rand}(0, 1) \times \frac{X_{elite}^k + X_{injure}^k}{2} - X_{SF(i)}^k \right), \quad (1)$$

$$\mu_k = 2 \times \text{rand}(0, 1) \times Pd - Pd, \quad (2)$$

$$Pd = 1 - \frac{\text{NumSP}}{\text{NumSF} + \text{NumSD}}. \quad (3)$$

Let $X_{SF(i)}^k$ be the preceding location of the i th sailfish, and μ_k indicates a coefficient created at k th iteration, using equation (2). To conserve the optimum solution of all the iterations, the sardine and sailfish with optimal fitness value are known as “elite” sailfish and “injured” sardine, respectively, and their location at iteration k is represented as X_{elite}^k and X_{injure}^k . Pd denotes the density of prey sardines that indicates the number of prey in all the iterations, as in equation (3). NumSF and NumSD stand for the population of sailfish and sardines [20], and the relation is $\text{NumSP} = \text{NumSD} \times \text{percent}$, in which percent characterizes the primary species of sailfish as a percentage of sardine populations.

A novel location of the sardines at k iteration is estimated as follows:

$$X_{SD(j)}^{k+1} = \text{rand}(0, 1) \times (X_{elite}^k - X_{SD(j)}^k + \text{ATK}), \quad (4)$$

$$\text{ATK} = A \times (1 - (2 \times \text{iter} \times \varepsilon)), \quad (5)$$

$$\alpha = \text{NumSD} \times \text{ATK}, \quad (6)$$

$$\beta = d \times \text{ATK}. \quad (7)$$

Here, $X_{SD(j)}^k$ signifies the preceding location of the j th sardine. iter denotes the amount of existing iterations. ATK means the sailfish attacking strength, i.e., decreased linearly on all the iterations given by equation (5). Once the $A = 4$ and $\varepsilon = 0.001$, if $\text{ATK} < 0.5$, the amount of sardines that upgrade the location (α) and the number of parameters of them (β) is evaluated by equations (6) and (7). When $\text{ATK} \geq 0.5$, each sardine gets upgraded.

For simulating the procedure of the sailfish catching sardines, when $f(\text{SD}_j) < f(\text{SF}_i)$, then the location of later can be substituted with the place of the sardine i , as follows:

$$X_{SF(i)}^k = X_{SD(j)}^k \text{ if } f(\text{SD}_j) < f(\text{SF}_i). \quad (8)$$

Chaotic mapping algorithms have both randomness and certainty and stochastic behavior and nonlinear motion. Chaos concept is the study of dynamic systems. The stimulating property of this system is that if there is a slight modification in the algorithm, the entire algorithm gets affected. The research has shown that the primary value of chaotic technique, the population of metaheuristic model, was initiated based on the relationship of chaotic mapping, and chaotic order was made, which could efficiently save the variety of populations and conquer the premature problems of traditional optimization method. Figure 1 illustrates the process flow of SFO technique.

The population initiation of sardines and sailfish in the SFO is a stochastic approach. It is based on population initiation while searching for an optimum solution. For enhancing the global searching capacity of the model and preventing the problems that the diversities of sardine and sailfish population reduce in late searches, hence we proposed a population initialization of sailfish and sardines using tent chaotic operator. The tent map can be described as follows:

$$T_{i+1} = \begin{cases} \frac{T_i}{0.7}, & T_i \leq 0.7, \\ \frac{1 - T_i}{0.3}, & T_i > 0.7. \end{cases} \quad (9)$$

In the equation, T_i denotes that the sequence of i th iteration ($T_i \in (0, 1)$) indicates the tent chaotic sequence distribution of T_n with the primary value $T_0 = 0.9$ in 200 iterations. Next, the sardine and sailfish populations are initiated:

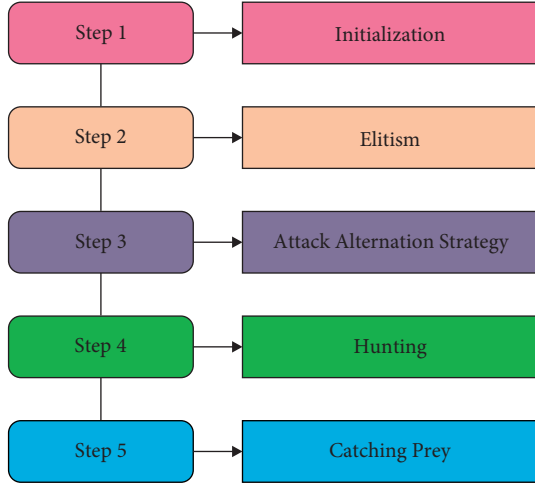


FIGURE 1: Process flow of SFO.

$$\begin{aligned} X_{SF(i+1)} &= T_{i+1} \times (X_{ub} - X_{lb}) + X_{lb}, \\ X_{SD(j+1)} &= T_{j+1} \times (X_{ub} - X_{lb}) + X_{lb}. \end{aligned} \quad (10)$$

While $X_{SF(i+1)}$ and $X_{SD(j+1)}$ indicate the location value of individual sardines and sailfish, X_{ub} and X_{lb} represent the upper and lower bounds of the individual sardines and sailfish in each dimension.

Assume the novel feature set be $\mathcal{F} = \{f_1, f_2, \dots, f_D\}$, where D implies the entire amount of features or dimension of feature set, and consider the class label be $C = \{c_1, \dots, c_l\}$, where l stands for the amount of classes. The FS technique determines a subset $S = \{s_1, \dots, s_m\}$, where $m < D$, $S \subset \mathcal{F}$, and S is minimal classification error rate than some other subsets of similar size or some appropriate subset of S . FS is the binary optimized issue, where the solution was restricted to binary values from 0 to 1. At this point, the solution has signified utilizing a binary vector where 1 refers that the equivalent feature was chosen and 0 demonstrates the equivalent feature is not chosen. The size of this vector was equivalent to the number of features from the original dataset. The ISFO was presented for solving continuous optimized issues in which the solution contains the real value. For mapping the continuous search space of typical ISFO to binary one, it can utilize a transfer function [21]. It can be utilized as a sigmoid transfer function and written as follows:

$$T(x) = \frac{1}{1 + e^{-x}}. \quad (11)$$

At this point, utilizing the probability value attained in equation (11), the present place of sailfish was upgraded by the following equation:

$$X^d(t) = \begin{cases} 1, & \text{if } \text{rnd} < T(X^d(t)), \\ 0, & \text{if } \text{rnd} \geq T(X^d(t)). \end{cases} \quad (12)$$

Usually, the FS is a multiobjective issue, with 2 objectives: (a) for achieving maximum classification accuracy (for

instance, maximized issue) and (b) for selecting minimal number of features (for instance, minimized issue). Using equation (15), these 2 objectives are joined and the FS issue was changed to single-objective issue.

$$\downarrow \text{Fitness} = \omega \gamma(S) + (1 - \omega) \frac{|S|}{D}, \quad (13)$$

where S stands for the chosen feature subset, $|S|$ defines the cardinality of chosen feature subset or the number of chosen features, $\gamma(S)$ signifies the classification error rate of S , D refers the novel dimensional of dataset, and $\omega \in [0, 1]$ signifies weight.

3.2. Design of the RSO-BiGRU-Based Classification Model.

During the classification process, the RSO-BiGRU model is applied to carry out the classification process. Learning is a continuous representation that is effective to control sequential data. An RNN is mostly appropriate to encoded sequential data. Figure 2 demonstrates the framework of BiGRU. During this analysis, it can utilize BiGRU for learning [22]. The computation of BiGRU was separated into 2 parts: forward and reverse order data broadcasts. To provide sentence $X = (x_1, x_2, \dots, x_n)$, $x \in \mathbb{R}^k$, x refers the concatenating vector of present word and place, and the forward GRU was computed as follows:

$$i = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i), \quad (14)$$

$$f = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f), \quad (15)$$

$$g = \tanh(W_{xg}x_t + W_{hg}(i \odot h_{t-1}) + b_g), \quad (16)$$

$$h_t = (1 - f) \odot h_{t-1} + f \odot g, \quad (17)$$

where W_* and b_* signify the weight matrix and bias vectors, respectively; σ refers the sigmoid functions; and \odot stands for the element-wise multiplication. x_t implies the input word vector at time steps \underline{t} , and h_t signifies the hidden state of current time step r . h_i and h_i demonstrate the outcome of forward and backward GRUs, respectively. The BiGRU output is represented as follows:

$$h_i^{\text{bi-gru}} = \left[\vec{h}_i; \overleftarrow{h}_i \right]. \quad (18)$$

To effectively tune the hyperparameters involved in the BiGRU model, the RSO is applied to it.

The rats are territory animals that live from the set of combined males and females. The performance of rats is very aggressive from several analyses that are outcome from the death of any animals. This aggressive performance is a vital simulation of this work but chase and fight with prey. The chasing and fighting behavior of the rats can be used to model the RSO algorithm and can be utilized to solve optimization problems. This subsection explains the performance of rats, for instance, chasing and fighting. Afterward, the presented RSO technique is summary.

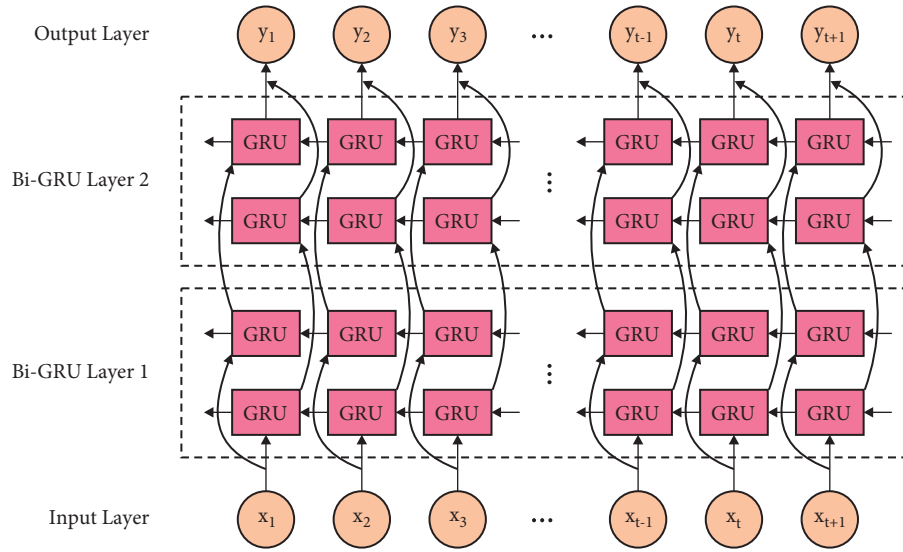


FIGURE 2: Structure of BiGRU.

3.2.1. Chasing the Prey. In general, the rats are social animals to chase the prey under the set with situation social agonistic efficiency. For defining this efficiency mathematically, it can be assumed that optimum search agents have skill of place of the prey. Another search agent has upgraded its places in terms of optimum search agents attained so far. The subsequent formulas are presented under this process:

$$\vec{P} = A \cdot \vec{P}_i(x) + C \cdot (\vec{P}_r(x) - \vec{P}_i(x)), \quad (19)$$

where $\vec{P}_i(x)$ demonstrates the places of rats and $\vec{P}_r(x)$ signifies the better optimum solutions.

However, A and C parameters were calculated as follows:

$$A = R - x \times \left(\frac{R}{\text{Max}_{\text{Iteration}}} \right) \text{ where, } x = 0, 1, 2, \dots, \text{Max}_{\text{Iteration}}, \quad (20)$$

$$C = 2 \cdot \text{rand}. \quad (21)$$

So, R and C imply the arbitrary numbers among $[1, 5]$ and $[0, 2]$, respectively. The parameters A and C are responsible for optimum exploration and exploitation over the course of rounds.

3.2.2. Fighting with Prey. For mathematically defining the fight procedure of rats with prey, the subsequent formula was projected:

$$\vec{P}_i(x+1) = \left| \vec{P}_r(x) - \vec{P} \right|, \quad (22)$$

where $\vec{P}_i(x+1)$ implies the upgraded next places of rat. It stores the optimum solution and upgrades the places of other search agents in terms of optimum search agent. The rat (A, B) upgraded their place nearby the place of prey (A^*, B^*). By altering the parameters as revealed in equations (20) and (21), the distinct amount of places is achieved on the present place [23]. Also, this technique is comprehensive

from n -dimensional environments. Consequently, the exploration and exploitation have been guaranteed using the value of parameters A and C . The projected RSO technique stores optimum solutions with many operators.

4. Performance Validation

This section inspects the PD classification result analysis of the presented IFSO-DL technique. The results are investigated against four datasets namely HandPD Spiral, HandPD Meander, Speech PD, and Voice PD [24–26]. Table 1 and Figure 3 offer the selected features attained by the IFSO-DL technique with other FS methods. The results show that the IFSO-DL technique has chosen the least number of features compared with other FS techniques on all test datasets. For instance, with the HandPD Spiral dataset with 13 features, the IFSO-DL technique has selected a set of 4 features, whereas the MGOA, MGWO, and OCFA techniques have chosen a total of 5, 7, and 8 features, respectively.

Likewise, with the HandPD Meander dataset with 13 features, the IFSO-DL system has selected a set of 6 features, whereas the MGOA, MGWO, and OCFA methods have chosen a total of 8, 8, and 7 features, respectively. Meanwhile, with the Speech PD dataset with 23 features, the IFSO-DL system has selected a set of 10 features, whereas the MGOA, MGWO, and OCFA techniques have chosen a total of 11, 12, and 13 features, respectively. Eventually, with the Voice PD dataset with 26 features, the IFSO-DL manner has selected a set of 7 features, whereas the MGOA, MGWO, and OCFA algorithms have chosen a total of 8, 9, and 17 features, respectively.

Table 2 offers a detailed comparative result analysis of the IFSO-DL technique with recent methods on the test HandPD Spiral dataset. The results show that the MGWO-KNN and MGOA-KNN techniques have obtained lower accuracy of 0.734 and 0.756, respectively. In line with this, the MGOA-DT technique has attained moderate accuracy of

TABLE 1: Selected features of existing with the proposed model.

Dataset	Total features	MGOA	MGWO	OCFA	IFSO-DL
HandPD Spiral	13	5	7	8	4
HandPD Meander	13	8	8	7	6
Speech PD	23	11	12	13	10
Voice PD	26	8	9	17	7

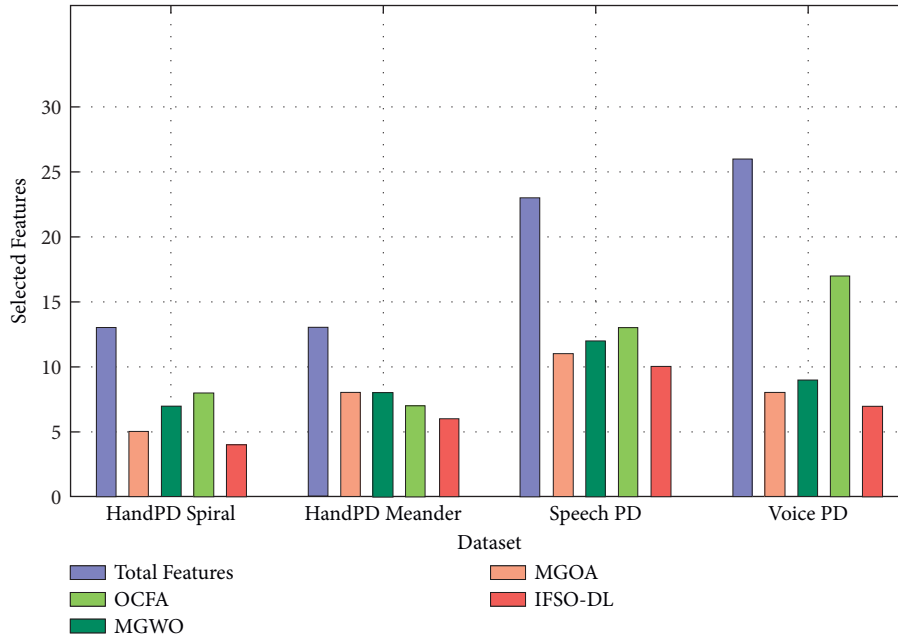


FIGURE 3: FS analysis of the IFSO-DL technique with 4 datasets.

TABLE 2: Result analysis of existing with the proposed IFSO-DL model on the HandPD Spiral dataset.

Methods	Accuracy	DR	FAR
MGOA-KNN	0.756	0.853	0.531
MGOA-RF	0.929	0.979	0.219
MGOA-DT	0.890	0.947	0.281
MGWO-KNN	0.734	0.819	0.500
MGWO-RF	0.924	0.940	0.119
MGWO-DT	0.924	0.940	0.119
IFSO-DL	0.933	0.982	0.080

0.890. At the same time, MGOA-RF, MGWO-RF, and MGWO-DT techniques have accomplished reasonable accuracy of 0.929, 0.924, and 0.924, respectively. However, the IFSO-DL technique has outperformed the other techniques with the maximum accuracy, DR, and FAR of 0.933, 0.982, and 0.080, respectively.

Figure 4 demonstrates the accuracy of graph analysis of the IFSO-DL technique on the test HandPD Spiral dataset. The figure portrays that the IFSO-DL technique has gained increased training and validation accuracies. It is noted that the IFSO-DL technique has accomplished improved validation accuracy over the training accuracy.

The loss graph analysis of the IFSO-DL technique is investigated in Figure 5. The figure shows that the IFSO-DL technique has accomplished enhanced outcomes with the lower validation loss compared with training loss. It also

demonstrates that the IFSO-DL technique has obtained reduced validation loss compared with training loss.

Table 3 suggests a detailed comparative outcome analysis of the IFSO-DL technique with recent approaches on the test HandPD Meander dataset. The results outperformed that the MGWO-KNN and MGOA-KNN systems have obtained minimum accuracy of 0.728 and 0.748, respectively. Afterward, the MGOA-DT manner has gained moderate accuracy of 0.890. Also, MGOA-RF, MGWO-RF, and MGWO-DT systems have accomplished reasonable accuracy of 0.937, 0.930, and 0.880, respectively. However, the IFSO-DL method has exhibited the other methodologies with the maximal accuracy, DR, and FAR of 0.940, 1.000, and 0.135, respectively.

Figure 6 reveals the accuracy graph analysis of the IFSO-DL manner on the test HandPD Meander dataset. The figure shows that the IFSO-DL technique has reached improved training and validation accuracies. It can be clear that the IFSO-DL algorithm has accomplished improved validation accuracy over the training accuracy.

The loss graph analysis of the IFSO-DL system is studied in Figure 7. The figure portrays that the IFSO-DL technique has accomplished enhanced outcomes with the lower validation loss related to training loss. It also outperforms that the IFSO-DL technique has gained lower validation loss related to training loss.

Table 4 provides a brief comparative outcome analysis of the IFSO-DL system with recent approaches on the test Speech

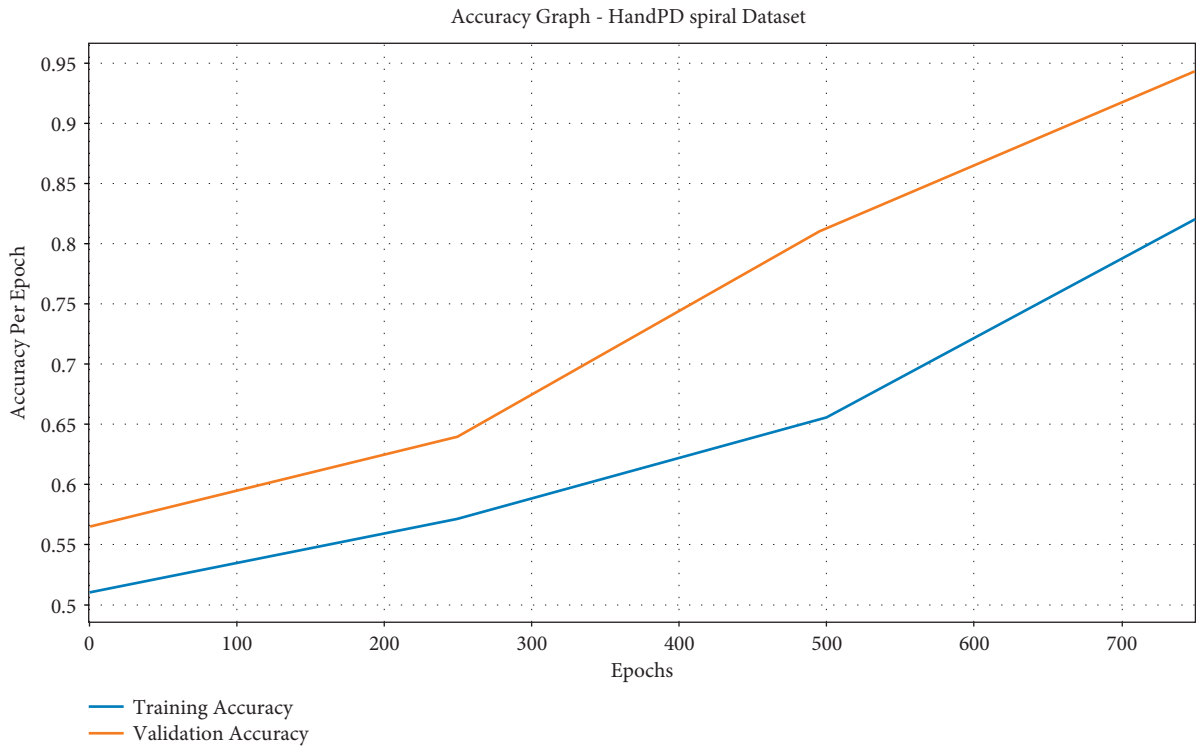


FIGURE 4: Accuracy analysis of IFSO-DL technique under the HandPD Spiral dataset.

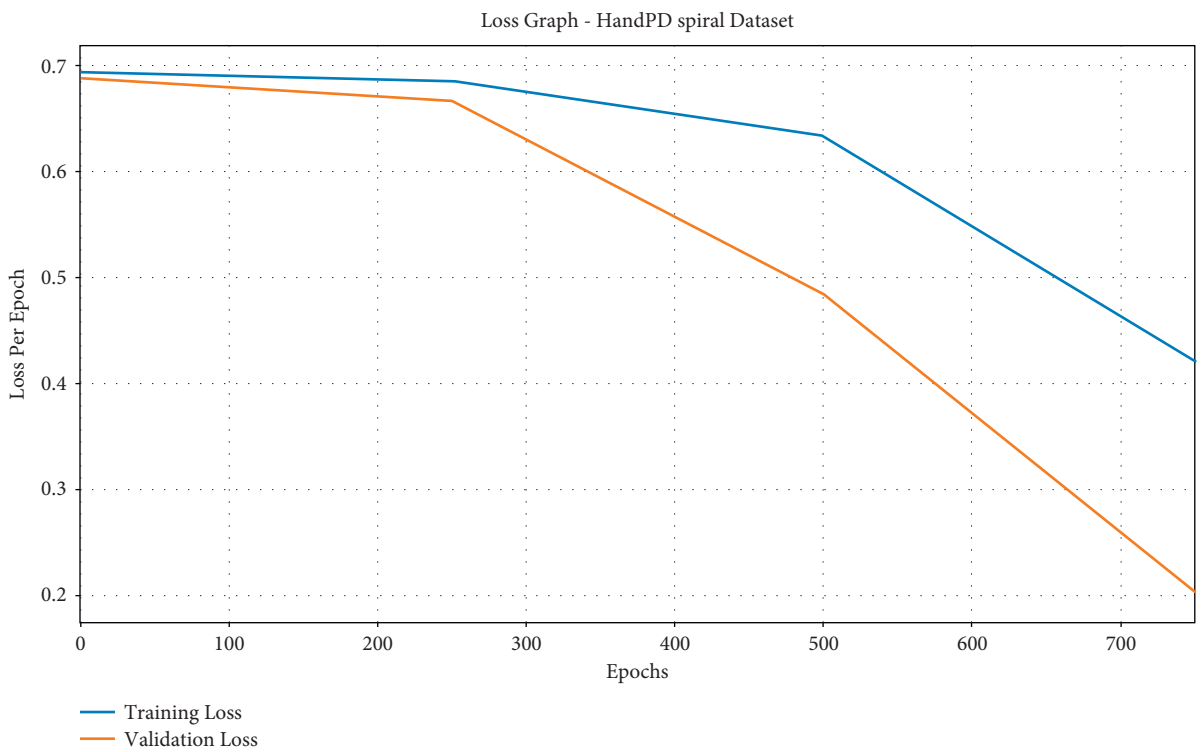


FIGURE 5: Loss analysis of IFSO-DL technique under the HandPD Spiral dataset.

TABLE 3: Result analysis of existing with the proposed IFSO-DL model on the HandPD Meander dataset.

Methods	Accuracy	DR	FAR
MGOA-KNN	0.748	0.858	0.476
MGOA-RF	0.937	1.000	0.191
MGOA-DT	0.890	0.918	0.167
MGWO-KNN	0.728	0.858	0.600
MGWO-RF	0.930	0.991	0.222
MGWO-DT	0.880	0.920	0.222
IFSO-DL	0.940	1.000	0.135

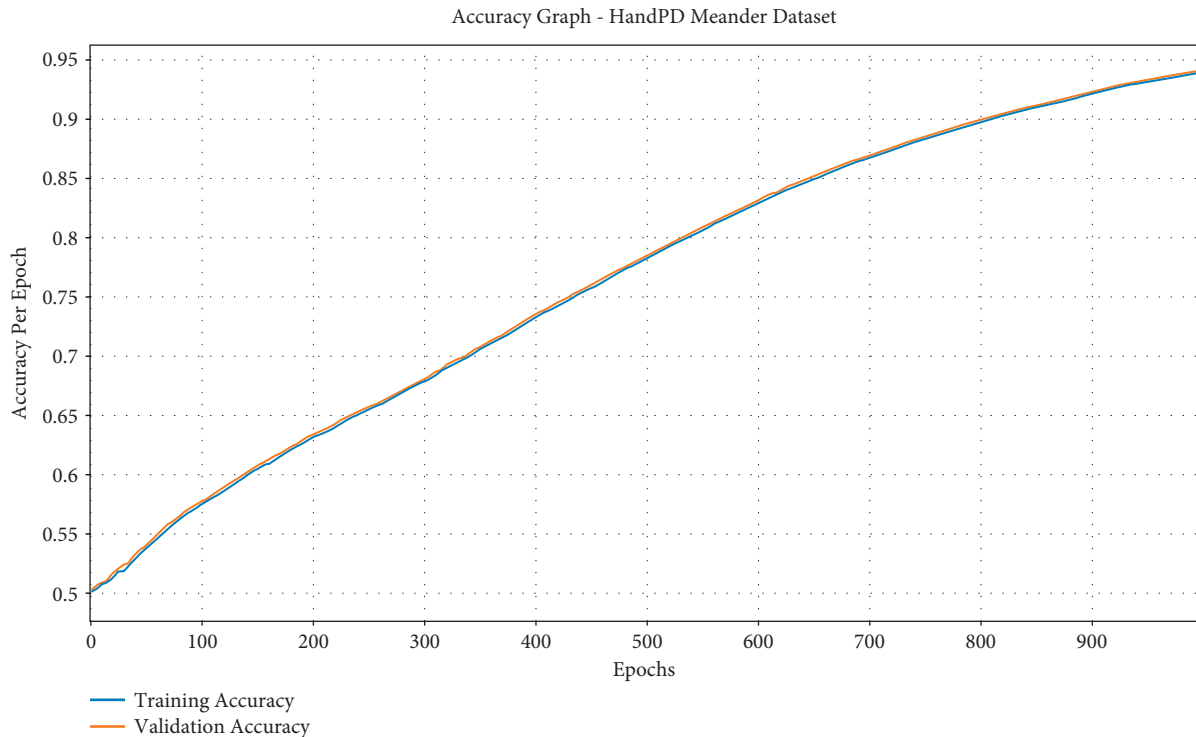


FIGURE 6: Accuracy analysis of IFSO-DL technique under the HandPD Meander dataset.

PD dataset. The results depicted that the MGWO-KNN and MGOA-KNN methods have obtained minimal accuracy of 0.918 and 0.897, respectively. Besides, the MGOA-DT technique has reached moderate accuracy of 0.846. Likewise, MGOA-RF, MGWO-RF, and MGWO-DT methods have accomplished reasonable accuracy of 0.949, 0.939, and 0.898, respectively. However, the IFSO-DL technique has shown the other algorithms with the maximal accuracy, DR, and FAR of 0.953, 1.000, and 0.185, respectively.

Figure 8 displays the accuracy graph analysis of the IFSO-DL approach on the test Speech PD dataset. The figure demonstrates that the IFSO-DL technique has achieved higher training and validation accuracies. It can be obvious that the IFSO-DL technique has accomplished increased validation accuracy over the training accuracy.

The loss graph analysis of the IFSO-DL algorithm is explored in Figure 9. The figure depicts that the IFSO-DL technique has accomplished superior results with the lower validation loss compared with training loss. It can also portray that the IFSO-DL technique has reached reduced validation loss related to training loss.

Table 5 provides a detailed comparative outcome analysis of the IFSO-DL manner with recent techniques on the test Voice PD dataset. The results demonstrated that the MGWO-KNN and MGOA-KNN methodologies have gained minimal accuracy of 0.858 and 0.918, respectively. Similarly, the MGOA-DT technique has achieved moderate accuracy of 1.000. Subsequently, MGOA-RF, MGWO-RF, and MGWO-DT methods have accomplished reasonable accuracy of 1.000, 1.000, and 1.000, respectively. Finally, the IFSO-DL technique has displayed the other algorithms with the higher accuracy, DR, and FAR of 1.000, 1.000, and 0.000, respectively.

Figure 10 exhibits the accuracy graph analysis of the IFSO-DL system on the test Voice PD dataset. The figure portrays that the IFSO-DL technique has reached increased training and validation accuracies. It is noticeable that the IFSO-DL methodology has accomplished higher validation accuracy over the training accuracy.

The loss graph analysis of the IFSO-DL approach is examined in Figure 11. The figure outperforms that the IFSO-DL method has accomplished enhanced outcomes with the lesser validation loss related to training loss. It also

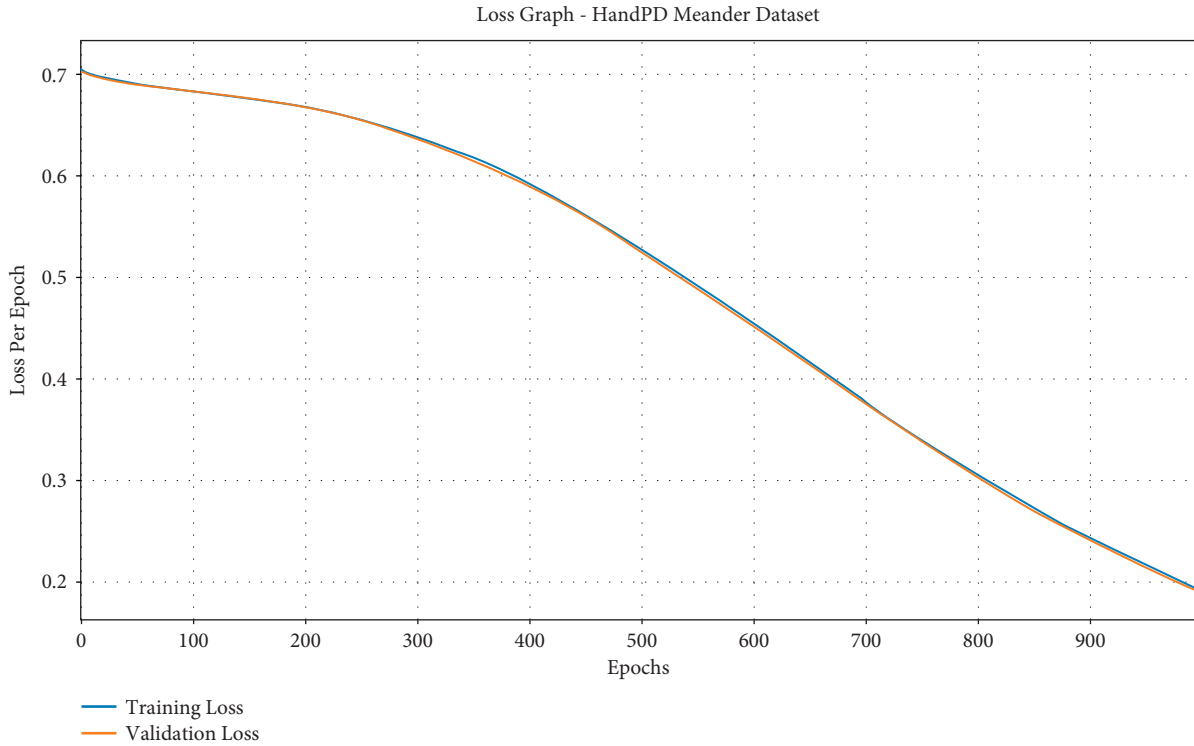


FIGURE 7: Accuracy analysis of IFSO-DL technique under the HandPD Meander dataset.

TABLE 4: Result analysis of existing with the proposed IFSO-DL model on the Speech PD dataset.

Methods	Accuracy	DR	FAR
MGOA-KNN	0.897	0.967	0.300
MGOA-RF	0.949	1.000	0.222
MGOA-DT	0.846	0.900	0.300
MGWO-KNN	0.918	0.974	0.300
MGWO-RF	0.939	1.000	0.300
MGWO-DT	0.898	0.949	0.300
IFSO-DL	0.953	1.000	0.185

shows that the IFSO-DL manner has obtained reduced validation loss connected to training loss.

Figure 12 shows the accuracy analysis of the IFSO-DL technique with other recent techniques on the four test datasets [27]. The figure portrays that the IFSO-DL technique has gained effective outcomes with the maximum accuracy values on all the test datasets.

Figure 13 illustrates the DR analysis of the IFSO-DL algorithm with other recent manners on the four test datasets. The figure shows that the IFSO-DL technique

has achieved effective outcomes with the maximal DR values on all the test datasets.

Figure 14 depicts the FAR analysis of the IFSO-DL method with other recent approaches on the four test datasets. The figure outperforms that the IFSO-DL system has reached effective outcomes with higher FAR values on all the test datasets. From the abovementioned tables and figures, it is apparent that the IFSO-DL technique has been found to be an effective tool for PD detection and classification.

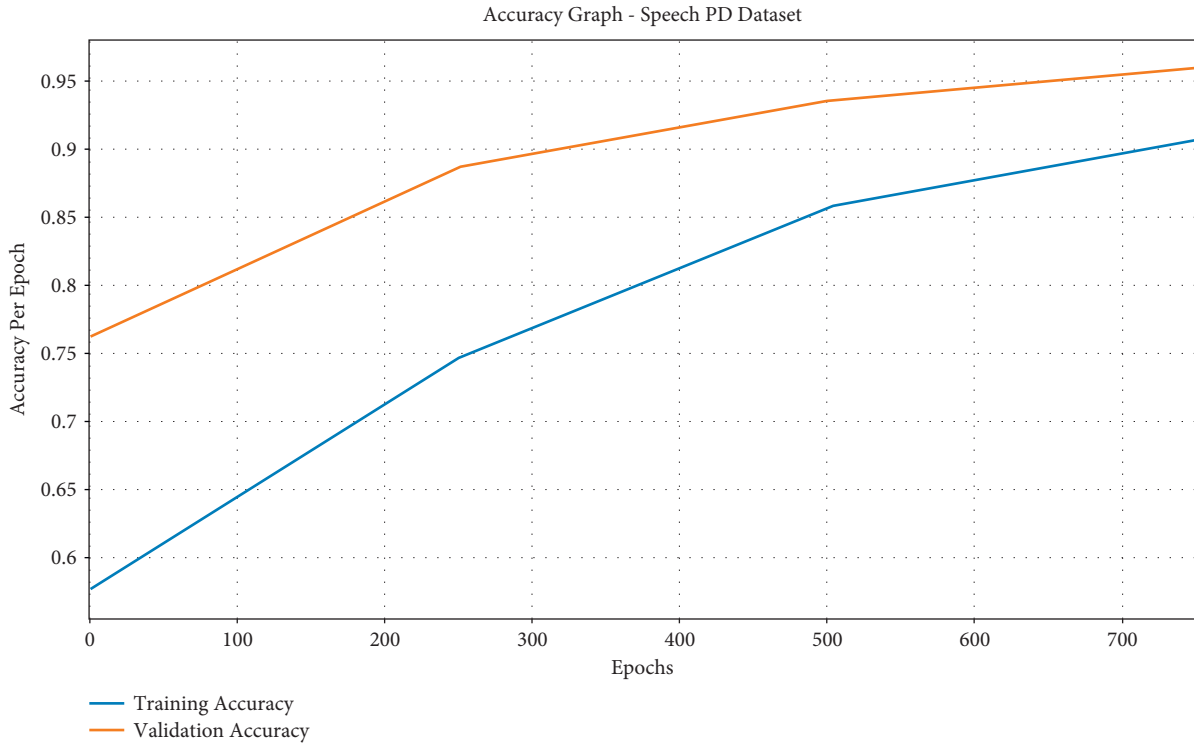


FIGURE 8: Accuracy analysis of IFSO-DL technique under the Speech PD dataset.

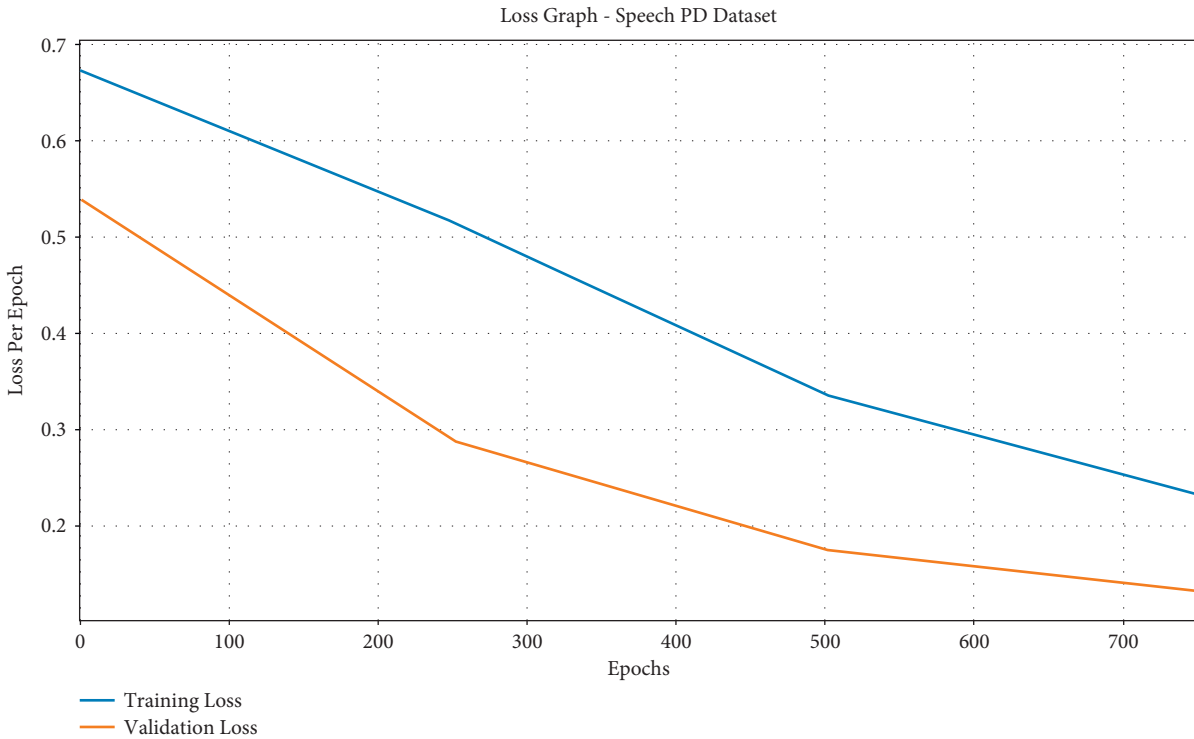


FIGURE 9: Loss analysis of IFSO-DL technique under the Speech PD dataset.

TABLE 5: Result analysis of existing with the proposed IFSO-DL model on the Voice PD dataset.

Methods	Accuracy	DR	FAR
MGOA-KNN	0.918	0.835	0.009
MGOA-RF	1.000	1.000	0.000
MGOA-DT	1.000	1.000	0.000
MGWO-KNN	0.858	0.803	0.081
MGWO-RF	1.000	1.000	0.000
MGWO-DT	1.000	1.000	0.000
IFSO-DL	1.000	1.000	0.000

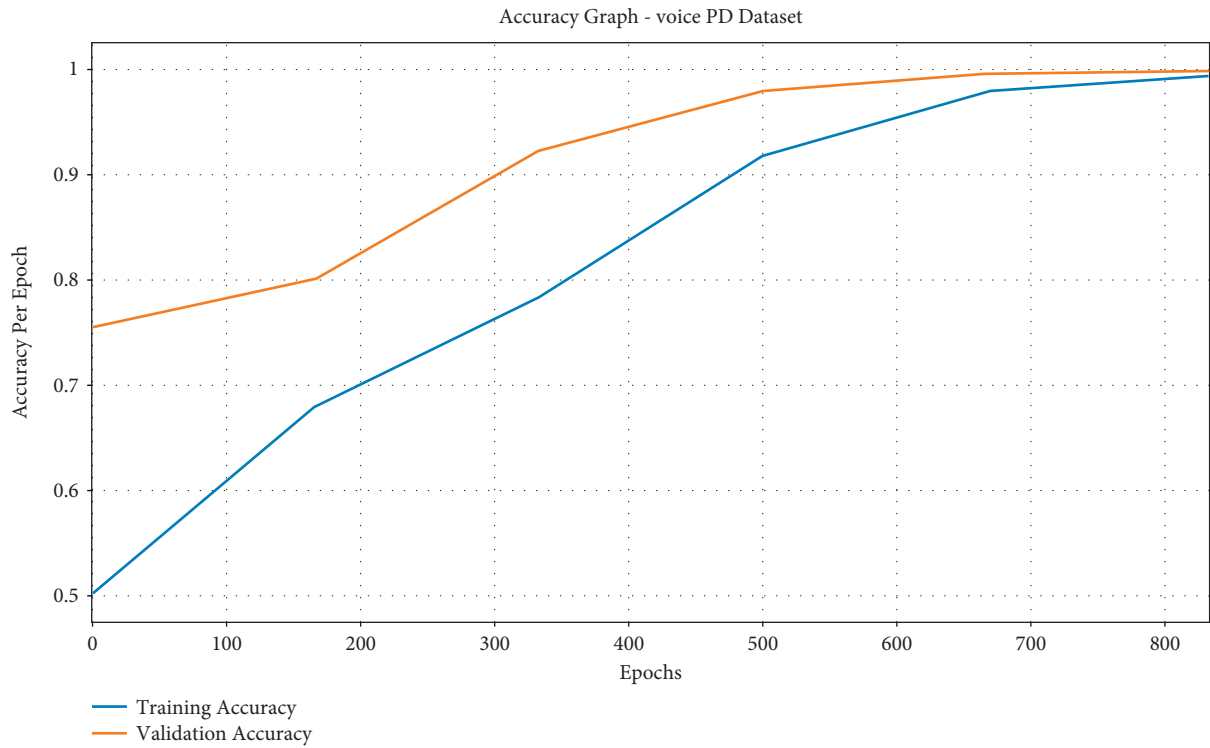


FIGURE 10: Accuracy analysis of IFSO-DL technique under the Voice PD dataset.

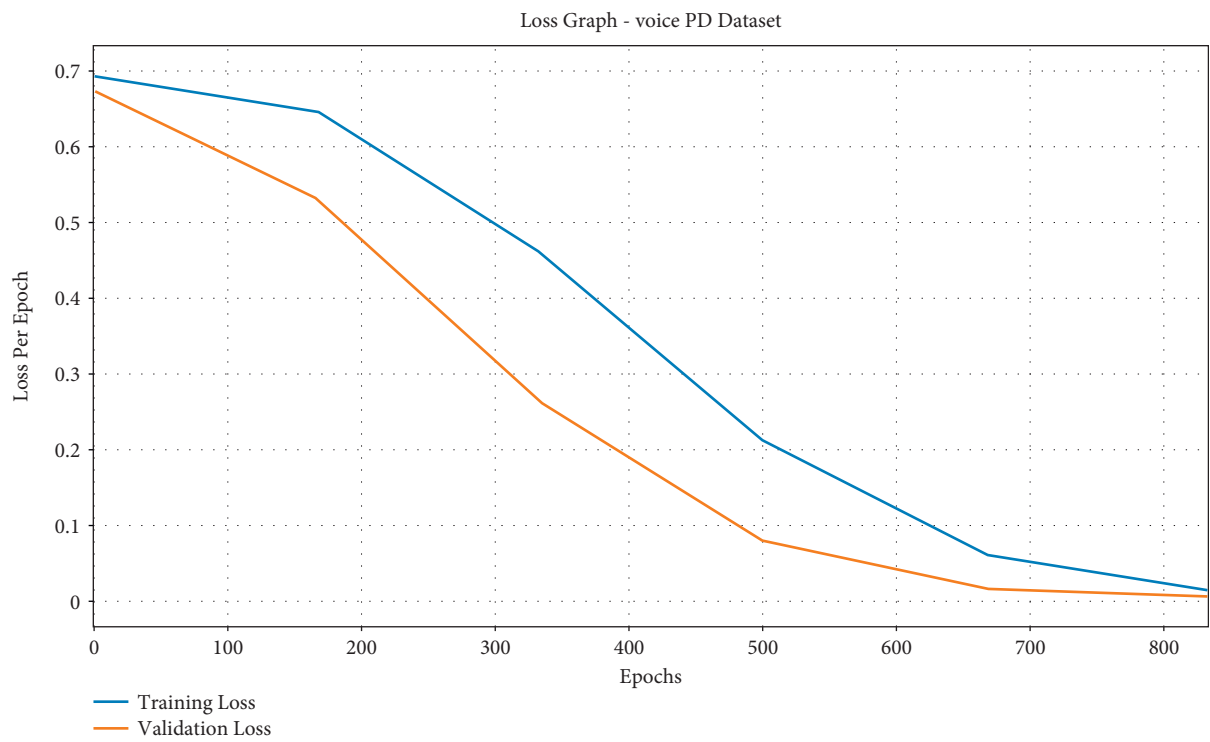


FIGURE 11: Loss analysis of IFSO-DL technique under the Voice PD dataset.

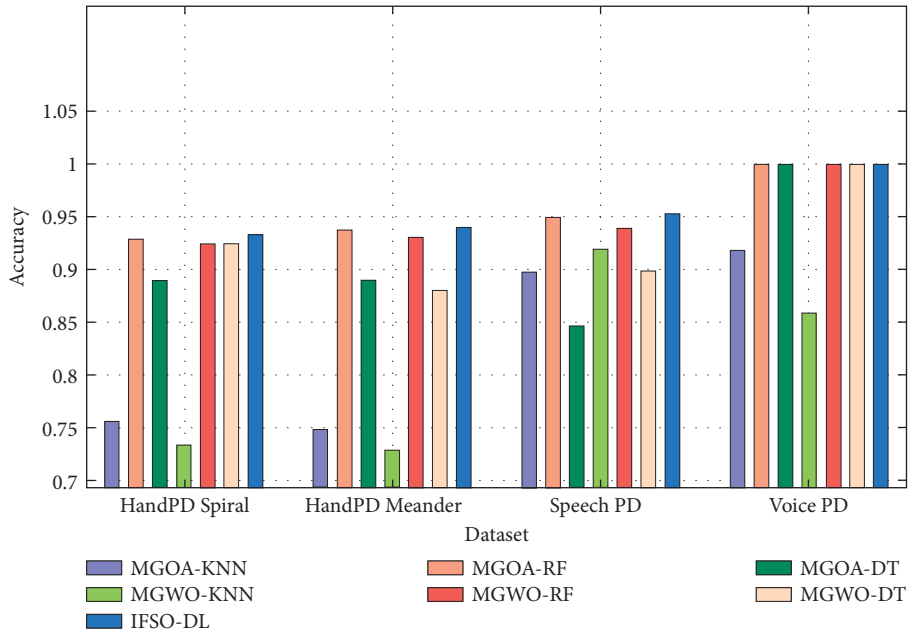


FIGURE 12: Accuracy analysis of IFSO-DL technique with existing approaches.

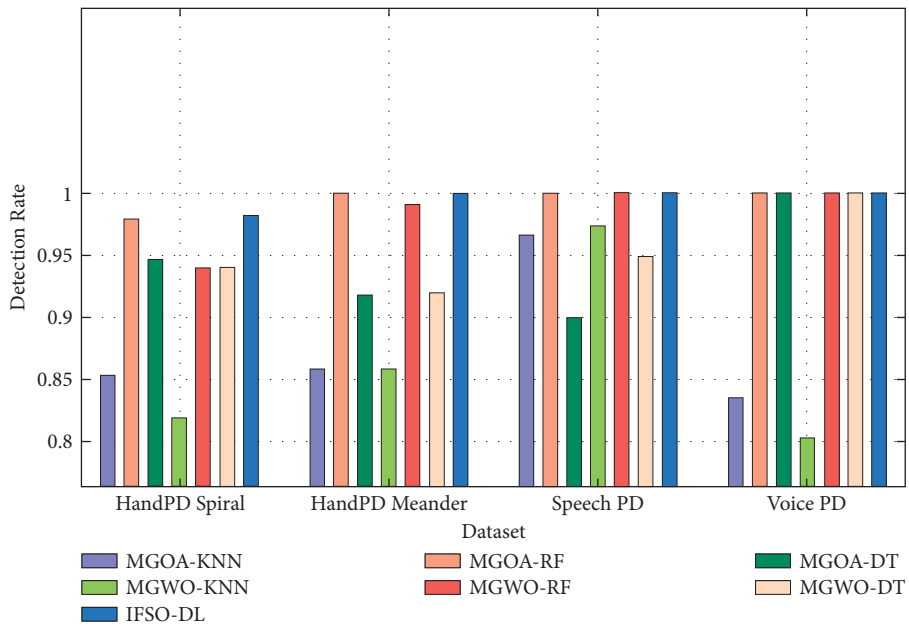


FIGURE 13: DR analysis of IFSO-DL technique with existing approaches.

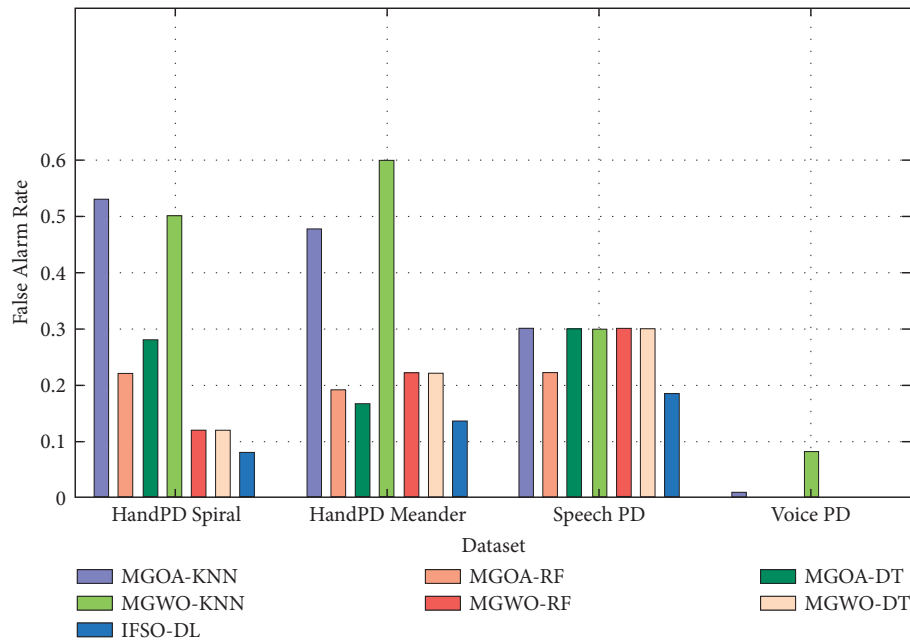


FIGURE 14: FAR analysis of IFSO-DL technique with existing approaches.

5. Conclusion

In this study, the ISFO-DL technique has been developed for PD detection and classification. The proposed ISFO-DL technique is mainly intended to determine PD and thereby enhance the survival rate of the person. The presented ISF-DL technique involves three major processes namely ISFO-based feature selection, BiGRU-based classification, and RSO-based hyperparameter optimization. The design of ISFO and RSO algorithms finds useful to significantly enhance the PD classification performance. The experimental validation of the IFSO-DL model is carried out using a benchmark Parkinson's dataset, and the results are inspected under several dimensions. The experimental results highlighted the enhanced classification performance of the ISFO-DL technique, and therefore, the proposed model can be employed for the earlier identification of PD. In future, the PD classification performance can be boosted by the use of outlier detection and clustering approaches.

Data Availability

The dataset used in this study is publicly available via the following link: <https://wwwp.fc.unesp.br/~papa/pub/datasets/Handpd/>.

Ethical Approval

This article does not contain any studies with human participants performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgments

This work was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant No. D-049-611-1439. The authors, therefore, gratefully acknowledge the DSR technical and financial support.

References

- [1] H. Gunduz, "Deep learning-based Parkinson's disease classification using vocal feature sets," *IEEE Access*, vol. 7, pp. 115540–115551, 2019.
- [2] M. Pramanik, R. Pradhan, P. Nandy, S. M. Qaisar, and A. K. Bhoi, "Assessment of acoustic features and machine learning for Parkinson's detection," *Journal of Healthcare Engineering*, vol. 2021, 2021.
- [3] S. Grover, S. Bhartia, A. Akshama, A. Yadav, and S. K.R., "Predicting severity of Parkinson's disease using deep learning," *Procedia Computer Science*, vol. 132, pp. 1788–1794, 2018.
- [4] M. Peker, B. Şen, and D. Delen, "Computer-aided diagnosis of Parkinson's disease using complex-valued neural networks and mRMR feature selection algorithm," *Journal of healthcare engineering*, vol. 6, no. 3, pp. 281–302, 2015.
- [5] A. Liu, H. Bi, Y. Li et al., "Galvanic vestibular stimulation improves subnetwork interactions in Parkinson's disease," *Journal of healthcare engineering*, vol. 2021, Article ID 6632394, 2021.
- [6] B. M. Eskofier, S. I. Lee, J. F. Daneault et al., "Recent machine learning advancements in sensor-based mobility analysis: deep learning for Parkinson's disease assessment," in

- Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 655–658, IEEE, Orlando, FL, USA, August 2016.
- [7] C. Che, C. Xiao, J. Liang, B. Jin, J. Zho, and F. Wang, “An RNN architecture with dynamic temporal matching for personalized predictions of Parkinson’s disease,” in *Proceedings of the 2017 SIAM international conference on data mining*, pp. 198–206, Society for Industrial and Applied Mathematics, Houston, Texas, USA, April 2017.
 - [8] T. T. Um, F. M. Pfister, D. Pichler et al., “Data Augmentation of wearable sensor data for Parkinson’s disease monitoring using convolutional neural networks,” in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, pp. 216–220, New York, NY, USA, November 2017.
 - [9] C. Stamate, G. D. Magoulas, S. Küppers et al., “Deep learning Parkinson’s from smartphone data,” in *Proceedings of the 2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 31–40, IEEE, Kona, HI, USA, March 2017.
 - [10] E. Huseyn, “Deep learning based early diagnostics of Parkinsons disease,” 2020, <https://arxiv.org/abs/2008.01792>.
 - [11] S. L. Oh, Y. Hagiwara, U. Raghavendra et al., “A deep learning approach for Parkinson’s disease diagnosis from EEG signals,” *Neural Computing & Applications*, vol. 32, no. 15, pp. 10927–10933, 2020.
 - [12] W. Wang, J. Lee, F. Harrou, and Y. Sun, “Early detection of Parkinson’s disease using deep learning and machine learning,” *IEEE Access*, vol. 8, pp. 147635–147646, 2020.
 - [13] A. H. Shahid and M. P. Singh, “A deep learning approach for prediction of Parkinson’s disease progression,” *Biomedical Engineering Letters*, vol. 10, no. 2, pp. 227–239, 2020.
 - [14] S. Kaur, H. Aggarwal, and R. Rani, “Hyper-parameter optimization of deep learning model for prediction of Parkinson’s disease,” *Machine Vision and Applications*, vol. 31, no. 5, pp. 1–15, 2020.
 - [15] S. Sivaranjini and C. M. Sujatha, “Deep learning-based diagnosis of Parkinson’s disease using convolutional neural network,” *Multimedia Tools and Applications*, vol. 79, no. 21, pp. 15467–15479, 2020.
 - [16] C. Quan, K. Ren, and Z. Luo, “A deep learning based method for Parkinson’s disease detection using dynamic features of speech,” *IEEE Access*, vol. 9, pp. 10239–10252, 2021.
 - [17] L. Sigcha, N. Costa, I. Pavón et al., “Deep learning approaches for detecting freezing of gait in Parkinson’s disease patients through on-body acceleration sensors,” *Sensors*, vol. 20, no. 7, p. 1895, 2020.
 - [18] K. H. Leung, S. P. Rowe, M. G. Pomper, and Y. Du, “A three-stage, deep learning, ensemble approach for prognosis in patients with Parkinson’s disease,” *EJNMMI Research*, vol. 11, no. 1, pp. 1–14, 2021.
 - [19] M. Masud, P. Singh, G. S. Gaba et al., “CROWD: crow search and deep learning based feature extractor for classification of Parkinson’s disease,” *ACM Transactions on Internet Technology*, vol. 21, no. 3, pp. 1–18, 2021.
 - [20] Y. Zhang and Y. Mo, “Dynamic optimization of chemical processes based on modified sailfish optimizer combined with an equal division method,” *Processes*, vol. 9, no. 10, p. 1806, 2021.
 - [21] K. K. Ghosh, S. Ahmed, P. K. Singh, Z. W. Geem, and R. Sarkar, “Improved binary sailfish optimizer based on adaptive β -hill climbing for feature selection,” *IEEE Access*, vol. 8, pp. 83548–83560, 2020.
 - [22] D. Zhao, J. Wang, H. Lin, Z. Yang, and Y. Zhang, “Extracting drug-drug interactions with hybrid bidirectional gated recurrent unit and graph convolutional network,” *Journal of Biomedical Informatics*, vol. 99, Article ID 103295, 2019.
 - [23] G. Dhiman, M. Garg, A. Nagar, V. Kumar, and M. Dehghani, “A novel algorithm for global optimization: rat swarm optimizer,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8457–8482, 2021.
 - [24] D. Gupta, A. Julka, S. Jain et al., “Optimized cuttlefish algorithm for diagnosis of Parkinson’s disease,” *Cognitive Systems Research*, vol. 52, pp. 36–48, 2018.
 - [25] M. Ragab, A. Algarni, A. Bahaddad, and R. Mansour, “Machine learning applied to problem-solving in medical applications,” *Computers, Materials & Continua*, vol. 69, no. 2, pp. 2277–2294, 2021.
 - [26] M. Ragab, S. M. Abo-Dahab, A. E. Abouelregal, and A. A. Kilany, “A thermoelastic piezoelectric fixed rod exposed to an axial moving heat source via a dual-phase-lag model,” *Int. J. Complex*, vol. 2021, Article ID 5547566, 232 pages, 2021.
 - [27] S. Sehgal, M. Agarwal, D. Gupta, S. Sundaram, and A. Bashambu, “Optimized grass hopper algorithm for diagnosis of Parkinson’s disease,” *SN Applied Sciences*, vol. 2, no. 6, pp. 1–18, 2020.