IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 3 October 2023, accepted 18 October 2023, date of publication 20 October 2023, date of current version 26 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3326369

RESEARCH ARTICLE

Hybrid Metaheuristics With Deep Learning-Based Fusion Model for Biomedical Image Analysis

MARWA OBAYYA[1], MUHAMMAD KASHIF SAEED[2], NUHA ALRUWAIS[3], SAUD S. ALOTAIBI[4], MOHAMMED ASSIRI[5], AND AHMED S. SALAMA[6]

[1]Department of Biomedical Engineering, College of Engineering, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[2]Department of Computer Science, Applied College at Muhayil, King Khalid University, Abha 61421, Saudi Arabia
[3]Department of Computer Science and Engineering, College of Applied Studies and Community Services, King Saud University, Riyadh 11495, Saudi Arabia
[4]Department of Information Systems, College of Computing and Information System, Umm Al-Qura University, Makkah 21421, Saudi Arabia
[5]Department of Computer Science, College of Sciences and Humanities-Aflaj, Prince Sattam bin Abdulaziz University, Al-Kharj 16273, Saudi Arabia
[6]Department of Electrical Engineering, Faculty of Engineering and Technology, Future University in Egypt, New Cairo 11845, Egypt

Corresponding author: Mohammed Assiri (m.assiri@psau.edu.sa)

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number (RGP2/117/44). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R203), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Research Supporting Project number (RSPD2023R608), King Saud University, Riyadh, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444). This study is partially funded by the Future University in Egypt (FUE).

ABSTRACT Biomedical image analysis has played a pivotal role in modern healthcare by facilitating automated analysis and interpretation of medical images. Biomedical image classification is the process of automatically labelling or categorizing medical images based on their content. In recent years, this field has received considerable attention because of the abundance of bio-medical image data and the potential for deep learning (DL) algorithms to assist medical staff in identifying diseases and making treatment decisions. DL methods are mostly convolutional neural networks (CNN) has illustrated outstanding performance in analyzing and classifying biomedical images. Therefore, this study presents a new Hybrid Metaheuristics with Deep Learning based Fusion Model Biomedical Image Analysis (HMDL-MFMBIA) technique. The HMDL-MFMBIA technique initially performs image pre-processing and Swin-UNet-based segmentation. Besides, a fusion of multiple DL-based feature extractors takes place using Xception and Residual Network (ResNet) model. Moreover, a hybrid salp swarm algorithm (HSSA) was employed for the optimal hyperparameter selection of the DL models. Finally, the gated recurrent unit (GRU) algorithm can be exploited for the detection and classification of bio-medical images. A widespread of simulated is conducted to establish the enhanced biomedical image classification results of the HMDL-MFMBIA method. The simulation outcomes inferred the greater outcome of the HMDL-MFMBIA algorithm over other DL models.

INDEX TERMS Biomedical image analysis, image classification, computer vision, fusion model, deep learning.

I. INTRODUCTION

Biomedical image analysis is an extensive field, which comprises a broad range of approaches and techniques for extracting, processing, and interpreting relevant data from biomedical images [1]. It acts an essential role in medical research, clinical practice, and healthcare management [2]. Biomedical image classification is referred to the tasks of automatically labelling or classifying medical images based on their content [3]. In recent times, the field has received considerable interest due to a bunch of medical imaging information and the possibility of Deep Learning (DL) methods to support medical specialists in making treatment decisions and diagnosing diseases [4]. Medical image analysis aims to give tools for effective diagnostic and treatment processes for physicians and radiologists [5]. Medical imaging devices

The associate editor coordinating the review of this manuscript and approving it for publication was Ines Domingues.

like CT, MRI, and X-ray can deliver functional and anatomical data about abnormalities and diseases inside the body non-destructively [6]. Since various medical images have artefacts, confounding details (in the absence of them), and imperfections, it is required to process them for extracting useful data from them [7]. Image processing techniques have done several contributions to medical applications; for instance, image-guided surgery, image segmentation, and image registration are extensively utilized in medical treatment and diagnoses [8]. In specific fields, like medical imaging, DL techniques outperform human specialists considerably. The DL methods utilize huge amounts of datasets for predicting complex tasks.

In Computer Vision (CV) tasks where interpreting the images, the DL technique has attained a remarkably higher level of accuracy [9]. Convolutional Neural Networks (CNNs) can be a special kind of NNs which is the most effective to analyze information in a grid-form topology [10]. Recently, the DL model, especially Deep CNN (DCNN) is the most efficient practice of medical imaging analysis [11]. In contrast to other Artificial Intelligence (AI) technologies, Deep Neural Networks (DNNs) provides highly improved outcomes in image classification tasks and can frequently exceed the human domain specialists. There has been a growing concern about the applications of AI technology for imaging segmentation and classification for bio-medical applications [12]. CNN is a recent model for image classification and segmentation, though encoder-decoder converter models are recommended [13].

This study presents a new Hybrid Metaheuristics with Deep Learning based Fusion Model Biomedical Image Analysis (HMDL-MFMBIA) technique. The HMDL-MFMBIA technique initially preprocesses the images and segments it using Swin-UNet model. Followed by, fusion of multiple DL-based feature extractors are carried out by Xception and Residual Network (ResNet) model. Besides, a hybrid salp swarm algorithm (HSSA) was employed for the optimal hyperparameter selection of the DL models. At last, the gated recurrent unit (GRU) algorithm was exploited for the detection and classification of bio-medical images. To highlight the better performance of the HMDL-MFMBIA algorithm, a widespread simulation analysis is made.

II. RELATED WORKS

Pradhan et al. [14] examine an intelligent approach for classifying CXR and MRI images as abnormal and normal classes for initial diagnosis of HD, COVID19, and AD dependent upon ensemble DNN. During the presented approach, CNN was utilized for automatic feature extraction in images, and LSTM was utilized for the last classifier. Furthermore, the Hill-Climbing Algorithm (HCA) was executed to determine an optimum feasible value for hyperparameters of CNN and LSTM like the dimensional of filters of CNN and the unit counts of LSTM but fitting the other parameters. Ahmed et al. [15] examine a new encoded-decoded structure termed as DOLG-NeXt, integrating 3 main improvements on the UNet-based variations. Primarily, the authors combine SE-Net-driven ConvNeXt phases as an encoded backbone to effectual extraction features. Secondarily, the authors utilize DOLG feature components from the decoded for retrieving fine-grained contextual feature representation. Barzekar and Yu [16] examine a new CNN structure comprised of the Concatenation of multiple Networks called as C-Net for classifying bio-medical images. The first 2 parts of the infrastructure comprise 6 networks, which assist as feature extractions to provide an Inner network for classifying the images with respect to benign and malignant. Mansour [17] introduces a new AI-based fusion method for CRC classification and analysis, termed AIFM-CRC. Moreover, the WOA-tuned deep SVM technique was utilized as a classification method for determining the presence of CRC.

A unique deep CNN approach for brain tumour classification was presented in [18]. A novel CNN technique could not need to be trained from scratch wasting days or weeks, but somewhat utilizing TL for data distillation. Besides, to improve the trained acceleration other optimizer approaches are employed, weights were initialized by the Gaussian initialization approach and then the ReLu activation function. Assad and Kiczales [19] develop a new deep feature extraction and classification system for bio-medical images termed Diagonal Bilinear Interpolated Deep Residual Network (DBI-DRSN). The DBI-DRSN approach integrates a balancing of information or feature using the DBI preprocessing approach and classifications the features using fine-tuned by the DRSN approach.

Pang et al. [20] introduce a novel fused CNN for developing a further correct and extremely effectual classification for biomedical images that integrates shallow and deep layer features in the presented DNN structure. Alanazi et al. [21] examine an intelligent DL-allowed oral squamous cell carcinoma detection and classification (IDL-OSCDC) approach utilizing bio-medical images. The projected IDL-OSCDC approach utilizes GF as the pre-processed stage for eliminating noise content. Additionally, an enhanced GOA (EGOA)-based DBN approach was utilized for OC recognition and classification. The hyper-parameter tuning of the DBN algorithm was carried out utilizing the EGOA approach that in turn increases the classifier outcomes.

Jaszcz et al. [22] the usage of heuristic red fox heuristic optimization algorithm (RFOA) for diagnostic image segmentation is proposed. The heuristics' operation was adapted to the analysis of 2D images, with the novel fitness function and the focus on equation modification. In [23], the Harris Hawks optimized convolution network (HHOCNN) is used for resolving these challenges. The brain magnetic resonance (MR) image was preprocessed, and the noisy pixels are eliminated to minimize the false tumor recognition rate. Then, the candidate region process is applied for identifying the cancerous area. Khan et al. [24] a lightweight deep learning method is presented for human gait recognition. Initially, two lightweight pretrained models are considered and finetuned with respect to additional layers and freezing

some middle layers. The fusion process can be done by the discriminating correlation analysis, which is enhanced by the moth-flame optimization algorithm.

In the realm of medical image classification, a notable research gap exists in the realm of hyperparameter tuning and feature extraction approaches. While DL algorithms have achieved great results in this domain, there remains room for development in enhancing feature extraction methods tailored to medical image features. Fusion-based feature extraction, which fuses data from different DL approaches presents an underlying potential for improving the discriminatory power of classification model. In addition, the systematic tuning of hyperparameters, involving those related to optimization algorithms and network architecture, holds the potential to further improve model performance for medical image classification task. Most of the existing works does not focus on the hyperparameter tuning using metaheuristics. Bridging this research gap may results in more robust and accurate classification techniques, ultimately benefiting clinical diagnoses and treatment planning in the medical field. Therefore, in this work, we have focused on the design and development of the fusion based feature extraction process comprising Xception and ResNet models. In addition, the hyperparameter tuning process can be carried out using the HSSA.

III. THE PROPOSED MODEL

In this article, we have presented the design of the automated biomedical image classification approach, named the HMDL-MFMBIA method. The main objective of the HMDL-MFMBIA method is to investigate biomedical images for the existence of diseases. To achieve this, the HMDL-MFMBIA technique comprises several phases of functions like pre-processing, Swin-UNet-based segmentation, fusion-based feature extraction, HSSA-based parameter tuning, and GRU-based classification. Fig. 1 depicts the workflow of the HMDL-MFMBIA approach.

A. IMAGE PRE-PROCESSING AND SEGMENTATION

Initially, the pre-processing of the biomedical images takes place. During the experimentation phase, Image processing methods including Morphological Operators, Histogram equalization, and Thresholding were explored [25]. In this work, Histogram Equalization was recognized as the most effective method to enhance the performance of DL techniques in feature extraction. The enhancement can be accomplished by the histogram equalization observed in sample images from normal and pneumonia data, where it efficiently enhances the overall quality and image contrast. Histogram equalization has a critical role to play in areas of lower contrast within an image. It can be done by rescaling the intensity level based on the frequency whereupon they occur. The crucial features are highlighted by expanding the intensity values, which makes it a good choice for pre-processing. The method efficiently increases the overall contrast of the image, allowing for better analysis and visualization of fundamental features.

Then, the Swin-UNet model can perform the segmentation of pre-processed images. Medical images often come with variations in terms of lighting, orientation, and quality. Swin-UNet's capability to learn and adapt to the variation can make it versatile and more robust in handling different medical image dataset. Swin-UNet integrates self-attention mechanism, which allows to focus on applicable area within the input images. In medical image segmentation, capturing fine details and subtle patterns can be crucial, and attention mechanism helps the model prioritize significant areas, resulting in more accurate segmentation. Swin-UNet exploits a hierarchical model to process image data at different scales. In medical images, structure of interest might differ in scale and size, and the ability to analyze the structure at dissimilar levels of granularity can be crucial for the accurate segmentation.

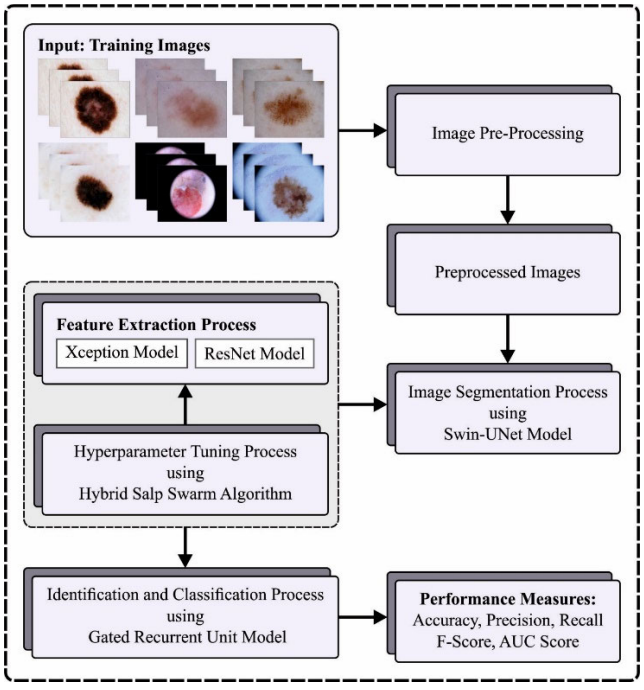

FIGURE 1. Workflow of HMDL-MFMBIA methodology.

The proposed method comprises a skip connection, bottleneck, encoder, and decoder. The building block of Swin-Unet is the Swin Transformer block. In the case of the encoded, the bio-medical image was divided as to non-overlapping patches with 4×4 patch size for transforming the input into sequence embedding. Using this partitioning method, the feature dimension of all the patches becomes $4 \times 4 \times 3 = 48$. Moreover, a linear embedded layer can be exploited to give the size of features as a random dimension. The transformed patch token passed over the patch merging layer and Swin Transformer block for generating the hierarchical feature representation. The study developed a symmetric transformer-based decoder based on

U-Net architecture. The decoder is made up of a patch-expanding layer and a Swin Transformer block. Finally, the patch expanding layer implemented $4\times up$-sampling for restoring the resolve of the mapping feature into the input ($W \times H$), and later the linear projection layer was performed on this up-sampled feature to the outcome of pixel-level segmentation prediction.

B. FUSION-BASED FEATURE EXTRACTION

In this phase, the fusion of two feature extractors takes place to derive feature vectors from the segmented images. Separable Convolution was replaced by the depthwise Separable Convolution in the Xception model [26]. Xception is an ''extreme'' variant of Inception architecture which outperform InceptionV3 on the ImageNet data and considerably surpasses it on large datasets with 17,000 classes. Compared to separable convolution, Depthwise Separable convolution need lesser computation. Therefore, Xception needs a lesser amount of parameters than other CNN variants. The depthwise 2D convolutions were slower than typical 2D convolutions while employing lesser memory. Notably, it takes a similar amount of model parameters as Inception, resulting in great computation performance. After the initial function, the absence or existence of nonlinearity. Nonlinearity was proposed in the Inception architecture by compressing and filtering input space, but Xception doesn't.

The ResNet18 version was used in this study for deep feature extraction. It has 18 layers. The ResNet18 architecture was initially pre-trained on the thousand classes of the ImageNet datasets, comprising the ResNet18 architecture with 11,511,784 trainable parameters. The main objective is to enable a massive amount of convolutional layers. Due to the presence of vanishing issues, the performance of the network becomes degraded or saturated. When there are multiple layers, a vanishing gradient arises, and the continuous multiplication leads to an even small gradient than the prior one; this situation results in the degradation of the network performance. In this work, a new concept was introduced for addressing the gradient vanishing problems that is ''skip connection''. The skip connection resolves the gradient vanishing problems by using the prior layer activation. Skip connection compresses the network, and the network begins to learn faster than before.

C. HYPERPARAMETER TUNING

The HSSA is used for adjusting the hyperparameter values of the DL models. The SSA is a method, which efficiently mimics the swimming and foraging procedures of a salp swarm (SS) from the ocean [27]. The SSA provide a valuable reference for the understanding and study of the group foraging behaviors of the salp swarm and simulates the movement and behavior of the salp swarm. In HSSA, the fusion of the parameter of SSA and the Levy flight (LF) reinforces the global search ability of the SSA and ensures the optimum value in time. This technique increases the diversity of the algorithm and enhances the search intensity of SSA. The optimization algorithm ensures that the algorithm can find the optimum value and prevent getting trapped in the local optima, and the algorithm has optimum global search capabilities due to its enhanced diversity.

Salp individual moves the same as jellyfish, the SSA is developed as an important device for researcher workers to search for its performance and movement. The most stimulating feature is the group-feeding performance of the tarantula. But, the major details of bioinformatics for this performance are not clear, many researchers are considered and initiated that this performance is optimum exploit quick co-ordination alters for achieving optimum movement trajectories for optimizing foraging performance. Finally, the SSA could not only mimic the movement and performance of SS and offer a valued reference for case and recognition of the group foraging performance of SS.

For modelling the movement trajectories of the SS population under the foraging, it is initially separated the SS population into 2 types such as leader and follower. The leader was placed at the top of the SS chain, but the other salps can be assumed followers. While the name recommends, the leaders of SS can be responsible to guide the group movement, and the followers follow. Related to other population intelligence optimizer methods, it is determined the positions of SS from the n-dimensional searching space as a model, but n is defined as dependent upon a provided problem. Therefore, the positions of every salp are saved in a 2D matrix.

During the search space, it can be utilized F as the food source (the FF that resolved) as the target to SS for searching for food.

The mathematical process of foraging movement trajectory of the SS chain is as:

$$\begin{cases} x_j^1 = \begin{cases} F_j + c_1 \times \left(\left(ub_j - lb_j\right) \times c_2 + lb_j\right), c_3 \geq 0 \\ F_j + c_1 \times \left(\left(ub_j - lb_j\right) \times c_2 + lb_j\right), c_3 < 0 \end{cases} \\ x_j^i = \dfrac{1}{2}\left(x_j^i - x_j^{i-1}\right), i \geq 2 \end{cases} \tag{1}$$

whereas X_j^1 represents the location of 1st bottle sea squirt in j dimensional, $i \geq 2, X_j^i$ signifies the location of i^{th} Nara sea squirt follower in j dimensional, F_j stand for the position of the food source from j dimensional, ub_j implies the maximal vector from the searching space, lb_j defines the minimal vector from the searching space, and c_1, c_2, and c_3 represents the arbitrary numbers.

Eq. (1) displays that the leader only upgrades their place to a food source. Most of the parameters c_1 refer to the very essential parameter in this method and are demonstrated as:

$$c_1 = 2e^{-(\frac{4L}{l})^2} \tag{2}$$

whereas l denotes the present iteration of the method and L represents the entire iteration counts of the method. c_2 and c_3 indicate the arbitrary numbers from the range of

zero and one. Actually, c_2 and c_3 define if the next position of SSs from the j^{th} dimensional can be positive/negative infinity and the step length of SSs.

The HSSA is designed by the combination of the SSA with the LF. It is a technique for representing the Levy distribution in arbitrary steps. Several analyses demonstrated that several insects and animals behave in a method which is typically a feature of LF. LF is a process of arbitrary step size, an inspired image of LF trajectory that continuously takes place with smaller step sizes but sometimes takes place with huge jumps. The formula for LF is as follows:

$$Levy \sim \mu = t^{-\lambda}, 1 < \lambda \leq 3 \tag{3}$$

The step s of LF are provided as:

$$s = \frac{\mu}{|v|^{\frac{1}{\beta}}} \tag{4}$$

whereas the parameters $\beta = 1.5, \mu = N(0, \sigma_\mu^2)$ and $v = N(0, \sigma_v^2)$ every refers to the gamma functions. The variance σ_μ of parameters are defined by the subsequent formula:

$$\sigma_\mu = \left[\frac{\Gamma(1+\beta) \times (\sin\pi \times \frac{\beta}{2})}{\Gamma\left[\frac{(1+\beta)}{2}\right] \times \beta \times 2^{\frac{(\beta-1)}{2}}} \right]^{\frac{1}{\beta}}, \sigma_v = 1 \tag{5}$$

While LF takes the property of enhancing the population diversity and growing the searching range, Eq. (1) is optimum whole the upgrade of leader position and create it simpler to jump of local optimal issue. Thus, an enhanced mathematical approach to the SSA technique was formulated by relating Eqs. (44) and (42) as follows:

$$x_j^1 = \begin{cases} F_j + c_1 \times ((ub_j - lb_j) + lb_j) \times Levy, c_3 \geq 0 \\ F_j - c_1 \times ((ub_j - lb_j) + lb_j) \times Levy, c_3 < 0 \end{cases} \tag{6}$$

The parameter c_1 in Eq. (6) was attained in Eq. (2) which allows the optimizer method to tighten the optimum value of the fitness function (FF) with enhancement from the iteration counts. The group of parameter c_1 and the LF supports the global searching ability of SSA and makes sure of the optimum value in time. This approach could not only enhances the searching SSA intensity and enhances the variety of method.

The HSSA method derives a FF to accomplish higher efficiency of classification. It described a positive integer to describe the greater outcomes of the candidate solution. The decay of the classification error rate is represented as FF.

$$\begin{aligned} fitness\,(x_i) &= ClassifierErrorRate\,(x_i) \\ &= \frac{No\ .of\ misclassified\ samples}{Total\ No\ .of\ samples} * 100 \end{aligned} \tag{7}$$

D. GRU-BASED CLASSIFICATION

Finally, the GRU approach is applied to the classification of images. The GRU network reconstructs the LSTM modules while retaining their quality [28]. All the GRU cells have two different gates namely the reset gate j_t and update gate b_t. Fig. 2 illustrates the infrastructure of GRU. Due to the GRU network being different from the LSTM network, it is utilized to resolve the problem of a long-dependency standard RNN network. The fundamental steps of GRU are given below:

Firstly, the hidden input x_t is the latest state determined by the prior cell h_{t-1} reset gate j_t afterwards describe gate v_t at current state (time t):

$$j_t = \sigma\left(w^j\,[h_{t-1}, x_t] + b^j\right) \tag{8}$$

$$v_t = \sigma\left(w^v\,[h_{t-1}, x_t] + b^Z\right) \tag{9}$$

where w^j and w^v denote the proper matrix of weight coefficient, b^j and b^Z refer to bias vectors, and $\sigma()$ shows the sigmoid function:

$$\tilde{h_t} = \tanh\left(w^h\,[(h_{t-1} * j_t), x_t] + b^h\right) \tag{10}$$

In Eq. (10), tanh tangent hyperbolic function w^h denotes the matrix weighted coefficient for the buried layer, b^h is the related bias vector, and $*$ denotes the matrix of dot progress among pairs.

At last, assume the existing latent state h_t output is linearly defined by integrating $\bar{h}_t$ and h_{t-1} existing and prior hidden state, total with the adjusted coefficient of 1:

$$h_t = (1 - v_t)^* \bar{h}_t + v_t^* h_{t-1} \tag{11}$$

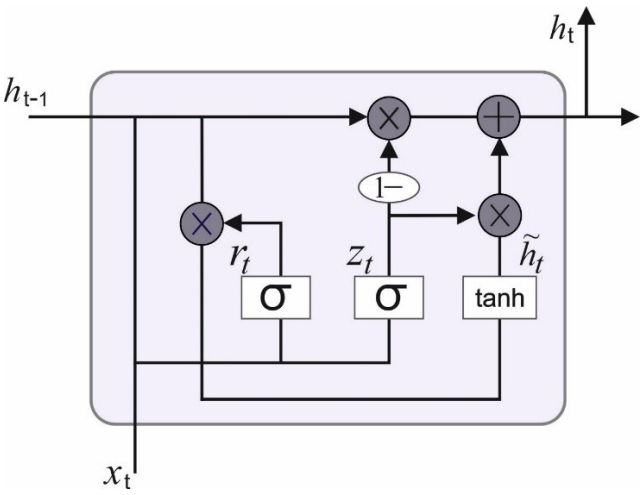

FIGURE 2. GRU structure.

IV. EXPERIMENTAL VALIDATION

In this section, the biomedical image classifier outcomes of the HMDL-MFMBIA technique are validated on two datasets: ISIC 2017 and ISIC 2020 dataset, available at https://challenge.isic-archive.com/data/. The ISIC (International Skin Imaging Collaboration) 2017 and ISIC 2020 are two skin lesion datasets utilized in the fields of dermatology and computer vision for the task of skin lesion classification and analysis. The ISIC 2017 dataset was created with the

aim of the ISIC 2017 Skin Lesion Classification Challenge. It aimed to advance the fields of automated skin lesion diagnoses through ML and CV techniques. The dataset has dermoscopy images of skin lesions, which are classified into several classes, involving nevus, melanoma, and seborrheic keratosis. All the images are obtained by ground truth annotations, which provide data about the type and features of the skin lesion. The ISIC 2020 dataset was created as part of the ISIC 2020 Challenge on Skin Lesion Analysis Towards Melanoma Detection. The challenge aimed to advance the state of the art in skin lesion analysis and melanoma recognition. ISIC 2020 comprises of a more diverse and larger collection of skin lesion images with different kinds of skin lesions, and they are available in different data splits, involving validation, training, and testing sets. These datasets comprise two classes such as benign and melanoma as depicted in Table 1.

TABLE 1. Description of two datasets.

Class	ISIC 2017 Dataset	ISIC 2020 Dataset	Total Samples
Benign	1500	1500	3000
Melanoma	1500	1500	3000

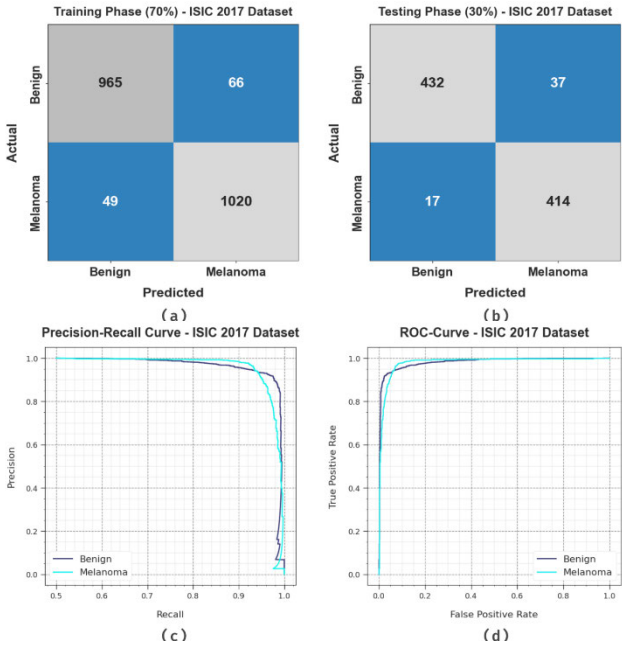

FIGURE 3. Performance on ISIC2017 database (a-b) 70:30 of TR set/TS set, (c) PR curve, and (d) ROC curve.

Fig. 3 establishes the classifier outcome of the HMDL-MFMBIA technique under the ISIC2017 dataset. Fig. 3a represents the confusion matrix offered by the HMDL-MFMBIA algorithm on 70% of the TR set. The outcome implied that the HMDL-MFMBIA approach has identified 965 instances of benign and 1020 instances of melanoma. Moreover, Fig. 3b exemplifies the confusion matrix offered by the HMDL-MFMBIA algorithm on 30% of the TS set. The outcome stated that the HMDL-MFMBIA methodology has recognized 432 instances of benign and 414 instances of melanoma. Also, Fig. 3c exhibits the PR curve of the HMDL-MFMBIA system. The result referred that the HMDL-MFMBIA system has acquired a higher PR curve on 2 classes. Likewise, Fig. 3d displays the ROC outcome of the HMDL-MFMBIA algorithm. The outcome demonstrated that the HMDL-MFMBIA methodology has managed to able outcomes with greater ROC values on 2 class labels.

In Table 2 and Fig. 4, the classifier outcome of the HMDL-MFMBIA technique is highlighted on the ISIC2017 dataset are given. The outcome implied that the HMDL-MFMBIA technique properly categorized benign and melanoma samples.

TABLE 2. Classifier outcome of HMDL-MFMBIA algorithm on ISIC2017 database.

ISIC 2017 Dataset					
Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{Score}	AUC_{Score}
Training Phase (70%)					
Benign	93.60	95.17	93.60	94.38	94.51
Melanoma	95.42	93.92	95.42	94.66	94.51
Average	94.51	94.55	94.51	94.52	94.51
Testing Phase (30%)					
Benign	92.11	96.21	92.11	94.12	94.08
Melanoma	96.06	91.80	96.06	93.88	94.08
Average	94.08	94.00	94.08	94.00	94.08

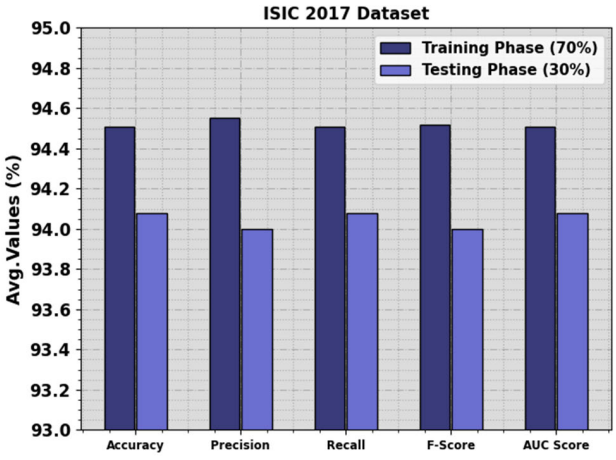

FIGURE 4. Average outcome of HMDL-MFMBIA methodology on ISIC2017 database.

On 70% of TR set, the HMDL-MFMBIA system provide average $accu_y$ of 94.51%, $prec_n$ of 94.55%, $reca_l$ of 94.51%, F_{score} of 94.52%, and AUC_{score} of 94.51%. Besides, on 30% of TS set, the HMDL-MFMBIA system provide average $accu_y$ of 94.08%, $prec_n$ of 94%, $reca_l$ of 94.08%, F_{score} of 94%, and AUC_{score} of 94.08%.

Fig. 5 demonstrates the training accuracy TR_accu_y and VL_accu_y of the HMDL-MFMBIA technique on the ISIC2017 dataset. The TL_accu_y is defined by the evaluation of the HMDL-MFMBIA algorithm on the TR dataset whereas

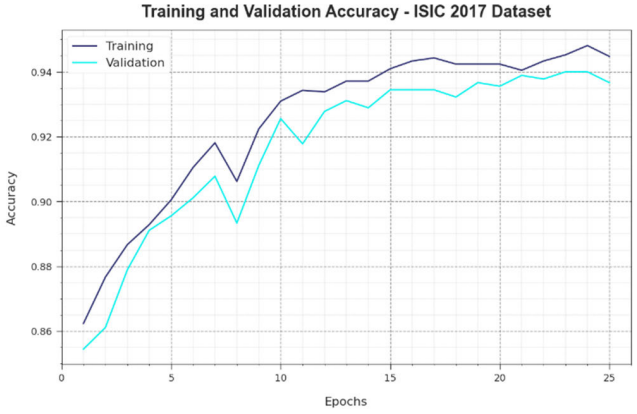

FIGURE 5. $Accu_y$ curve of HMDL-MFMBIA methodology on ISIC2017 database.

the VL_accu_y is computed by estimating the performance on a separate testing dataset. The outcomes reveal that TR_accu_y and VL_accu_y are higher with an increase in epochs. Consequently, the outcome of the HMDL-MFMBIA technique gets enhanced on the TR and TS dataset with a rise in several epochs.

In Fig. 6, the *T_loss* and *VR_loss* results of the HMDL-MFMBIA algorithm on the ISIC2017 dataset are exposed. The *TR_loss* defines the error among the predictive performance and original values on the TR data. The *VR_loss* represents the measure of the outcome of the HMDL-MFMBIA technique on individual validation data. The results indicate that the *TR_loss* and *VR_loss* tend to reduce with increasing epochs. It portrayed the improved performance of the HMDL-MFMBIA method and its ability to create an accurate classification. The reduced value of *TR_loss* and *VR_loss* establishes the greater performance of the HMDL-MFMBIA method in capturing patterns and relationships.

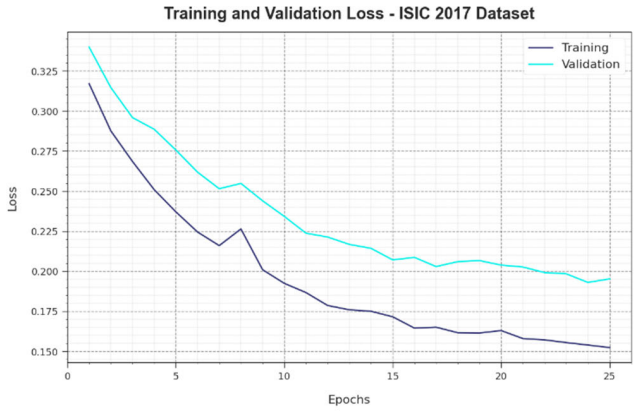

FIGURE 6. Loss curve of HMDL-MFMBIA methodology on ISIC2017 database.

The comparison results of the HMDL-MFMBIA algorithm on the ISIC2017 database are provided in Table 3 and Fig. 7. The results show that the ResNet18, Inceptionv3, AlexNet, SVM Model, and Ensemble Model have reported least classification results. At the same time, the MCUN-DCNN technique has accomplished considerable performance with $accu_y$, $prec_n$, and $reca_l$ of 88.20%, 78.50%, and 87.80% respectively. Nevertheless, the HMDL-MFMBIA technique exhibited improved performance with $accu_y$, $prec_n$, and $reca_l$ of 94.51%, 94.55%, and 94.51% correspondingly

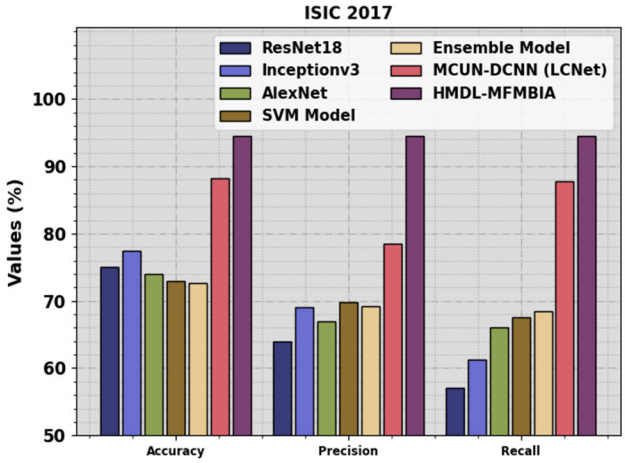

FIGURE 7. Comparative outcome of HMDL-MFMBIA algorithm on ISIC2017 database.

TABLE 3. Comparative outcome of the HMDL-MFMBIA method with other approaches on the ISIC2017 database.

ISIC 2017			
Approach	$Accu_y$	$Prec_n$	$Reca_l$
ResNet18	75.00	64.00	57.10
Inceptionv3	77.40	69.10	61.20
AlexNet	74.00	67.00	66.00
SVM Model	72.93	69.83	67.55
Ensemble Model	72.65	69.2	68.48
MCUN-DCNN (LCNet)	88.20	78.50	87.80
HMDL-MFMBIA	94.51	94.55	94.51

Fig. 8 shows the classifier outcomes of the HMDL-MFMBIA methodology under the ISIC2020 database. Fig. 8a illustrates the confusion matrix gained by the HMDL-MFMBIA approach on 70% of the TR set. The outcome pointed out that the HMDL-MFMBIA method has detected 999 instances of benign and 1004 instances of melanoma. Moreover, Fig. 8b portrays the confusion matrix accomplished by the HMDL-MFMBIA algorithm on 30% of the TS set. The outcome inferred that the HMDL-MFMBIA method has recognized 432 instances of benign and 423 instances of melanoma. Likewise, Fig. 8c shows the PR curve of the HMDL-MFMBIA methodology. The outcome stated that the HMDL-MFMBIA approach has attained enhanced PR outcomes in 2 class labels. At last, Fig. 8d displays the ROC analysis of the HMDL-MFMBIA methodology. The outcome depicted that the

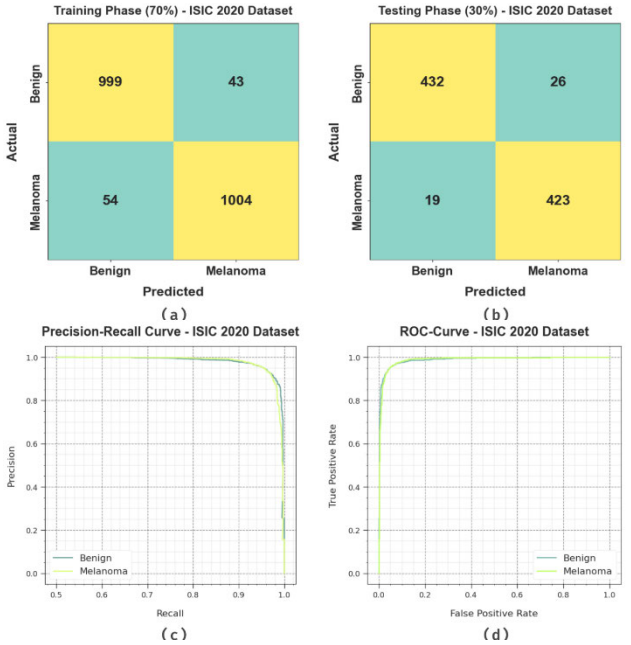

FIGURE 8. Performance on ISIC2020 database (a-b) 70:30 of TR set/TS set, (c) PR_curve, and (d) ROC_curve.

TABLE 4. Classifier outcome of HMDL-MFMBIA system on ISIC2020 database.

ISIC 2020 Dataset					
Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{Score}	AUC_{Score}
Training Phase (70%)					
Benign	95.87	94.87	95.87	95.37	95.38
Melanoma	94.90	95.89	94.90	95.39	95.38
Average	95.38	95.38	95.38	95.38	95.38
Testing Phase (30%)					
Benign	94.32	95.79	94.32	95.05	95.01
Melanoma	95.70	94.21	95.70	94.95	95.01
Average	95.01	95.00	95.01	95.00	95.01

HMDL-MFMBIA approach has led to capable outcomes with higher ROC values on 2 classes.

In Table 4 and Fig. 9, the classifier outcomes of the HMDL-MFMBIA algorithm are highlighted on the ISIC2020 dataset are given. The outcome implied that the HMDL-MFMBIA algorithm properly categorized benign and melanoma samples. On 70% of TR set, the HMDL-MFMBIA system offered average $\boldsymbol{accu_y}$ of 95.38%, $\boldsymbol{prec_n}$ of 95.38%, $\boldsymbol{reca_l}$ of 95.38%, $\boldsymbol{F_{score}}$ of 95.38%, and $\boldsymbol{AUC_{score}}$ of 95.38%. In addition, on 30% of TS set, the HMDL-MFMBIA method provided average $\boldsymbol{accu_y}$ of 95.01%, $\boldsymbol{prec_n}$ of 95%, $\boldsymbol{reca_l}$ of 95.01%, $\boldsymbol{F_{score}}$ of 95%, and $\boldsymbol{AUC_{score}}$ of 95.01%.

Fig. 10 illustrates the training accuracy TR_accu_y and VL_accu_y of the HMDL-MFMBIA approach on the ISIC2020 dataset. The TL_accu_y is defined by the estimate of the HMDL-MFMBIA system on the TR dataset whereas the VL_accu_y is computed by evaluating the outcome on a separate testing dataset. The results exhibit that TR_accu_y and VL_accu_y are higher with a rising in epochs. Accordingly, the performance of the HMDL-MFMBIA technique gets

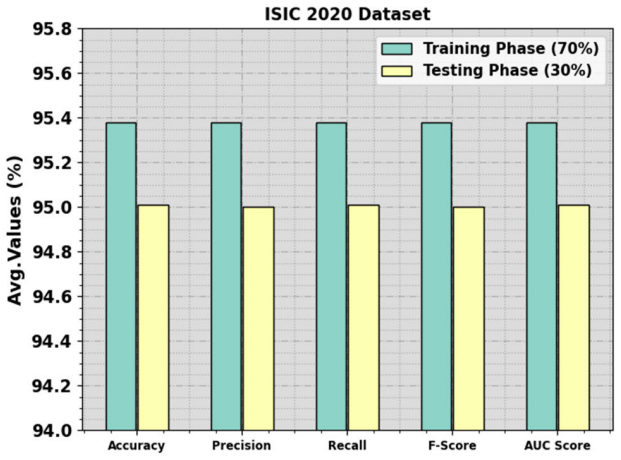

FIGURE 9. Average outcome of HMDL-MFMBIA system on ISIC2020 database.

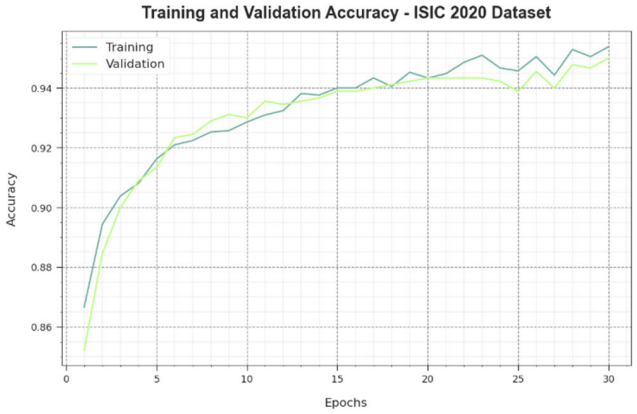

FIGURE 10. $Accu_y$ curve of HMDL-MFMBIA methodology on ISIC2020 database.

enhanced on the TR and TS dataset with an increase in the count of epochs.

In Fig. 11, the *TR_loss* and *VR_loss* outcomes of the HMDL-MFMBIA approach on the ISIC2020 dataset are

FIGURE 11. Loss curve of HMDL-MFMBIA methodology on ISIC2020 database.

shown. The *TR_loss* demonstrates the error among the predictive performance and original values on the TR data. The *VR_loss* signifies the measure of the outcome of the HMDL-MFMBIA approach on individual validation data. The outcomes implied that the *TR_loss* and *VR_loss* tend to be lesser with increasing epochs. It portrayed the enhanced result of the HMDL-MFMBIA method and its ability to generate accurate classification. The reduced value of *TR_loss* and *VR_loss* demonstrates the higher performance of the HMDL-MFMBIA methodology in capturing patterns and relationships.

The comparison outcomes of the HMDL-MFMBIA system on the ISIC2020 dataset are provided in Table 5 and Fig. 12. The outcome depicted that the ResNet18, Inceptionv3, AlexNet, SVM, and Ensemble approaches have reported minimal classifier outcomes. Likewise, the MCUN-DCNN algorithm has accomplished considerable performance with $accu_y$, $prec_n$, and $reca_l$ of 90.40%, 90.40%, and 90.30% correspondingly. But, the HMDL-MFMBIA method displayed better performance with $accu_y$, $prec_n$, and $reca_l$ of 95.38%, 95.38%, and 95.38% correspondingly.

TABLE 5. Comparative outcome of the HMDL-MFMBIA system with other methodologies on the ISIC2020 database.

ISIC 2020			
Approach	$Accu_y$	$Prec_n$	$Reca_l$
HMDL-MFMBIA	95.38	95.38	95.38
MCUN-DCNN (LCNet)	90.40	90.40	90.30
ResNet18	90.80	89.80	88.80
Inceptionv3	48.60	29.70	49.20
AlexNet	75.40	69.10	68.50
SVM Model	74.58	76.57	74.16
Ensemble Model	74.41	76.34	77.36

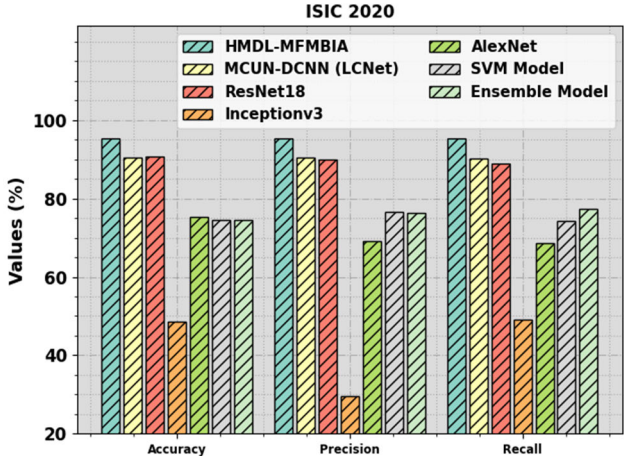

FIGURE 12. Comparative outcome of HMDL-MFMBIA algorithm on the ISIC2020 dataset.

These outcomes illustrated the enhanced classification outcome of the HMDL-MFMBIA method on biomedical images. The enhanced performance of the HMDL-MFMBIA method is due to the integration of the feature fusion and hyperparameter tuning processes. Fusion-based feature extraction enables the model to capture intricate relationships and patterns within medical images, leveraging complementary information from different DL techniques. Moreover, the systematic hyperparameter tuning ensures that the model works at its optimum configuration, efficiently balancing tradeoffs between various parameters and finetuning the learning process. This comprehensive optimization process collectively contributed to the model's superior performance, which makes it a valuable tool for reliable and accurate medical image classification, ultimately aiding healthcare professionals in making more informed decisions for patient care.

V. CONCLUSION

In this study, we have focused on the design of the automated biomedical image classification approach, named the HMDL-MFMBIA approach. The major drive of the HMDL-MFMBIA algorithm is to investigate biomedical images for the existence of diseases. To achieve this, the HMDL-MFMBIA technique comprises several phases of functions like pre-processing, Swin-UNet-based segmentation, fusion-based feature extraction, HSSA-based parameter tuning, and GRU-based classification. To demonstrate the enhanced biomedical image classification results of the HMDL-MFMBIA system, a widespread of simulated outcomes are applied. The simulation outcomes inferred the enhanced efficiency of the HMDL-MFMBIA methodology with DL approaches with maximum accuracy of 94.51% and 95.38% on ISIC 2017 and ISIC 2020 datasets, respectively. In future, the computational complexity of the HMDL-MFMBIA model can be examined on large scale real time datasets. Besides, future work can focus on integrating information from multiple imaging modalities (e.g., MRI, CT, X-ray) for classification tasks holds promise for improved accuracy. Finally, future work can explore interpretable DL models and techniques for generating saliency maps or explanations can enhance trust and clinical adoption.

ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number (RGP2/117/44). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R203), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Research Supporting Project number (RSPD2023R608), King Saud University, Riyadh, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444). This study is partially funded by the Future University in Egypt (FUE).

REFERENCES

[1] S. Nazir, D. M. Dickson, and M. U. Akram, "Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks," *Comput. Biol. Med.*, vol. 156, Apr. 2023, Art. no. 106668.

[2] M. Obayya, M. Alamgeer, J. S. Alzahrani, R. Alabdan, F. N. Al-Wesabi, A. Mohamed, and M. I. A. Hassan, "Artificial intelligence driven biomedical image classification for robust rheumatoid arthritis classification," *Biomedicines*, vol. 10, no. 11, p. 2714, Oct. 2022.

[3] S. Chakraborty and K. Mali, "An overview of biomedical image analysis from the deep learning perspective," in *Research Anthology on Improving Medical Imaging Techniques for Analysis and Intervention*. 2023, pp. 197–218.

[4] M. Di Giammarco, F. Mercaldo, F. Martinelli, and A. Santone, "Explainable deep learning methodologies for biomedical images classification," in *Proc. IEEE 42nd Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jul. 2022, pp. 1262–1264.

[5] J. Yang, R. Shi, D. Wei, Z. Liu, L. Zhao, B. Ke, H. Pfister, and B. Ni, "MedMNIST v2—A large-scale lightweight benchmark for 2D and 3D biomedical image classification," *Sci. Data*, vol. 10, no. 1, p. 41, Jan. 2023.

[6] J. Zhu, Y. Ma, J. Huang, and L. Wang, "Image segmentation combining pulse coupled neural network and adaptive glowworm algorithm," *Inf. Technol. Control*, vol. 52, no. 2, pp. 487–499, Jul. 2023.

[7] F. Li, X. Du, L. Zhang, and A. Liu, "Image feature fusion method based on edge detection," *Inf. Technol. Control*, vol. 52, no. 1, pp. 5–24, Mar. 2023.

[8] S. Maqsood and R. Damasevicius, "Multiclass skin lesion localization and classification using deep learning based features fusion and selection framework for smart healthcare," *Neural Netw.*, vol. 160, pp. 238–258, Mar. 2023.

[9] A. Shrivastava, M. Chakkaravathy, and M. A. Shah, "A comprehensive analysis of machine learning techniques in biomedical image processing using convolutional neural network," in *Proc. 5th Int. Conf. Contemp. Comput. Informat. (ICI)*, Dec. 2022, pp. 1363–1369.

[10] D. Sharma, I. Mishra, R. Parepalli, and S. Jayanth, "Biomedical image classification using convolutional neural networks," in *Proc. Int. Conf. Intell. Innov. Technol. Comput., Electr. Electron. (IITCEE)*, Jan. 2023, pp. 840–845.

[11] S. Dash, P. Parida, and J. R. Mohanty, "Illumination robust deep convolutional neural network for medical image classification," *Soft Comput.*, pp. 1–13, Feb. 2023.

[12] K. VanHorn and M. C. Çobanoglu, "Democratizing AI in biomedical image classification using virtual reality," *Virtual Reality*, vol. 26, no. 1, pp. 159–171, Mar. 2022.

[13] S. Patel, R. Patel, N. Ganatra, and A. Patel, "Spatial feature fusion for biomedical image classification based on ensemble deep CNN and transfer learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 5, 2022.

[14] A. K. Pradhan, K. Das, D. Mishra, and P. Chithaluru, "Optimizing CNN-LSTM hybrid classifier using HCA for biomedical image classification," *Exp. Syst.*, vol. 40, no. 5, Jun. 2023, Art. no. e13235.

[15] M. R. Ahmed, M. A. I. Fahim, A. K. M. M. Islam, S. Islam, and S. Shatabda, "DOLG-NeXt: Convolutional neural network with deep orthogonal fusion of local and global features for biomedical image segmentation," *Neurocomputing*, vol. 546, Aug. 2023, Art. no. 126362.

[16] H. Barzekar and Z. Yu, "C-Net: A reliable convolutional neural network for biomedical image classification," *Exp. Syst. Appl.*, vol. 187, Jan. 2022, Art. no. 116003.

[17] R. F. Mansour, N. M. Alfar, S. Abdel-Khalek, M. Abdelhaq, R. A. Saeed, and R. Alsaqour, "Optimal deep learning based fusion model for biomedical image classification," *Exp. Syst.*, vol. 39, no. 3, Mar. 2022, Art. no. e12764.

[18] G. Habib and S. Qureshi, "Biomedical image classification using CNN by exploiting deep domain transfer learning," *Int. J. Comput. Digit. Syst.*, vol. 10, pp. 2–11, Jul. 2020.

[19] M. B. Assad and R. Kiczales, "Deep biomedical image classification using diagonal bilinear interpolation and residual network," *Int. J. Intell. Netw.*, vol. 1, pp. 148–156, Jan. 2020.

[20] S. Pang, A. Du, M. A. Orgun, and Z. Yu, "A novel fused convolutional neural network for biomedical image classification," *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 107–121, Jan. 2019.

[21] A. A. Alanazi, M. M. Khayyat, M. M. Khayyat, B. M. E. Elnaim, and S. Abdel-Khalek, "Intelligent deep learning enabled oral squamous cell carcinoma detection and classification using biomedical images," *Comput. Intell. Neurosci.*, vol. 2022, Jun. 2022, Art. no. 7643967.

[22] A. Jaszcz, D. Polap, and R. Damasevicius, "Lung X-ray image segmentation using heuristic red fox optimization algorithm," *Sci. Program.*, vol. 2022, pp. 1–8, Jul. 2022.

[23] S. Z. Kurdi, M. H. Ali, M. M. Jaber, T. Saba, A. Rehman, and R. Damasevicius, "Brain tumor classification using meta-heuristic optimized convolutional neural networks," *J. Personalized Med.*, vol. 13, no. 2, p. 181, Jan. 2023.

[24] M. A. Khan, H. Arshad, R. Damasevicius, A. Alqahtani, S. Alsubai, A. Binbusayyis, Y. Nam, and B.-G. Kang, "Human gait analysis: A sequential framework of lightweight deep learning and improved moth-flame optimization algorithm," *Comput. Intell. Neurosci.*, vol. 2022, Jul. 2022, Art. no. 8238375.

[25] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, "Swin-UNet: UNet-like pure transformer for medical image segmentation," in *Computer Vision—ECCV*. Tel Aviv, Israel: Springer, 2023, pp. 205–218.

[26] M. T. Aziz, S. M. H. Mahmud, M. F. Elahe, H. Jahan, M. H. Rahman, D. Nandi, L. K. Smirani, K. Ahmed, F. M. Bui, and M. A. Moni, "A novel hybrid approach for classifying osteosarcoma using deep feature extraction and multilayer perceptron," *Diagnostics*, vol. 13, no. 12, p. 2106, Jun. 2023.

[27] L. Liu and Y. Zeng, "Intelligent ISSA-based non-singular terminal sliding-mode control of DC–DC boost converter feeding a constant power load system," *Energies*, vol. 16, no. 13, p. 4973, Jun. 2023.

[28] R. S. Kartha, "Localized skin disease detection and classification using an optimized hybrid deep learning model with attention mechanism," Tech. Rep.