A Novel Architecture of Depthwise Separable CNN and Multi-Level Pooling for Detection and Classification of Myopic Maculopathy

Alaa E. S. Ahmed
College of Computer and Information Sciences,
Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia
Shoubra Faculty of Engineering, Benha University, Cairo, Egypt

***Abstract*—Myopic maculopathy (MM), also known as myopic macular degeneration, is the most serious, irreversible, vision-threatening complication and the leading cause of visual impairment and blindness. Numerous research studies demonstrate that the convolutional neural network (CNN) outperforms many applications. Current CNN designs employ a variety of techniques, such as fixed convolutional kernels, the absolute value layer, data augmentation, and domain knowledge, to enhance performance. However, some network structure designing hasn't received much attention yet. The intricacy of the MM categorization and definition system makes it challenging to employ deep learning (DL) technology in the diagnosis of pathologic myopia lesions. To increase the detection precision of MM's spatial domain, the proposed work first concentrates on creating a novel CNN network structure then improve the convolution kernels in the preprocessing layer. The number of parameters is decreased, and the characteristic of a small local region is modeled using the smaller convolution kernels. Next channel correlation of the residuals with separable convolutions is employed to compress the image features. Then, the local features using the spatial pyramid pooling (SPP) technique is combined, which improves the features' capacity to be represented by multi-level pooling. The use of data augmentation is the final step in enhancing network performance. Compress the residuals in this paper to make use of the channel correlation. The accuracy achieved by the model was 95%, F1-score of 96.5% and AUC of 0.92 on augmented MM-PALM dataset. The paper concludes by conducting a comparative study of various deep-learning architectures. The findings highlight that the hybrid CNN with SPP and XgBoost (Depthwise-XgBoost) architecture is the ideal deep learning classification model for automated detection of four stages of MM.**

Keywords*—Retinograph; ophthalmologists; computer-aided diagnosis; vision loss; deep learning; retinograph images; myopic maculopathy

I. Introduction

Due to its fast-rising incidence internationally [1] and the risk to eyesight, myopia is presently a major public health issue. By 2050, it is predicted that 50% of the world's population will be myopic, with 10% of them having severe myopia [2]. Cataracts, glaucoma, retinal detachment, and myopia maculopathy can all be brought on by myopia (MM). As a result, organizations in the health sector like WHO, are confident that myopia might cause visual impairment. MM puts a hardship on patients, their families, and society as a whole. According to Naidoo et al., the global productivity loss resulting from MM might be $6 billion, and in 2050, myopia could impact nearly half of the world's population. This financial stress will probably get worse shortly. There is no recognized cure for MM as of yet. Preventive treatment, however, lessens ocular headaches and should be taken into account for all myopic patients. The International Photographic and Grading System for Myopic Maculopathy [3] identifies and categorizes myopic maculopathy. According to the severity of the condition, pathologic myopia was divided into five categories: category 1, just tessellated fundus, category 2, diffuse chorioretinal atrophy, category 3, patchy chorioretinal atrophy, category 2, and category 0, no macular lesions. In addition, characteristics including lacquer cracks, Fuchs spots, and choroidal neovascularization, are utilized to classify diseases. Additionally, the posterior staphyloma offers more details on the illness. In this study, myopic maculopathy is taken into consideration when a fundus picture image falls into category 2 or above.

The rapid advancement of artificial intelligence [4] is essential for the automation of challenging medical diagnoses and the analysis of clinical data. The most sophisticated category of AI is deep learning [5]. It uses deep artificial neural networks to solve feature-dependent issues while simulating the functioning of the human brain. The deep learning system (DLS) surpasses board-certified professionals in medical settings [6], [7]. The employment of DSL-based diagnosis software in ophthalmology's clinical and public healthcare settings has proved effective. Artificial intelligence (AI)-based medical imaging, such as retinal fundus pictures, is a valuable and effective option for managing and diagnosing MM. However, automated diagnostics based on CT scans are thought to be an image analysis challenge, which may be solved by labelling the data and applying machine learning techniques like deep learning. The rapid advancement of artificial intelligence [4] is essential for the automation of challenging medical diagnoses and the analysis of clinical data. The most sophisticated category of AI is deep learning [5]. It uses deep artificial neural networks to solve feature-dependent issues by simulating the functioning of the human brain. The deep learning system (DLS) surpasses board-certified professionals in medical settings [6], [7]. The employment of DSL-based diagnosis software in ophthalmology's clinical and public healthcare settings has

proved effective. Artificial intelligence (AI)-based medical imaging, such as retinal fundus pictures, is a valuable and effective option for managing and diagnosing MM. However, automated diagnostics based on CT scans are thought to be an image analysis challenge, which may be solved by labelling the data and applying machine learning techniques like deep learning.

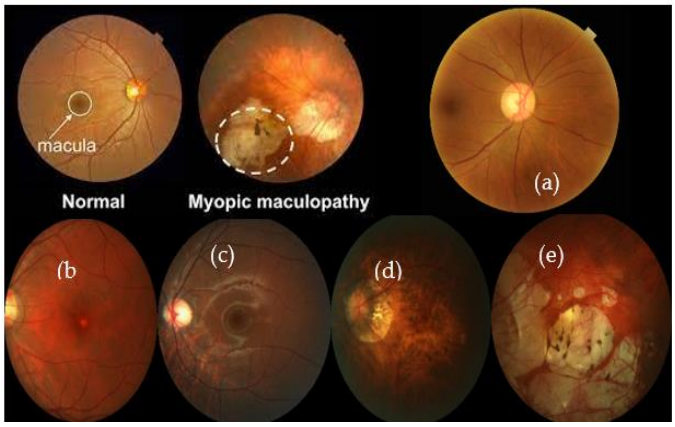

Fig. 1. Grading of myopic maculopathy, where figure (a) Shows the category 0:No macular lesions, figure (b) Category 1: Tessellated fundus, figure (c) Shows category 2: Diffuse chorioretinal atrophy, figure (d) Category 3: Patchy chorioretinal atrophy and figure (e) Represents category 4: macular atrophy.

Due to the intricacy of the categorization and characterization of the PM system, using deep learning approaches in PM lesion scanning is still difficult [11]. There was a lengthy period of disagreement on the precise definition of PM until a classification for MM was suggested by pathologic myopia (META-PM) meta-analysis. The severity of eyes with MM is approximately equal to or greater than that of eyes with spreading choroidal atrophy (Category 2), or eyes with at least one "plus" lesion are considered to have PM at this level of categorization. With such a categorization system in mind, creating an AI program to automatically recognize the PM and aid doctors in making a precise diagnosis is advantageous. Sufficient high-resolution PM retinal fundus picture dataset resources and a highly qualified staff are needed to do this. This study aims to construct and train DLSs that can automatically identify PM and categorize MM utilizing a beautiful dataset of color retinal fundus pictures gathered from the hospital's ophthalmology facilities. A visual example of the stages of MM is represented in Fig. 1.

An original model is offered in this paper. Convolutional neural networks and the cycle generative adversarial network (CycleGAN) [12] are combined to optimize the convolutional neural network (CNN). The suggested technique can locate lesion locations with less initial training data and can identify retinal disorders. With cycle consistency, CycleGAN can provide more trustworthy and realistic pictures. Adopting the discriminator and generator adversarial results in the best solution. Additionally, to differentiate the domain pictures, the classifier and generator cooperate [13]. A unique res-guided sampling block strategy is proposed using the combination of learnable residual features and pixel-adaptive convolutions. As a generator, a res-guided U-Net [14] is created, and conventional convolution is used in place of res-guided sampling blocks. Large training datasets are frequently required for supervised learning to account for all potential variances. However, gathering a lot of training data can be time-consuming, especially for medical imaging, where hand annotation is necessary. DepthCNN-XgBoost is one method for solving this issue since it takes a lot less training data than the standard method of using vast quantities of data [15]. Several variations [16], which can be broken down into the three main views of data, method, and model, were used to carry out the DepthCNN-XgBoost learning. The dataset was enhanced by the data-driven algorithm, which employed previous knowledge. The space is constrained by model-oriented approaches like embedding. Finally, from the perspective of an algorithm, it is comparable to tweaking the network weights by looking at data from a fresh sample. Therefore, rather than referring to specific learning algorithms, FSL refers to a general understanding of algorithms (such as supervised or unsupervised learning principles). In addition, different configuration stages, modelling, and formulation were required when applying FSL to various deep-learning classifiers.

As we previously explained, supervised learning models are used to train the deep learning models used for the area segmentation of MM RETINGRAPH images. These models are mostly based on DepthCNN-XgBoost and FCN structures. Therefore, their weight cannot be changed dynamically. There is a risk of problems if a large data sample is required for training. By suggesting a DepthCNN-XgBoost learning model in this study, where just a small sample of the network would be taught dynamically, we are able to get around this constraint. Our primary focus is pretrained learning-based classification, and we constantly update and improve weights by incorporating fresh sample data. Fig. 2 explains this DepthCNN-XgBoost learning approach. To the best of our knowledge, the dynamic updating of model weights is a new and original method. The DepthCNN-XgBoost scheme, which has been shown to be particularly helpful for detection of MM when diagnosis using retinograph images.

As of now, the MM eye-related disease has a unified region in retinograph images using a modified depthwise separable CN and XgBoost classifier. The findings were then evaluated by a domain expert during testing in order to categorize the output. Some of the samples were then picked for additional training. Due to the small amount of fresh ground truth data utilized as a training set, the deep model was able to learn to update its behavior dynamically with little modification to the learned behavior.

The following are our main contributions to myopia detection.

1) A novel deep learning technique is developed that recognizes the presence of myopia and categorizes it.

2) To decrease the number of parameters and enhance local features, we reduce the size of the convolution kernel in the preprocessing layer and initialize the kernels using 30 SRM basic filters [12]. Additionally, the suggested "forward-backward-gradient descent" approach is used to optimize the

convolution kernel in order to improve accuracy and hasten network convergence.

3) To replace the conventional convolution layer, we utilize two separable convolution blocks. In order to enhance accuracy and boost the signal to noise ratio, separable convolution may be used to extract the spatial and channel correlation of residuals.

4) Before feeding the feature maps to the segment of the network that is completely linked, we condense them using spatial pyramid pooling [19]. Through multi-level pooling, spatial pyramid pooling may enhance feature expressions and map feature maps to set lengths.

5) A loss function is created to increase the distance between the PM and HM classes by combining the triple loss with the binary cross-entropy loss (BCE loss). Our technique consistently produces the greatest outcomes and performs at the highest level when compared to deep learning models, machine learning models, and other methods. It ensures physicians' convenience and accuracy in clinics.

The primary contributions of this study are summarized below to address these issues: (1) The iChallengePM dataset will be used to create 12 DAMFs. To our knowledge, all of the operations employed in the present DA are covered by these DA approaches. Our goal is to enhance data features, control sample imbalance, and significantly boost dataset quality. (2) A variety of optimizers, loss functions, and learning rates are built using the AlexNet, VGG-16, GoogleLeNet, and ResNet-50 models as a foundation. Using training data from 12 datasets, the model with the highest accuracy will be used as the main learner. This approach will improve the model's capacity for generalization. (3) Following the training of the fusion model to create the final model, the main learner prediction indicated above will be utilized as a new input and added to the hard voting model. Without transfer learning, the model optimized by the aforementioned processes achieves great accuracy. More importantly, by utilizing the augmented dataset and the model fusion technique, we successfully avoid overfitting and enhance the model's generalization capability when processing different types of data, which further enhances the model's expressive capability. As a result, the model's ability to recognize complicated and uncommon case pictures will be much improved.

II. Research Background

In the early studies [17], images were segmented using methods based on edges, regions, clusters, and thresholds. These traditional techniques include manually extracting features, which are subsequently put to use for background separation, among other things. Additionally, the segmentation results are influenced by the feature quality, and this method is occasionally time- and labor-intensive. However, in recent years, research has evolved away from deep learning algorithms and toward traditional neural networks, particularly in the area of semantic picture segmentation [19, 20]. Additionally, as time has gone on, the recognition and forecast accuracy of these approaches have significantly increased. They were the first to use deep convolutional neural networks to segment semantic images. To create FCN, they switched out the convolution layer with a fully connected layer. One of the finest prototypes for the encoder-decoder architecture used for pixel-level image categorization is the FCN (Fully Convolutional Neural Network). Upsampling and transposed convolution might be used to reconstruct a whole segmented image with categorized pixels. Researchers now have the chance to train deeper and bigger neural networks thanks to the introduction of new GPUs and better algorithms. Compared to the original FCN, the suggested DeconvNet [20] is a more comprehensive decoder. The aforementioned encoder and decoder have the same number and size of features. In addition to deconvolution, the DeconvNet decoder employs unpooling layers to enhance the outcomes. Due to the encoder's fully linked layers, the DevconvNet also uses a lot more memory than FCS.

The settings and memory should be optimized. The SegNet [21], which is similar to the VGG-16 but different from the FCN and DeconvNet in up-sampling and convolution, thereby doing away with deconvolution, is introduced by Badrinarayanan et al. The feature maps are extremely well managed by SegNet. Inference, however, calls for additional memory. Generative adversarial networks (GANs) have recently achieved great success in a variety of applications [22] (e.g., DCGAN [18], SRGAN [17], and Pix2Pix [19]). The generative adversarial loss is calculated to determine the difference between the real and generated data distribution. The GAN was formally proposed by Goodfellow in 2014, and since then, it has operated on five adversarial processes that alternate between faking and identifying. Several researchers have discovered generative adversarial loss to be beneficial for improving network performance. In response to the success of GANs in image translation [23], a powerful GAN network for picture semantic segmentation is developed. It most closely resembles the approach put forth by Luc et al., in which adversarial networks help with semantic segmentation training. But there is no improvement over the starting point. Global data is included in fully linked CRFs (FullCRFs) by Deeplab as an independent post-processing step to further enhance CNN. Two orders of magnitude improve the speed of inference and training with this technique. Additionally, the incorporation of learnable transformations together with learnable Gaussian features outperforms and transforms a significant chunk of the inference into convolutions with the development of ConvCRFs, enabling efficient implementation on GPUs.

In various investigations [14, 15–16], clinicians used the CT scan to identify illnesses related to SMM. This study has two key benefits: (a) early viral infection patterns may be shown [15, 16], and (b) in 70% of patients, viral pneumonia-related CT abnormalities can be detected before laboratory testing [15]. As a result, early SMM infection identification is greatly aided by CT imaging. Detecting SMM in chest X-ray pictures has also been the subject of several investigations [7, 17]. We prioritize work involving CT scans nonetheless. According to SMM study findings, clinical symptoms do not typically present until after CT abnormalities [17].

Furthermore, asymptomatic people's chest CTs commonly show abnormalities that are consistent with viral pneumonia. On the one hand, certain patterns target unilateral, multifocal, and peripherally based ground-glass opacities. However, symptomatic groups were more likely to have

lymphadenopathy, pleural effusion, bronchiectasis, round cystic alterations, nodules, thickening of the surrounding pleura, and interlobular septal thickening.

The visual detection method should concentrate on identifying prominent lung abnormality patterns such as GGOs, crazy-paving patterns, consolidation, and linear opacities. However, the density and appearance of the sickness varied depending on the stage of the illness. The illness should manifest after nine days of early symptoms [14]. Deep learning-based algorithms are frequently used for detection, identification, or segmentation in medical imaging [18] and biomedical applications [19]. Researchers are looking at a number of strategies to assist medical personnel in SMM detection in this area. To categorize the many CT slices, convolutional neural network variants are first used [13]. With a ROCAUC value of 0.95, the applied approach may detect a viral infection; a score of 1.00 indicates a flawless classic. Even with a high detection rate, it proved challenging to distinguish between viral pneumonia using a simple CT scan. For coronavirus diagnosis, CNN variants have been proposed [20]. This method aids in the differentiation between instances of SMM, non-infection, and other viral infections. The findings indicate a good detection rate, much better than RT-PCR analysis. The accuracy of CNN is increased in the following stage by combining it with long-term memory networks [21]. The inf-Net parallel partial decoder, which integrates high-level features to produce a global map, has also been introduced [22]. Hierarchies of convolution are used for this.

Another choice to consider is U-Net structures. It will only be used for medicinal purposes [23]. The multistage technique includes the segmentation and categorization of SMM and other viral diseases [24]. Additionally, it aids in tracking the development of advanced illnesses. The methods utilized for SMM picture segmentation, which are based on U-Net topologies, are briefly detailed in [6]. The region of interest is first separated from the lung scan using U-Net. The categorization of SMM or other situations is then updated using a pretrained Resnet-50 [25]. AdaResU-Net [26], a multi-object adaptive CNN with the capacity to automatically adapt to new datasets and residual learning paradigms, was suggested in the following phase. For the purpose of SMM detection on high-resolution CT images, U-Net++ [8], a U-Net-based model, was also applied. Additionally, SMM's detection has been evaluated using Xception, ResNet-18, ResNet-50, ResNet-101, SqueezeNet, GoogleNet, VGG-16, VGG-19, and ResNet-19 [28]. ResNet-101 and Xception outperform the competition. In a different article, AlexNet and Inception-V4 were also used for SMM detection [29]. Additionally, to identify SMM, CNN and an Artificial Neural Network Fuzzy Inference System (ANNFIS) are used [30]. In a different study [31] proposes a Stack Hybrid Classification (SHC) approach based on ensemble learning.

Additionally, object-detection techniques are taken into account [32] for SMM diagnoses, and in another study, VGA variations were also employed to find symptomatic lung regions [33]. The suggested approach can differentiate between community-acquired pneumonia (CAP) and SMM (CAP). The Naive Bayes classifier, discrete wavelet transformations, and evolutionary algorithms are employed in study [34] for SMM identification. A suggested approach for MM RETINGRAPH image segmentation is integrated with super-pixel-based fuzzy-modified flower pollination and a type 2 fuzzy clustering method in a segmentation-based study [35]. For SMM image segmentation, volumetric medical image segmentation networks, or V-Nets [36], provide an option. Similar to this, V-Net was employed in a different research project to concurrently segment every MRI slice [37]. The quantitative findings support the viability and efficacy of infection-region marking. As we have already stated [38], deep learning techniques were crucial in the segmentation of lung CT images. They can now measure the degree of infection and judge the severity of the condition [40, 41]. Table I lists the deep learning methods applied to CT picture segmentation and SMM identification.

A large ground-truth dataset for training is a fundamental prerequisite of deep learning-based approaches, which can sometimes be quite challenging. Additionally, annotating the vast volumes of data is a labor- and time-intensive task. Due to these restrictions, deep learning techniques can only be used to solve real-world issues. A relatively small number of papers, including [22], where a semi-supervised learning strategy was applied with multiclass segmentation to identify the infected zone, have begun to examine this problem. However, the results of this strategy were subpar. In this work, we suggested a depthwise separable CNN method with XGBoost that enables a system to classify various stages of MM. This method does away with the requirement for a sizable dataset. Additionally, this system interacts with subject-matter experts to dynamically alter the settings. On the other hand, the weights in the existing models cannot be adjusted after training.

III. Related Work

Pathological myopia [8], often known as nearsightedness, is one of the severe forms of myopia. Because it might cause blindness, pathological myopia is also known as degenerative myopia. One might spot pathological myopia by looking at the diseases that develop in the posterior of the eye. Pathological myopia can cause several eye conditions, such as posterior staphyloma, vitreous opacities, Weiss' reflex, liquefaction, macular degeneration, cystoid degeneration, liquefaction, Foster-Fuchs' spot, etc. In this study, an automated method for the diagnosis of problematic myopia based on fundus pictures is constructed using a deep learning approach known as a convolutional neural network.

TABLE II. SUMMARY OF STATE-OF-THE-ART MACHINE LEARNING TECHNIQUES EMPLOYED FOR DETECTION OF PATHOLOGIC MYOPIA IN RETINAL FUNDUS IMAGES

Reference	Image Processing	Techniques/Models	Results	Advantage
Rauf et al. [8]	Grayscale, histogram, Red channel, Shuffle	The preprocessed images are then fed to the designed CNN model. The CNN model automatically extracts the features from the input images and classifies the images, i.e., normal image or pathological myopia.	AUC: 0.9845	Detect different stages of PM.
Li et al. [9]	Color histogram distribution	A dual-stream DCNN (DCNN-DS) model that perceives features from both original images and corresponding processed images.	sensitivities of 90.8% and 97.9% and specificities of 99.1% and 94.0% for detecting PM	Detect PM and TF
Devda et al. [10]	Morphological edge detection	A deep learning model with convolutional neural networks (CNN) is employed for classification, and the U-net model does image segmentation.	ACC of 97.8%	Detect PM
Hemelings et al. [24]	CNN and Semantic Segmentation	CNN is combined with lesion segmentation. Furthermore, domain knowledge is incorporated by using Optic Nerve Head (ONH)-based prediction to improve the segmentation of atrophy and fovea.	AUC of 0.9867 for PM detection, Euclidean distance of 58.77 pixels for fovea localization	Detect PM and fovea localization
Li Lu et al. [25]	NA	The author proposed a series of deep learning systems to detect myopic macular lesions and PM in accordance with the international photographic classification system (META-PM) using color fundus images.	AUC of 0.989	
Du et al. [26]		Deep Learning (DL) algorithms are proposed to identify the key features.	AUC values were 0.970, 0.978, 0.982, and 0.881	diffuse atrophy, 87.22% for patchy atrophy, 85.10% for macular atrophy, and 37.07% for choroidal neovascularization
Zhang et al. [27]		In [27], using ultra-wide field of view (UWF) fundus color imaging, a screening system named DeepUWF was developed, which can diagnose three kinds of fundus diseases (diabetic retinopathy, retinal tear, retinal detachment, and pathological myopia). This system is composed of CNN and two customer classifiers.		three kinds of fundus diseases (retinal tear & retinal detachment, diabetic retinopathy and pathological myopia)
Shi et al. [28]		A Myopia Detection Network (MDNet) is proposed that combines the advantages of dense connection and Residual Squeeze-and-Excitation attention to detect myopia in optos fundus images.	Mean Absolute Error of the Spherical Equivalent detected by this network can reach 1.1150 D	
Freire et al. [29]		First, different Deep Learning techniques are applied on fundus images, and, then transfer learning is applied on all tasks using Xception.		algorithms to diagnosis Pathological Myopia (PM) and detection of retinal structures and lesions such asOptic Disc (OD), Fovea, Atrophy and Detachment

The CNN was invented by Spyder. The characteristics are automatically extracted from the photos and categorized for pathological myopia. The metrics AUC = 0.9845 and validation loss = 0.1457 show the CNN model's excellent performance. The identification of pathological myopia from fundus pictures is therefore possible in the medical field using CNN.

MM, pathologic myopia (PM), and tessellated fundus were classified using a dual-stream DCNN (DCNN DS) model in [9]. (TF). It functions by taking characteristics out of the original image and applying them to an image that has been color histogram. The DCNN-DS model achieved sensitivities of 93.3% and 91.0%, specificities of 99.6% and 98.7%, and an AUC of 0.988 and 0.994 for identifying PM. According to the author's claims in [10], the suggested algorithm is trustworthy and has high sensitivity, specificity, and AUC to discriminate against various levels of MM on fundus images. A deep learning model using convolutional neural networks (CNN) is used for classification, and a DepthCNN-XgBoost model handles image segmentation. Devda et al. concentrate on segmenting lesions (atrophy and detachment), classifying nonpathological and pathological myopia images, detecting the fovea, and localizing the optical disc. Positive outcomes are produced by combining CNN with DepthCNN-XgBoost.

CNN and lesion segmentation are integrated in [24]. The segmentation of atrophy and fovea is further enhanced by applying optic nerve head (ONH)-based prediction, which incorporates domain knowledge. Segmentation, as opposed to detection or regression models, is used in this work to locate the fovea. Euclidean distance for fovea localization, AUC for PM detection, and F1 and Dice for semantic segmentation are some of the metrics that are used to evaluate the outcomes (optic disc, retinal atrophy, and retinal detachment). The model successfully

localizes the fovea at a distance of 58.77 Euclidean pixels and detects PMs with an AUC of 0.9867. The optic disc lesion, retinal detachment lesion, and retinal atrophy lesion F1 and Dice metrics for semantic segmentation of lesions are observed at .9303 and 0.9869, 0.8073 and 0.7059, and 0.8001 and 0.9135, respectively. To identify myopic macular lesions and PM in line with the worldwide photographic classification system (META-PM) using color fundus photos, the author of [25] presented a number of deep learning algorithms. Both the test and external validation.

Datasets are said to have robust performance. The identification of the relevant traits is presented using deep learning (DL) methods in [26]. Additionally, these models are used to create a meta-analysis for pathologic myopia (META-PM) classifying system (CS) by adding a specific layer. The DL models' sensitivity to choroidal neovascularization was 37.07%, 87.22%, 84.44%, and 85.10%, respectively, as were their sensitivity to patchy atrophy, diffuse atrophy, and macular atrophy. These are the relevant AUC values: 0.970, 0.978, 0.982, and 0.881. The META-PM study CS had an overall accuracy of 87.53%, with rates of 90.18%, 95.28%, 97.50%, and 91.14% for each kind of lesion, respectively.

A screening technique known as DeepUWF was created in [27], utilizing ultra-wide field of view (UWF) fundus color imaging, and it can diagnose three different fundus disorders (diabetic retinopathy, retinal tear, retinal detachment, and pathological myopia). This system is built using CNN and two customer classifiers. Six different image preparation approaches are also used to fix the low contrast problem with UAF photos. These preprocessing steps improve the networks' ability to learn and help them achieve high levels of sensitivity and specificity. The benefits of dense connection and residual squeeze-and-excite attention are combined in [28] to present a myopia detection network (MDNet) that can identify myopia in a fundus image. Following the extraction of the region of interest using the optical disc identification approach, the dataset is expanded using the data augmentation method. This network's capacity to recognize spherical equivalents with a mean absolute error of 1.1150 D (diopters) demonstrates the utility of this approach. In [29], transfer learning is used to complete all tasks using Xception after various deep learning algorithms are initially applied to fundus pictures. The optical disc segmentation algorithm pipeline also employs the YOLO design. The model is assessed using the following metrics: AUC-ROC, F1-Score, Mean Dice Score, and Mean Euclidean Distance. The approach has so far shown positive outcomes.

IV. Research Methodology

1) Data acquisition: The training dataset for this study was acquired from the event hosted by the International Symposium on Biomedical Imaging (ISBI-2019) in Italy. It contains 400 labelled funds images. The dataset consists of 239 pathological myopic eye images and 161 normal eye images. The image size is 1444×1444×3 (RGB image). The database is available at https://palm.grand-challenge.org/. Fig. 2 shows the preprocessing step.

2) Proposed method: The initial step involves preprocessing the image (as illustrated in subsection A) to emphasize particular patterns, which aids in effectively training Deep Learning models for classification purposes. Overall steps of proposed system is described in Algorithm 1. This algorithm outlines the steps for extracting features using Depthwise Separable CNN and Multi-Level Pooling. It applies depthwise separable convolution to capture spatial features efficiently and then performs multi-level pooling to reduce dimensionality and retain important information. The extracted features are then used to train a CNN model for classification.

The proposed framework for steganography based on a convolutional neural network (CNN) is illustrated in Fig. 3. The proposed CNN architecture consists of several layers that take a 256 x 256 input image and generate two class labels, namely "Normal" and "Pathological Myopia". The network includes an image preprocessing layer, two separable convolutions (sepconv) blocks, four fundamental feature extraction blocks, a spatial pyramid pooling (SPP) module, and two fully connected layers followed by a softmax function. The convolutional blocks consist of four blocks known as "Basic Blocks 1" to "Basic Blocks 4," which perform operations to capture the spatial relationships between feature maps and transmit this information to the fully connected layer for classification using the XgBoost classifier. Each Basic Block carries out a set of actions to achieve this.

A. Image Preprocessing

In the initial stage of processing, we adjust the size of the convolution kernel and employ 30 fundamental SRM filters [12] to set the kernels to minimize parameters and enhance local features. We also use the "forward-backward-gradient descent" approach to optimize the convolution kernel, thereby improving accuracy and speeding up the network's convergence.

B. Convolution Layer

Instead of using bigger convolution kernels like 55, employed in earlier publications [18], [20], we use compact convolution kernels like 33 in our CNN design to limit the number of parameters. The number of parameters is decreased while the extraction of local characteristics is made effective by the use of small convolution kernels. As a result, we decided to use a convolutional kernel size of 3 with 32 channels for each of the first four Basic Blocks. The performance of the network and computational complexity are carefully analyzed to determine the number of channels for each fundamental block.

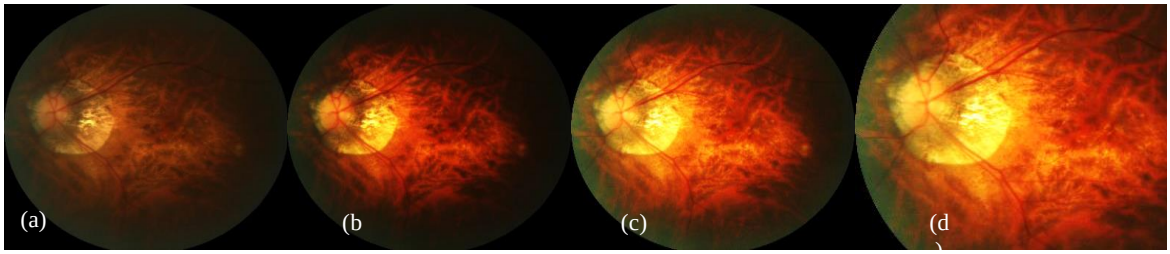

Fig. 2. A preprocessing step to enhance the contrast while adjusting the light illumination, where figure (a) shows input image, (b) enhance the contrast, and (c) light adjustment. Also, the figure (d) shows the region-of-interest (ROI).

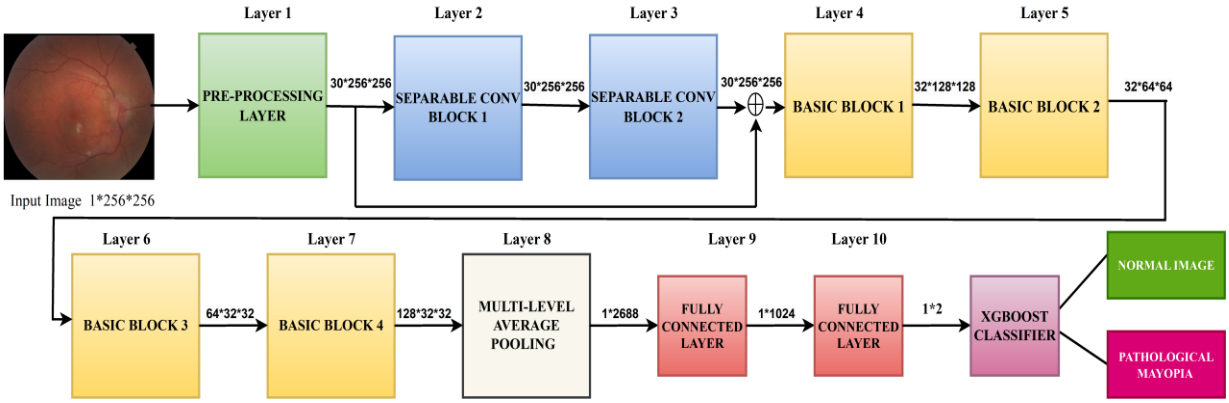

Fig. 3. The systematic flow diagram of proposed system.

C. Batch Normalization (BN) Layer

Batch normalization [25] is a technique commonly employed during training to normalize the distribution of each mini batch, typically resulting in a zero mean and unit variance. According to study [25], incorporating a BN layer in deep neural networks can prevent the issues of gradient vanishing or explosion and overfitting. Moreover, it allows for a reasonably high learning rate, which helps in achieving faster convergence. After conducting experiments, we observed that networks similar to Ye-Net that do not have BN are highly vulnerable to inadequate parameter initialization and may not reach convergence. These findings were noted in the study [20]. Therefore, BN is utilized in the proposed approach.

D. Non-Linear Activation Function

We use the traditional rectifying linear unit (ReLU) as the activation function for each block in the Zhu-Net to avoid gradient vanishing or exploding issues, hasten network convergence, and achieve several additional aims. Utilizing ReLU on neurons during training can teach them to respond exclusively to inputs that carry significant signals, which can enhance the creation of more efficient features. The ReLU function is useful and makes computing back-propagation gradients easier. In our research, we made use of the network shown in Figure to assess the performance of other activation functions, including the truncated linear unit (TLU) suggested in Ye-Net, for purposes alongside ReLU. ReLU is used as the activation function to train the entire model, comprising all of its layers and building components. The utilization of ReLU results in enhanced performance and accelerated convergence.

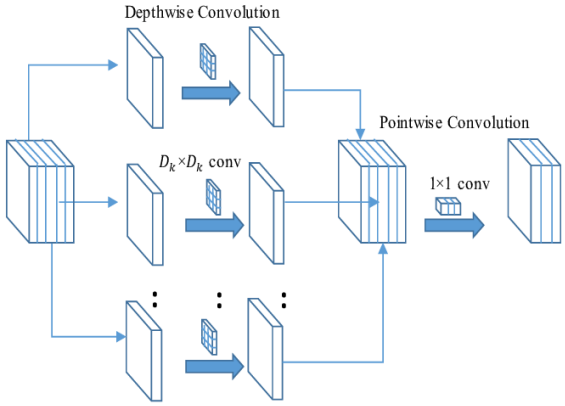

Fig. 4. Depth-wise separable convolution layer.

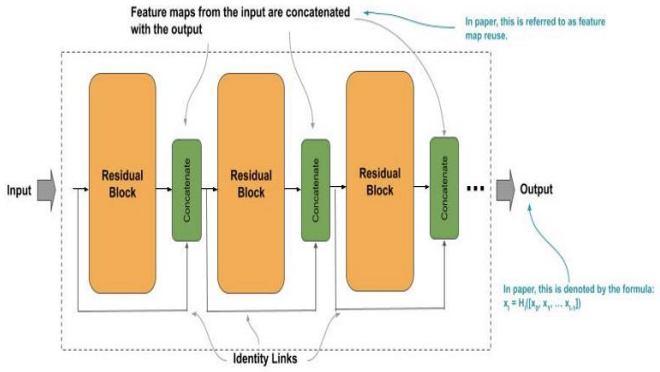

Fig. 5. A visual architecture of residual block is utilized in this work to build the model.

E. Average Pooling Layer

Average pooling layers, which can improve receptive fields, shrink feature maps, and improve image feature abstraction, are included in the first three essential building blocks. Moreover, average pooling improves the network's generalization capacity. To prevent information loss, the network's first block does not use pooling. Furthermore, we utilize separable convolution blocks (Sepconv Blocks 1 and 2) to enhance the SNR (the ratio of the signal-to-noise in the stereo signal and the resultant image) and efficiently manage spatial and channel correlations. In the last stage, we utilize an SPP module to improve the extraction of features. Through the application of multi-level pooling, the SPP module enhances the representation of features. At the conclusion of the suggested approach, a trio of fully connected layers are implemented, consisting of 2688, 1024, and 2 neurons in each layer. The ultimate layer with complete connectivity employs a softmax activation function to establish scores for the two class labels.

1) Separable convolution architecture: Recent achievements in computer vision projects like Inception [29], Xception [30], and other architectures have been made possible by separable convolution. In Fig. 5, we can observe Xception, which is a modified version of the Inception module (a). In this particular Inception variation, known for its extreme approach, the interdependence between channels is entirely eliminated, resulting in a boost in model expressiveness and storage efficiency. When the layer has been preprocessed, we create the relevant sepconv blocks using two separable convolution blocks made up of a 1x1 and a 3x3 convolution. This is done in order to use the leftover information from the normal and pathological myopia images more efficiently (as shown in Fig. 1). In our configuration, we presume that the residual correlations in the spatial and channel domains are independent of one another. Each feature map produced by the high-pass filter can be subjected to group convolution using the sepconv block. Fig. 4 depicts the structure of sepconv blocks. A sepconv block consists of three repetitions of both a 1x1 pointwise convolution and a 3x3 depthwise convolution.

In order to extract spatial correlations, a convolution operation with a depth of 3 x 3 is first carried out, employing a total of 30 groups. Pytorch uses the "groups" argument to implement separable convolution. After that, a pointwise convolution is performed in a sepconv block to get rid of any remaining channel correlations. After the initial 1-1 convolutional layer of sepconv block 1, we add an ABS layer [26] to help our model recognize the symmetry present in the noise residual. The two sepconv blocks integrate residual connections to improve classification performance and prevent gradient vanishing or explosion. Fig. 5 represents the visual architecture example used in our proposed model. It's important to note that the second SepConv block lacks an activation mechanism. We chose to employ the ABS layer in the first sepconv block even though depthwise separable convolutions are typically applied without nonlinearities. It somehow boosts the network performance. The optimized kernel and the hyper parameter description are as follows:

F. Optimizing Kernels

Modeling the residuals rather than the pixel values will produce more robust characteristics. The convolution kernels in the preprocessing layer are constant during training for the Yedroudj-Net and Xu-Net architectures. We built a preprocessing method termed "forward-backward-gradient descent" to improve the SRM feature sets that were manually created using domain-specific expertise.

To determine the residual, we follow this method: Firstly, we take each image X = Xij and compute the residual R = Rij as

$$\mathrm{Rij} = \mathrm{X\,pred\,(Nij)} - \mathrm{cXij}, \quad (1)$$

In this equation, c is an integer that represents the residual order, Nij denotes the neighboring pixels of Xij, and X pred (.) is a predictor of cXij based on the values of Nij. Generally, we utilize high-pass filters to obtain X pred ()

Algorithm 1: An algorithm for feature extraction using Depthwise Separable CNN and Multi-Level Pooling
Input:
- Images dataset (X) with corresponding labels (Y)
- Hyperparameters: number of layers (L), filter sizes (F), pool sizes (P), depthwise separable convolution parameters (D)
- Number of classes (C)
Output:
- Extracted features (X_features)
- Updated labels (Y)
Algorithm:
1. Initialize an empty list X_features.
2. Initialize an empty list Y.
3. For each image x and its corresponding label y in the dataset:
- Perform depthwise separable convolution on x with parameters D, resulting in feature maps.
- Perform multi-level pooling on the feature maps with pool sizes P, resulting in pooled feature maps.
- Flatten the pooled feature maps to obtain a 1D feature vector.
- Add the feature vector to X_features.
- Add the label y to Y.
Convert X_features and Y to numpy arrays.
Split X_features and Y into training and testing sets.
Initialize a depthwise separable CNN model.
Add L convolutional layers to the model, each with filter size F.
Add a fully connected layer with C neurons for classification.
Compile the model with an appropriate loss function and optimizer.
Train the model using X_features_train and Y_train, and validate it using X_features_test and Y_test using XGBoost classifier as described in section 4.4.
Evaluate the model's performance metrics such as accuracy, precision, recall, etc.
Return the trained model for future predictions.

During the backpropagation step of each iteration, we utilize the stochastic gradient descent (SGD) algorithm to update filter weights. Earlier studies have demonstrated that a Convolutional Neural Network (CNN) with weights initialized randomly typically fails to converge. To address this, previous research has often utilized SRM kernels to initialize the weights of the initial layers in order to generate a group of prediction errors based on pixel values, which can enhance the performance of the CNN.

G. Hyper-Parameters

We utilize mini-batch stochastic gradient descent (SGD) as the training approach for the CNN networks, while setting the momentum and weight degradation values to 0.9 and 0.0005, respectively. Due to limitations in GPU memory, the training mini-batch size has been established as 16, comprising 8 Normal/Myopia pairs. After that, the networks undergo training to reduce the cross-entropy loss using the variables mentioned earlier. During training, we modified the learning rate, which was initially set to 0.005. At certain predetermined steps throughout the training process, this modification entails dividing the learning rate by five. Specifically, during the 400-epoch training of a CNN, the learning rate will decline during epochs 50, 150, and 250. In the final phases of training, using a slower learning rate can significantly decrease training loss and boost accuracy. To prevent over-fitting, stopping training before reaching 400 epochs is common. This means that training is stopped when the cross-entropy loss on the training set continues to decrease, but the accuracy on the validation set begins to decrease. Specifically, during the 400-epoch training of a CNN, the learning rate will decline during epochs 50, 150, and 250. In the final phases of training, using a slower learning rate can significantly decrease training loss and boost accuracy. The testing accuracy was used to evaluate the performance. The proposed DSC-XGBOOST structure for identifying myopia anomalies is shown in Fig. 3.

2) Classification: The XGBoost algorithm has become the preferred tool for many data scientists, as it is a highly sophisticated algorithm capable of managing any kind of data abnormalities. Crafting a model using XGBoost is effortless, but enhancing it with XGBoost is arduous, at least in my experience. There are several things to take into account when using this strategy. To improve the performance of the model, it is essential to modify certain parameters. Nonetheless, it can be challenging to come up with a satisfying response to practical queries such "What is the perfect parameter setup for optimal results?"

3) Regularization: While XGBoost is known as a "regularized boosting" method, standard GBM lacks regularization, which helps to avoid overfitting. Moreover, XGBoost uses parallel processing, which accelerates performance compared to GBM. Yet, because the boosting process is sequential, it begs the question of how parallelization is even conceivable. What prevents us from building a tree employing all cores at once if each tree can only be formed after the one before it?

The XGBoost algorithm tries multiple approaches to handle missing data in each node while the user inputs a unique value as a parameter. In GBM, the tree pruning strategy is employed to prevent further division of that node if split results in a loss. In comparison to GBM, the XGBoost algorithm is more greedy because it prunes the tree backward and eliminates splits that don't offer any additional benefits. Moreover, XGBoost allows positive loss splits even after negative loss splits, something that GBM does not. For instance, XGBoost would continue and preserve both divides if they resulted in a total effect of +8, while GBM would stop at a split of -2.

Built-in Cross-Validation: XGBoost simplifies the process of obtaining the perfect number of boosting iterations by enabling users to perform cross-validation at every stage of boosting. This is different from GBM, which requires a grid search and only permits a limited number of variables to be analyzed. Using the latest iteration of an XgBoost model as the starting point for training can be extremely advantageous in specific contexts. The GBM implementation in sklearn includes the same capability, so both XgBoost and GBM are equally equipped. However, XgBoost may produce unstable models due to overfitting on the training set. To avoid this, regularization techniques can be employed to consider the model's complexity and prevent overfitting. In XgBoost, including a term that measures the model's complexity can modify the cost function. The two parameters used for regularization in XgBoost are alpha and lambda, which correspond to L1 regularization (Manhattan distance) and L2 regularization (squared Euclidean distance), respectively [1]. In order to implement L2 regularization, we need to assign a value to the reg lambda parameter in XgBoost.

The term "extreme gradient boosting," abbreviated as "XgBoost," refers to a method of gradient boosting that has been rigorously analyzed and parallelized to minimize the training time of the entire boosting procedure drastically. Instead of the traditional approach of creating the best possible model based on the data and then selecting it, we train numerous models on various subsets of the training dataset and choose the one that performs the best by gathering the results from all the models. XGBoost is often superior to standard gradient-boosting techniques in various scenarios. A vast array of key parameters can be adjusted for improved precision and accuracy by utilizing the Python implementation.

Consider a function or an approximation, and then generate a sequence of values based on the gradients of the function. The subsequent formula models a particular form of gradient descent. The loss function indicates the direction of the function's descent, which needs to be minimized. The fitted change rate is equivalent to the learning rate used in gradient descent. It is expected to match the behavior of the loss function accurately.

$$F_{x_i} = F_{x_i} + \alpha_{x_i} \frac{\partial}{\partial x}(x_t) \tag{2}$$

To find the best definition of the model, we need to describe the formula as a sequence and find a function that efficiently converges to its minimum. This function will be used as an error metric to help us reduce loss and maintain performance over time. Eventually, the sequence will reach the minimum of the function. This notation defines the error function for assessing a gradient boosting regressor.

$$f(x, \theta) = \sum l(f(x_i, \theta), y_i) \tag{3}$$

The following are the steps involved in the XgBoost algorithm:

XgBoost algorithm classifier for severity level of Myopic Maculopathy.	
Steps	Given training data from the instance space
1	Space S_1= $\{(x_1,y_1)\}$, where $S = \{S_1, S_2, S_3, \ldots, S_n\}$
2	[Initialize], $D_1(i) = \frac{1}{m}$
3	Repeat: For i= 1,2,3,…, n do
4	Train a weak leaner $h_i: x \to R$ using distribution D_i
5	Update the distribution over the training set:
6	$D_{i+1}(k) = \frac{D_i(k)e^{-\alpha_i}}{Z_i}$ (4)
7	Where Z_i is a normalization factor D_{i+1} chosen so that D_{i+1} will be a distribution
8	[end for]
9	$f(x) = \sum_{i=0}^{n} \alpha_t h_t(x)$ *and* $H(x) = sign\,(f(x)$ (5)

V. Experimental Results

A. Data Augmentation

The augmentation approach is used to make the balance among classes of MM. The network becomes resistant to certain alterations in this way. The integration of spatial information, which is essential for image segmentation tasks, is a strength of CNNs and DepthCNN-XgBoost in particular, although they are not equally resilient to transformations like scaling and rotation. The network may get the necessary invariance and resilience properties through the use of rotations and flips, two data augmentation techniques. The data augmentation also included shears, a derivation of elastic deformations recommended as a general best practice for convolutional neural networks and flips and rotations. The ImageDataGenerator function of Keras is used to implement the augmentation. Fig. 6 is visually displayed the distribution after performing data augmentation.

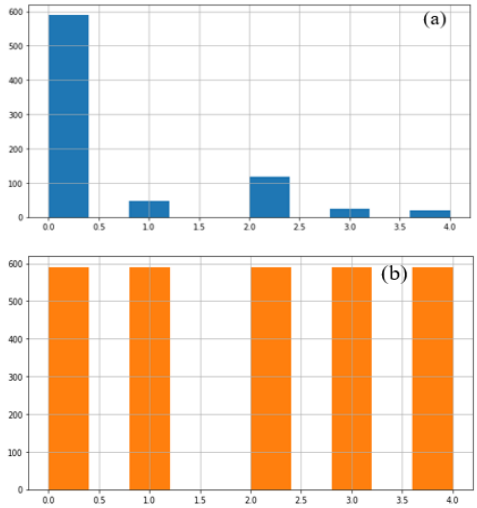

Fig. 6. Sample distribution of PALM dataset, where figure (a) shows the original images, whereas figure (b) shows the number of images after data augmentation.

B. Experimental Setup

All networks were developed in Python using the Keras and TensorFlow packages. The models were trained on an NVIDIA Tesla P4 GPU supplied by Google Colab. The test was run on a computer with an 8-core AMD FX-8320 CPU running at 3.5 GHz and 8 GB of RAM.

C. Assessment Criteria

The false positive rate (FPR) is the proportion of times a biometric system incorrectly accepts a fake subject. The false negative rate (FNR) is a biometric system that wrongly rejects the percentage of times a legitimate subject. Finally, the proportion at which FPR and FNR are identical is referred to as the equal error rate (EER). The binary classification error rates are shown graphically by the detection error trade-off (DET) curve. The FPR is on the x-axis, and the corresponding FNR is on the y-axis in this curve. The system's effectiveness is evaluated using a verification system comprising EER and DET curves. Since it tries to match the biometrics provided by a person with the precise biometrics already enrolled, it is sometimes referred to as a "1-to-1 matching system."

The identification system is represented as a 1-to-n matching system, in contrast to verification systems, where n is the total number of records in the database. Here, rank-1 IR and the CMC curve are used to evaluate the framework's performance. The rank-k identification rate is the proportion of times the true subject's match score appears in the top k matches (IR). A 1:1 identification system may have its performance evaluated using the cumulative match curve (CC). Plotting a curve between rank-k IR on the y-axis and rank-k on the x-axis illustrates it. Using a number of other assessment metrics, such as specificity, sensitivity, F1-score, accuracy, recall, and precision, the effectiveness of the suggested approach is measured in numbers.

Accuracy (ACC) is one of the most frequent and fundamental performance indicators. It is simply the likelihood that a randomly chosen example (positive or negative) will be true. In this measure, the diagnostic test shows how likely it is that the correct result will happen or how likely it is that the diagnosis is correct.

$$Accuracy\ (ACC) = \frac{TP+TN}{FP+FN+TP+TN} \quad (6)$$

The ability to properly identify positive categories within whole expected positive classes is referred to as precision, and it is stated as a ratio of all successfully predicted positive categories to all correctly expected positive categories:

$$Precision\ (PR) = \frac{TP}{TP+FP} \quad (7)$$

Sensitivity (SEN), Recall, True Positive Rate, Hit Rate: It is a measure of a model's capability to detect all positive instances and is represented as:

$$Recall(RE) = \frac{TP}{TP+FN} \quad (8)$$

It's worth noting that the above equation implies that a low false-negative rate almost always accompanies a high recall.

Specificity (SPE): Ratio of true negatives to total negatives in the data. Mathematically can be represented as follows:

$$Specificity = \frac{TN}{TN+FP} \quad (9)$$

F1-score: It is not as straightforward as accuracy, but this metric is useful in determining the classifier's exact and robustness. The F1 score, which is a key metric that considers both recall and precision for performance testing, it could be represented as follow:

$$F1 - score = 2.\frac{Precision \times Recall}{Precision \times Recall} \times 100\% \quad (10)$$

Where TN (true negative) and TP (true positive) are accurately predicted negative and positive outcomes, respectively. FN (false negative) and FP (false positive) do not predict negative and positive human identification cases correctly.

AUC: This stands for area under the receiver operating characteristics (AUC). The AUC is a graphical representation or plotting the diagnostic ability of any machine learning classifier using all thresholds.

D. Hyper-parameters Fine-tune

To determine the optimal hyper-parameter values for the optimizer and initial learning rate, a grid search was conducted. This search involved considering five different optimizers: stochastic gradient descent (SGD), SGD-Momentum, Nesterov Accelerated GD, RMSProp, and ADAM. The initial learning rate was varied within the range of 10^-1 to 10^-4. The training process involved each benchmark CNN being initially trained with the hyper-parameter values specified in their respective papers. However, due to factors such as the relatively small size of the training dataset compared to the larger datasets they were originally trained on ImageNet and the differences in discriminative features among classes, these initial attempts resulted in poor learning.

For each combination of optimizer and initial learning rate in the grid, the CNN models were trained for 20 epochs using the modified EyePACS train-set. Specific parameter values were set for each optimizer: a momentum of 0.9 for SGD and Nesterov Accelerated GD, a discounting factor (ρ) of 0.9 and a stability factor (ε) of 0.1 for RMSProp, and exponential decay rates β1 and β2 of 0.9 and 0.999 respectively, along with a stability factor (ε) of 1e-7 for ADAM. The best-performing optimizer and initial learning rate pair, which resulted in the highest training accuracy within the 40 epochs, was selected as the optimal combination of hyper-parameters for each benchmark CNN.

E. Result Analysis

The experimental results took numerous classification-related performance metrics into account, including the identification skills and the average computational time required by a trained network to fully annotate a CT image. Five performance metrics are typically considered when assessing a classifier: accuracy, precision, recall, F1-score, and AUC. The DepthCNN-XgBoost model beat the other deep learning models in terms of classification accuracy for detecting the MM-infected regions. It has been found that class imbalance may be used to explain the difference in accuracy, F1-score, and AUC. The majority class (no detections) was almost always classified properly. On the edges of infected regions in photos, false-negative detections were discovered when MM symptoms were plainly discernible. Nevertheless, the minority class (MM symptomatic regions) was discernible because the F1-score and AUC were both reasonably high. Model loss and accuracy curves are visually displayed in Fig. 7.

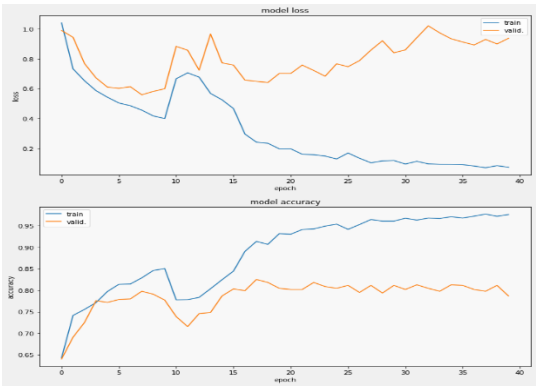

Fig. 7. Accuracy and validation loss curve for the proposed architecture.

The DepthCNN-XgBoost model's higher generalization capabilities when compared to the CNN, LSTM and CNN-LSTM. The current work aimed to decrease false positives since erroneous detections in medical imaging applications are crucial (normal areas are diagnosed as symptomatic). Additionally, the pandemic has raised the need for chest CT scan interpretation. With this in mind, we concentrated on minimizing radiologists' burden by striving for a high proportion of true positives (symptomatic areas diagnosed as symptomatic). In this situation, it is important to investigate the techniques that may result in extremely high accuracy and appropriate recall scores. DepthCNN-XgBoost and CNN-LSTM occasionally outperformed the traditional CNN because of their high-precision scores, which included both FP and TP values, even though all three models produced the same results for the F1-score, AUC, and accuracy. The result of AUC curve in Fig. 8 shows that the higher AUC value of 0.92 achieved by the proposed Depthwise-XgBoost with data augmentation compared other classifiers. Similarly, Fig. 9 shows the confusion matrix of the proposed system without data augmentation.

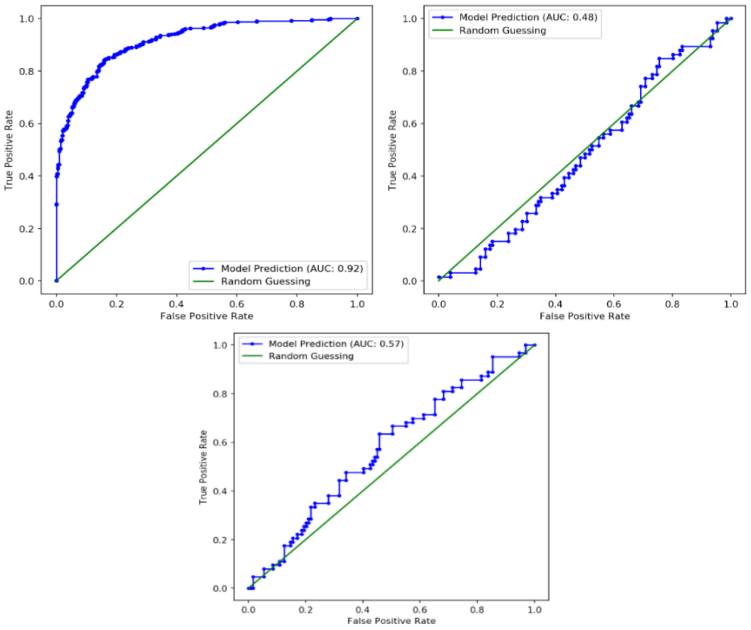

Fig. 8. Three AUC curves for figure (a) Proposed depthwise-XgBoost with data augmentation, (b) Original depthwise separable CNN, and (c) XgBoost classifier.

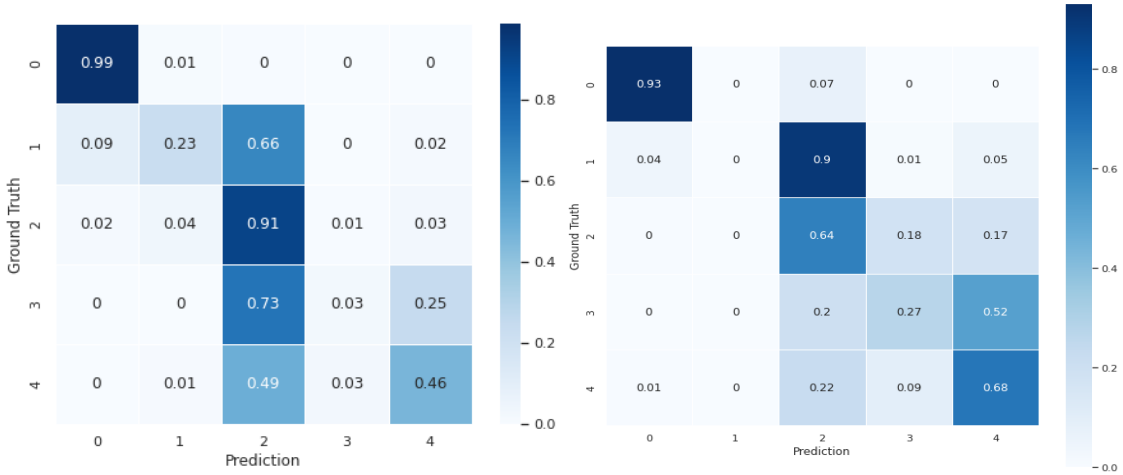

Fig. 9. Confusion matrix of proposed model for recognition of each class of MM.

Here we include a detailed comparison of various machine learning classifiers' performance in detecting and classifying stages of Myopic Maculopathy, using three different train-test partition strategies, Table II, Table III and Table IV. Each table assesses the classifiers based on Accuracy (ACC), Precision (PR), Recall (RE), and F1-Score. In the 70%-30% train-test partition strategy, the classifiers including CNN, LSTM, CNN-LSTM, Depthwise Separable, and the proposed Depthwise-XgBoost, all showcase high performance with the metrics mostly in the mid-90s percentile. The proposed Depthwise-XgBoost model exhibits a competitive edge with a 96.5% F1-Score. When the partition strategy shifts to 80%-20%, the classifiers show similar or slightly improved performance. Notably, the Depthwise-XgBoost stands out with the highest precision of 98% and maintains a robust F1-Score of 96.5%.

This indicates a consistency in the model's performance even as the data partitioning varies. The 90%-10% partition further underscores this consistency and, in some cases, an increase in accuracy and other metrics for all classifiers. The CNN-LSTM model achieves the highest accuracy at 97%, while the Depthwise-XgBoost maintains its high precision and F1-Score, emphasizing its reliability and effectiveness across different data distributions. Overall, the comparisons indicate that the advanced machine learning techniques, particularly the proposed Depthwise-XgBoost, are highly effective in diagnosing Myopic Maculopathy. The consistent performance of the Depthwise-XgBoost across various partition strategies highlights its potential as a robust and reliable model for medical diagnostic purposes. Each classifier demonstrates strengths in different metrics, but collectively they underscore the capability of deep learning architectures in enhancing the accuracy and reliability of medical diagnoses in ophthalmology.

TABLE III. CLASSIFICATION PERFORMANCE OF THE PROPOSED METHOD WITH OTHER MACHINE LEARNING CLASSIFIERS USING 70%-30% TRAIN-TEST PARTITION STRATEGY

Classifier	ACC	PR	RE	F1-Score
CNN	95%	95%	95%	96%
LSTM	94%	93%	96%	95%
CNN-LSTM	96%	94%	97%	96%
Depthwise Separable	95%	94%	97%	96%
Proposed Depthwise-XgBoost	95%	96%	97%	96.5%

TABLE IV. CLASSIFICATION PERFORMANCE OF THE PROPOSED METHOD WITH OTHER MACHINE LEARNING CLASSIFIERS USING 80%-20% TRAIN-TEST PARTITION STRATEGY

Classifier	ACC	PR	RE	F1-Score
CNN	95%	96%	94%	95%
LSTM	94%	95%	94%	95%
CNN-LSTM	96%	95%	96%	96%
Depthwise Separable	95%	95%	95%	95%
Proposed Depthwise-XgBoost	95%	98%	97%	96.5%

TABLE V. CLASSIFICATION PERFORMANCE OF THE PROPOSED METHOD WITH OTHER MACHINE LEARNING CLASSIFIERS USING 90%-10% TRAIN-TEST PARTITION STRATEGY

Classifier	ACC	PR	RE	F1-Score
CNN	96%	94%	96%	96%
LSTM	95%	95%	95%	95%
CNN-LSTM	97%	96%	96%	97%
Depthwise Separable	96%	96%	96%	96%
Proposed Depthwise-XgBoost	95%	98%	97%	96.5%

In addition to comparing various machine learning classifiers, the article presents Table V, which contrasts the performance of the proposed method with other existing studies in the field of Myopic Maculopathy detection and classification. The comparison is based on four key metrics: Recall, Precision, F1-score, and Accuracy. The table lists several methods from different researchers, including Rauf et al., Li et al., Devda et al., Li Lu et al., Du et al., Zhang et al., along with the proposed method. Each method's performance is quantified, demonstrating a range of effectiveness in diagnosing Myopic Maculopathy. For instance, Rauf et al. show balanced performance across all metrics at 94%. Li et al. have a notably high precision of 97% but lower accuracy at 88%. Other methods like Devda et al. and Li Lu et al. present a balanced mix of recall, precision, and accuracy, reflecting the diversity in effectiveness and approach among different studies. The proposed method distinguishes itself at the end of the table, demonstrating superior recall (97%), precision (98%), and an F1-score of 96.5% with an accuracy of 95%. These numbers indicate a high level of reliability and precision in detecting and classifying Myopic Maculopathy, surpassing the other methods listed. This comparison not only underscores the proposed method's robust performance but also contextualizes it within the broader landscape of existing research, highlighting its potential as a significant advancement in the field. A visual result of the proposed system is also displayed in Fig. 10 to detect different classes of MM.

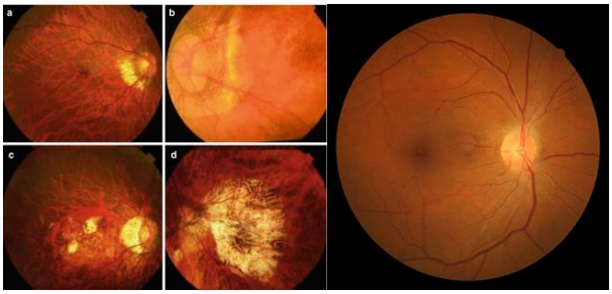

Fig. 10. Color fundus photographs showing the worsening levels of myopic macular degeneration; (a) Category 1, (b) Category 2, (c) Category 3 (c), and (d) Category.

TABLE VI. COMPARISON OF THE PROPOSED METHOD WITH OTHER EXISTING STUDIES

Method	Recall	Precision	F1-score	Accuracy
Rauf et al [8]	94%	94%	94%	94%
Li et al /pol	82%	97%	89%	88%
Devda et al [10]	86%	96%	91%	94%
Li Lu et al [25]	90%	92%	91%	87%
Du et al [26]	83%	89%	82%	93%
Zhang et al [27]	94%	96%	95%	95%
Proposed Method	**97%**	**98%**	**96.5%**	**95%**

VI. DISCUSSION

This work suggests the identification and classification of myopia maculopathy (MM) from retinograph pictures, utilizing multi-layer deep learning and pretrained learning techniques. In reality, a number of conditions, such as myopia maculopathy, can be followed by cataracts, glaucoma, retinal detachment, and other conditions (MM) as described in Table VI.

The World Health Organization consequently recognizes myopia as a significant factor in visual impairment if it is not completely treated. On patients, their families, and society as a whole, MM imposes a heavy cost. For MM, there is presently no effective therapy. For all myopic individuals, preventive treatment can lessen ocular problems. According to the International Photographic Classification and Grading System for Myopic Maculopathy [3], myopic maculopathy was identified and categorized. Myopia is classified according to its severity. In this study, identifying myopic maculopathy for fundus pictures in categories 2 and above is explored. Deep learning algorithms have lately been the subject of several academic studies aimed at segmenting MM-infected areas. In pixel-based segmentation for medical pictures, fully convolutional networks and U-shaped convolutional networks perform exceptionally well. So, when separating the MM-infected part of the retinal fundus picture, both are given top priority.

The adoption of deep learning (DL) technologies in identifying pathologic myopia (PM) lesions remains a difficulty

due to the complexity of the PM classification and definition system. However, enough resources can achieve objectives, such as high-quality PM retinal fundus picture collections and high-caliber expert teams. This study aims to create and train DLs to recognize PM as well as the categories. In this article, we used a novel deep learning model based on depthwise separable convolution layer for the detection of MM using fundus images. Motivated by the DSC model's outstanding results in various research disciplines. They perform well on small number of samples and allow.

In this instance, we employed the idea of depth-wise separable. A small number of training examples are used at a time in this online learning process. According to this process, new fundus pictures are fed to the model with the approval of subject-matter experts, and the model outputs are also assessed by experts throughout the testing phase to detect findings that were incorrectly categorized. Despite the fact that this technique improved segmentation performance and provided a foundation for online learning, Nevertheless, it required expert input throughout the algorithm's testing stage. As a result, human participation is required throughout the learning process with this technique. Although unlikely, errors in judgment made by the medical experts would have led to a decline in network performance. The same problem arises when an incorrectly labeled dataset is introduced to the network in the supervised learning paradigm. But because the supervised learning ground truth data were generated offline, there was plenty of time to evaluate the accuracy of the annotation. The dynamic weight adjustments in the recommended learning technique prevent the expert from having time to reconsider their choice. In order to condense the training dataset and accommodate fresh training examples, a forgetting mechanism is used. The image is carefully examined at every level, from coarse to fine, in order to grasp its features. In the first stage, classification will be done, and an image's MM infection will be looked at.

The FCN model, on the other hand, initially performs multi-scale image processing, in which feature maps are created at several sizes. As the name suggests, an FCN model is built using locally linked layers, including convolution, pooling, and up-sampling [42]. Fig. 4, which contrasts FCN processing with conventional CNN structure processing, serves as an illustration of this. A down-sampling path is in charge of obtaining semantic and contextual data, and an up-sampling path is in charge of extracting spatial data. Together, these two components make up an FCN's topology. Due to the absence of a thick layer in this architecture, the number of parameters required and the associated computational expense are reduced [44, 45]. Implementing a skip connection action, which bypasses at least one layer, can minimize any downsides related to information loss due to pooling or down-sampling layers. An FCN model is compelled by this structure to operate inside a global-local data processing architecture.

It is clear that global-local analysis, as opposed to local-based ones like CNN, offers a superior classification framework for MM RETINGRAPH image segmentation. DepthCNN-XgBoosts is another design that may maintain the local-data features during the upsampling process and is similar to FCNs [23]. As a result, in this work, the MM segmentation of CT images is performed using the DepthCNN-XgBoost model. Ranneberger et al. (2015) demonstrated extremely strong performance when segmenting arterial brain arteries in a patient with cerebrovascular disease using a modified version of DepthCNN-XgBoost. This achievement motivates the development of vessel segmentation techniques for computer-aided diagnosis of cerebrovascular illness. Deep learning-based networks do not require unique feature engineering or selection, in contrast to earlier "rule-based" non-neural network techniques. While DepthCNN-XgBoost outperforms the traditional graph-cut-based segmentation approach by effectively extracting the pertinent features during training.In the second phase, the MM region is localized and labeled, and a bound box is created around the area of interest. This will help specialists focus on the diagnosis. However, for many purposes, bounding boxes are inadequate (for example, precise tumor diagnosis). In such cases, we need extremely detailed "pixel-based segmentation," or information at the pixel level. Semantic segmentation is aimed at achieving this. In this case, each pixel in a picture is assigned to a certain class. But due to time restraints, computational limitations, and low false-negative detection limits, semantic segmentation is restricted.

TABLE VII. COMPARISONS WITH STATE-OF-THE-ART APPROACHES

Cited.	Methodology	Dataset	Results	Limitations
[38]	The detection and segmentation of PM using semantic adversarial networks (SAN) and few-short learning (FSL), respectively. Unlike DL methods, conventional segmentation techniques employ supervised learning models.	PALM	sensitivity (SE) of 95%, specificity (SP) of 96%, and area under the receiver operating curve (AUC) of 98%	Preprocessing steps are required and applied on a limited dataset. In addition, fixed data augmentation parameters are required.
[39]	Fundus images are first preprocessed and then images are fed to the designed CNN model.	PALM	AUC score of 0.9845	CNN architecture is not optimized and generalize solution for detecting of different type of MM.
[40]	A dual-stream DCNN (DCNN-DS) model that perceives features from both original images and corresponding processed images by color histogram distribution optimization method was designed for classification of no MM, tessellated fundus (TF), and pathologic myopia (PM).	PALM	Sensitivities of 90.8% and 97.9% and specificities of 99.1% and 94.0%	
[41]	CNN bundles lesion segmentation and PM classification	PALM	AUC of 0.9867	No multiclass categorization of different types of MM and so limited capability

[42]	four convolutional neural network (CNN) architectures, namely DenseNet201, ResNet50, VGG16, and Xception. The CNN architectures were evaluated in the test dataset and their performances were compared. Xception had the best metrics compared with the other architectures in all three tasks	META-PM	--	Multiclass MM classification but without preprocessing and data augmentation.
[43]	DL models were able to recognize the lesions of myopic maculopathy	META-PM	AUC of 0.970	No multiclass recognition of MM and no generalize tool. No preprocessing to adjust the pixels.
[44]	Combination of dense connection and Residual Squeeze-and-Excitation attention is proposed in this paper to detect myopia automatically	Private	--	No multiclass recognition of MM and no generalize tool. No preprocessing to adjust the pixels.
[45]	three five-classification models based on Vision Outlooker for Visual Recognition (VOLO), EfficientNetV2, and ResNet50 for detecting myopic maculopathy were trained with data-augmented images	Meta-PM	SE of 96.43	No generalize tool. No preprocessing to adjust the pixels.
[46]	The efficientNet model was utilized to recognize multi-classes of MM.		AUC of 0.98	No preprocessing to adjust the pixels.
[47]	Image Processing and feature fusion approach were developed.	PALM	AUC of 0.9981	No multiclass recognition of MM and no generalize tool. No preprocessing to adjust the pixels.

A medical expert can evaluate the segmentation quality in addition to a quantitative evaluation. U-Visual Net's analysis performance was therefore shown to be much better. However, compared to smaller arteries, huge vessels can be seen very well, which can be enhanced in the future. As a result, it demonstrates the excellent performance of the DepthCNN-XgBoost architecture in the clinical area. Utilizing more recent segmentation topologies, such as the MS-net (Shah et al., 2018), can result in even greater performance. We should examine the problem's constraints before beginning any implementation technique. Two key criteria in deep learning methods are data availability and data imbalance, which both affect the choice of classification model and topological complexity. The fundus samples' positive-to-negative ratio (492:447) is reasonable; however, the pixel ratio between MM and non-MM is unbalanced. This is because the infected eye part is smaller than the healthy one (see Fig. 3 and Fig. 6). As a result, the first step was to implement a training data balancing strategy that included undersampling the majority class (non-MM regions) [48]. To do this, 492 photographs with a positive annotation ratio of 0.01 to 59% of the total pixels are supplied to deep networks for training, while the 447 images with negatively annotated pixels are excluded from the training process. A visual example of heatmaps show in Fig. 11 about the success of classifying different patterns in MM retinograph images.

4) Limitations of current study: The study acknowledges the significance of proper hyper-parameter tuning and the size and quality of the training dataset in achieving successful learning of a CNN model. In the context of this research, the focus was on evaluating various benchmark CNNs for MM classification tasks specifically using retinal fundus images. For future investigations, it is suggested that hybrid variations of the architectural concepts from the best-performing benchmark models, combined with attention mechanisms and spatial pooling, could be explored. This approach aims to synthesize a robust and accurate MM classification model by leveraging the strengths of different architectures and incorporating attention mechanisms to enhance the model's ability to focus on important regions or features in the images.

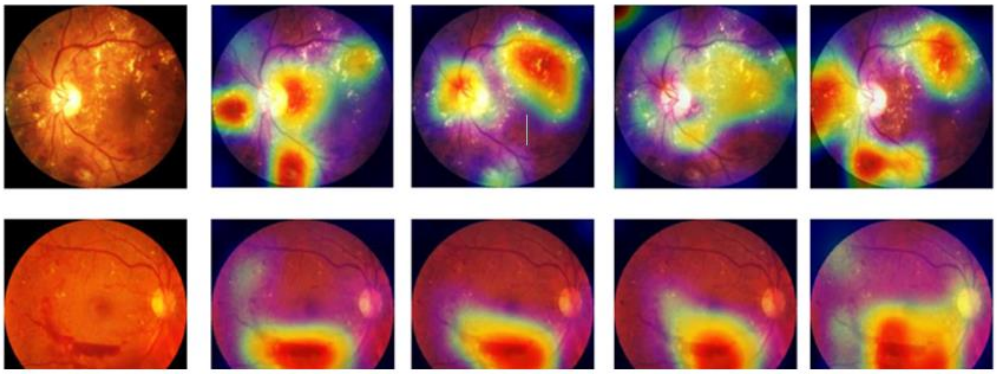

Fig. 11. Three AUC curves for figure (a) Proposed depthwise-XgBoost with data augmentation, (b) Original depthwise separable CNN, and (c) XgBoost classifier.

VII. Conclusion and Future Directions

This study presents a pretrained learning technique for segmenting SMM-infected regions. The DepthCNN-XgBoost framework was used to construct this model. Based on a small number of newly received samples, it modifies the network dynamically. This retraining method reduced the loss of existing knowledge while allowing the model to trust the approaching

new incoming data to the greatest extent possible. The recommended solution differs from conventional methods in that it employs an online learning paradigm using spatial pyramid pooling technique. This novel approach, called "few-shot powered DepthCNN-XgBoost," is reportedly effective and persuasive in the segmentation of SMM-infected areas. Experimental results show the effectiveness of the suggested few-shot learning strategy in combination with a DepthCNN-XgBoost model for finding and characterizing infectious SMM regions. The few-shot powered DepthCNN-XgBoost is a possible artificial intelligence (AI) framework for medical imaging, particularly beneficial for locating pathogenic SMM regions when compared to deep learning models like convolutional neural networks, fully convolutional networks, and traditional DepthCNN-XgBoost structures. The proposed few-shot DepthCNN-XgBoost model exhibited an IoU increase of 5.388% (3.046% for all test data utilizing 4-fold cross-validation results from the various classifiers) compared to a regular DepthCNN-XgBoost. The F1-Score also increases by 5.394, or 3.015%. We found increases in accuracy and recall of 1.162, 2.137%, and 4.409, 4.790%, respectively. The Kruskal-Wallis test p-value on the F1-score and IoU values between the proposed few-shot DepthCNN-XgBoost model and the traditional model was 0.026 (below 0.05). This indicates that there is a significant difference between the metrics of the two techniques, with a 95% confidence level. The recommended model needed around eight photos and a small number of new incoming samples in order to change its behavior effectively. Due to the fact that new data was combined with older samples to improve the network's generalization skills, the suggested few-shot DepthCNN-XgBoost model has a similar level of computational complexity to the traditional DepthCNN-XgBoost model. The combination of few-shot learning with other deep models and learning techniques like transformers [48] is quite appealing for future development. According to a recent study, transformer-based models beat other types of networks, such as recurrent and convolutional structures, in a variety of benchmarks for visual information.

REFERENCES

[1] Modjtahedi, B. S., Abbott, R. L., Fong, D. S., Lum, F., Tan, D., Ang, M., ... & Zadnik, K. (2021). Reducing the global burden of myopia by delaying the onset of myopia and reducing myopic progression in children: the Academy's Task Force on Myopia. Ophthalmology, 128(6), 816-826.

[2] Sankaridurg, P., Tahhan, N., Kandel, H., Naduvilath, T., Zou, H., Frick, K. D., ... & Resnikoff, S. (2021). IMI impact of myopia. Investigative ophthalmology & visual science, 62(5), 2-2.

[3] Lu, L., Ren, P., Tang, X., Yang, M., Yuan, M., Yu, W., ... & Han, W. (2021). AI-Model for Identifying Pathologic Myopia Based on Deep Learning Algorithms of Myopic Maculopathy Classification and "Plus" Lesion Detection in Fundus Images. Frontiers in cell and developmental biology, 2841.

[4] Jotterand, F., & Bosco, C. (2022). Artificial Intelligence in Medicine: A Sword of Damocles?. Journal of Medical Systems, 46(1), 1-5.

[5] Li, Y., Foo, L. L., Wong, C. W., Li, J., Hoang, Q. V., Schmetterer, L., ... & Ang, M. (2022). Pathologic myopia: advances in imaging and the potential role of artificial intelligence. British Journal of Ophthalmology.

[6] Abbas, Q., Qureshi, I., Yan, J., & Shaheed, K. (2022). Machine Learning Methods for Diagnosis of Eye-Related Diseases: A Systematic Review Study Based on Ophthalmic Imaging Modalities. Archives of Computational Methods in Engineering, 1-58.

[7] Qureshi, I., Ma, J., & Abbas, Q. (2021). Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning. Multimedia Tools and Applications, 80(8), 11691-11721.

[8] Rauf, N., Gilani, S. O., & Waris, A. (2021). Automatic detection of pathological myopia using machine learning. Scientific Reports, 11(1), 1-9.

[9] Li, J., Wang, L., Gao, Y., Liang, Q., Chen, L., Sun, X., ... & Xie, L. (2022). Automated detection of myopic maculopathy from color fundus photographs using deep convolutional neural networks. Eye and Vision, 9(1), 1-12.

[10] Devda, J., & Eswari, R. (2019). Pathological myopia image analysis using deep learning. Procedia Computer Science, 165, 239-244.

[11] Zhang, C., Zhao, J., Zhu, Z., Li, Y., Li, K., Wang, Y., & Zheng, Y. (2022). Applications of Artificial Intelligence in Myopia: Current and Future Directions. Frontiers in Medicine, 9.

[12] Zhang, Z., Ji, Z., Chen, Q., Yuan, S., & Fan, W. (2021). Joint optimization of CycleGAN and CNN classifier for detection and localization of retinal pathologies on color fundus photographs. IEEE Journal of Biomedical and Health Informatics, 26(1), 115-126.

[13] You, A., Kim, J. K., Ryu, I. H., & Yoo, T. K. (2022). Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey. Eye and Vision, 9(1), 1-19.

[14] Abbas, Q., Qureshi, I., & Ibrahim, M. E. (2021). An Automatic Detection and Classification System of Five Stages for Hypertensive Retinopathy Using Semantic and Instance Segmentation in DenseNet Architecture. Sensors, 21(20), 6936.

[15] Sun, L., Li, C., Ding, X., Huang, Y., Chen, Z., Wang, G., ... & Paisley, J. (2022). Few-shot medical image segmentation using a global correlation network with discriminative embedding. Computers in biology and medicine, 140, 105067.

[16] Tian, Y., & Fu, S. (2020). A descriptive framework for the field of deep learning applications in medical images. Knowledge-Based Systems, 210, 106445.

[17] Wang, Z., Ma, B. & Zhu, Y. Review of Level Set in Image Segmentation. Arch Computat Methods Eng 28, 2429–2446 (2021). https://doi.org/10.1007/s11831-020-09463-9.

[18] Yang R and Yu Y (2021) Artificial Convolutional Neural Network in Object Detection and Semantic Segmentation for Medical Imaging Analysis. Front. Oncol. 11:638182. doi: 10.3389/fonc.2021.638182.

[19] L., & Wu, Y. Z. (2022). Semantic segmentation of pancreatic medical images by using convolutional neural network. Biomedical Signal Processing and Control, 73, 103458.

[20] Lu, H., Tian, S., Yu, L., Liu, L., Cheng, J., Wu, W., ... & Zhang, D. (2022). DCACNet: Dual context aggregation and attention-guided cross deconvolution network for medical image segmentation. Computer Methods and Programs in Biomedicine, 214, 106566.

[21] Gao, N., Xue, H., Shao, W., Zhao, S., Qin, K. K., Prabowo, A., ... & Salim, F. D. (2022). Generative adversarial networks for spatio-temporal data: A survey. ACM Transactions on Intelligent Systems and Technology (TIST), 13(2), 1-25.

[22] You, A., Kim, J. K., Ryu, I. H., & Yoo, T. K. (2022). Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey. Eye and Vision, 9(1), 1-19.

[23] Zhan, B., Xiao, J., Cao, C., Peng, X., Zu, C., Zhou, J., & Wang, Y. (2022). Multi-constraint generative adversarial network for dose prediction in radiotherapy. Medical Image Analysis, 77, 102339.

[24] Hemelings, R., Elen, B., Blaschko, M. B., Jacob, J., Stalmans, I., & De Boever, P. (2021). Pathological myopia classification with simultaneous lesion segmentation using deep learning. Computer Methods and Programs in Biomedicine, 199, 105920.

[25] Lu, L., Zhou, E., Yu, W., Chen, B., Ren, P., Lu, Q., ... & Han, W. (2021). Development of deep learning-based detecting systems for pathologic myopia using retinal fundus images. Communications biology, 4(1), 1-8.

[26] Du, R., Xie, S., Fang, Y., Igarashi-Yokoi, T., Moriyama, M., Ogata, S., ... & Ohno-Matsui, K. (2021). Deep learning approach for automated detection of myopic maculopathy and pathologic myopia in fundus images. Ophthalmology Retina, 5(12), 1235-1244.

[27] Zhang, W., Zhao, X., Chen, Y., Zhong, J., & Yi, Z. (2020). DeepUWF: an automated ultra-wide-field fundus screening system via deep learning. IEEE Journal of Biomedical and Health Informatics, 25(8), 2988-2996.

[28] Shi, Z., Wang, T., Huang, Z., Xie, F., & Song, G. (2021). A method for the automatic detection of myopia in Optos fundus images based on deep learning. International Journal for Numerical Methods in Biomedical Engineering, 37(6), e3460.

[29] Freire, C. R., Moura, J. C. D. C., Barros, D. M. D. S., & Valentim, R. A. D. M. (2020). Automatic lesion segmentation and pathological myopia classification in fundus images. arXiv preprint arXiv:2002.06382.

[30] Jia, S., Jiang, S., Lin, Z., Li, N., Xu, M., & Yu, S. (2021). A survey: Deep learning for hyperspectral image classification with few labeled samples. Neurocomputing, 448, 179-204.

[31] Voulodimos, A., Protopapadakis, E., Katsamenis, I., Doulamis, A., & Doulamis, N. (2021). A few-shot U-net deep learning model for COVID-19 infected area segmentation in CT images. Sensors, 21(6), 2215.

[32] Feng, Y., Gao, J., & Xu, C. (2022). Learning Dual-Routing Capsule Graph Neural Network for Few-shot Video Classification. IEEE Transactions on Multimedia.

[33] Abdelaziz, M., & Zhang, Z. (2022). Multi-scale kronecker-product relation networks for few-shot learning. Multimedia Tools and Applications, 1-20.

[34] Zhu, Q., Mao, Q., Jia, H., Noi, O. E. N., & Tu, J. (2022). Convolutional relation network for facial expression recognition in the wild with few-shot learning. Expert Systems with Applications, 189, 116046.

[35] Korshunov, P., & Marcel, S. (2022). Improving Generalization of Deepfake Detection with Data Farming and Few-Shot Learning. IEEE Transactions on Biometrics, Behavior, and Identity Science.

[36] Li, W., Gao, Y., Zhang, M., Tao, R., & Du, Q. (2022). Asymmetric Feature Fusion Network for Hyperspectral and SAR Image Classification. IEEE Transactions on Neural Networks and Learning Systems.

[37] Singh, R., Bharti, V., Purohit, V., Kumar, A., Singh, A. K., & Singh, S. K. (2021). MetaMed: Few-shot medical image classification using gradient-based meta-learning. Pattern Recognition, 120, 108111.

[38] Wang, S. Y., Liao, W. S., Hsieh, L. C., Chen, Y. Y., & Hsu, W. H. (2012). Learning by expansion: Exploiting social media for image classification with few training examples. Neurocomputing, 95, 117-125.

[39] Xian, Y., Korbar, B., Douze, M., Schiele, B., Akata, Z., & Torresani, L. (2020, August). Generalized many-way few-shot video classification. In European Conference on Computer Vision (pp. 111-127). Springer, Cham.

[40] Javaria Amin, Muhammad Almas Anjum, Muhammad Sharif (2022, May). Fused information of DeepLabv3+ and transfer learning model for semantic segmentation and rich features selection using equilibrium optimizer (EO) for classification of NPDR lesions, Knowledge-Based Systems.

[41] Abbas, Qaisar, Abdul Rauf Baig, and Ayyaz Hussain. "A Semantic Adversarial Network for Detection and Classification of Myopic Maculopathy." CMC-COMPUTERS MATERIALS & CONTINUA 75, no. 1 (2023): 1483-1499.

[42] N. Rauf, S. O. Gilani and A. Waris, “Automatic detection of pathological myopia using machine learning,” Scientific Reports, vol. 11, no. 1, pp. 1–9, 2021.

[43] J. Li, L. Wang, Y. Gao, Q. Liang, L. Chen et al., “Automated detection of myopic maculopathy from color fundus photographs using deep convolutional neural networks,” Eye and Vision, vol. 9, no. 1, pp. 1–12, 2022.

[44] R. Hemelings, B. Elen, M. B. Blaschko, J. Jacob, I. Stalmans et al., “Pathological myopia classification with simultaneous lesion segmentation using deep learning,” Computer Methods and Programs in Biomedicine, vol. 199, pp. 1–18, 2021.

[45] L. Lu, E. Zhou, W. Yu, B. Chen, P. Ren et al., “Development of deep learning-based detecting systems for pathologic myopia using retinal fundus images,” Communications Biology, vol. 4, no. 1, pp. 1–8, 2021.

[46] R. Du, S. Xie, Y. Fang, T. I. Yokoi, M. Moriyama et al., “Deep learning approach for automated detection of myopic maculopathy and pathologic myopia in fundus images,” Ophthalmology Retina, vol. 5, no. 12, pp. 1235–1244, 2021.

[47] Z. Shi, T. Wang, Z. Huang, F. Xie and G. Song, “A method for the automatic detection of myopia in optos fundus images based on deep learning,” International Journal for Numerical Methods in Biomedical Engineering, vol. 37, no. 6, pp. 1–10, 2021.

[48] Sun, Y., Li, Y., Zhang, F., Zhao, H., Liu, H., Wang, N., & Li, H. (2023). A deep network using coarse clinical prior for myopic maculopathy grading. Computers in Biology and Medicin.