



A Semantic Adversarial Network for Detection and Classification of Myopic Maculopathy

Qaisar Abbas¹, Abdul Rauf Baig^{1,*} and Ayyaz Hussain²

¹College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, 11432, Saudi Arabia

²Department of Computer Science, Quaid-i-Azam University, Islamabad, 44000, Pakistan

*Corresponding Author: Abdul Rauf Baig. Email: abbaig@imamu.edu.sa

Received: 28 September 2022; Accepted: 04 November 2022

Abstract: The diagnosis of eye disease through deep learning (DL) technology is the latest trend in the field of artificial intelligence (AI). Especially in diagnosing pathologic myopia (PM) lesions, the implementation of DL is a difficult task because of the classification complexity and definition system of PM. However, it is possible to design an AI-based technique that can identify PM automatically and help doctors make relevant decisions. To achieve this objective, it is important to have adequate resources such as a high-quality PM image dataset and an expert team. The primary aim of this research is to design and train the DLs to automatically identify and classify PM into different classes. In this article, we have developed a new class of DL models (SAN-FSL) for the segmentation and detection of PM through semantic adversarial networks (SAN) and few-shot learning (FSL) methods, respectively. Compared to DL methods, the conventional segmentation methods use supervised learning models, so they (a) require a lot of data for training and (b) fixed weights are used after the completion of the training process. To solve such problems, the FSL technique was employed for model training with few samples. The ability of FSL learning in UNet architectures is being explored, and to fine-tune the weights, a few new samples are being provided to the UNet. The outcomes show improvement in the detection area and classification of PM stages. Betterment in the result is observed by sensitivity (SE) of 95%, specificity (SP) of 96%, and area under the receiver operating curve (AUC) of 98%, and the higher F1-score is achieved using 10-fold cross-validation. Furthermore, the obtained results confirmed the superiority of the SAN-FSL method.

Keywords: Artificial intelligence; cardiovascular; vision loss; deep learning; few-shot learning; semantic segmentation; myopic maculopathy



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

Cardiovascular disease (CV) is the main chronic disease and the major contributor to the global burden of disease. The CV health is directly related to all the organs of your body, and it can directly affect blood vessels due to high blood pressure [1]. As a result, the vision is lost if it is not treated early. Myopia is currently a public health problem due to its rapidly growing prevalence globally and the threat to vision. It has been estimated that 50 percent of the world's population could be myopic in 2050, and 10 percent of them suffer from severe myopia [2]. The identification and classification of myopic maculopathy are done based on the international photographic and grading system for Myopic Maculopathy (MM) [3]. Pathologic myopia (PM) was classified with respect to severity as shown in Fig. 1.

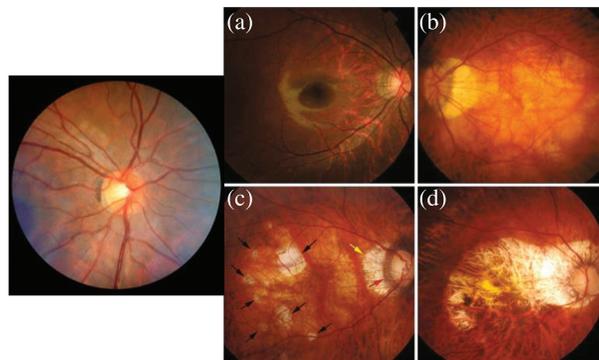


Figure 1: Myopic maculopathy (a) category 1: Fundus tessellation temporal to the optic disc, (b) category 2: Diffuse chorioretinal atrophy, (c) category 3: patchy chorioretinal atrophy and (d) category 4: macular atrophy

The rapid development of artificial intelligence (AI) [4] plays a vital role in the automation of clinical data processing and complicated medical diagnostics. Deep learning systems (DLS) are the most advanced class of AI [5]. It emulates the human brain at work and utilizes neural networks (NN) to solve feature-dependent problems. In medical practices, the DLS [6–11] outperforms in many cases. In ophthalmology, DLS-based diagnosis software has been successfully used in clinical and public healthcare departments. AI technologies is a worthwhile and efficient alternative for the control and diagnostics of MM. On the other hand, automated diagnosis using images is considered an image analysis problem, which can be figured out by labeling the data and using machine learning (ML) algorithms such as DL. However, the use of deep learning techniques in MM lesion scanning remains a challenge [12] because of the classification complexity and definition of the MM system. The objective of this research is to design and train the DLS, which is capable of automatically detecting MM and categorizing the MM using a handsome dataset of color retinal fundus images, which are collected from the ophthalmic centers of hospitals. A visual example of MM fundus images is displayed in Fig. 1.

To address this problem, there is a CycleGAN [13] technique, which uses the convolutional neural networks (CNN) model to detect Myopic maculopathy (MM). In practice, CycleGAN has the capability to generate more realistic and reliable images with cycle consistency. Moreover, the classifier and generator work together to discriminate the domain images [14]. A res-guided UNet [15] is developed as a generator and res-guided sampling blocks are replaced by traditional convolution. Supervised learning usually requires a large training dataset to incorporate all possible variations.

However, it might be possible that the large dataset is not available, and if it is available, then it is a tedious task in the domain of medical imaging. This problem is solved if we can employ few-shot learning (FSL) algorithms [16,17] that require a smaller training dataset. Additionally, there is a chance of issues when a large training dataset is required.

In this paper, we overcome this limitation by proposing a few-shot learning (FSL) model, where only a few samples of the network will be dynamically trained. Our focus is on FSL classification as we are updating and refining weights dynamically by feeding new sample data. To the best of our knowledge, dynamic updating of model weights is a novel and unique approach up to this point. There is no use of the FSL-based UNet (SAN-FSL) scheme (in the testing phase of dynamic learning), which has also proven to be very useful for the detection and classification of MM-infected regions. The SAN-FSL technique is newly proposed by integrating a semantic adversarial network (SAN) based on FSL and UNet model. As a result, this SAN-FSL has the capability to dynamically tune its parameters in accordance with user feedback. In particular, this process is enhanced the segmentation outcomes when there is a poor performance.

Our major contributions towards myopia detection are as follows.

- 1) To develop this SAN-FSL method, we have also proposed a new preprocessing step to enhance Myopic Maculopathy (MM) regions.
- 2) A novel DL method is proposed by integrating semantic adversarial networks (SAN) and few-shot learning (FSL) techniques to detect and recognize the category of myopia.
- 3) In a comparison among deep learning models, machine learning models, and our method, the best performance is demonstrated by our SAN-FSL. It guarantees the accuracy and convenience of clinics for doctors.

2 Literature Review

In the early research, images were segmented using techniques based on edges, regions, clusters, and thresholds [18]. In these conventional methods, features are extracted manually and then used for tasks like background separation. Moreover, the quality of features affects the segmentation results, and sometimes this approach is labor-intensive and time-consuming. However, in recent years, research's focus has shifted toward conventional neural networks and deep learning algorithms, especially in semantic image segmentation [19,20]. Moreover, with the passage of time, the accuracy of these methods has also been considerably improved in recognition and prediction. The DeconvNet [21] is a more extensive decoder than the original FCN. The mentioned decoder is balanced in number and feature size with the encoder. Aside from the deconvolution, to improve the results, the DeconvNet decoder uses un-pooling layers. The DevconvNet also consumes large amounts of memory as compared to FCS due to fully connected layers in the encoder. To optimize the parameters and memory, the authors introduced the SegNet [22] model, which is like the VGG-16 but differs from the FCN and DeconvNet in up-sampling and convolution. This technique effectively eliminates the deconvolution process. The SegNet manages the feature maps very efficiently. However, it requires more memory.

Recently, generative adversarial networks (GANs) [23] have proved very successful in different applications. Many researchers have found that generative adversarial loss is helpful for enhancing network performance. Motivated by the success of GANs on image translation [24], an efficient GAN network for image semantic segmentation is designed. It is closest to the adversarial networks, which aid in training for semantic segmentation. But it does not improve the overall baseline. This research has two significant advantages, such as (a) a clear pattern of viral infections can be observed at an early

stage, and (b) CT abnormalities related to viral pneumonia can be identified before laboratory tests in 70% of cases [16]. Therefore, a CT investigation is very helpful in the early detection of MM infection. There are also many studies on MM detection in chest X-ray images. However, we concentrate on work related to CT scans. U-Net structures are another option to use. It is intended only for medical applications. It is a multistage method involving the segmentation and classification of MM and other viral infections [25]. It also helps in monitoring advanced disease progression.

A pretrained Resnet-50 is updated for the classification of MM or other cases [26]. In the next phase, multi-object adaptive CNN, called AdaResU-Net [27], was proposed with the ability to automatically adapt to new datasets and residual learning paradigms. U-Net++ [28], a U-Net-based model, was also used on high-resolution CT images for MM detection. Moreover, the MM's detection has been tested with many variants of transfer learning [29] algorithms. Among many TL architectures, the ResNet-101 and Xception models were outperformed. Other studies [30–32] use TL methods to detect MM eye diseases. Moreover, object-detection methods are considered [33] for MM diagnostics and, in another study, VGA variations are also used for the detection of symptomatic lung areas [34]. The proposed method can classify SMM from non-pneumonia (NP) and community-acquired pneumonia (CAP). In [35], for MM detection, the Nave Bayes classifier, discrete wavelet transforms, and genetic algorithms are used. In a segmentation-based study, super-pixel-based type 2 fuzzy clustering methods are combined into a proposed algorithm [36]. As an alternative, the V-Nets [37] are used for MM image segmentation. Likewise, in another study, V-Net was used to segment all the slices of MRI simultaneously [38]. Deep learning methods [39] were critical in the segmentation of pulmonary CT images. Now they have gained the ability to evaluate the severity and quantify the infection levels of diseases [40,41].

The basic requirement of DLS methods is a large ground-truth dataset for training, which is very difficult sometimes. Moreover, annotating large amounts of data is a labor-intensive and time-consuming job. These limitations confine the applicability and utility of DLS to real-world problems. This limitation has started to be discussed in a limited number of studies where a semi-supervised learning approach was used with multiclass segmentation to identify the infected region. However, this method showed poor output. In this study, we proposed a few-shot learning process with U-Net that supports dynamic re-training mechanisms. This approach eliminates the need for a large dataset. It dynamically updated the UNet parameters with a few samples. Furthermore, this scheme dynamically adjusts the parameters with the interaction of domain experts. On the other hand, current models are not able to tune the weights once trained.

3 Proposed Methodology

3.1 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 labeled fund images. The dataset consists of 239 pathological myopic eye images and 161 normal eye images. The image size is $1444 \times 144 \times 3$ (RGB image). The dataset is online available at <https://palm.grand-challenge.org/>.

3.2 Preprocessing Method

The Myopic Maculopathy (MM) images are captured in different environments. As a result, there is a need to improve the color fundus image by preprocessing to adjust light and enhance the contrast so that the MM-infected region can be highlighted. An example of low light is shown in Fig. 1. The preprocessing step is used to enhance the patterns. Moreover, the processed images make

the training task easy for Deep Learning (DL) models to extract features and detect the infected regions. In practice, the color fundus images can be depicted in various color spaces such as RGB (red, green, blue), HSV (hue, saturation, value), CIELUV, etc. However, the CIECAM02 color space is closer to human perception, and it is an uniform color space compared to other non-uniform color spaces. Therefore, the preprocessing step utilizes the CIECAM02 color appearance model to enhance the MM-infected region compared to other parts of the eye fundus image. This step can assist ophthalmologists in the accurate diagnosis of Myopic Maculopathy (MM). Furthermore, the uniform color space is very crucial for the image enhancement of some applications, whereas the RGB and HSV are not uniform, therefore they are not used for image enhancement.

The CIECAM02 color space also has an issue with white adaptation, which may result in a poor image, even after enhancement. In this paper, we used the CIECAM02 color space for image preprocessing. Therefore, the image is converted from RGB to CIECAM02 uniform color space to adjust the light and enhance the contrast of the MM-infected region. The CIECAM02 color appearance model provides JCh (J: luminance, C: chroma, and h: hue) channels to process the fundus image. The J-component of this uniform color space is used with a discrete-wavelet transform (DWT) approach to update or enhance this channel of uniform color space. The preprocessing follows three steps: Firstly, the J-component of the color space coefficients of DWT is modified by Ben Graham's method [19] to increase the illumination so that perception augmentation and contrast adjustment can be done. Secondly, the inverse transform is used to adjust J coefficients for good visualization and reconstruction without generating artifacts. In addition, Gaussian Filtering is applied to enhance Myopic Maculopathy (MM) regions compared to other regions of fundus images. The enhancement results are visually displayed in Fig. 2.

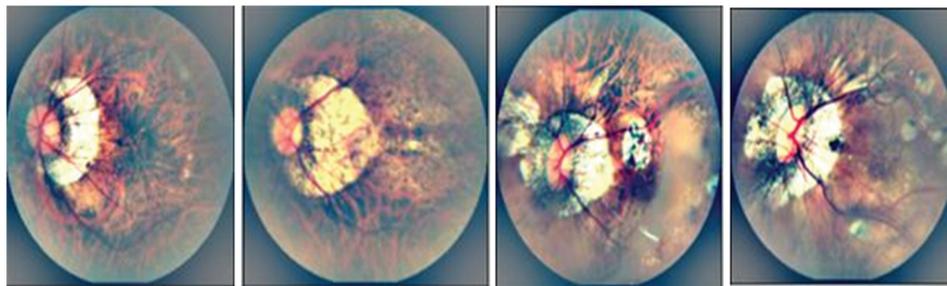


Figure 2: A visual example of myopic maculopathy in case of preprocessing to enhance the infected region

3.3 Semantic Adversarial Network Method

In this step, a Generative Adversarial Network (GAN) is used for image semantic segmentation. The architecture of the proposed algorithm is shown in Fig. 3. In contrast to the original GAN model, the framework of the proposed GAN is established on two separate deep convolutional neural networks, the discriminator network D and the segmentation network S, whose joint work can outperform for a given input image. The proposed SegGAN network is designed to generate a reasonable result by segmenting the given input image. The anatomy of a generative network is based on the UNet framework. In comparison to known fully-connected network (FCN) methods [41], the major variations are connection-skipping between paths and symmetry.

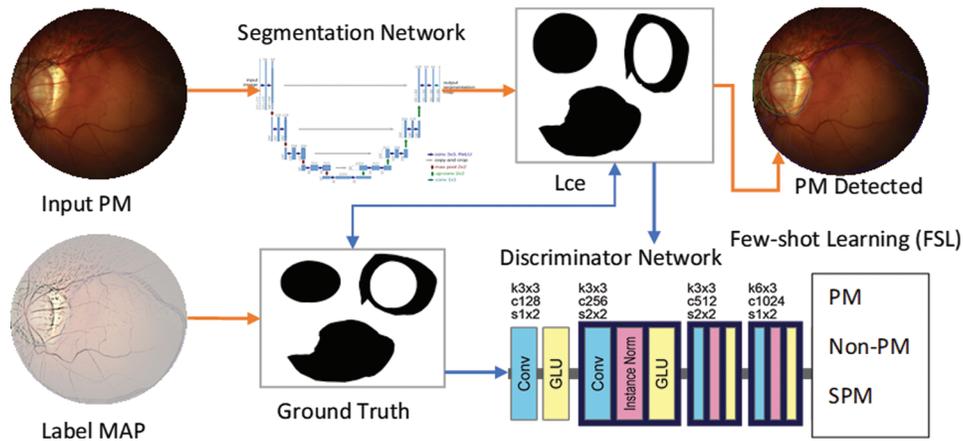


Figure 3: A systematic flow diagram of proposed SAN-FSL system for recognition of stages of myopic maculopathy

The UNet architecture is depicted in Fig. 2. The framework consists of two paths. The left side is the contraction path. The encoder part is a conventional stack of convolutional and max-pooling layers that scale down the height and width of the input images. The right side is a symmetrical, expansive path. The decoding part restores the original dimensions of the input images, which helped in the precise localization using transposed convolutions. Each box in the figure represents the multi-channel feature map. The dotted box shows the concatenated feature maps from the contractive part. The arrows in the figure represent the operations listed in the legend. The image dimension is displayed on the left side of the box, and the number of channels is displayed on the top of the box. The half UNet has a similar architecture, except it halved the number of channels in the network. Therefore, it is referred to as half UNet. The 2D image patches are fed into the network, and it returns the 2D segmentation probability map for every patch. More specifically, by applying the same model to inputs at multiple scales, the network can have access to information with different levels of detail. Thus, semantic segmentation can be successfully performed by merging features obtained using different input scales. However, this type of model presents an important drawback: they do not scale well for deep convolutional neural networks (CNNs) due to factors such as limitations in graphical processing unit (GPU) memory (in semantic segmentation, it is not uncommon to work with images of large size). This sort of method is sometimes referred to as an image pyramid.

Because of its success in image generation, generative adversarial loss (GAN) has been widely used in learning generative models. Because GANs solve the issues of other models by introducing the theory of adversarial learning between the generator and the discriminator, the sampling in other models is inaccurate and slow, whereas GANs can simply sample the generated data. Due to these benefits, GAN was selected to deal with the image semantic segmentation problem. Furthermore, CNN training can aid in the stabilization of GAN training. The proposed Seg-GAN is composed of two feed-forward CNNs: the segmentation part S and the discriminator part D. The Seg-GAN can also suggest the guidelines for architecture, in which the segmentation network consists of convolutional neural networks and the discriminator consists of fully connected CNNs combined with cascaded ConvCRFs. Batch normalization (BN), Leaky ReLU (rectified linear units), and ReLU activation functions are applied for the segmentation network part, and a discriminator is utilized to stabilize

the GAN training. The segmentation network labeled the segmented image and assigned a ground-truth image y to each input image x . Furthermore, $S(x)$ is urged to have data distribution like that of a ground-truth image. The discriminator network D determines the difference between the data distribution of the segmentation outcome and the respective ground-truth image. The networks S and D compete to attain their respective goals, coining the term adversarial. For discriminator network training, the cross-entropy loss DL is minimized in the respective classes. The loss DL can be expressed as (1):

$$L_D = - \sum_{h,w} (1 - y_n) \log(1 - D(S(X_n))^{h,w}) + y_n \log(D(y_n)^{h,w}) \quad (1)$$

From Eq. (1), The parameter X_n is the input image with size of (h.w.3). The segmentation network is denoted as S , which has a corresponding output $S(X_n)$. For our fully convolutional discriminator, we denote it as D . y_n is the corresponding ground-truth label. Where $y_n = 1$ when the image is drawn from the ground-truth label, and $y_n = 0$ when the image is generated from the segmentation network. We propose to train the segmentation network via minimizing a multi-task loss function:

$$L_{seg} = L_{ce} + \mu L_{adv} \quad (2)$$

where L_{ce} , L_{adv} denote the multi-class cross entropy loss, the adversarial loss, respectively. The parameter μ represents a hyperparameter for balancing the proportion of the L_{ce} in multi-task loss function. L_{ce} and L_{adv} are respectively obtained by the following Equations and other parameters are derived as follows.

$$L_{ce} = - \sum_{h,w} \sum_{c \in C} Y_n^{h,w,c} \log(s(X_n)^{h,w,c}) \quad (3)$$

$$L_{adv} = - \sum_{h,w} \log(D(s(X_n))^{h,w}) \quad (4)$$

where c represents categories in the dataset. With the adversarial loss, the segmentation network first tries to train the network and then cheats the discriminator by maximizing the probability of the segmentation prediction being considered as the ground truth distribution.

3.4 Few-Shot Learning (FSL) for Classification

The main issue with deep-learning-based (DL) studies is that they require many ground-truth datasets for training, whereas collecting large numbers of labeled datasets is time-consuming, labor-intensive, and sometimes impossible due to the nature of the data. This requirement limits the utilization of these approaches. The alternate solution is a few-shot [31–40], which is better than the semi-supervised learning paradigm. In this methodology, the huge dataset problem is overcome by adopting a few-shot learning process for U-Net with a dynamic re-training mechanism. In this way, the model's parameters are dynamically adjusted using a small, labeled dataset with expert feedback.

When a convolutional neural network (CNN) is used on a local dataset for training and diagnostic of eye diseases, it classified the normal eye as infected while testing, despite accurate segmentation. whereas the fully-connected network (FCN) and UNet models showed better performance. This is due to the multiscale capability of algorithms and the availability of local and global information at the same time. The UNet has been found to be a better classifier than FCN for the medical imaging domain since it represents local information more efficiently. However, UNet failed to segment the infected areas. Therefore, a new architecture is proposed and named the few-shot U-Net model. The performance of the model is increased by improving the initial parameter training. Furthermore, only

specific data samples were provided to the few-shot model for retraining where low performance was observed.

Algorithm 1. The mini batch stochastic gradient descent algorithm for training adversarial semantic network

Initialization:	$k=1$, is a hyperparameter to find best expensive option
Output:	Extracted gradient-based updates by input training samples
Step 1.	Data Augmentation
Step 2.	For number of training iterations do
Step 3.	For k steps do
	a) Simple batch of m noise samples $\{z(1), \dots, z(m)\}$ from noise prior $P_g(z)$
	b) Sample mini batch of m examples $\{x_1, \dots, x_m\}$ from data generating distribution $P_{data}(x)$
	c) Update the discriminator by ascending its stochastic gradient:
Step 4.	$\nabla_{\theta_d} = \frac{1}{m} \sum_{i=1}^m [\log D(x^{(i)}) + \log(1 - D(G(z^{(i)})))]$
Step 5.	End for
Step 6.	Simple batch of m noise samples $\{z(1), \dots, z(m)\}$ from noise prior $P_g(z)$ Sample mini batch of m examples $\{x_1, \dots, x_m\}$ from data generating distribution $P_{data}(x)$ Update the discriminator by ascending its stochastic gradient:
Step 7.	$\nabla_{\theta_d} = \frac{1}{m} \sum_{i=1}^m [\log(1 - D(G(z^{(i)})))]$
Step 8.	End for

The framework of the few-shot UNet model is described in Fig. 3. We specifically assume that a training set $S_{init} = (x_{new}(n), t_{new}(n))$ is in pairs of input-target association at time instance n . Variable $x_i(n)$ represents the annotated input data (fundus image), whereas targets t_i are supervised (desired) outputs (the segmentation of the MM), given by the medical experts. From a traditional learning paradigm and this initial training set, a set of parameters for the UNet structure were obtained. Then, the segmentation performance of the UNet was examined over a test data set by experts (medical professionals). The new augmented training set was developed after assessing the low accuracy of the segmentation results by the experts.

$$S_{update} = \{x_{new}(n), t_{new}(n)\} \cup S_{init} \quad (5)$$

The network is trained with a new augmented training set. It is assumed that samples were picked by the domain experts one at a time. The network began its training from the place where the learning algorithm had been stopped, and only a couple of training epochs were taken into consideration. The FSL-based UNet network is trained in this manner to trust the new incoming (few) samples, while also encountering the least amount of deterioration of the prior network knowledge. Finally, the softmax classifier is employed to classify eye diseases into various classes.

4 Experiment and Results

4.1 Experimental Setup

The TensorFlow and Keras libraries were used to create all networks in Python. A Google Colab-provided NVIDIA Tesla P4 GPU was used to train the models. The test was conducted on a PC with an 8-core CPU (AMD FX-8320 @ 3.5 GHz) and 8GB of RAM. The suggested CNN classifier and the FCN-8 adopted topologies, which displayed in Figs. 3 and 4, respectively. In addition, Fig. 3 also shows the architecture of the new FSL-based U-Net that has been proposed with the main objective of

enhancing the existing accuracy of U-Net network. Whereas the architecture of UNet model is shown in Fig. 4. It is observed that the FSL-based U-Net model needs only 8MB of storage space.

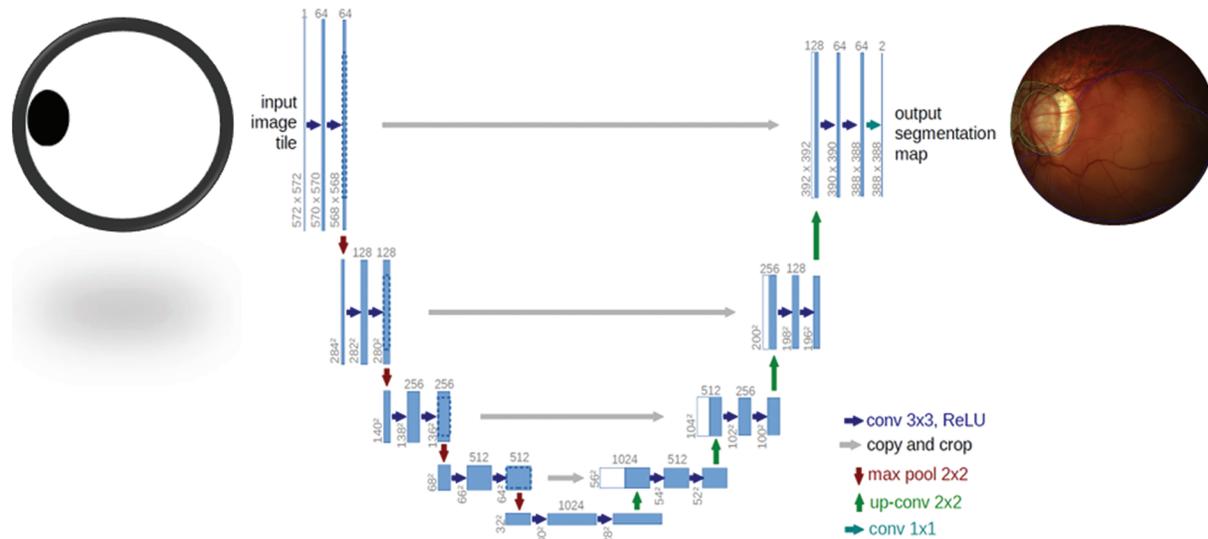


Figure 4: U-Net segmentation model used to detect Myopic maculopathy region

Several other assessment measures, such as specificity (SP), sensitivity (SE), F1-score, accuracy (ACC), recall (RE), and precision (PR), are used to quantitatively evaluate the performance of the proposed technique. In the accuracy (ACC) metric, the most common and one of the basic performance measures is accuracy, which is essentially the probability that a randomly selected example (negative or positive) would be accurate. The diagnostic test in this metric shows the probability of a correct result, i.e., the probability that the diagnosis will be correct. The precision metric refers to the capacity to accurately identify positive categories among entire expected positive classes, which is expressed as a proportion of all accurately expected positive categories to all correctly predicted positive categories. The sensitivity (SE) is a measure of a model's capability to detect all positive instances and represent them. It's worth noting that the above equation implies that a low false-negative rate almost always accompanies a high recall. The specificity (SP) provides the ratio of true negatives to total negatives in the data. The F1-score is useful in determining the classifier's exactness and robustness. The F1 score, which is a key metric that considers both recall and precision for performance testing, could be represented as follows. Whereas the AUC stands for area under the receiver operating characteristics. The AUC curve is a graphical representation or plot of the diagnostic ability of any machine learning classifier using all thresholds.

4.2 Data Augmentation

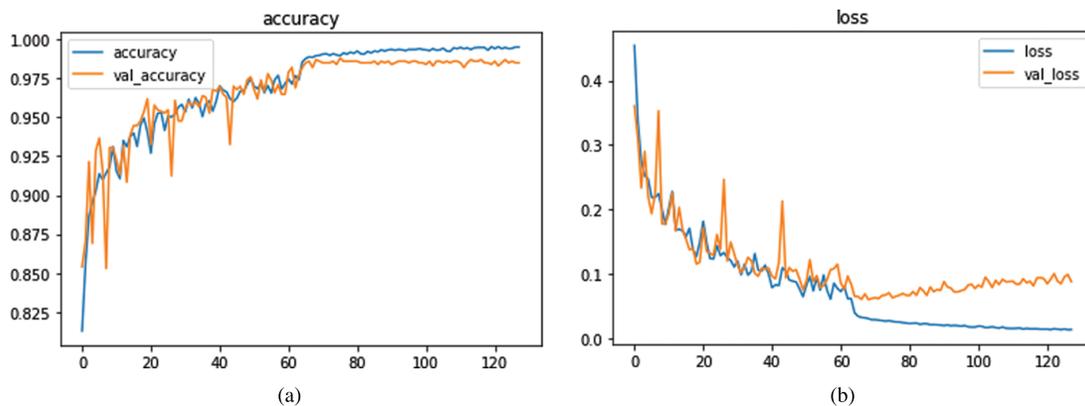
To enhance the variance in the training data, the augmentation method is employed in this paper. In this way, the network becomes invariant to certain transformations. The augmentation is applied using the following parameters, which are described in Table 1.

Table 1: Data augmentation with defined parameters

Rotation-range	Width_shift-range	Shear-range	Zoom-range	Horizontal-flip	Vertical-flip
22	0.2	0.4	0.4	True	False

4.3 Result Analysis

The experimental results considered both the identification capabilities and the computational average time needed by a trained network to fully annotate a fundus image by using several classification-related performance criteria. The average execution time per sample for the FCN and UNet models varied from 0.364 s to 0.970 s. According to what was seen, the U-Net structure had the lowest computing cost when compared to the other models, whereas the CNN has a high computational cost and is more time-consuming. It required 13 s per image. On the other hand, the UNet model is computationally efficient in retraining too, because just a small amount of new data is used for retraining. Therefore, the proposed few-shot UNet structure is more computationally productive than the native UNet model. On 10-fold cross validation set, the proposed SAN-FSL model loss vs. accuracy graphs are visually displayed in Fig. 5. For prediction stages of MM, the Fig. 6 represents the confusion metric for classification of four stages of myopic maculopathy (MM).

**Figure 5:** The proposed model loss versus accuracy based on 10-fold cross validation set

The segmentation results of the MM-infected areas for the three deep learning models with other state-of-the-art comparisons are shown in Table 5. It is noticed that CNN performed over-segmentation by including regions that were not MM-infected compared to the proposed model. However, the FCN outperformed with UNet segmentation. Here, we give a sign of poor segmentation. Furthermore, the detection failure may be due to no annotation at all or partial annotation of the area, whereas symptoms can be seen in the fundus image. Hence, if identification is not successful in earlier images, it can be done in adjacent images. In this way, valuable help can be given to medical professionals. A model named the few-shot UNet model is proposed, and its performance is demonstrated in the next paragraphs.

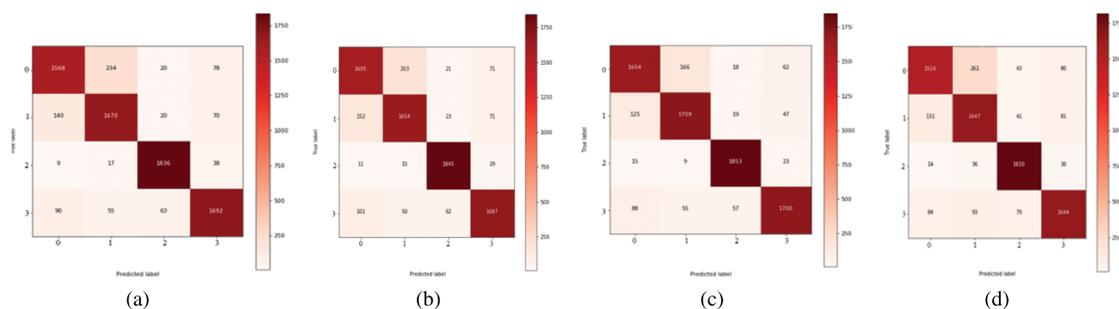


Figure 6: Confusion metric for classification of four stages of Myopic maculopathy

Tables 2 to 5 illustrates the U-Net model’s performance scores before and after rectification. The rectified UNet outperformed the original test dataset (the test dataset is reduced by eight images). Additionally, it was noticed that the deep network kept up with its high-performance metrics on the training and validation datasets, in this manner, preventing it from overfitting on the enlarged training data.

Table 2: Classification performance of the proposed method with other machine learning classifier using 70%–30% train-test partition strategy

Feature extraction	Classifier	ACC	SPE	SEN	F1-Score
Proposed	CNN	95%	95%	95%	96%
	LSTM	94%	93%	96%	95%
	CNN-LSTM	96%	94%	97%	96%
	Semantic Seg (UNet)	95%	94%	97%	96%
	Proposed Classifier (UNet-GAN)	97%	96%	97%	97%

Table 3: Classification performance of the proposed method with other machine learning classifier using 80%–20% train-test partition strategy

Feature extraction	Classifier	ACC	SPE	SEN	F1-Score
Proposed	CNN	95%	96%	94%	95%
	LSTM	94%	95%	94%	95%
	CNN-LSTM	96%	95%	96%	96%
	Semantic Seg (UNet)	95%	95%	95%	95%
	Proposed Classifier (UNet-GAN)	97%	97%	95%	96%

Table 4: Classification performance of the proposed method with other machine learning classifier using 90%–10% train-test partition strategy

Feature extraction	Classifier	ACC	SPE	SEN	F1-Score
Proposed	CNN	96%	94%	96%	96%
	LSTM	95%	95%	95%	95%
	CNN-LSTM	97%	96%	96%	97%
	Semantic Seg (Unet)	96%	96%	96%	96%
	Proposed Classifier (UNet-GAN)	98%	97%	98%	98%

Table 5: Comparison of proposed method with other existing studies

Method	Recall	Precision	F1-score	Accuracy
Rauf et al. [8]	94%	94%	94%	94%
Li et al. [9]	82%	97%	89%	88%
Devda et al. [11]	86%	96%	91%	94%
Lu et al. [26]	90%	92%	91%	87%
Du et al. [27]	83%	89%	82%	93%
Zhang et al. [28]	94%	96%	95%	95%
Proposed method	97%	97%	97%	97%

Tables 2–4 represent the classification performance of the Haar-Like feature extraction method with different ML classifiers in terms of ACC, SEN, SPE, and F1-Score using three different train-test partition strategies. Our proposed classifier gives us 97%, 97%, and 98% accuracy on three different splits. It is noteworthy to see other metrics such as specificity, sensitivity, and F1-score. We have noted that our proposed method outperforms other classifiers, and the performance of these metrics is not less than 96% on all data splits using our proposed classifier. Fig. 5 depicts the confusion matrix of the proposed method and other machine learning classifiers for different training test splits. These confusion matrixes were depicted to evaluate the performance of the different classifiers better. It was noticed clearly from these confusion matrixes that our proposed classifier consistently performs better on each training test split strategy. Our proposed technique wrongly classified 5 samples, 20 samples, and 24 samples using three different train test splits strategies (90%–11%, 80%–20%, and 70%–30%), which is much less than other classifiers. This further confirms our proposed classifier’s effectiveness in identifying three stages of MM.

5 Discussions

In this paper, Myopia maculopathy (MM) detection and classification from retinograph images are proposed by using multi-layer deep-learning and few-short learning (FSL) algorithms. In practice, numerous diseases, including cataracts, glaucoma, retinal detachment, and myopia maculopathy, can be followed by myopia maculopathy (MM). Myopia is therefore acknowledged by the World Health Organization as a substantial contributor to visual impairment if it is not entirely cured.

The burden of MM on patients, their families, and society is enormous. Currently, there is no proven treatment for MM. However, prevention therapy can reduce ocular complications for all myopic patients. The existence of myopic maculopathy was characterized and grouped in view of the International Photographic classification and grading System for myopic maculopathy (MM) [3]. Myopia is categorized based on severity. In this research, fundus images in category 2 and above are considered for diagnosing myopic maculopathy.

Many scientific studies have recently been published on using deep learning models for MM-infected area segmentation. The fully-connected convolution network and the U-shaped convolution network are very successful in pixel-based segmentation for medical images. As a result, both are prioritized for segmentation of the PM-infected area in the retinal fundus image. Due to the complexity of the PM or MM classification and definition systems, the implementation of deep learning (DL) technologies in detecting pathologic myopia (PM) lesions remains a challenge. However, with adequate resources such as high-quality PM retinal fundus image datasets and high-standard expert teams, goals can be met. The goal of this research is to develop and train DLs to identify PM as well as the categories. In this study, we examine the efficacy of deep learning models for segmenting PM in fundus images by utilizing semantic adversarial networks (SAN). In contrast to DL models, which require a large dataset and fixed (static) network weights, a few-shot UNet deep learning approach is explored for detecting MM regions. A few-shot can train the network model using a minimal number of samples and enables the network weights to be dynamically adjusted as new samples are introduced into the UNet. According to experimental findings, the segmentation accuracy for identifying MM-infected zones (as shown in Fig. 6) has improved. Moreover, few-shot learning entails a generic pool of algorithms with various modifications depending on the task, degree of user involvement, data complexity, and so on.

In this case, we used the concept of few-shot learning with a UNet model and GAN for MM image segmentation and classification as shown in Fig. 7. In addition, the proposed deep model is correctly classified MM into four classes as visually displayed in Fig. 8. It is an online learning methodology that uses only a few samples at a time to train. In this methodology, the new fundus images with the consent of domain experts are fed into the model. Next, the model results are evaluated by experts in the testing phase to identify the misclassified results. Even though this methodology enhanced segmentation performance and offered an online learning framework, it needed expert interaction throughout the testing phase. Therefore, in this approach, human involvement is necessary throughout the learning phase. Mistakes in the medical experts' decisions, albeit unlikely, would have resulted in a decrease in overall network performance.

An identical issue is experienced in the supervised learning paradigm where an incorrectly labeled dataset is fed to the network. However, the supervised learning ground-truth data was created offline, so there was ample time to assess the annotation correctness. In the suggested learning approach, weights are adjusted dynamically, so the expert has no time to reevaluate the decision. Secondly, a forgetting mechanism is adopted to reduce the training dataset and entertain new training samples. To understand the image details, it is analyzed in detail from a coarse-grained to a fine-grained level. In the first step, classification will be done, and MM infection will be checked in the image. In the second step, the MM area is localized, labeled, and a bound box is drawn around the area of interest. This will aid professionals in narrowing down the diagnosis. However, bounding boxes are insufficient for many applications (for example, precise tumor diagnosis). In such situations, we require highly comprehensive pixel-level information, commonly known as pixel-based segmentation. This is the target of semantic segmentation. In this scenario, all pixels of an image are labeled with the respective class. However, semantic segmentation is constrained by time constraints, limited computational

capabilities, and low false-negative detection limits. In future work, we hope to incorporate some other segmentation-based networks [42,43] to detect MM regions from fundus images.

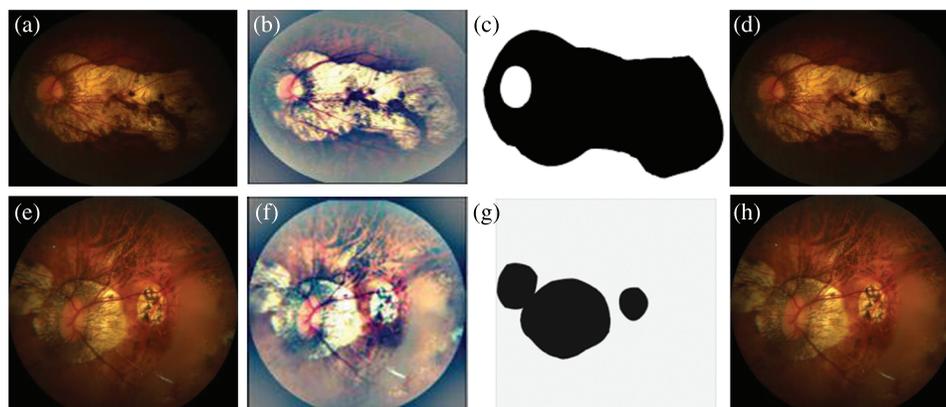


Figure 7: A Visual example of proposed model for segmentation of MM-infected region, where figure (a) shows the input image, (b) preprocessed image, (c) segmented MM-infected region and (d) classify as category 1: tessellated. Also, the figures (e) represents the input image, (f) preprocessed image, (g) segmented MM-infected region and (h) classify as category 2: diffuse chorioretinal atrophy



Figure 8: A visual example of proposed model for classification of different stages of myopic maculopathy with stages such as Category 1: tessellated fundus, Category 2: diffuse chorioretinal atrophy, Category 3: patchy chorioretinal atrophy, Category 4: macular atrophy from left to right

6 Conclusions

The basic aim of this research is to design and train the DLs to automatically identify myopic maculopathy (MM) or pathologic myopia (PM) and its categories. In this article, we have developed a new strength of DL models for the segmentation and detection of PM through semantic adversarial networks (SAN) and few-shot learning (FSL). In this research, a few-shot learning method for the detection and classification of pathologic myopia (PM) infected areas is presented. This model was built using the UNet framework. It dynamically updates the network based on a few incoming new samples. This retraining technique permitted the model to trust however much as could reasonably be expected of the approaching new incoming samples while minimizing the loss of current understanding. The suggested method differs from standard approaches in that it uses an online learning paradigm, in contrast to the static learning of UNet. This innovative strategy is known as few-shot learning (FSL) powered by UNet to perform segmentation of PM or MM-infected regions. The efficiency of the proposed FSL technique in conjunction with a UNet model is demonstrated by

experimental data. In comparison to deep learning models like CNN, the few-shot learning (FSL) powered by UNet is a potential artificial intelligence (AI) framework for medical imaging.

Acknowledgement: The authors extend their appreciation to the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) for funding and supporting this work through Research Partnership Program no. RP-21-07-04.

Funding Statement: The authors extend their appreciation to the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) for funding and supporting this work through Research Partnership Program no. RP-21-07-04.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] B. S. Modjtahedi, R. L. Abbott, D. S. Fong, F. Lum, D. Tan *et al.*, “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*, vol. 128, no. 6, pp. 816–826, 2021.
- [2] P. Sankaridurg, N. Tahhan, H. Kandel, T. Naduvilath, H. Zou *et al.*, “IMI impact of myopia,” *Investigative Ophthalmology & Visual Science*, vol. 62, no. 5, pp. 1–12, 2021.
- [3] L. Lu, P. Ren, X. Tang, M. Yang, M. Yuan *et al.*, “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*, vol. 2841, pp. 1–9, 2021.
- [4] F. Jotterand and C. Bosco, “Artificial intelligence in medicine: A sword of damocles?” *Journal of Medical Systems*, vol. 46, no. 1, pp. 1–5, 2022.
- [5] Y. Li, L. L. Foo, C. W. Wong, J. Li, Q. V. Hoang *et al.*, “Pathologic myopia: Advances in imaging and the potential role of artificial intelligence,” *British Journal of Ophthalmology*, pp. 1–23, 2022.
- [6] Q. Abbas, I. Qureshi, J. Yan and K. Shaheed, “Machine learning methods for diagnosis of eye-related diseases: A systematic review study based on ophthalmic imaging modalities,” *Archives of Computational Methods in Engineering*, vol. 29, pp. 3861–3918, 2022.
- [7] Q. Abbas, M. E. A. Ibrahim and A. Rauf Baig, “Transfer learning-based computer-aided diagnosis system for predicting grades of diabetic retinopathy,” *Computers, Materials & Continua*, vol. 71, no. 3, pp. 4573–4590, 2022.
- [8] 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.
- [9] 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.
- [10] R. Lu, W. Zhu, X. Cheng and X. Chen, “Choroidal atrophy segmentation based on deep network with deep-supervision and EDT-auxiliary-loss,” *International Society for Optics and Photonics*, vol. 12, no. 10, pp. 1–10, 2020.
- [11] J. Devda and R. Eswari, “Pathological myopia image analysis using deep learning,” *Procedia Computer Science*, vol. 165, pp. 239–244, 2019.
- [12] C. Zhang, J. Zhao, Z. Zhu, Y. Li, K. Li *et al.*, “Applications of artificial intelligence in myopia: Current and future directions,” *Frontiers in Medicine*, vol. 9, pp. 1–22, 2022.
- [13] Z. Zhang, Z. Ji, Q. Chen, S. Yuan and W. Fan, “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*, vol. 26, no. 1, pp. 115–126, 2021.

- [14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley *et al.*, “Generative adversarial networks,” *Communications of the ACM*, vol. 22, no. 63, pp. 139–44, 2020.
- [15] Q. Abbas, I. Qureshi and M. E. Ibrahim, “An automatic detection and classification system of five stages for hypertensive retinopathy using semantic and instance segmentation in DenseNet architecture,” *Sensors*, vol. 21, no. 20, pp. 1–20, 2021.
- [16] L. Sun, C. Li, X. Ding, Y. Huang, Z. Chen *et al.*, “Few-shot medical image segmentation using a global correlation network with discriminative embedding,” *Computers in Biology and Medicine*, vol. 140, pp. 1–21, 2022.
- [17] Y. Tian and S. Fu, “A descriptive framework for the field of deep learning applications in medical images,” *Knowledge-Based Systems*, vol. 210, pp. 1–14, 2020.
- [18] L. R. Ren, J. X. Liu, Y. L. Gao, X. Z. Kong and C. H. Zheng, “Kernel risk-sensitive loss based hyper-graph regularized robust extreme learning machine and its semi-supervised extension for classification,” *Knowledge-Based Systems*, vol. 227, no. 107226, pp. 1–14, 2021.
- [19] S. Irtaza Haider, K. Aurangzeb and M. Alhussein, “Modified anam-net based lightweight deep learning model for retinal vessel segmentation,” *Computers, Materials & Continua*, vol. 73, no. 1, pp. 1501–1526, 2022.
- [20] A. Deepak Kumar and T. Sasipraba, “Multilevel augmentation for identifying thin vessels in diabetic retinopathy using unet model,” *Intelligent Automation & Soft Computing*, vol. 35, no. 2, pp. 2273–2288, 2023.
- [21] H. Lu, S. Tian, L. Yu, L. Liu, J. Cheng *et al.*, “DCACNet: Dual context aggregation and attention-guided cross deconvolution network for medical image segmentation,” *Computer Methods and Programs in Biomedicine*, vol. 214, pp. 1–12, 2022.
- [22] N. Gao, H. Xue, W. Shao, S. Zhao, K. Qin *et al.*, “Generative adversarial networks for spatio-temporal data: A survey,” *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 13, no. 2, pp. 1–25, 2022.
- [23] A. You, J. K. Kim, I. H. Ryu and T. K. Yoo, “Application of generative adversarial networks (GAN) for ophthalmology image domains: A survey,” *Eye and Vision*, vol. 9, no. 1, pp. 1–19, 2022.
- [24] B. Zhan, J. Xiao, C. Cao, X. Peng, C. Zu *et al.*, “Multi-constraint generative adversarial network for dose prediction in radiotherapy,” *Medical Image Analysis*, vol. 77, pp. 1–20, 2022.
- [25] 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.
- [26] 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.
- [27] 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.
- [28] W. Zhang, X. Zhao, Y. Chen, J. Zhong and Z. Yi, “DeepUWF: An automated ultra-wide-field fundus screening system via deep learning,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 8, pp. 2988–2996, 2022.
- [29] 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.
- [30] C. R. Freire, J. C. D. C. Moura, D. M. D. S. Barros and R. A. D. M. Valentim, “Automatic lesion segmentation and pathological myopia classification in fundus images,” *ArXiv Preprint*, pp. 1–15, 2020.
- [31] S. Jia, S. Jiang, Z. Lin, N. Li, M. Xu *et al.*, “A survey: Deep learning for hyperspectral image classification with few labeled samples,” *Neurocomputing*, vol. 448, pp. 179–204, 2021.
- [32] A. Voulodimos, E. Protopapadakis, I. Katsamenis, A. Doulamis and N. Doulamis, “A Few-shot U-net deep learning model for COVID-19 infected area segmentation in CT images,” *Sensors*, vol. 21, no. 6, pp. 1–21, 2021.

- [33] Y. Feng, J. Gao and C. Xu, "Learning dual-routing capsule graph neural network for few-shot video classification," *IEEE Transactions on Multimedia*, pp. 1–12, 2022.
- [34] M. Abdelaziz and Z. Zhang, "Multi-scale kronecker-product relation networks for few-shot learning," *Multimedia Tools and Applications*, vol. 81, pp. 6703–6722, 2022.
- [35] Q. Zhu, Q. Mao, H. Jia, O. E. N. Noi and J. Tu, "Convolutional relation network for facial expression recognition in the wild with few-shot learning," *Expert Systems with Applications*, vol. 189, pp. 1–20, 2022.
- [36] P. Korshunov and S. Marcel, "Improving generalization of deepfake detection with data farming and few-shot learning," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, pp. 1–21, 2022.
- [37] W. Li, Y. Gao, M. Zhang, R. Tao and Q. Du, "Asymmetric feature fusion network for hyperspectral and SAR image classification," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–11, 2022.
- [38] R. Singh, V. Bharti, V. Purohit, A. Kumar, A. K. Singh *et al.*, "MetaMed: Few-shot medical image classification using gradient-based meta-learning," *Pattern Recognition*, vol. 120, pp. 1–15, 2021.
- [39] S. Y. Wang, W. S. Liao, L. C. Hsieh, Y. Y. Chen and W. H. Hsu, "Learning by expansion: Exploiting social media for image classification with few training examples," *Neurocomputing*, vol. 95, pp. 117–125, 2012.
- [40] Y. Xian, B. Korbar, M. Douze, B. Schiele, Z. Akata *et al.*, "Generalized many-way few-shot video classification," in *Proc. of European Conf. on Computer Vision*, Glasgow, UK, Springer, Cham, pp. 111–127, 2020.
- [41] A. Javaria, A. A. Muhammad and M. Sharif, "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*, vol. 249, pp. 1–12, 2022.
- [42] W. Sun, G. Z. Dai, X. R. Zhang, X. Z. He and X. Chen, "TBE-Net: A three-branch embedding network with part-aware ability and feature complementary learning for vehicle re-identification," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 14557–14569, 2022.
- [43] W. Sun, L. Dai, X. R. Zhang, P. S. Chang and X. Z. He, "RSOD: Real-time small object detection algorithm in UAV-based traffic monitoring," *Applied Intelligence*, vol. 52, no. 8, pp. 8448–8463, 2022.