IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 27, 2020, accepted October 4, 2020, date of publication October 13, 2020, date of current version October 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3030787

Automated Pterygium Detection Using Deep Neural Network

N. SYAHIRA M. ZAMANI[1], WAN MIMI DIYANA WAN ZAKI[1], AQILAH BASERI HUDDIN[1], (Member, IEEE), AINI HUSSAIN[1], HALIZA ABDUL MUTALIB[2], AND AZIAH ALI[3], (Member, IEEE)

[1]Department of Electrical, Electronic, and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
[2]Optometry and Vision Sciences Programme, Faculty of Health Sciences, School of Healthcare Sciences, Universiti Kebangsaan Malaysia, Kuala Lumpur 50300, Malaysia
[3]Faculty of Computing & Informatics, Multimedia University, Cyberjaya 63100, Malaysia

Corresponding author: Wan Mimi Diyana Wan Zaki (wmdiyana@ukm.edu.my)

This work was supported in part by the Ministry of Higher Education, Malaysia, under Grant FRGS/1/2017/TK04/UKM/02/4, and in part by the Research University Grants from Universiti Kebangsaan Malaysia under Grant DIP-2018-020.

ABSTRACT Ocular imaging has developed rapidly and plays a critical role in clinical care and ocular disease management. Development of image processing technologies pertinent to ocular diseases has paved the way for automated diagnostic systems including detection techniques using deep learning (DL) approaches. The prevalence of an abnormal tissue layer in the conjunctiva, known as pterygium eye disease, is increasing due to lack of awareness. Despite the non-cancerous/benign nature of pterygium, a clinical diagnosis from an ophthalmologist is still required to prevent the pterygium tissues from extending into the pupil, which would result in blurred vision. However, current diagnostic methods are mostly dependent on human expertise. Automated detection can potentially serve as an assistive method to reduce diagnosis time by applying a DL approach. Considering the lack of comprehensive research work on pterygium detection using DL, we propose a new architecture consisting of an improved CNN-based trained network named VggNet16-wbn that is derived from VggNet16, a pre-trained CNN algorithm. This paper presents an overview of the DL as a core approach to the transfer learning (TL) concept, as well as current efforts towards automated ocular detection approaches. A new architecture of a CNN-based trained network was proposed based on a network assessment from six CNN pre-trained networks to detect pterygium. This work consists of two main modules, namely, data acquisition and DCNN classification. The proposed trained network, VggNet16-wbn, shows the best performance with 99.22% accuracy, 98.45% sensitivity, and a perfect score on specificity and area under the curve metrics. This work has high potential for creating a pterygium screening system that can be used as a baseline for fully automated detection using a DL approach.

INDEX TERMS Ocular imaging, pterygium detection, deep learning, transfer learning.

I. INTRODUCTION

Diseases caused by ultraviolet (UV) exposure affect 1.6 million people and account for an estimated 0.1% of the total diseases in the world based on disability-adjusted life years (DALYs) as measured by WHO's Global Burden of Disease (GBD) project [1], [2]. Ocular diseases, such as pterygium, which involves the anterior structure of the eye, are caused by excessive UV exposure and normally result in eye dryness, irritation, and discomfort. From a worldwide view of pterygium prevalence that started early in the year 1965, Cameron [3] concluded that this disease is more common in tropical regions. Approximately one-tenth of Norfolk Islanders in Australia are positive for pterygium [4]. Moreover, the 9.5% prevalence of pterygium in South India, especially in the tropical region, is associated with a lifetime of high UV exposure [5]. The Singapore Malay Eye Study has reported that the number of Malay adults with pterygium is higher than that of Chinese adults by 12.3% [6]. Early screening and detection of pterygium tissue can be performed through clinical diagnosis by a trained ophthalmologist. Examinations are executed using a slit lamp to detect the

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

presence of pterygium. The presence of pterygium tissue can also be seen clearly in the overview of a topographic map of the anterior eye through the use of corneal topography machines [7]. Current pterygium diagnosis methods rely heavily on human expertise involving the use of expensive diagnostic tools that takes time to execute.

The effort to achieve a fully automated pterygium detection method using a digital image processing (DIP) approach has been reported in previous research literature in which anterior segment photographed images (ASPIs) of the eyes are captured using a slit lamp or regular digital camera [8]–[11]. The three-step frame differencing technique has been introduced to segment the cornea areas before the regions of interest are fed to the feature extraction and cornea classification modules [9]. Technological advancement using machine learning (ML) leads to increase in development of automatic pterygium detection methods while reducing diagnostic time and assisting ophthalmologists in the pterygium screening process. ML implementation is based on learning image features through a neural network (NN) of ASPIs input into the system to predict the output image class from a probability calculated by the classifier function during the classification stage. However, the implementation of this method requires several conventional image processing steps by manual execution.

Automated detection, which has been widely used with deep neural networks (DNN), exists as an alternative approach to reduce diagnostic time. The implementation of DNN requires large-scale data, thus becoming a challenge for certain domains. As a solution, the concept of transfer learning (TL) has been applied to overcome this constraint. Based on a literature review, comprehensive research related to pterygium detection using DNN is still currently limited. DL method has long been applied in classification, tracking and detection tasks for various diseases [12]–[17]. Therefore, pterygium detection using TL in the DNN approach is proposed in this study.

This paper is divided into five sections, starting with an introduction of the entire research interest area of ocular diseases, with an emphasis on pterygium. An overview of DL in medical imaging including an evolutionary study of a pretrained convolutional neural network (CNN), an introduction of pterygium disease and a review of implemented pterygium detection methods using the DL through a TL approach are discussed in Section 2. Section 3 presents the methods to construct the proposed network, which is followed by a discussion of the experimental results in Section 4. Finally, Section 5 discusses the limitations of this work and recommendations for future studies to make the proposed pterygium detection method more portable, accurate and efficient, thus concluding this research conference.

II. BACKGROUND

In this section, we provide an overview of the DL approach as a core approach to TL using CNN pretrained network, followed by an introduction to pterygium and its medical assessment.

A. OVERVIEW OF DEEP LEARNING

ML is a type of artificial intelligence (AI) that allows computers to learn by experience from a data sample without being explicitly programmed. Through previous learning experiences, the ML method searches for natural features in input data that help in achieving good feature extraction results. ML is widely used in the field of image processing and computer vision and in various applications such as face recognition and tumour detection [12]. There are three main categories of ML, namely, supervised, unsupervised learning, and semi-supervised learning [18]–[20]. Supervised learning refers to the development of a predictive network based on information from labelled input and output data. In unsupervised learning, collected data is translated based on information gathered from unlabelled input data. Semi-supervised learning is a combination of feature learning from labelled data (supervised learning) and unlabelled data (unsupervised learning). These types of learning can be applied to classification predictions in which input data are classified into appropriate categories for medical imaging, voice recognition and credit scoring [21]. In addition, regression predictions are used in supervised learning to predict continuous responses, such as changes in temperature or power, stock prices, and electricity projection forecasts [20]. Clustering techniques, including object identification, sequential gene analysis, and market research, are commonly used in data analysis to find embedded patterns in input data by unsupervised learning.

NNs are biologically-inspired sub techniques in ML that aim to resemble the human nervous system. The visual cortex contains an easy and complex alternate layer of cells in the brain that inspired the creation of an artificial neural network (ANN) architecture [22]. The NN architecture consists of three processing layers, namely, the input, hidden, and output layers, that are represented by several neuron units in each layer. Generally, the NN data classification or identification process undergoes three phases, namely, training, validation, and testing. With advancements in data processing, the DL method has been applied to achieve effective performance.

Deep learning in NN, or in a deep neural network (DNN), is a study of a high level of data information that uses an architectural hierarchy that starts from a low level and layer by layer forms a specific feature. Deep learning processes require more hidden layers than traditional NN [23]–[25]. DL requires an advanced computer system technology to process data in a DL structure, a deeper learning process and large scales of input data to achieve maximum performance [26]. DL techniques include recurrent neural network (RNN) for voice recognition and natural language processing [27], deep belief network (DBN) as proposed by Hinton *et al.* [28] in 2006 for use in stock price index forecasts [29], [30], and CNN as proposed by LeCun *et al.* [31] in 1989, which is also known as DNN, or more precisely, deep CNN (DCNN) [24], [30], [32]–[34]. DCNN algorithms are typically used for image classification with DL processes [14], [17], [33], [35], [36], in which the hidden layers will process more layers; DCNN algorithms are quite complex.

CNN or ConvNet is a class of DNN that is most commonly used in image analysis. It is a specific network used for classification through learning directly from input data sources, such as images, videos, texts or voice [37], [38]. A DCNN eliminates the need to deploy manual feature extractions as documented in previous works [8], [9], [11], and the image features can be extracted automatically. Image processing is executed automatically during feature extraction via CNN; the feature extraction mechanism is embedded within the hidden network layer [39]. Feature extraction by CNN has been proven to be more effective [17], [40] than manual techniques in DIP.

B. PRETRAINED CNN

Implementation of DL techniques require large-scale databases for maximum performance [41]. In some domains, particularly in medical imaging, acquisition of large-scale images or datasets is a challenging and difficult task because of data-acquisition factors and other requirements. The image classification task is based on DL methods employed in the common phases of training, validation, and testing. The DNN training process can be implemented using training datasets of beginning or existing CNN-trained networks [42]. Learning from scratch refers to a process of training a data set from the beginning [42], [43], which requires a manual network configuration and a crucial understanding of the DNN structure. Moreover, it requires a large quantity of datasets so the DNN can study each feature in the new input data. It therefore requires more effort and a longer training time to perform early network training. Hence, TL becomes an initiative step in training a new dataset without needing a manual configuration of the network layer and data deficits.

TL is a common approach related to DL in computer vision and natural language processing tasks; it was introduced following the emergence of ML in network training [44]. Instead of training data from scratch, which requires large-scale data, TL is an alternative for small data representations in DL [41], [45]. TL is referred to as the transfer of relevant knowledge and information that has been learned into a new task.

C. PTERYGIUM AND ITS MEDICAL ASSESSMENT

Pterygium refers to the abnormal tissue layer on the conjunctiva known as wing-shaped conjunctivitis [6], [46], [47], as illustrated in Fig. 1 (a). This ocular disease appears as a triangular shape tissue, growing near the nasal area and invading the cornea, as shown in Fig. 1 (c)[47]–[51]. The conjunctiva is extremely susceptible to damage by UV rays, because it is located at the outer lining, which protects the inner part of the eye. Excessive UV rays are among the primary influences on the prevalence of pterygium [46], [52]–[54].

Pinguecula is a common ocular disease similar to pterygium, as depicted in Fig. 1 (b). As highlighted in Fig. 1 (d), the pinguecula tissue is a yellowish white spot that appears on the scleral area because of excessive UV exposure, as in pterygium [46], [55]. Unlike pterygium, pinguecula does not spread to the corneal area. However, it can potentially develop into pterygium if the tissue continues to grow on the sclera and enter the corneal region [51]. According to Viso *et al.* [54], this conjunctivitis tissue will lead to vision disorders that are categorised as noncancerous; it is not harmful [56] but should be of concern to prevent the disease from causing vision disturbances. Nevertheless, pterygium and pinguecula do not require serious treatment unless they cause extreme discomfort, which would require consultation with an ophthalmologist for further treatment.

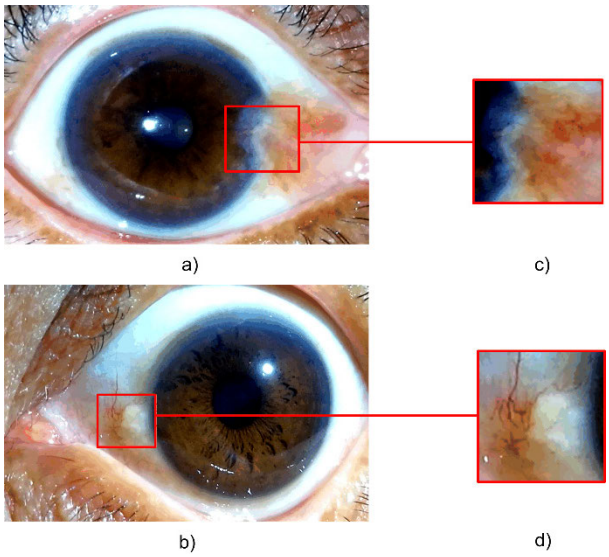


FIGURE 1. ASPI of (a) pterygium, (b) pinguecula and appearance of (c) pterygium tissue, (d) pinguecula tissue.

The abnormal cells of conjunctivitis are associated with gene anomalies that cause interference in eye cells. Continuous UV radiation, especially UVB, disrupts the structure of the eye cell DNA [56]. In addition to excessive sunlight exposure, wind, dust, and excessively dry conditions can also trigger abnormal tissue formation in the scleral area [57]. Tropical countries along the equatorial region, such as Malaysia, Brazil, and Australia, which receive the highest UV intensities, are found to have higher prevalence of pterygium and pinguecula [6], [47], [50], [55], [58]. According to the Journal of Health Sciences, 88.9% of the farmers working in the Cameron Highlands, Malaysia and 94.4% in Bachok and Pasir Puteh, Kelantan, Malaysia are positive for pterygium [59]. The formation of pterygium based on UV radiation exposure factors has also been proven by recent studies conducted in Padang Terap, Kedah, Malaysia; a majority of patients suffer from pterygium and pinguecula [10]. This can be attributed to the daily routines of the residents that are consistently exposed to UV radiation for long periods. Most of them are working as farmers, bus drivers, and fishermen [9], [10].

1) MANUAL ASSESSMENT

Early detection of any disease is usually achieved through clinical tests in hospitals or clinics where an ophthalmologist will manually check for the presence of any abnormalities

in patient's eyes. During the diagnostic process, the use of a UV fluorescent liquid, a yellow-coloured liquid, is focused on abnormal tissue [60], which results in the manifestation of a difference between the pterygium tissue and the normal tissue in the eye. Abnormal tissues are visible in the early stages of the pterygium progression, which spreads across the corneal area of the eye. Nevertheless, the accuracy of detection and diagnosis depends on the expertise and experience of the doctors and the special instruments used. Additionally, the use of UV fluorescent liquids can be cost inhibitive and may have side effects on the patients [8], [9].

Corneal topography is an advanced medical assessment of ocular diseases through a medical instrument that aims to assist ophthalmologists in their diagnosis. It is used to map the surface of the corneal tomographic image to visualise the irregular abnormal tissues of pterygium. These instruments are normally expensive and bulky, and human expertise is needed to integrate the diagnosis report. Thus, clinicians are now aware of the need for an automated detection approach to assist with the diagnosis of ocular diseases.

2) TOWARDS AUTOMATED DETECTION

Previous studies have suggested appropriate methodologies for the detection of many ocular diseases such as pterygium, pinguecula, cataracts, glaucoma, and diabetic retinopathy (DR) using DIP approaches in computer vision for medical imaging. Image processing techniques applied on ASPIs produced by video keratography machines provide a clear picture of abnormal tissues for pterygium cases [9]. A video keratography machine or corneal topography machine is usually used by an eye specialist to verify the presence of pterygium. Corneal topography uses a special device that radiates a series of corneal light rings and is normally integrated with CAD techniques to diagnose pterygium tissues [7]. In addition, tomographic analysis is usually used to obtain information on abnormal corneal surface qualities for use in image processing methods [9], [61]–[63].

Mesquita and Figueiredo have applied the Circular Hough technique (CHT) and the Otsu thresholding method to detect the advancement of pterygium tissue in the ASPIs of the eyes [64]. Their proposed method achieved 63.4% accuracy and proved that the DIP approach could be used to investigate pterygium cases. Another study conducted by Gao *et al*. [65] applied Daugman's algorithm in retro-illumination images to detect pterygium and achieved 85.38% accuracy. In addition, early detection of pterygium disease using adaptive sigmoid enhancement techniques performed well on the ASPI database collected from [64]. This technique was successful in detecting pterygium with 93.53% accuracy and 88.18% specificity [8]. A study of pterygium detection in 2017 by [66], where the ASPI was obtained from a slit-lamp camera, achieved 90% sensitivity.

Recently, in 2018, Zaki *et al.* [9] applied image enhancements using sigmoid techniques, as in their previous study [8] for pterygium disease detection. Prior to that, Hue, Saturation, Value (HSV) contrast stretching had been applied to diminish luminosity and the effects of artefact reflections on the image. Corneal segmentation has also been applied, where the corneal region is segmented as a foreground and the non-corneal region as a background. Their studies have yielded high accuracy, specificity, sensitivity, and AUC metrics of 91.27%, 88.3%, 88.7%, and 95.6%, respectively, using ANN classifiers [9]. Nevertheless, this performance showed a slight decline compared to their previous study in 2015 [9]. 33Thus, based on previous work (see [8] and [9]), there have been efforts to find an advanced approach for better automatic pterygium detection using deep learning.

Detection techniques through image processing are geared towards the development of automatic detection techniques using machines in DL approaches. The DL network has demonstrated outstanding performance in computer vision applications because of the ability to extract the appropriate features automatically [67], [68], specifically for image classifications. The classification of images has been conducted in the context of computer vision through tasks such as detection, classification, recognition, segmentation, and localisation [26], [69], [70]. To the best of our knowledge, the implementation of DL approaches for pterygium ocular disease detection or screening is very limited. Recently, pterygium detection has been conducted using TL in a DNN approach by Zulkifley *et al*. [71] in which a Vgg network was adopted to detect and localise pterygium tissues. While the proposed method successfully achieved detection with 95% sensitivity, 98.3% specificity and 81.1% localisation, the size of the database used for validation was rather limited with only 120 images. This work aims to improve the performance of their proposed method by introducing a batch normalisation layer and then performing validation using a larger number of images.

III. METHODOLOGY

A proposed block diagram of the pterygium detection using a deep neural network is depicted in Fig. 2. It consists of two main modules, namely, data acquisition and DCNN classification. A total of 386 pterygium and normal ASPIs are obtained from local and non-local databases. Prior to the development of the classification module, all ASPIs were uniformly resized according to the input size of the six selected CNN-trained networks. Then, the six selected CNN-trained networks, namely, AlexNet [35], VggNet16, VggNet19 [72], GoogLeNet [36], ResNet101 [73] and DenseNet201 [74] were trained and evaluated with some hyper-parameter settings using 5- and 10-fold cross-validation (CV) techniques. Based on these preliminary results, the network with the best performance was selected as the input to the proposed network.

Fig. 3 shows a proposed trained network known as VggNet16-wbn, in which additional batch normalisation layers have been integrated with the developed network. The addition of a batch normalisation (BNorm) layer aims to normalise the input layer to the next layer. This layer is located between the convolutional layer and the rectified

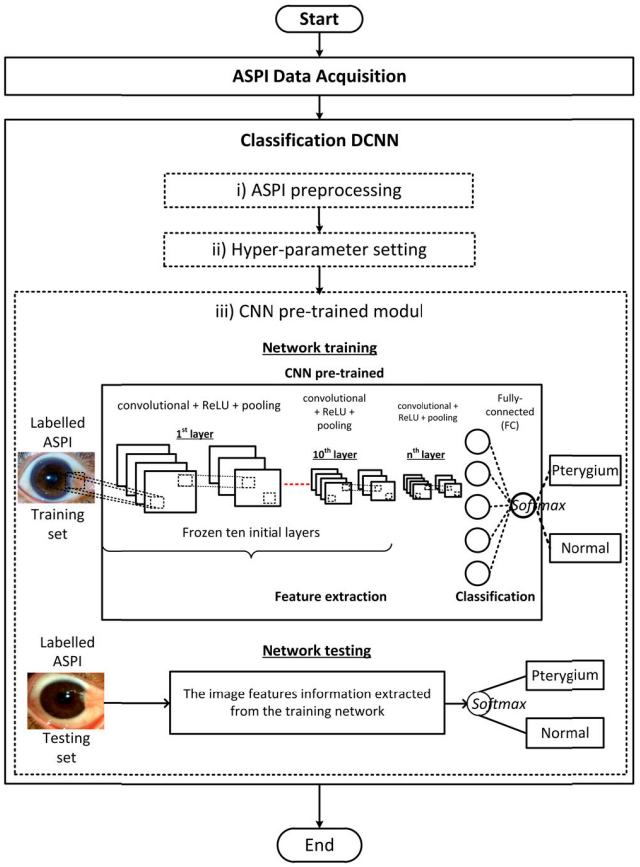


FIGURE 2. Proposed block diagram of pterygium detection using a deep neural network.

linear unit (ReLU) activation function in order to reduce the sensitivity at the beginning of the network on the next layer. The BNorm is the statistical calculation of the subgroups for the whole set of exercises through the mini-batch set before the network training is conducted. These thirteen additional layers are highlighted with a yellow marker in Fig. 3.

The proposed method was implemented on an Intel Core i7 personal computer with 8 GB of RAM and a 3.40 GHz CPU. MATLAB©version 2018b (license no: 40699855) was used to develop the CNN-trained network algorithms and evaluate their statistical performance.

A. DATA ACQUISITION

In this work, 193 normal and 193 pterygium ASPIs have been used that were acquired from local and nonlocal databases. A set of 254 ASPIs from the total ASPIs have been captured using a high-resolution Leica camera from a Huawei P9 smartphone; they are from a local database acquired during the 'Operasi Khidmat Masyarakat Optometri ke-26 (OPKOM-26)' held in Padang Terap Community Centre, Kedah. The OPKOM-26 was organised by the Department of Optometry and Health Vision in the Faculty of Health Science, UKM from 30th March to 1st April 2018, and these images were verified and validated by an ophthalmologist. During the data collection process, subjects were seated on an adjustable chair with their eye, forehead and chin placed towards the chinrest. A nonlocal database consisting of 66 pterygium ASPIs contributed by [64] from the Centre of Informatics, Federal University of Pernambuco, Recife Brazil and 66 normal ASPIs were from an online database UBIRIS [75], [76] were also used in this study.

B. DCNN CLASSIFICATION

The classification of ASPIs into either pterygium or normal classes were conducted by first pre-processing an ASPI and then setting the neural network parameter to produce a pre-trained CNN model for classification.

1) PRE-PROCESSING ASPI

The input size in NNs were fixed by standardisation throughout the network training. Raw input data were depicted in the form of coloured images of red, green, and blue (RGB) represented by three matrices that entered the CNN-trained networks with different image sizes. Hence, all the ASPI images were resized according to fixed input sizes on the pre-trained CNN.

In addition, all resized ASPIs then underwent data augmentation, which is a common technique for data management in DL. This technique scales images by up to 10% vertically and horizontally and then rotates the image randomly. Augmentation techniques are applied to generalise the input data size specified in the CNN-trained networks. However, this process does not create a new database, instead, the data are generated only during the training process. The network will save all the extracted features without confronting overfitting issues; the network training process takes some time to complete.

2) NEURAL NETWORK PARAMETER SETTING

Network training tunes several parameters while training the dataset in CNN-trained networks and applies optimisation algorithms to minimise losses throughout the training process. The general optimisation methods used were stochastic gradient descent with momentum (SGDM), Adam and RMSProp. An SGDM optimiser is applied to train datasets in trained networks to observe the methods that affect the accuracy, specificity, sensitivity, and area under the curve (AUC) metrics of the obtained results. Gradient descent algorithms update network parameters, such as weight and bias, by taking steps through a negative gradient loss to minimise the loss using (1),

$$\theta_{\ell+1} = \theta_\ell - \alpha \nabla E(\theta_\ell) \tag{1}$$

where ℓ stands for the iteration number, θ is the parameter vector, $\alpha > 0$ is the learning rate, and $E(\theta)$ is the abbreviation for the loss function. The standard linear gradient algorithm assesses the entire set of data at once, and the loss function at the gradient $\nabla E(\theta)$ is assessed as a whole training set. A CNN-trained network algorithm evaluates the slope and updates the parameters using a set of hyper-parameter training sets known as a mini-batch of 10. Each gradient

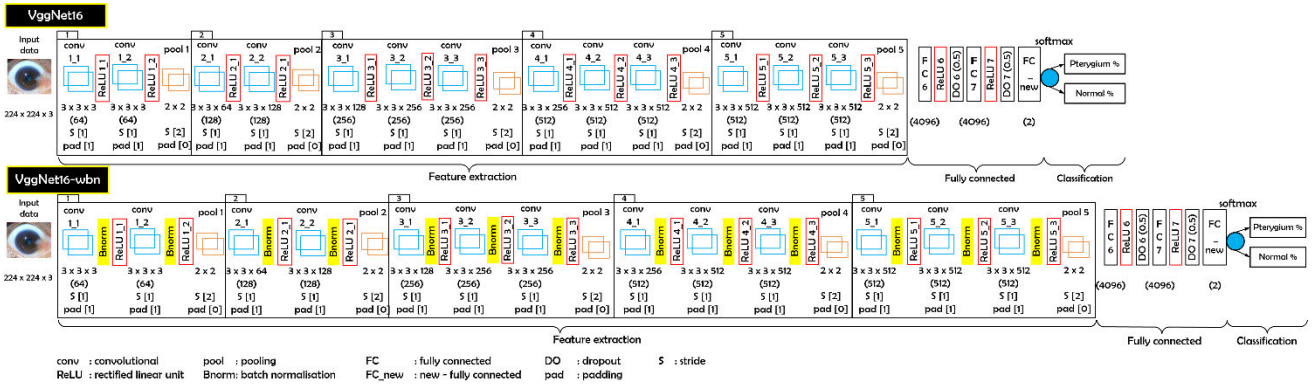


FIGURE 3. Proposed network architecture of VggNet16-wbn. These network shows the revolution of purposed network sourced by VggNet16 [72].

assessment that uses a mini-batch is an iteration. A full pass on the training algorithm for the whole dataset using a mini-batch is an epoch, which is the whole cycle of training with 20 cycles. The loss function is minimised in every iteration by the optimisation algorithm throughout the network training. A large probability stochastic gradient will oscillate along the steep slope towards the optimisation of the training network. This problem is overcome by adding the term momentum to the parameter update as a measure to reduce the swing at the slope [77], as shown with (2),

$$\theta_{\ell+1} = \theta_\ell - \alpha \nabla E(\theta_\ell) + \gamma(\theta_\ell - \theta_{\ell-1}) \tag{2}$$

where the addition of γ from the original (1) determines the merging of the previous gradient step to the current iteration. The learning rate is the proportion of the learning machine in an NN as a major hyper-parameter that controls the speed of learning. The smaller the value of learning rate, the more precise the classification in the learning process. The learning rate value used to train the training dataset in this study was lowered to the value of 0.0001, but it took much longer to train compared with using the value of 0.01. Nevertheless, performance can be improved with the use of a small learning value. These hyper-parameters were tested empirically using an AlexNet network with a single central processor unit (CPU). Regardless, the network structure of AlexNet is a base reference for all the CNN-trained networks.

3) PRE-TRAINED CNN MODEL

The network assessment on six CNN-trained networks was conducted through an evaluation with k-fold cross-validation (CV) techniques to acquire the generalisation of CNN networks towards the unseen samples data. A total of 386 ASPIs have been trained and tested for 5- and 10-fold CVs to obtain the data generalisations among the databases. There was some disagreement regarding the implementation of CV techniques for small-scale samples on whether they will weaken or reduce the network accuracy in identifying image classification, as stated by Varoquaux [81]. This is because the partition of the sample on the small part will increase the level of loss with one false prediction. However, Varoquaux studies are based on neuroimaging images that are more complex than pterygium and normal ASPIs. In this experimental work, ASPI samples are divided into five partition with 20% (5-fold); 4 (training):1 (testing) partition, and ten partition with 10% (10-fold); 9 (training):1 (testing) partition, alternately. The CV techniques implement a training and testing network, where one portion of samples was used for testing and the rest for training. These partitions alternate for 5- and 10-fold training and testing accordingly. Therefore, the proposed network was conducted on a 10-fold CV after comparison with well-known six CNN pre-trained networks.

Training and Testing Network: The network accepts an image that appears as a collection of numbers in a matrix where each number represents the intensity of light at the focal points known as pixels. Raw input data in the form of coloured images of red, green, and blue (RGB) represent three matrices in colour intensity with augmented sizes according to a specific network with a default input size.

Generally, feature extraction in a convolutional network is performed by a feature extractor in the convolutional layer. DCNN has a much deeper network layer compared to a basic convolutional network. The training network starts with input data on the network, followed by several feature extraction layers of convolution, ReLU, pooling, and additional layers, and a classification layer. Additional layers of processes during feature extraction, such as batch normalisation, depth concatenation, normalised local response, and dropout, are added for enhancement in characteristic learning. The implementation of these additional layers is different for each CNN pre-trained network. NN processing starts with input data, which can be in the form of images, videos, text or audio sources, among which images are used as the input data in the CNN pre-trained network.

The training network used a set of training images that correspond to a 5- and 10-fold CV. As we implemented the TLs, which reuse the pretrained network, the weight of the ten initial layers of the proposed network and six CNN pre-trained networks were frozen to retain the trained weight for pterygium and normal ASPIs classification. It is recommended to freeze the initial layer for a small-scale database.

TABLE 1. Performance of AlexNet network tested using each CHP values.

Hyper-parameter			Matric assessment			
Size of mini-batch	Learning rate	Number of epoch	Accuracy	Sensitivity	Specificity	AUC
10	0.01	10	-	-	-	-
	0.001		50.00	0.00	100.00	50.00
	0.0001		94.60	89.64	99.48	99.63
32	0.01		50.00	0.00	100.00	50.00
	0.001		96.60	93.78	99.48	96.63
	0.0001		95.10	100.00	91.04	99.29
64	0.01		50.00	100.00	0.00	50.00
	0.001		97.70	95.85	99.48	97.67
	0.0001		94.30	88.60	100.00	98.87
10	0.01	**20**	-	-	-	-
	0.001		50.00	100.00	0.00	50.26
	0.0001		97.20	96.37	97.93	99.61
32	0.01		50.00	100.00	0.00	50.00
	0.001		95.60	91.19	100.00	98.98
	0.0001		95.30	92.75	97.93	99.32
64	0.01		-	-	-	-
	0.001		97.70	95.34	100.00	99.67
	0.0001		94.80	92.23	97.41	99.07

The non-frozen layers will proceed the feature extraction process. However, the final layer of fully connected layers was changed to a new layer according to the classification of the pterygium and a normal ASPI, and the classification layer was changed as well. A new set of images was tested to classify the pterygium and normal classes using the features extracted throughout the training progress.

IV. RESULTS AND DISCUSSION

In this experimental work, the proposed trained network has been evaluated against six established CNN pre-trained networks. They are AlexNet with eight layers, VggNet16 with 16 layers, VggNet19 with 19 layers, GoogLeNet with 22 layers, ResNet101 with 101 layers and DenseNet201 with 201 layers. The layers represent feature extraction processes. The six selected networks are evaluated with some hyper-parameter settings using 5- and 10-fold cross-validation (CV) techniques. Based on these preliminary results, a network with the best performance is selected as input into the proposed network.

For the hyper-parameter setting, each hyper-parameter value is tested on the ASPIs myMata database using the AlexNet network. In this regard, the AlexNet network is the basis of the CNN pre-trained network architecture since the hyper-parameter settings can meet the needs of other CNN pre-trained networks during training. The ASPI database is divided into three sets, where 70% of all the images were used as the training set, 10% as the validation set and 20% as the test set. These hyper-parameter values were tested to determine the suitable value for 5-folds and 10-folds CV network training. Table 1 shows the performance of each combination of the hyper-parameter (CHP) values. Evaluations were conducted on mini-batch sizes of 10, 32, and 64, in accordance with the prevalence of the CNN networks. The network is trained using three different learning rates of 0.01, 0.001, and 0.0001 with the number of epochs of 10 and 20.

Accuracy is one of the most significant performance metrics in assessing how good an image classifier is in predicting real class of the test images. Based on the observations in Table 1, the highest accuracy value was achieved at 97.70% when trained on the first CHP with a 64 mini-batch size and 20 epochs at a 0.001 learning rate. However, the batch size of 64 uses much computer memory space, making it difficult for the network training process to converge because of the limited computer memory space. Since this study only used CPU chips in all the training and testing of the network, a size of 64 could not be set as the hyper-parameter of the mini-batch because of the limited processing power of the CPU.

Accordingly, the second CHP value setting was identified with an accuracy value of 97.20% by using a 10 mini-batch size trained at 0.0001 for 20 epochs. The mini-batch size of 10 is sufficient to train a small-scale database using the TL approach. To avoid an underfitting problem during the training process, feature extraction was performed at small learning rate (0.0001) by 20 epochs. The ASPI features can be extracted in a more detailed manner, although it takes longer for network training. The number of epochs is determined when network performance is no longer increased during the training process. In addition, this CHP also recorded good sensitivity of 96.37%. Although other CHP have achieved 100% sensitivity, the overall performance of the combination failed to train and test the network. This can be seen with 50% accuracy and AUC and 0% specificity assessment. These could be attributed to overfitting and underfitting as almost all predictions fall short of real expectations. This problem also occurs when the network is trained at a high rate of learning at 0.01 with a small mini-batch size of 10. Poor performance may occur when the learning rate is high and fast, resulting

TABLE 2. Average matric assessment of 5-fold and 10-fold CVs.

CNN-trained networks	Accuracy (Acc), %		Sensitivity (Se), %		Specificity (Sp), %		Area under the curve (AUC), %	
	5-fold	10-fold	5-fold	10-fold	5-fold	10-fold	5-fold	10-fold
AlexNet	90.76	97.42	91.30	96.37	97.11	98.47	96.55	99.73
VggNet16	90.76	98.70	97.40	98.42	84.10	98.97	98.36	99.95
VggNet19	90.50	97.44	91.30	96.95	97.73	97.92	95.39	99.89
GoogLeNet	90.50	95.55	90.78	98.42	97.10	97.95	96.38	99.86
ResNet101	91.52	94.02	91.30	96.89	98.54	91.13	94.50	99.85
DenseNet201	91.52	96.39	91.30	96.95	98.50	95.84	98.41	99.85

TABLE 3. Performance based on training time and complexity.

CNN-trained networks	Training time		Error, %	
	5-fold	10-fold	5-fold	10-fold
AlexNet	94 m 1 s	693 m 37 s	9.24	2.58
VggNet16	331 m 28 s	802 m 18 s	9.24	1.30
VggNet19	880 m 1 s	2873 m 10 s	9.50	2.56
GoogLeNet	81 m 92 s	404 m 14 s	9.50	4.45
ResNet101	608 m 27 s	2082 m 39 s	8.48	5.98
DenseNet201	457 m 47 s	1653 m 32 s	8.48	3.61

m = minute, s = second

in the network not being able to extract the image features properly.

Therefore, a learning rate of 0.0001 on a second CHP is recommended for network training. In fact, this selection is in line with the hypothesis that the smaller the rate of learning, the higher the chance to achieve best performance, even though it may take much longer for the network to train images. A mini-group size of 10 was chosen to train and test the network at a learning rate of 0.0001. The number of the epoch was chosen as 20 because the performance of the network stopped increasing after 20 epochs. This evaluation was trained by using a stochastic gradient descent with momentum (SGDM) optimiser to achieve a minimum gradient during network training.

The classification performance of pterygium and normal ASPIs was evaluated using accuracy (Acc), sensitivity (Se), specificity (Sp), and area under the curve (AUC) metrics, as depicted in Table 2. As stated by Varoquaux [81], a small-scale sample is not relevant for 10-fold CV. By contrast, the 10-fold CV outperformed the 5-fold CV with higher statistical measures, as shown in Table 2. Whereby, a 10-fold CV of VggNet16 outperformed AlexNet, VggNet19, GoogLeNet, ResNet101, and DenseNet201 with higher statistical measures of Acc, Se, Sp, and AUC of 98.70%, 98.42%, 98.97%, and 99.95%, respectively. Furthermore, VggNet16 needed less training time and complexity, as seen in (Table 3), and was 802 minutes and 18 seconds faster than VggNet19, ResNet101, and DenseNet201 but not much longer than AlexNet and GoogLeNet. The VggNet16 training times are neither longer nor shorter, justifying it being a better network than other CNN pre-trained networks. This is because of its fewer hidden layers and complexity, with only 16 layers of feature learning with a small filter size of 3 by 3, which captures image features more efficiently.

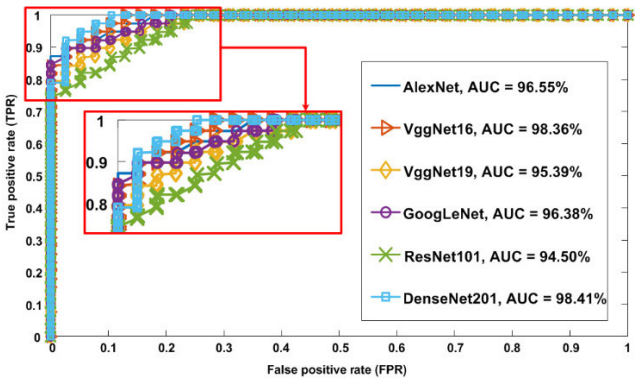


FIGURE 4. Average 5-fold CV ROC curve.

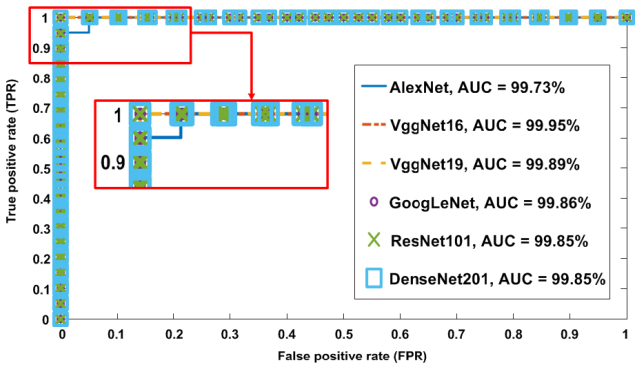
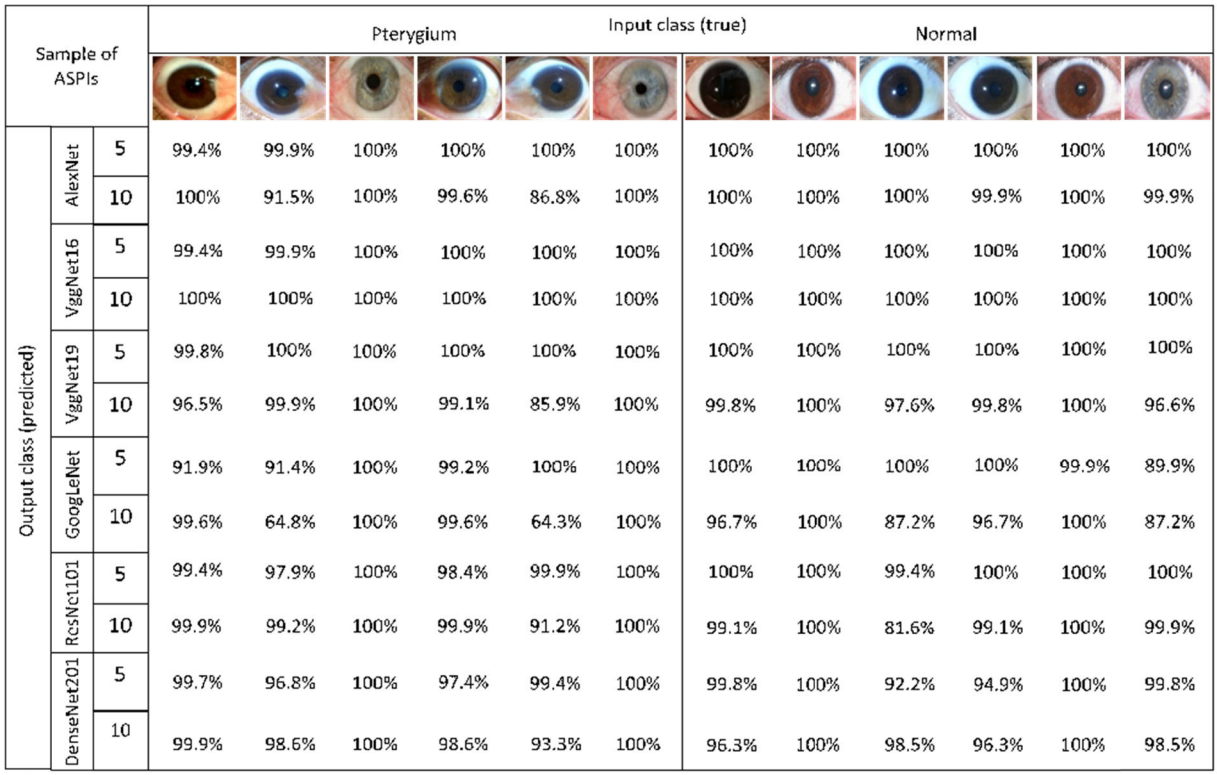


FIGURE 5. Average 10-fold CV ROC curve.

The AUC performance represented by the receiver operating characteristic (ROC) curve, are illustrated in Fig. 4 for the 5-fold CV and Fig. 5 for the 10-fold CV, which denotes the overall performance of six CNN pre-trained networks. According to the ROC curve representation for the 5-fold CV in Fig. 4, the curve for algorithms with better performance



Output class (predicted)		Sample of ASPIs	Pterygium						Normal					
			Input class (true)											
AlexNet	5		99.4%	99.9%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	10		100%	91.5%	100%	99.6%	86.8%	100%	100%	100%	100%	99.9%	100%	99.9%
VggNet16	5		99.4%	99.9%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	10		100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
VggNet19	5		99.8%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	10		96.5%	99.9%	100%	99.1%	85.9%	100%	99.8%	100%	97.6%	99.8%	100%	96.6%
GoogLeNet	5		91.9%	91.4%	100%	99.2%	100%	100%	100%	100%	100%	100%	99.9%	89.9%
	10		99.6%	64.8%	100%	99.6%	64.3%	100%	96.7%	100%	87.2%	96.7%	100%	87.2%
RcsNct101	5		99.4%	97.9%	100%	98.4%	99.9%	100%	100%	100%	99.4%	100%	100%	100%
	10		99.9%	99.2%	100%	99.9%	91.2%	100%	99.1%	100%	81.6%	99.1%	100%	99.9%
DenseNet201	5		99.7%	96.8%	100%	97.4%	99.4%	100%	99.8%	100%	92.2%	94.9%	100%	99.8%
	10		99.9%	98.6%	100%	98.6%	93.3%	100%	96.3%	100%	98.5%	96.3%	100%	98.5%

FIGURE 6. Some examples of pterygium and normal ASPI classification for 5-fold and 10-fold CV.

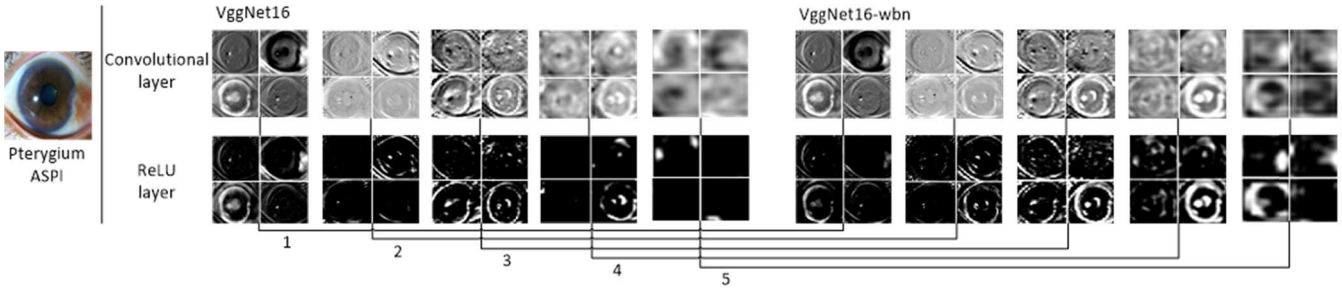


FIGURE 7. Comparison of qualitative performance between VggNet16 and VggNet16-wbn on sample of pterygium ASPI. Activation feature map of convolutional layer (first row) and ReLU layer (second row).

will be closer to the upper corner edge. DenseNet201 curve shows a perfect classification, followed by AlexNet with only a slight difference of 1.86% at 98.41% AUC. While the performance of the 10-fold CV in Fig. 5 shows improvement of the AUC represented by ROC curve with a higher value of 99.95% based on the performance of VggNet16. The performance of VggNet16 shows the majority of high performances and high increments from 5-fold to 10-fold CVs. These performances were based on a numerical result representation of the confusion matrix from the sample in Fig. 6, as scored by the SoftMax classifier function. This result concludes that the VggNet16 network is the best network and that it outperformed the others due to its low network complexity.

The VggNet16 network succeeded in overcoming five other networks at the 10-fold CV. Therefore, the VggNet16-wbn network was proposed to be the best. The performance of the VggNet16-wbn network was evaluated qualitatively and quantitatively. For the qualitative evaluation, Fig. 7 and Fig. 8 show the improvement of the extracted features layer by layer performed by VggNet16-wbn when compared to VggNet16. These activation feature maps were presented by a complete set of feature extractions, whereby VggNet16 contains 5 complete feature extraction sets. Based on the evaluations of Fig. 7, the first layer of convolutions on the first feature extraction set shows the same output for both networks. This is because there is no improvement in the feature extraction performed in the first convolutional

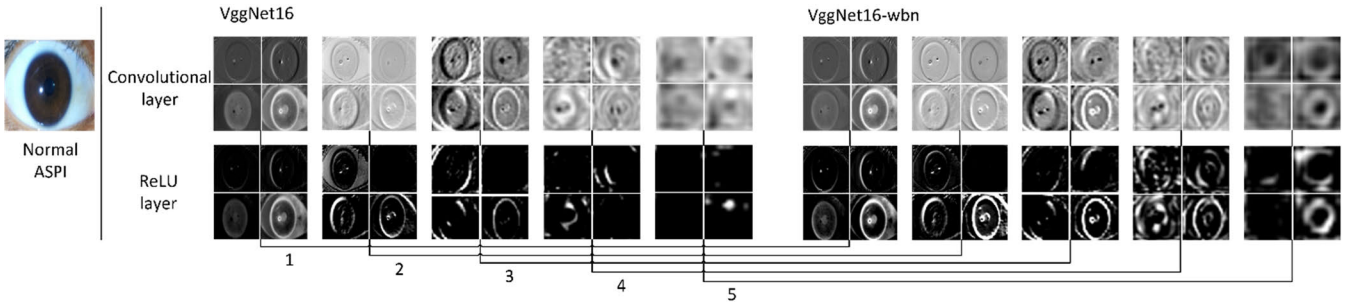


FIGURE 8. Comparison of qualitative performance between VggNet16 and VggNet16-wbn on sample of normal ASPI. Activation feature map of convolutional layer (first row) and ReLU layer (second row).

TABLE 4. Performance of VggNet16-wbn compared to six CNN pre-trained networks on 10-fold.

Matrices	VggNet16-wbn	AlexNet	VggNet16	VggNet19	GoogleNet	ResNet101	DenseNet201
Acc, %	99.22	97.42	98.70	97.44	95.55	94.02	96.39
Se, %	98.45	96.37	98.42	96.95	98.42	96.89	96.95
Sp, %	100	98.47	98.97	97.92	92.68	91.13	95.84
AUC, %	100	99.73	99.95	99.89	99.86	99.85	99.85
Training time	922 m 28 s	693 m 37 s	802 m 18 s	2873 m 10 s	404 m 14 s	2082 m 39 s	1653 m 32 s
Error, %	0.78	2.58	1.30	2.56	4.45	5.98	3.61

layer. However, the contra pixel of an imperfect corneal circle and the pterygium tissue can be seen clearly after the feature map is activated by ReLU from the first to the last set of feature extraction layers in VggNet16-wbn compared to VggNet16. This is because of the normalisation layer that helped in reducing the pixel sensitivity that will be activated by ReLU. The same effect was also applied to the sample of a normal ASPI in Fig. 8, where the perfect corneal circle can be clearly seen. Therefore, the circularity of the cornea and the presence of pterygium tissue improves the network's ability to discriminate and classify a pterygium ASPI and a normal ASPI.

Table 4 shows the quantitative evaluation of the average network performance of VggNet16-wbn versus six CNN pre-trained networks. Performance findings recorded an increase in the Acc, Se, Sp, and AUC metrics compared to other networks with values of 0.52%, 0.03%, 1.03%, and 0.05%, respectively. Regardless, the performance of the VggNet16-wbn network increased over the performance of the VggNet16 proposed by Simonyan & Zisserman [72] and corresponded to the qualitative performance.

Moreover, the time taken for the network to undergo the training process is 922 minutes 28 seconds and is a worthwhile performance time. The performance of the proposed network by using BNorm could resolve the gradient vanishing problem due to a backpropagation process caused by the small derivatives of the activation function, resulting in a small gradient being updated in weights and biases in the initial layers. Furthermore, the layer reduces this problem by normalising the input to a certain range across the processing layer in the backpropagation by enclosing the gradient variance so that the derivative is not too small. Thus, this result concludes that the VggNet16-wbn network with the addition of a BNorm layer performs better than the other networks.

Recently, a deep neural network named Pterygium-Net [71] has been proposed to perform pterygium detection. Like our VggNet16-wbn, the architecture of the proposed Pterygium-Net implements Vgg network, but it only utilizes three initial VggNet layers. This could be due to the limited dataset being used in their pterygium study which uses a detection of only 120 ASPIs comprising 60 images of pterygium and 60 images of normal. To address the data limitation issue, the authors proposed the use of data augmentation and obtained 95% sensitivity and 98.30% specificity. On the contrary, our transfer learning approach via the VggNet16-wbn network has outperformed their Pterygium-Net both in terms of sensitivity and specificity. Tested using datasets acquired from both local and nonlocal databases, Table 5 shows that the pterygium ASPIs detection performance of the VggNet16-wbn (98.45%) surpassed the Pterygium-Net (95%) in terms of sensitivity and obtained perfect score specificity wise. These findings suggest that the developed network of VggNet16-wbn has superior discrimination ability with the addition of batch normalization processing layer. The batch normalization layer helps the network to find every possible features of pterygium tissues. Furthermore, our network uses the hyper-parameter setting of 0.0001 learning rate and mini-batch size of 10 which is well-suited with the datasets and hardware being used as compared to 0.0005 learning rate and 1024 mini-batch size which would need high hardware requirement considering the amount of available dataset. Hence, our VggNet16-wbn has higher potential in detecting pterygium at low cost and minimal hardware requirements.

TABLE 5. Comparison of the performance of sensitivity and specificity between VggNet16-wbn and Pterygium-Net.

Matrices	VggNet16-wbn	Pterygium-Net
Se, %	98.45	95.00
Sp, %	100	98.30

TABLE 6. Comparison of the performance of Softmax and SVM classifiers on Vggnet16-wbn and Vggnet16 networks.

Matrices	VggNet16	VggNet16-SVM	VggNet16-wbn	VggNet16-wbnSVM
Acc, %	98.70	97.68	99.22	96.41
Se, %	98.42	97.95	98.45	96.89
Sp, %	98.97	97.42	100	95.92

The performance of VggNet16-wbn using a Softmax classifier has shown to outperform the VggNet16 network on qualitative and quantitative evaluations. In addition to this result, a simple experimental work has been conducted to compare the performance of the Softmax and support vector machine (SVM) classifiers on VggNet16-wbn and VggNet16. SVM is a well-known and robust conventional ML classifier compared to other ML classifiers [82], [83]. From this experimental work, it was shown that the performance of SVM in DNN is not as expected from previous ML implementations. Table 6 compares the performance of Softmax and SVM classifiers, whereby the former implementation on the VggNet16-wbn network gave the best Acc, Se, and Sp results compared to SVM. This result is based on the fact that the Softmax classifier is more compatible for high data processing than the SVM classifier. Therefore, the Softmax classifier succeeded in detecting pterygium and normal ASPI of the eyes.

V. CONCLUSION AND FUTURE WORK

Automated detection associated with the classification of pterygium using the deep learning approach performs well, which is in line with the state-of-the-art methods. TL, as an alternative way to solve data deficit issues, is the core approach proposed in this work. Currently, only a few studies on DL implementations in medical imaging focus on ocular disease detection because of the limitations of labelled datasets. This limitation is the main challenge faced by researchers in terms of data acquisition, regardless of the requirements of a large-scale database. Data acquisition for medical images needs verification from medical experts in the labelling of disease images, and some other procedures may be required. Moreover, the training network used must be appropriate for classification or detection to maximise performance to enable an accurate and efficient ocular disease detection platform that is fully automated. On average, ocular disease detection or screening from clinics using a fully automated system could show a large improvement that may exceed human expectation with the increase in DL in ML approaches. The performance of VggNet16-wbn using the Softmax classifier to detect pterygium outperforms other comparative networks with 99.22% accuracy, 98.45% sensitivity, and perfect scores for the specificity and area under the curve metrics using the 10-fold CV. This excellent performance proves that the additional batch normalisation layers integrated with the VggNet16-wbn have successfully reduced the sensitivity of the initial layer by adjusting the activation layer.

In addition, it makes the network generalisation process on the 10-fold CV more efficient while minimising overfitting problems. In conclusion, a pterygium detection approach has been successfully developed using an improved CNN pre-trained network, VggNet16-wbn. In future, the proposed network can be used as a baseline for the development of an automated pterygium screening system in myMata application for the convenience of rural community.

ACKNOWLEDGMENT

The authors would like to thank the Optometry and Vision Sciences Programme, Faculty of Health Sciences, School of Healthcare Sciences, Universiti Kebangsaan Malaysia, for providing the facilities and resources and the Padang Terap Society for voluntarily participating as subjects for image data acquisition.

REFERENCES

[1] C. J. L. Murray and A. D. Lopez, "The global burden of disease?: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020," *World Heal. Organ.*, vol. 1, pp. 1–43, Feb. 1996, doi: 10.1088/1742-6596/707/1/012025.

[2] R. M. Lucas, A. J. Mcmichael, B. K. Armstrong, and W. T. Smith, "Estimating the global disease burden due to ultraviolet radiation exposure," *Int. J. Epidemiology*, vol. 37, no. 3, pp. 654–667, Jun. 2008, doi: 10.1093/ije/dyn017.

[3] M. E. Cameron, *Pterygium Throughout the World*. Geneseo, IL, USA: Springf. Thomas, 1965.

[4] J. C. Sherwin, A. W. Hewitt, L. S. Kearns, L. R. Griffiths, D. A. Mackey, and M. T. Coroneo, "The association between pterygium and conjunctival ultraviolet autofluorescence: The norfolk island eye study," *Acta Ophthalmologica*, vol. 91, no. 4, pp. 363–370, Jun. 2013, doi: 10.1111/j.1755-3768.2011.02314.x.

[5] R. Asokan, R. S. Venkatasubbu, L. Velumuri, V. Lingam, and R. George, "Prevalence and associated factors for pterygium and pinguecula in a South Indian population," *Ophthalmic Physiolog. Opt.*, vol. 32, no. 1, pp. 39–44, Jan. 2012, doi: 10.1111/j.1475-1313.2011.00882.x.

[6] H. Cajucom-Uy, L. Tong, T. Y. Wong, W. T. Tay, and S. M. Saw, "The prevalence of and risk factors for pterygium in an urban malay population: The singapore malay eye study (SiMES)," *Brit. J. Ophthalmology*, vol. 94, no. 8, pp. 977–981, Aug. 2010, doi: 10.1136/bjo.2008.150847.

[7] American Medical Association. (2017). *Medical Policy: Corneal Topography*. Accessed: Oct. 25, 2018. [Online]. Available: http://jfs.ohio.gov/

[8] S. R. Abdani, W. M. D. W. Zaki, A. Hussain, and A. Mustapha, "An adaptive nonlinear enhancement method using sigmoid function for iris segmentation in pterygium cases," in *Proc. Int. Electron. Symp. (IES)*, Sep. 2015, pp. 53–57, doi: 10.1109/ELECSYM.2015.7380813.

[9] W. M. D. W. Zaki, M. M. Daud, S. R. Abdani, A. Hussain, and H. A. Mutalib, "Automated pterygium detection method of anterior segment photographed images," *Comput. Methods Programs Biomed.*, vol. 154, pp. 71–78, Feb. 2018, doi: 10.1016/j.cmpb.2017.10.026.

[10] N. S. M. Zamani, L. A. Ramlan, W. M. D. W. Zaki, A. Hussain, and H. A. Mutalib, "Mobile screening framework of anterior segment photographed images," *Int. J. Eng. Technol.*, vol. 7, no. 4, pp. 85–89, 2018.

[11] S. N. A. Ahmad, W. M. D. W. Zaki, and N. S. M. Zamani, "Sistem saringan penyakit pterigium untuk imej mata terangkum hadapan," *J. Kejuruter.*, vol. 31, no. 1, pp. 99–105, 2019.

[12] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018, doi: 10.1109/ACCESS.2017.2788044.

[13] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017, doi: 10.1038/nature21056.

[14] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," *Neurocomputing*, vol. 191, pp. 214–223, May 2016, doi: 10.1016/j.neucom.2016.01.034.

[15] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation," *Expert Syst. Appl.*, vol. 42, no. 3, pp. 990–1002, Feb. 2015, doi: 10.1016/j.eswa.2014.09.020.

[16] B. Sahiner, H.-P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 598–610, Oct. 1996, doi: 10.1109/42.538937.

[17] X. Yang, S.-Y. Yeo, J. M. Hong, S. T. Wong, W. T. Tang, Z. Z. Wu, G. Lee, S. Chen, V. Ding, B. Pang, A. Choo, and Y. Su, "A deep learning approach for tumor tissue image classification," in *Proc. Biomed. Eng.*, 2016, pp. 1–7, doi: 10.2316/P.2016.832-025.

[18] K. Nguyen, C. Fookes, and S. Sridharan, "Improving deep convolutional neural networks with unsupervised feature learning," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 2270–2274.

[19] P. S. Grewal, F. Oloumi, U. Rubin, and M. T. S. Tennant, "Deep learning in ophthalmology: A review," *Can. J. Ophthalmology*, vol. 53, no. 4, pp. 309–313, Aug. 2018, doi: 10.1016/j.jcjo.2018.04.019.

[20] X. Zhu and A. B. Goldberg, "Introduction to semi-supervised learning," *Synth. Lectures Artif. Intell. Mach. Learn.*, vol. 3, no. 1, pp. 1–130, Jan. 2009, doi: 10.2200/S00196ED1V01Y200906AIM006.

[21] A. Lawrynowicz and V. Tresp, "Introducing machine learning," *Perspect. Ontol. Learn.*, vol. 2, no. 1, pp. 35–48, 2011, doi: 10.1111/j.2041-210X.2010.00056.x.

[22] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol.*, vol. 160, no. 1, pp. 106–154, Jan. 1962, doi: 10.1113/jphysiol.1962.sp006837.

[23] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, Apr. 2016, doi: 10.1016/j.neucom.2015.09.116.

[24] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015, doi: 10.1016/j.neunet.2014.09.003.

[25] H. Yi, S. Shiyu, D. Xiusheng, and C. Zhigang, "A study on deep neural networks framework," in *Proc. IEEE Adv. Inf. Manage., Communicates, Electron. Autom. Control Conf. (IMCEC)*, Oct. 2016, pp. 1519–1522, doi: 10.1109/IMCEC.2016.7867471.

[26] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," 2015, *arXiv:1503.02531*. [Online]. Available: https://arxiv.org/abs/1503.02531, doi: 10.1063/1.4931082.

[27] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern Recognit. Lett.*, vol. 119, pp. 3–11, Mar. 2019, doi: 10.1016/j.patrec.2018.02.010.

[28] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.

[29] G. Ghazali, J. Jondri, and M. Si, "Prediksi saham menggunakan DBN," *eProc. Eng.*, vol. 4, no. 1, pp. 1258–1273, 2017.

[30] L. Deng, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, 2014, doi: 10.1561/2000000039.

[31] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989, doi: 10.1162/neco.1989.1.4.541.

[32] A. Mishra and H. Cheng. (2017). *Advanced CNN Architectures*. [Online]. Available: https://pdfs.semanticscholar.org/f109/f538152c3a5485ad52d8c0c67b50b12d2a14.pdf

[33] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis.*, Cham, Switzerland: Springer, 2014, pp. 818–833, doi: 10.1007/978-3-319-10590-1_53.

[34] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with_newline deep neural networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2369–2380, Nov. 2016, doi: 10.1109/TMI.2016.2546227.

[35] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105, Accessed: Oct. 25, 2018. [Online]. Available: http://code.google.com/p/cuda-convnet/

[36] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Sep. 2014, pp. 1–9, Accessed: Oct. 23, 2018. [Online]. Available: http://arxiv.org/abs/1409.4842

[37] T. Liu, S. Fang, Y. Zhao, P. Wang, and J. Zhang, "Implementation of training convolutional neural networks," 2015, *arXiv:1506.01195*. [Online]. Available: https://arxiv.org/abs/1506.01195

[38] W. G. Hatcher and W. Yu, "A survey of deep learning: Platforms, applications and emerging research trends," *IEEE Access*, vol. 6, pp. 24411–24432, 2018, doi: 10.1109/ACCESS.2018.2830661.

[39] C. Affonso, A. L. D. Rossi, F. H. A. Vieira, and A. C. P. D. L. F. de Carvalho, "Deep learning for biological image classification," *Expert Syst. Appl.*, vol. 85, pp. 114–122, Nov. 2017, doi: 10.1016/j.eswa.2017.05.039.

[40] D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1–19, doi: 10.1109/CVPR.2012.6248110.

[41] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," in *Proc. Int. Conf. Artif. Neural Netw.*, Cham, Switzerland: Springer, 2018, pp. 270–279, Accessed: Oct. 31, 2018. [Online]. Available: https://arxiv.org/pdf/1808.01974.pdf

[42] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/tmi.2016.2528162.

[43] Mathworks. (2016). *Introducing Machine Learning*. Accessed: Oct. 25, 2018. [Online]. Available: https://www.mathworks.com/content/dam/mathworks/tag-team/Objects/i/88174_92991v00_machine_learning_section1_ebook.pdf

[44] L. Torrey and J. Shavlik, "Transfer learning," *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods and Techniques*. Hershey, PA, USA: IGI Global, pp. 242–264, vol. 2010, doi: 10.1016/j.jbi.2011.04.009.

[45] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.

[46] J. Panchapakesan, F. Hourihan, and P. Mitchell, "Prevalence of pterygium and pinguecula: The blue mountains eye study," *Aust. N. Z. J. Ophthalmol.*, vol. 26, pp. 52–55, May 1998, doi: 10.1111/j.1442-9071.1998.tb01362.x.

[47] J. C. S. Yam and A. K. H. Kwok, "Ultraviolet light and ocular diseases," *Int. Ophthalmology*, vol. 34, no. 2, pp. 383–400, Apr. 2014, doi: 10.1007/s10792-013-9791-x.

[48] J. Khoo, S.-M. Saw, K. Banerjee, S.-E. Chia, and D. Tan, "Outdoor work and the risk of pterygia: A case-control study," *Int. Ophthalmol.*, vol. 22, pp. 293–298, Sep. 1998.

[49] D. J. Moran and F. C. Hollows, "Pterygium and ultraviolet radiation: A positive correlation.," *Brit. J. Ophthalmology*, vol. 68, no. 5, pp. 343–346, May 1984, doi: 10.1136/bjo.68.5.343.

[50] J. S. Paula, F. Thorn, and A. A. V. Cruz, "Prevalence of pterygium and cataract in indigenous populations of the Brazilian amazon rain forest," *Eye*, vol. 20, no. 5, pp. 533–536, May 2006, doi: 10.1038/sj.eye.6701917.

[51] W. Tarkowski, J. Moneta-Wielgoś, and D. Młocicki, "Do demodex mites play a role in pterygium development?" *Med. Hypotheses*, vol. 98, pp. 6–10, Jan. 2017, doi: 10.1016/j.mehy.2016.09.003.

[52] S. Young, "Sun and the eye: Prevention and detection of light-induced disease," *Clin. Dermatol.*, vol. 16, no. 4, pp. 477–485, Jul. 1998, doi: 10.1016/S0738-081X(98)00021-2.

[53] M. Coroneo, "Ultraviolet radiation and the anterior eye," *Eye Contact Lens, Sci. Clin. Pract.*, vol. 37, no. 4, pp. 214–224, Jul. 2011, doi: 10.1097/ICL.0b013e318223394e.

[54] E. Viso, F. Gude, and M. T. Rodríguez-Ares, "Prevalence of pinguecula and pterygium in a general population in Spain," *Eye*, vol. 25, no. 3, pp. 350–357, Mar. 2011, doi: 10.1038/eye.2010.204.

[55] Q. Le, J. Xiang, X. Cui, X. Zhou, and J. Xu, "Prevalence and associated factors of pinguecula in a rural population in shanghai, eastern China," *Ophthalmic Epidemiology*, vol. 22, no. 2, pp. 130–138, Mar. 2015, doi: 10.3109/09286586.2015.1012269.

[56] H. R. Taylor, "The biological effects of uv-B on the eye," *Photochem. Photobiol.*, vol. 50, no. 4, pp. 489–492, Oct. 1989, doi: 10.1111/j.1751-1097.1989.tb05553.x.

[57] S. Serrins, V. Mathur, and J. Serrins. *Pingueculae & Pterygia*. Accessed: Nov. 27, 2018. [Online]. Available: https://www.drserrins.com/eye-health/eye-conditions/pingueculae-pterygia/

[58] H. R. Taylor, "The long-term effects of visible light on the eye," *Arch. Ophthalmol.*, vol. 110, no. 1, pp. 99–104, Jan. 1992, doi: 10.1001/archopht.1992.01080130101035.

[59] Z. A. Hamid, U. K. Malaysia, Z. Harun, S. H. Lubis, N. Mohamed, I. Ishak, H. F. Othman, N. Z. Mohd Saat, J. Rafaai, M. R. M. Noor, and S. Z. Jamil, "Adoption of the mobile health screening programme for farming communities: A study among pesticide-exposed farmers from north east of peninsular Malaysia," *Jurnal Sains Kesihatan Malaysia*, vol. 12, no. 2, pp. 63–69, Dec. 2014. [Online]. Available: http://journalarticle.ukm.my/8541/1/Adoption_of_the_Mobile_Health_Screening_Programme.pdf

[60] K. Walsh, "UV radiation and the eye," *Vis. Care Inst. Johnson Johnson Med. Ltd.*, vol. 237, no. 6204, pp. 26–33, 2009, Accessed: Oct. 23, 2018. [Online]. Available: https://www.jnjvisioncare.co.uk/sites/default/files/public/uk/documents/tvci_uv_radiation_and_the_eye.pdf

[61] S. Jaffar, U. Dukht, and F. Rizvi, "Impact of pterygium size on corneal topography," *Rawal Med. J.*, vol. 34, no. 2, pp. 145–151, 2009, Accessed: Oct. 24, 2018. [Online]. Available: https://pdfs.semanticscholar.org/c3fe/1615b3d7935cd674eea400fd45c97e41770c.pdf

[62] J. D. Twelker, I. L. Bailey, M. J. Mannis, and W. A. Satariano, "Evaluating pterygium severity: A survey of corneal specialists," *Cornea*, vol. 19, no. 3, pp. 292–296, May 2000, doi: 10.1097/00003226-200005000-00007.

[63] J. D. Twelker and I. L. Bailey, "Clinical evaluation of pterygium," in *Pterygium*, H. R. Taylor, Ed. The Hague, The Netherlands: Kugler Publication, 2000, pp. 57–69.

[64] R. G. Mesquita and E. M. N. Figueiredo, "An algorithm for measuring pterygium's progress in already diagnosed eyes," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 733–736, doi: 10.1109/ICASSP.2012.6287988.

[65] X. Gao, D. W. K. Wong, A. W. Aryaputera, Y. Sun, C.-Y. Cheng, C. Cheung, and T. Y. Wong, "Automatic pterygium detection on cornea images to enhance computer-aided cortical cataract grading system," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 4434–4437, doi: 10.1109/EMBC.2012.6346950.

[66] M. A. Woodward, D. C. Musch, C. T. Hood, J. B. Greene, L. M. Niziol, V. S. E. Jeganathan, and P. P. Lee, "Tele-ophthalmic approach for detection of corneal diseases: Cccuracy and reliability HHS public access," *Cornea*, vol. 36, no. 10, pp. 1159–1165, 2017, doi: 10.1097/ICO.0000000000001294.

[67] Y. LeCun, "Learning invariant feature hierarchies," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer-Verlag, 2012, pp. 496–505, doi: 10.1088/0022-3727/19/12/002.

[68] D. Li, "A tutorial survey of architectures, algorithms, and applications for deep learning," *Trans. Signal Inf. Process.*, vol. 3, no. 2014, pp. 1–29, 2014, doi: 10.1017/atsip.2013.9.

[69] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[70] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1725–1732, doi: 10.1109/CVPR.2014.223.

[71] M. A. Zulkifley, S. R. Abdani, and N. H. Zulkifley, "Pterygium-Net: A deep learning approach to pterygium detection and localization," *Multimedia Tools Appl.*, vol. 78, no. 24, pp. 34563–34584, 2019.

[72] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent.*, 2014, pp. 1–14, doi: 10.2146/ajhp170251.

[73] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[74] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708, doi: 10.1109/CVPR.2017.243.

[75] H. Proença and L. A. Alexandre. (2004). *Ubiris Iris Image Database*. Accessed: Sep. 9, 2018. [Online]. Available: http://iris.di.ubi.pt/

[76] H. Proença and L. A. Alexandre, "UBIRIS: A noisy iris image database," in *Proc. Int. Conf. Image Anal. Process.*, 2005, pp. 970–977, doi: 10.1007/11553595_119.

[77] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.

[78] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, "Transfusion: Understanding transfer learning for medical imaging," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 3347–3357. [Online]. Available: http://arxiv.org/abs/1902.07208

[79] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets," in *Proc. Brit. Mach. Vis. Conf.*, 2014, pp. 1–11, doi: 10.5244/c.28.6.

[80] A. P. Carneiro, G. Nascimento, and J. Bradley, "Unregistered multiview mammogram analysis with pre-trained deep learning models," in *Proc. Int. Conf. Med. Image Comput. Comput. Intervent.* Cham, Switzerland: Springer, 2015, pp. 652–660, doi: 10.1007/978-3-319-24574-4.

[81] G. Varoquaux, "Cross-validation failure: Small sample sizes lead to large error bars," *NeuroImage*, vol. 180, pp. 68–77, Oct. 2018, doi: 10.1016/j.neuroimage.2017.06.061.

[82] Jayadeva, R. Khemchandani, and S. Chandra, "Twin support vector machines for pattern classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 5, pp. 905–910, 2007, doi: 10.1109/TSMCB.2006.887427.

[83] J. Y. Choi, T. K. Yoo, J. G. Seo, J. Kwak, T. T. Um, and T. H. Rim, "Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database," *PLoS ONE*, vol. 12, no. 11, pp. 1–16, 2017, doi: 10.1371/journal.pone.0187336.

N. SYAHIRA M. ZAMANI received the B.Eng. degree in computer and communication engineering from Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia, in 2017, where she is currently pursuing the M.S. degree in system engineering. She is also working as a Graduate Research Assistant with the Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, UKM. Her research interests include image processing and artificial intelligence systems.

WAN MIMI DIYANA WAN ZAKI received the bachelor's degree in electronics engineering, the master's degree (Engr. Sc.), and the Ph.D. degree from Multimedia University (MMU), Cyberjaya, Malaysia, in 2000, 2005, and 2013, respectively.

She is currently an Associate Professor with the Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia, which she joined in 2008. Her research specialization is in biomedical engineering and bio-informatics. Her research interests include intelligent systems, image processing, and the IoT related healthcare technology.

AQILAH BASERI HUDDIN (Member, IEEE) received the B.Eng. degree (Hons.) in electrical and electronics engineering and the Ph.D. degree in electrical and electronics engineering from The University of Adelaide, Australia, in 2007 and 2015, respectively.

She is currently a Senior Lecturer with the Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia. Her research interests include image processing and artificial intelligence.

AINI HUSSAIN received the B.Sc. degree in electrical engineering from Louisiana State University, USA, the M.Sc. degree in systems and control from the University of Manchester Institute of Science and Technology (UMIST), U.K., and the Ph.D. degree in electrical and electronics engineering from Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia. She is currently a Professor with the Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, UKM. Her research interests include AI, intelligent systems, machine learning, and signal and image processing. In the big data era, she is also looking into research involving predictive analytics, cloud computing, IoT applications, and so on.

HALIZA ABDUL MUTALIB received the bachelor's degree (Hons.) in optometry and the M.S. degree from Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia, in 1993 and 1998, respectively, and the Ph.D. degree from UMIST, Manchester, U.K., in 2000.

She is currently an Associate Professor of optometry with Universiti Kebangsaan Malaysia. She specializes in contact lenses and corneal morphology with extensive research. She has authored a few books and also a few chapters in books. She is also actively involved in community work and is the coordinator of the UKM Mobile Optometry Clinic.

AZIAH ALI (Member, IEEE) is currently a Lecturer with the Faculty of Computing & Informatics, Multimedia University, Malaysia. Her research interests include image and signal processing focusing on biomedical field. More specifically, her work examines how image and signal processing techniques could contribute towards better and more efficient ways of analysing and understanding medical images and signals. She has worked on different medical image modalities, including ultrasound, x-ray, and retinal images in order to develop more efficient and accurate ways of extracting useful health indicators from the images.