IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 8, 2020, accepted June 23, 2020, date of publication June 26, 2020, date of current version July 7, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3005152

Hyperparameter Tuning Deep Learning for Diabetic Retinopathy Fundus Image Classification

K. SHANKAR[1], YIZHUO ZHANG[2], YIWEI LIU[2], LING WU[2], AND CHI-HUA CHEN[2], (Senior Member, IEEE)

[1]Department of Computer Applications, Alagappa University, Karaikudi 630003, India
[2]College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China

Corresponding authors: Ling Wu (wuling1985@fzu.edu.cn) and Chi-Hua Chen (chihua0826@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61906043, Grant 61877010, Grant 11501114, and Grant 11901100, in part by the Science Foundation of the Fujian Province, China, under Grant 2019J01243, in part by the Funds of Education Department of Fujian Province under Grant JAT190026, and in part by the Fuzhou University under Grant 510872/GXRC-20016, Grant 510730/XRC-18075, Grant 510809/GXRC-19037, Grant 510649/XRC-18049, and Grant 510650/XRC-18050. The work of K. Shankar was supported by RUSA—Phase 2.0 Grant Sanctioned Vide Letter, Policy (TNMulti-Gen), Department of Education, Government of India, in 2018, under Grant F.24-51/2014-U.

ABSTRACT Diabetic retinopathy (DR) is a major reason for the increased visual loss globally, and it became an important cause of visual impairment among people in 25-74 years of age. The DR significantly affects the economic status in society, particularly in healthcare systems. When timely treatment is provided to the DR patients, approximately 90% of patients can be saved from visual loss. Therefore, it becomes highly essential to classify the stages and severity of DR for the recommendation of required treatments. In this view, this paper introduces a new automated Hyperparameter Tuning Inception-v4 (HPTI-v4) model for the detection and classification of DR from color fundus images. At the preprocessing stage, the contrast level of the fundus image will be improved by the use of contrast limited adaptive histogram equalization (CLAHE) model. Then, the segmentation of the preprocessed image takes place utilizing a histogram-based segmentation model. Afterward, the HPTI-v4 model is applied to extract the required features from the segmented image and it subsequently undergoes classification by the use of a multilayer perceptron (MLP). A series of experiments take place on MESSIDOR (Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology) DR dataset to guarantee the goodness of the HPTI-v4 approach and the obtained results clearly exhibited the supremacy of the HPTI-v4 model over the compared methods in a significant way.

INDEX TERMS Diabetic retinopathy, image classification, hyperparameter, deep learning.

I. INTRODUCTION

In past decades, diabetes is caused because of the excess growth of glucose in the blood. If the same condition is retained for a long time, it results in severe blood vessel damage. A person affected with diabetes is vulnerable to kidney failure, loss of eyesight, bleeding teeth, lower limb confiscation, nerve failures, and so on. It also leads to a heart attack as well as stroke in diabetic affected individuals. The nephrons present in the kidney are damaged and leads to diabetic neuropathy while neurons present in the brain get damaged, and cause diabetic retinopathy (DR) which results in the retinal infection. According to the survey of the World Health Organization (WHO) has been detected that diabetes would be in the 7th place which comes under fatal disease [1]. Furthermore, the additional information states that around 108 million diabetic patients were in an earlier period which has been increased to 422 million respectively. Based on the statistics, the diabetic affected persons have been maximized above 18 years of age from 4.7% to 8.5% is obtained. Some of the poor people have a greater option to be affected by diabetes. In India, people around 61.3 million between the age of 25–75 are found to be diabetic individuals. Alternatively, it states that they might be raised to 101.2 million by 2030 [2]. If there is a frequent improvement in blood glucose level which affects the retina prior. The maximum enhancement in glucose level greatly influences the blood vessels that cause oozing of blood from eyes and weakens the human visual

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

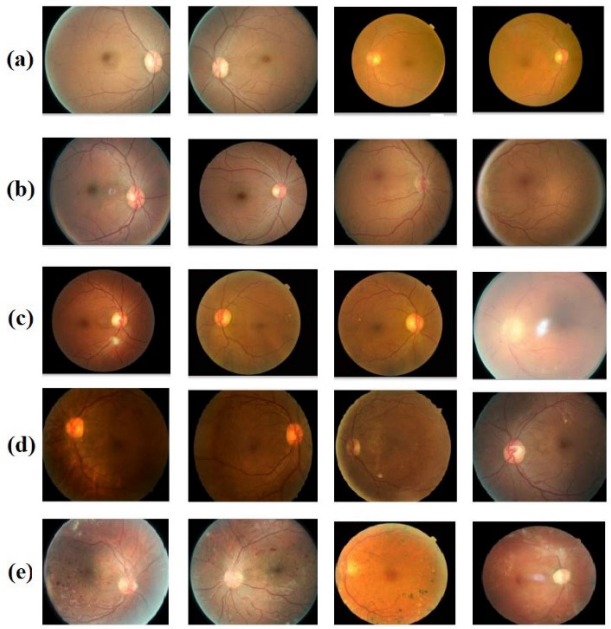

FIGURE 1. Stages involved in DR: (a) normal; (b) mild; (c) moderate; (d) severe; (e) PDR.

system. However, humans are inherited with the capability of curing the disease. If the brain detects the blood leakage, it tends to stimulate the adjacent tissues to manage the current status. As a result, it leads to the random development of new blood vessels [3], but the formed cells are anemic in nature. At the initial stage, the vision gets affected slowly. Consequently, it is compulsory for diabetic patients must undertake a systematic eye checkup where the retina should be observed by an ophthalmologist. There are various techniques for detecting the affected eye, some of them are slit-lamp biomicroscopy, optical coherence tomography (OCT), fundus fluorescein angiography (FFA), and fundus images.

The intensity of disease would be identified from the abnormal size of any part that is affected. There are a few forecasting models, namely venous beading, microaneurysms, hemorrhage, etc. are considered to be significant principles applied. Here, microaneurysms denote the size of a blood clot which is 100–120μm with a circular shape. Hemorrhage is caused due to the enormous amount of blood leakage from the affected blood vessel. Furthermore, the uneven enlargement of blood vessels is called neovascularization. DR is classified into non-proliferative DR (NPDR) and Proliferative DR (PDR). Hence, a DR sample implies different levels are shown in Figure 1.

DR plays a significant role in preventing vision loss only when the disease is predicted in the earlier stage. If the anomaly is detected, the patient must undergo regular treatment every 6 months to examine the disease [4]. To overcome the above short comes, a new method has been designed to predict and classify fundus images, which is helpful in degrading the problem of vision loss due to the contribution of DR. Diverse works have provided different models to obtain exact DR identification. While performing the identification and classification of DR takes place under the segmentation of fundus images as small parts to find the presence of exudates, lesions, microaneurysms, etc.

A method to compute different features of abnormal foveal zone, as well as microaneurysms, has been presented in [4]. It often applies curvelet coefficients which are attained from fundus images and angiograms. There is a three-stage classification technique applied where the experiment has been performed, 70 diabetic patients. Therefore, the projected method reaches an optimal sensitivity value of 100%. A DR image classification approach has been presented in [5] which is processed according to the previous microaneurysms. Furthermore, circularity as well as region of microaneurysms has been used to perform the feature extraction. Some of the datasets like DIARETDB1, DRIVE, and ROC have been applied in this research work. Furthermore, the presented method attains a better classification in terms of maximum sensitivity and specificity than alternate techniques. It is also employed with the principal component analysis (PCA) model to divide the images of optic disc (OD) from fundus photography. Under the application of enhanced minimum distance discriminant (MDD) classifier technique, it attains a maximum classification result including optimal detecting value. Subsequently, diverse images are employed to perform the above process along with a classification process which denotes the normal and abnormal images.

Cunha-Vaz [6] measured the practical result of support vector machine (SVM) from three standard datasets as well as to attain maximum accuracy rate. Simultaneously, texture attributes have been derived from the local binary pattern (LBP) [7] to predict the presence of exudates to attain the best classification process. The alternate dual classification process, designed by Omar *et al.* [8], which has a bootstrapped decision tree (DT) is applied for fundus image segmentation.

Here, Gabor filtering and SVM classifier has been projected in [9] to perform the DR classification. Before the application of classification technique, circular Hough transform (CHT) and contrast limited adaptive histogram equalization (CLAHE) methods have been generated for input images which help to reach the optimal detecting value in the STARE dataset. Some of the morphological processes use the intensity level of image as threshold values for segmenting images [10]. Bhatkar and Kharat [11] mentioned CNN features with data augmentation technology in classifying DR images. Therefore, DR images undergo 5 phases that are tested using the Kaggle dataset.

Partovi *et al.* [12] established an error based independent system that helped to divide images. Deep CNN (DCNN) offers advanced feature extracting method to classify medicinal images.

Followed by, Xu *et al.* [13] utilized DCNN to reduce the human inventions and attempts in maximum feature representation of histopathological colon for classifying cancer images and other images. Shen *et al.* [14] developed a technique of multi-crop pooling which is applied in DCNN to

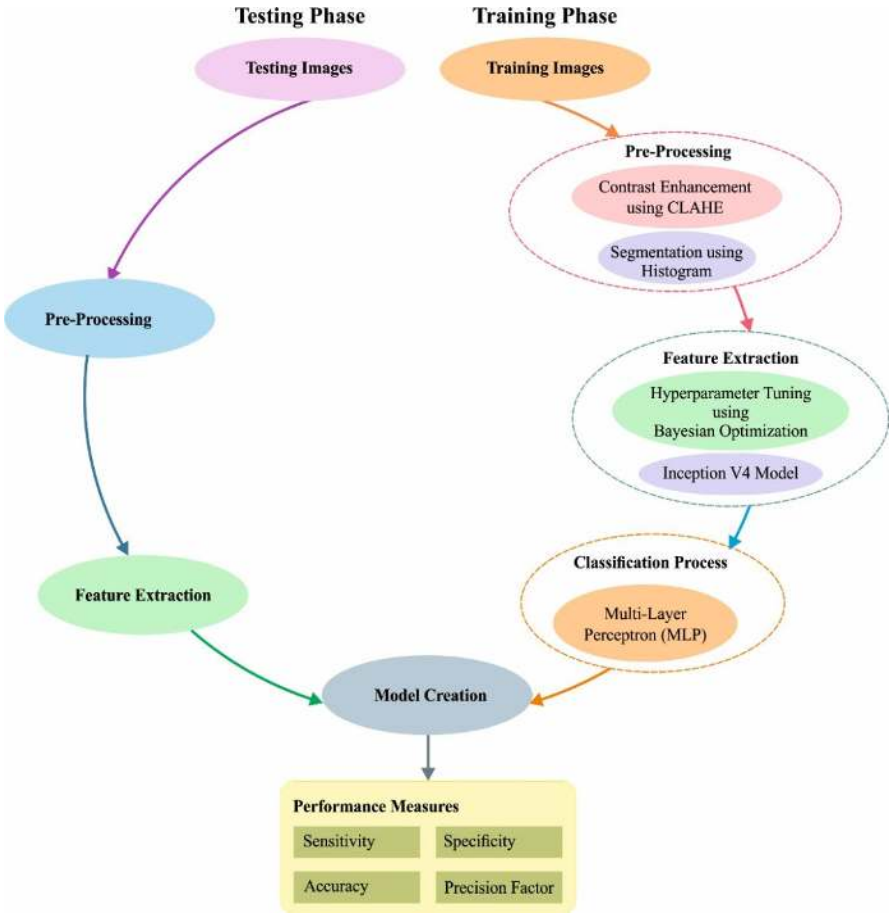

FIGURE 2. Overall process involved in the HPTI-v4 model.

capture object salient information for categorizing the lung cancer with the consumption of CT photographs. In Shanthi and Sabeenian [15], a new DR classification model using modified AlexNet architecture has been presented. The presented model using CNN with appropriate pooling, softmax, and ReLU layers to attain maximum detection performance. The performance of the model is assessed on Messidor dataset. The proposed Modified AlexNet architecture has reached to a maximum average accuracy of 96.25%. In Shankar *et al.* [16], an effective DR classification model has been presented by the use of Inception-ResNet v2 model-based feature extraction and deep neural network with moth search optimization (DNN-MSO) algorithm. The presented model is validated using Messidor dataset and the results exhibited better results with the maximum accuracy, sensitivity, and specificity of 99.12%, 97.91%, and 99.47%

Though various DL models are available for DR classification, the hyperparameter tuning process of DL models are not extensively addressed. The hyperparameter tuning technique helps to properly select the parameter values and leads to better classification performance. In this view, this paper introduces a new automated Hyperparameter Tuning Inception-v4 (HPTI-v4) model for the prediction and classification of DR from color fundus images. The presented HPTI-v4 model is composed of diverse sub-processes like preprocessing, segmentation, feature extraction, and classification. Here, the segmentation process takes place by histogram-based segmentation, and the inception-v4-based feature extraction process takes place. For tuning the hyperparameters in inception-v4, the Bayesian optimization technique is involved. Finally, the classification process takes place under the application of multilayer perceptron (MLP).

The structure of this study includes four sections. The presented HPTI-v4 model is neatly discussed in Section II. The validation of the HPTI-v4 model takes place in Section III, and the contributions of this study are concluded in Section IV.

II. THE PROPOSED HPTI-V4 MODEL FOR DR CLASSIFICATION

The overall working principle of the HPTI-v4 model is shown in Figure 2. As shown, the input image will be preprocessed to improve the contrast level by the use of CLAHE model. Then, the segmentation of the preprocessed image takes place by utilizing the histogram-based segmentation model. After-

ward, the HPTI-v4 model is used for feature extraction. The tuning of hyperparameter is carried out by the use of the Bayesian optimization technique. At last, the MLP classifier model is applied for the classification of diverse DR levels in color fundus pictures.

A. PREPROCESSING: CONTRAST ENHANCEMENT

In order to achieve an accurate segmentation and DR classification of images, the CLAHE method is applied to improve image contrast. The extended version of CLAHE is helpful in avoiding unwanted noise amplification. Furthermore, various histograms intensities are computed under the application of CLAHE, which every node relies on standard image area as well as the histogram is distributed to prevent additional amplification as well as re-mapping the intensity measures which is carried out with the help of shared histogram. When it comes to medical images, the currently developed CLAHE methods are deployed to improvise the image contrast.

To reveal the efficiency of CLAHE approach, few steps have been provided as follows.

- Producing entire input: An image, count of parts present in all rows and columns directions, bin number of histograms are employed to develop image transformation and clip boundary application to limit the contrast from 0 to 1.
- Input pre-processing: In ordinary values, find a clip boundary to add the image at the primary stage before processing the image partition.
- Determining every background area that tends to create mapping to grayscale: Producing whole image region, deploy a histogram region under the application of specified bin count, histogram clip applies a clip boundary as well as offers an optimal mapping.
- For collecting final CLAHE image, arbitrate gray level mapping: Extraction of cluster with 4 adaptive mapping operations, compute image regions which result in two-third of overlapping for all mapping tiles, then extract all pixels, applies 4 mappings to pixel, and interpolates among the outcome to accomplish a maximum efficiency from the whole image.

B. HISTOGRAM-BASED SEGMENTATION PROCESS

Here, segmentation is said to be the division of images into the group of non-overlapping areas respectively. The segmentation model mainly focuses on dividing the regions with objects of interest that are adopted from the portion of image concerning background. The above process could be performed with the application of similarity as well as homogeneity. In addition, the thresholding mechanism is simple and satisfying which helps to process a specific constraint of foreground against the background segmenting process. Some of thresholding methods are simple, perceptive, and based on various intensities of diverse images. From the thresholding mechanism, the collection of better thresholds can be obtained which is relied on many useful regions that help to discover the image. Several thresholding approaches have been established with a complete study that is provided by [16]. By identifying optimal thresholds, an effective segmentation process could be attained where other models are based on histogram computation. Hence, the histogram of an image denotes the comparative frequency of various color occurrence. A digital image exists with a value of L colors, and then the histogram would be a discrete function as Equation (1).

$$h(g_k) = \frac{n_k}{N}, \quad k = 0, 1, \ldots, L-1. \tag{1}$$

The number of pixels n_k includes color g_k which is assumed to be the fraction of entire pixels N. Images with bi-modal histograms, a global threshold T has been applied for portioning the image histogram. Therefore, the distinct condition of having objects as well as background is denoted in image histogram through a deep and narrow valley. Thus, threshold value could be selected from the end of valley. The other technique for computing the threshold is easy and simple which does not acquire more information about the image which is strong enough in contrast to image noise. Furthermore, it can be adaptive for iterations under the application of mean intensity values that is obtained from segmented images, yet to attain some modifications in a successive process that is small.

Generally, the global thresholding is processed for the peaks in image histogram, with respect to objects of interest as well as to background has been divided. Thus, it is proven that the adapted technique is not applicable for images that are strongly illuminated and the valleys of histogram are said to be lengthy as well as broad with different measures of heights. It enables the thresholding process for an image including a histogram that lacks in original peaks. Furthermore, the image is divided into sub-images under the help of specified technique as well as the application of mean value for sub-image. When a definite feature of an image is obtained then, the process is named as priori, where the task of selecting threshold would be simple and the main purpose of selecting threshold is to assure the satisfaction in the process. A new technique termed as P-tile applies these kinds of data like objects may be dark when compared with other backgrounds and consumes a definite fraction i.e. percentile ($1/p$) is the overall image area. The threshold can be identified by examining the intensity level where the target fraction of image pixels is lower than the given value. Hence, the intensity level could be discovered with the help of cumulative histogram by Equation (2).

$$c(g_k) = \sum_{i=0}^{k} g_i. \tag{2}$$

The value of threshold T is fixed as $c(T) = 1/p$. In case of dark background, it might be set as $c(T) = 1 - 1/p$.

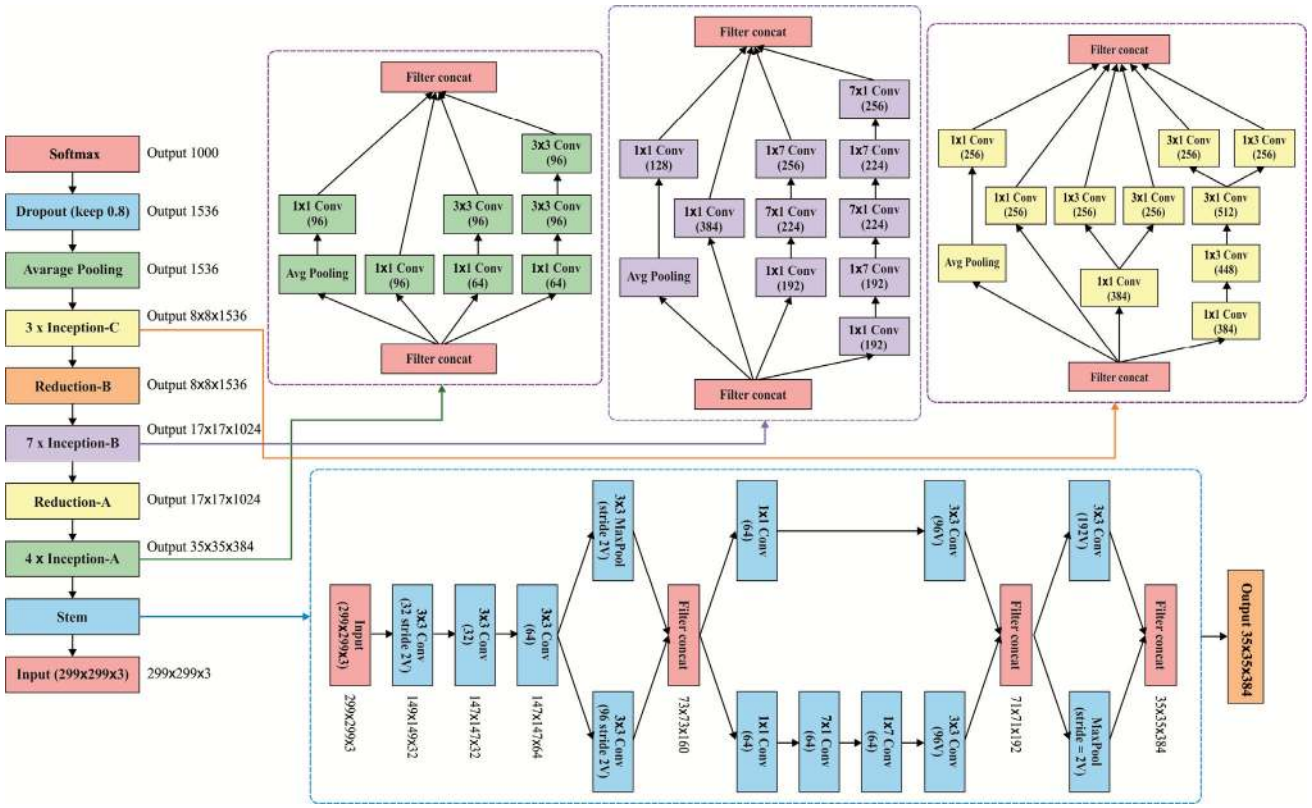

FIGURE 3. The architecture of inception-v4 model.

C. HYPERPARAMETER OPTIMIZATION OF INCEPTION-V4 USING THE BAYESIAN OPTIMIZATION FOR FEATURE EXTRACTION

The basic model of inceptions is applied in training different portions such that all monotonous blocks are segmented as several sub-networks which is applicable to allow the complete method in memory. Therefore, the inception method could be tuned in an easy manner that denotes the probability of altering number of filters from unique layers that do not affects the quality of complete trained network. In order to enhance the speed of training process, the size of layer should be tuned in a proper way to reach the tradeoff among different sub-networks. In contrast, by applying TensorFlow, advanced inception methods are designed without repeated partitioning. This is because of using smart memory optimizing for backpropagation (BP) which is attained by enabling tensors that are essential in calculating gradient as well as determining reduced numbers. Furthermore, inception-v4 has been proposed to eliminate the unnecessary operation which makes similar methods for Inception blocks in all grid sizes [17], [18]. The whole architecture of inception-v4 technique is shown in Figure 3.

1) RESIDUAL INCEPTION BLOCKS

Here, inception blocks are used by the filter-expansion layer which is used to improve the dimensionality of filter bank, before totaling the depth of input. This process is essential in replacing the dimensionality cutback tempted through the inception block. There are diverse types of inception, inception-v4 is slow, due to the presence of massive layers. The extra modification between residual and non-residual. It is named as batch-normalization that is applied to conventional layers. However, the model of batch-normalization in TensorFlow takes maximum storage space as well as it is essential to reduce the total number of layers if batch normalization is to be used in any place.

2) SCALING OF THE RESIDUALS

In this study, when a filter count is more than 1000, residual models show that it is unstable and the network is terminated at the initial stage of training which signifies the target layer before the pooling layer starts to produce zeros from various iterations. However, it could not be eliminated by minimizing the training measure.

In addition, the reduced values are appended before the activation layer is identified and stable in learning. In general, some of scaling factors are within the range of 0.1 to 0.3 that is used in scaling the accumulated layer activations.

3) PARAMETER TUNING

Hyperparameters can be tuned to control the performance of the applied techniques. The process of the selection of hyperparameters is the major characteristic of the DL models which nearly explain particular hyperparameters that handles various aspects like memory and computation complexity.

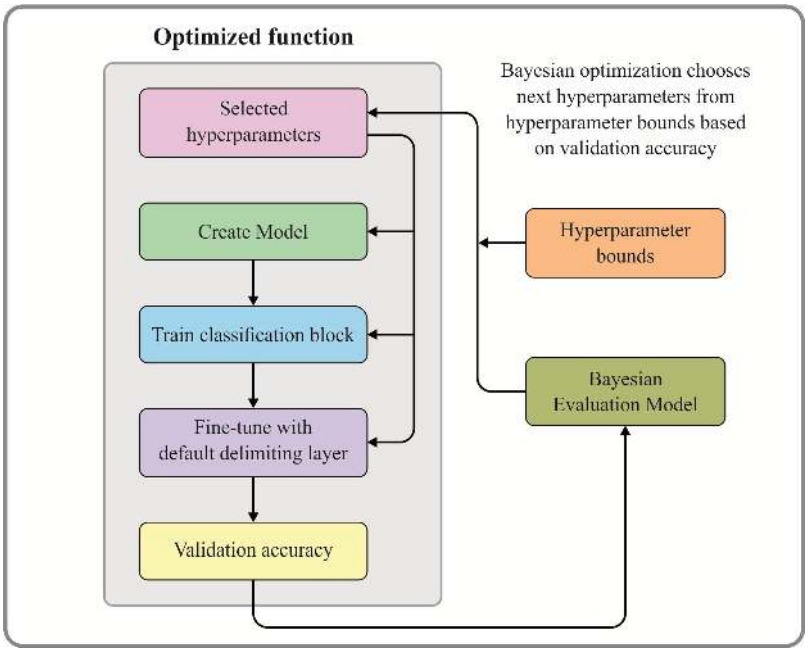

FIGURE 4. Bayesian optimization process.

Therefore, the additional hyperparameters are represented to settle in a definite technique on a particular scenario. Generally, more time is needed to tune the hyperparameters for obtaining effective results.

Hyperparameter optimization can be defined as follows:

$$x^* = arg \min_{x \in X} f(x) \tag{3}$$

where f(x) denotes the objective score to reduce the error assessed on the validation set, x∗ is the collection of hyperparameters which produces minimum score value, and x could take on any value in the domain X. In other words, the aim is to determine the model hyperparameters which produce high score on the validation set metric. There are two ways exist to select them such as manual and automatic. All these methods are officially feasible and the decision normally defines a tradeoff among the deep interpretation of MLP method needed for choosing hyperparameters in a manual way versus the maximum computation cost needed by automated selection models. Here, the hyperparameters in the inception-v4 model are tuned by the use of Bayesian optimization model (shown in Figure 4) [19].

The Bayesian optimization model analyzes the previous validation outcome in which it utilizes to create a probabilistic model, which will map the hypermeters to a probability of a score. This technique is termed as a "surrogate" for objective function $p(b|a)$. The surrogate can be easily optimized compared to the objective function.

Bayesian optimization is a well-known and simple approach that is mainly applied for sequential optimization of unlabeled objective functions and defined as, $x^* = argmax_{x \in X} f(x)$, where $X \subseteq R^D$. Unfortunately, the function estimations could be filled with artefacts with $y = f(x) + \epsilon$ where $\epsilon \sim N(0, \sigma^2)$. Gaussian process (GP) (Williams and Rasmussen 2006) is a familiar choice to specify the smoothly varying functions. Assume x is the D dimensional observation and X as matrix with p whereas y denotes the parallel function values in these observations. With mo loss of generality, the prior mean function is assigned as zero function which makes the Gaussian process completely defined by covariance function as: $p(f|X) = N(f|0, K)$ where f represents the corresponding latent parameters and K refers to the kernel matrix with $K_{i,j} = k(x_i, x_j)$. Here, it selects the well-known squared exponential kernel as the choice of covariance function.

The above-defined model is defined as, $k(x, x') = exp(-\frac{1}{2\theta} \|x - x'\|^2)$ where θ signifies the length-scale parameter of a kernel. Under the application of GP features, incomplete derivatives of a GP is still Gaussian as the differentiation is said to be linear process. It enables to enclose the derivative of unknown objective function by applying the provided criteria for covariance functions as:

$$k\left[\frac{\partial f^{(i)}}{\partial x_d^{(i)}}, f^{(j)}\right] = \frac{\partial}{\partial x_d^{(i)}} k\left[f^{(i)}, f^{(j)}\right] \tag{4}$$

$$k\left[\frac{\partial f^{(i)}}{\partial x_d^{(i)}}, \frac{\partial f^{(j)}}{\partial x_g^{(j)}}\right] = \frac{\partial^2}{\partial x_d^{(i)} \partial x_g^{(j)}} k\left[f^{(i)}, f^{(j)}\right] \tag{5}$$

where $d, g \in [1, ..., D]$. The newly developed model applies the acquisition function which helps in searching optimal objective function. Here, Expected Improvement (EI) has been employed for acquiring the advantages of the given approach. But, it is notified that is capable of working with

any kind of acquisition function. EI is depicted as:

$$\alpha(x_p+1) = (\mu(x_p+1)-f(x^+))\Phi(z)+\sigma(x_p+1)\phi(z) \quad (6)$$

where

$$z = (\mu(x_p+1) - f(x^+))/\sigma\left(x_p+1\right) \quad (7)$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are CDF as well as PDF of a standard normal distribution. The Bayesian optimization model is effective because it selects the subsequent hyperparameters in an informed way. The core concept is to allocate slightly higher time to choose the next hyperparameters for making less number of objective function. The efficiency of this approach can be identified only with minimum iterations. The primary commitment of HPTI-v4 model is used for the classification of the provided input image into different classes of DR.

D. MLP-BASED CLASSIFICATION

Here, MLP is considered as a feed-forward artificial neural network (ANN) technique that is mapped with the collection of input data to set of output data. Therefore, it is assumed as logistic regression (LR) classifier, and the conversion of primary input is carried out with the help of non-linear transformation. Any MLP along with a single hidden layer is a discriminant of non-linear transformation.

$$y = f_\theta(x) = \sigma\left(b + \mathbf{B} \times \sigma\left(a + \mathbf{A} \times x\right)\right). \quad (8)$$

where the parameter θ indicates two vectors a and b; and $\mathbf{A}$ and $\mathbf{B}$ indicate two matrices, and $\sigma(\cdot)$ is the element-wise sigmoid function of a vector. These parameters undergo optimization using stochastic gradient descent (SGD), minimizing few objective functions involving y and the ground truth classification information. Some of the reducing functions are composed of y where element becomes optimized for the whole network along with SGD as well as original data classification. The function $c(\theta)$ is an objective function. Hence, an upcoming effective technique is employed to attain maximum optimization. Usually, the function f_θ has several local minima. In order to train a little ANN, the local minima should be analyzed consecutively with maximum problems like XOR, spiral as well as parity. Due to the existence of f_θ symmetry, an individual local minimum introduces higher local minima from the parameter vector. In various cases, $c(\theta)$ often has a lower hierarchy. To train the ANN model, many techniques have been attained by the above study. For lower scales, the momentum is embedded from local minima, and programs for learning measures are arranged with the application of minima hierarchy in various scales of same method to repeated annealing.

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

A series of experimental analysis takes place to verify the performance of the proposed method. The optimal set of hyperparameters chosen by Bayesian optimization is epoch count: 500, learning rate: 0.001, and momentum: 0.9.

TABLE 1. Dataset details.

DR stages	Details	Labels
Healthy	Zero abnormalitoes	Normal
Mild NPDR	Microaneurysms	Stage_1
Moderate NPDR	Few microaneurysms	Stage_2
Severe NPDR	Venous beading + intraretinal microvascular abnormality (IRMA)	
PDR	Vitreous/ Pre-retinal hemorrhage	Stage_3

A. DATASET

To compute the performance of the presented HPTI-v4 model, a benchmark MESSIDOR dataset is applied [20] which has color fundus images with appropriate annotation. The images present in the dataset can undergo classification into a set of four stages as shown in Table 1. The 1200 eye fundus color images of the posterior pole of the Messidor database were gathered by 3 ophthalmologic departments through a color video 3CCD camera placed on the Topcon TRC NW6 non-mydriatic retinograph with a 45° field of view. The images are taken by the use of 8 bits per color plane at 1440∗960, 2240∗1488 or 2304∗1536 pixels. The images which have some microaneurysms belong to Stage 1. The image in which both microaneurysms and hemorrhages belong to Stage 2 and the images involve high microaneurysms and hemorrhages belong to Stage 3. For experimentation, 10 fold cross-validation process is employed to divide the dataset into training and testing sets.

B. PERFORMANCE MEASURES

The set of 4 estimation parameters were applied to evaluate the function of HPTI-v4 models in terms of sensitivity, specificity, accuracy, and precision factor. The given equations are provided to compute the above-mentioned factors as provided in Equations (9)-(12). The sensitivity is measured according to true-positive (T_P) and false-negative (F_N)(shown in Equation (9)). The specificity is measured according to true-negative (T_N) and false-positive (F_P)(shown in Equation (10)). The accuracy is measured according to true-positive (T_P), false-positive (F_P), true-negative (T_N), and false-negative (F_N)(shown in Equation (11)). The precision factor is measured according to true-positive (T_P) and false-positive (F_P)(shown in Equation (12)).

$$Sen = \frac{T_P}{T_P + F_N}. \quad (9)$$

$$Spe = \frac{T_N}{T_N + F_P}. \quad (10)$$

$$Acc = \frac{T_P + T_N}{T_P + F_P + T_N + F_N}. \quad (11)$$

$$Pre = \frac{T_P}{T_P + F_P}. \quad (12)$$

C. RESULT ANALYSES

A sample visualization results attained by the HPTI-v4 model is shown in Figure 5. As shown, Figure 5(a) shows the sample input test color fundus images which undergo effective

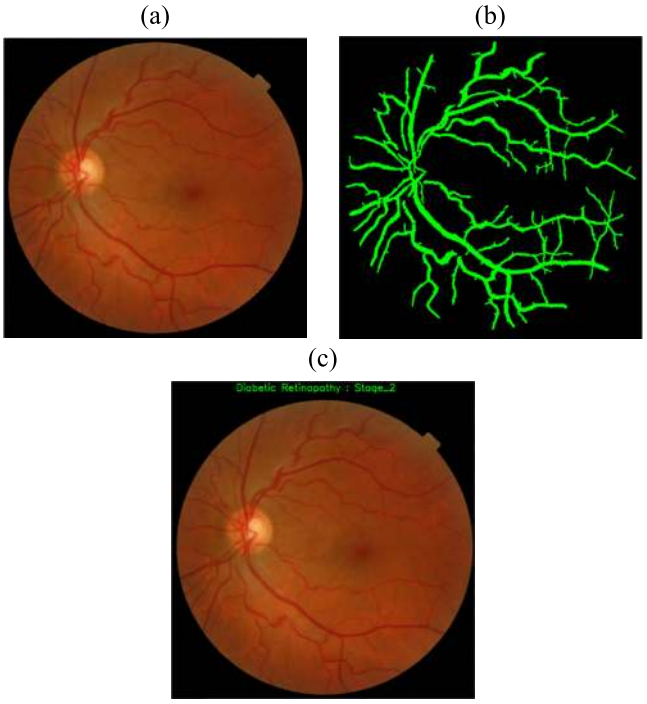

FIGURE 5. (a) Original image; (b) segmented image; (c) classified image.

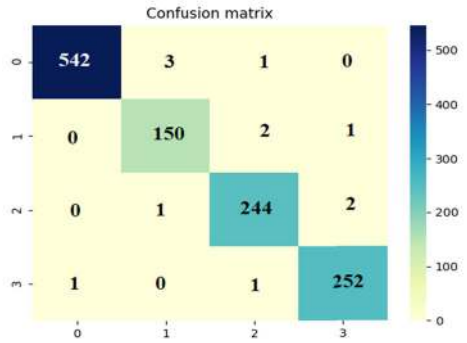

FIGURE 6. Generated confusion matrix for DR.

segmentation and classification model. Figure 5(b) and Figure 6(c) show the segmented and classified output of the used test image.

Figure 6 visualizes the confusion matrix generated by HPTI-v4 model in the classification of the different stages of DR images. It is transformed into an understandable form in Table 2. The values in Table 2 indicated that a total of 542 images were classified into the class label 'Normal', 150 images were classified into the class label 'Stage_1', 244 images were classified into the class label 'Stage_2' and 252 images were classified into the class label 'Stage_3'.

Table 3 provide the confusion matrix of the HPTI-v4 model in terms of T_P, T_N, F_P, and F_N for every stage of DR. It is the preliminary process for the determination of the classification performance of the presented HPTI-v4 model. The figure shows that the values of T_P and T_N are significantly higher compared to F_P and F_N. These values itself

TABLE 2. Confusion matrix.

Input Label	Different Level of DR				Total No. of Images
	Normal	Stage_1	Stage_2	Stage_3	
Normal	542	3	1	0	546
Stage_1	0	150	2	1	153
Stage_2	0	1	244	2	247
Stage_3	1	0	1	252	254
Total No. of Images	543	154	248	255	1200

TABLE 3. Manipulation from confusion matrix.

DR Level	Normal	Stage_1	Stage_2	Stage_3
TP	542	150	244	252
TN	646	1038	944	936
FP	1	4	4	3
FN	4	3	3	2

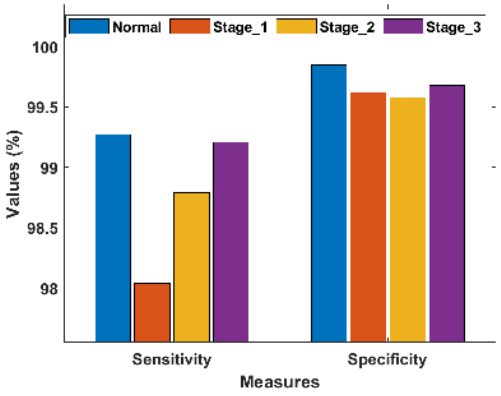

FIGURE 7. Sensitivity and specificity analyses of the HPTI-v4 model.

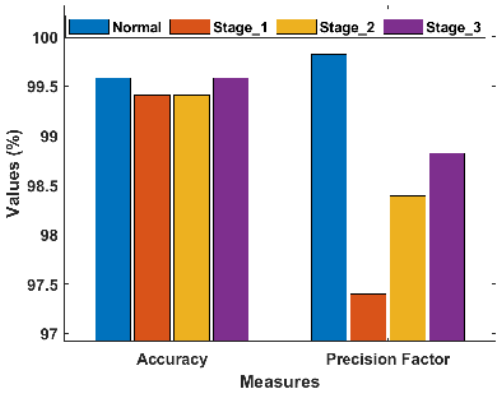

FIGURE 8. Accuracy and precision factor analyses of the HPTI-v4 model.

clearly portrayed the effective classification outcome of the HPTI-v4 model.

Figure 7 and Figure 8 show the classifier results of the HPTI-v4 technique in terms of accuracy, sensitivity, specificity, and precision factor. The figure depicted that the DR images under the grade 'Normal' are effectively classified with the sensitivity, specificity, accuracy, and precision factor of 99.27%, 99.85%, 99.58%, and 99.82% respectively. Similarly, the DR images under the grade 'Stage_1' are effectively classified with the sensitivity, specificity, accuracy, and precision factor of 98.04%, 99.62%, 99.42%, and 97.40% respectively. Likewise, the DR images under the grade 'Stage_2'

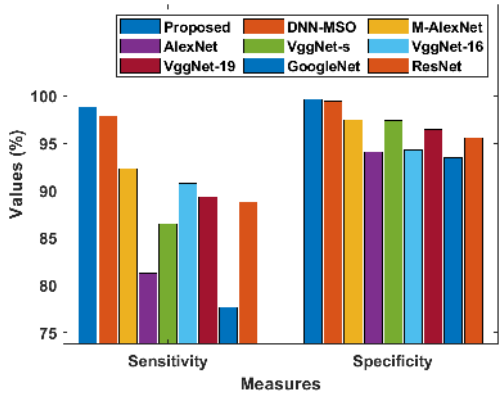

FIGURE 9. Comparison of HPTI-v4 model with diverse methods in terms of sensitivity and specificity.

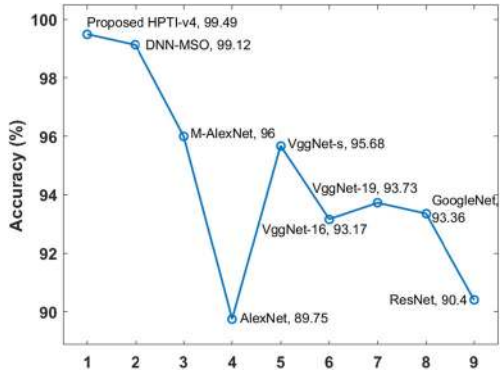

FIGURE 10. Comparison of HPTI-v4 model with diverse methods in terms of accuracy.

are effectively classified with the sensitivity, specificity, accuracy, and precision factor of 98.79%, 99.58%, 99.42%, and 98.39% respectively. At last, the DR images under the grade 'Stage_3' are effectively classified with the sensitivity, specificity, accuracy, and precision factor of 99.21%, 99.68%, 99.58%, and 98.82% respectively.

Figure 9 and Figure 10 offer a detailed comparative analysis of different models [15], [16] in terms of accuracy, sensitivity, and specificity. On measuring the classifier outcome in terms of accuracy, it is noticeable that the AlexNet technique attains an inferior classifier outcome with minimal accuracy of 89.75%. At the same time, the ResNet model offers slightly better classification with the moderate accuracy of 90.40%. On continuing with, the VggNet-16, VggNet-19, and GoogleNet models exhibit manageable and near identical results with the accuracy values of 93.17%, 93.73%, and 93.36% correspondingly. In line with, the VGGNet-s model shows better results of those previous models with an accuracy value of 95.68%. Then, the modified AlexNet (M-AlexNet) and deep neural network with moth swarm optimization (DNN-MSO) models show competitive results and attain higher accuracy values of 96.00% and 99.12%. However, the HPTI-v4 model showcased extraordinary classification performance and obtained the highest accuracy of 99.49%.

During the classification performance assessment in terms of sensitivity, it is reported that the GoogleNet model achieves worse classification with the minimal sensitivity of 77.66%. Together, the AlexNet model has resulted in a moderate sensitivity of 81.27%. On continuing with, the VggNet-s, VggNet-16, and ResNet models exhibit controllable and closer results with the sensitivity values of 86.47%, 86.47%, and 88.78% correspondingly. In line with, the VggNet-16 technique shows improved results over previous models with a sensitivity value of 90.78%. Afterward, M-AlexNet and DNN-MSO models display competitive results and attain higher sensitivities of 92.35% and 97.91%. Nevertheless, the HPTI-v4 model showcased extraordinary classification performance and obtained the highest sensitivity of 98.83%.

On computing the classifier result in terms of specificity, it is conveyed that the GoogleNet model accomplishes inferior classification with a minimal specificity of 93.45%. At the same time, the AlexNet and VggNet-16 models offer somewhat better classification with reasonable specificity values of 94.07% and 94.32% correspondingly. Besides, the VggNet-19 and ResNet models exhibit manageable and near identical results with the specificity values of 96.49 and 95.56 correspondingly. Then, the M-AlexNet and VggNet-s models demonstrate viable results and attain higher specificity values of 97.45% and 97.43% respectively. It is also noted that the DNN-MSO algorithm has reached a competitive result with a higher specificity of 99.47%. The HPTI-v4 model showcased extraordinary classification performance and obtained a specificity of 99.68%. By looking into the above tables and figures, it is apparent that the presented HPTI-v4 model shows extraordinary results with the maximum accuracy, sensitivity, and specificity of 99.49%, 98.83%, and 99.68%, respectively. Therefore, the presented model can be employed as an automated diagnostic tool for the classification of DR images.

IV. CONCLUSION AND FUTURE WORK

An efficient DR disease diagnosis model called the HPTI-v4 model has been presented. The presented HPTI-v4 model involves the segmentation process by the feature extraction processes based on histogram and Inception v4. For tuning the hyperparameters in Inception v4, the Bayesian optimization technique is involved. Finally, the classification processes are performed by the use of MLP. The experimental outcomes stated that the presented HPTI-v4 model showed extraordinary results with the maximum accuracy, sensitivity, and specificity of 99.49%, 98.83%, and 99.68% respectively. Therefore, the HPTI-v4 model can be employed as an automated diagnostic tool for the classification of DR images. In the future, the proposed model can be further improved with the application of classification models.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[2] D. R. Whiting, L. Guariguata, C. Weil, and J. Shaw, "IDF diabetes atlas: Global estimates of the prevalence of diabetes for 2011 and 2030," *Diabetes Res. Clin. Pract.*, vol. 94, no. 3, pp. 311–321, Dec. 2011.

[3] S. Hajeb Mohammad Alipour, H. Rabbani, and M. R. Akhlaghi, "Diabetic retinopathy grading by digital curvelet transform," *Comput. Math. Methods Med.*, vol. 2012, Sep. 2012, Art. no. 761901.
[4] S. Sb and V. Singh, "Automatic detection of diabetic retinopathy in non-dilated RGB retinal fundus images," *Int. J. Comput. Appl.*, vol. 47, no. 19, pp. 26–32, Jun. 2012.
[5] N. Singh and R. C. Tripathi, "Automated early detection of diabetic retinopathy using image analysis techniques," *Int. J. Comput. Appl.*, vol. 8, no. 2, pp. 18–23, Oct. 2010.
[6] B. S. P. J. G. Cunha-Vaz, "Measurement and mapping of retinal leakage and retinal Thickness–Surrogate outcomes for the initial stages of diabetic retinopathy," *Current Medicinal Chem.-Immunol., Endocrine Metabolic Agents*, vol. 2, no. 2, pp. 91–108, Jun. 2002.
[7] H. Anandakumar and K. Umamaheswari, "Supervised machine learning techniques in cognitive radio networks during cooperative spectrum handovers," *Cluster Comput.*, vol. 20, no. 2, pp. 1505–1515, Jun. 2017.
[8] M. Omar, F. Khelifi, and M. A. Tahir, "Detection and classification of retinal fundus images exudates using region based multiscale LBP texture approach," in *Proc. Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, Apr. 2016, pp. 6–8.
[9] R. A. Welikala, M. M. Fraz, T. H. Williamson, and S. A. Barman, "The automated detection of proliferative diabetic retinopathy using dual ensemble classification," *Int. J. Diagnostic Imag.*, vol. 2, no. 2, pp. 72–89, Jun. 2015.
[10] A. Haldorai, A. Ramu, and C.-O. Chow, "Editorial: Big data innovation for sustainable cognitive computing," *Mobile Netw. Appl.*, vol. 24, no. 1, pp. 221–223, Feb. 2019.
[11] A. P. Bhatkar and G. U. Kharat, "Detection of diabetic retinopathy in retinal images using MLP classifier," in *Proc. IEEE Int. Symp. Nanoelectronic Inf. Syst.*, Indore, India, Dec. 2015, pp. 21–23.
[12] M. Partovi, S. H. Rasta, and A. Javadzadeh, "Automatic detection of retinal exudates in fundus images of diabetic retinopathy patients," *J. Anal. Res. Clin. Med.*, vol. 4, no. 2, pp. 104–109, May 2016.
[13] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, and E. I.-C. Chang, "Deep learning of feature representation with multiple instance learning for medical image analysis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 4–9.
[14] W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, Y. Zang, and J. Tian, "Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification," *Pattern Recognit.*, vol. 61, pp. 663–673, Jan. 2017.
[15] T. Shanthi and R. S. Sabeenian, "Modified alexnet architecture for classification of diabetic retinopathy images," *Comput. Electr. Eng.*, vol. 76, pp. 56–64, Jun. 2019.
[16] K. Shankar, E. Perumal, and R. M. Vidhyavathi, "Deep neural network with moth search optimization algorithm based detection and classification of diabetic retinopathy images," *Social Netw. Appl. Sci.*, vol. 2, no. 4, p. 748, Mar. 2020.
[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. 26th Annu. Conf. Neural Inf. Process. Syst. (NIPS)*, Stateline, NV, USA Dec. 2012, pp. 1097–1105.
[18] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell. (AAAI)*, San Francisco, CA, USA, Feb. 2017, pp. 1–12.
[19] M. Pelikan, D. E. Goldberg, and E. Cantú-Paz, "BOA: The Bayesian optimization algorithm," in *Proc. 1st Annu. Conf. Genetic Evol. Comput.*, Orlando, FL, USA, Jul. 1999, pp. 525–532.
[20] (2020). *Messidor-ADCIS*. Accessed: May 5, 2020. [Online]. Available: http://www.adcis.net/en/third-party/messidor/

K. SHANKAR is currently a Postdoctoral Fellow with Alagappa University, Karaikudi, India. He has authored/coauthored over 48 ISI Journal articles (total Impact Factor 122.404) and more than 100 Scopus Indexed Articles. He has delivered several invited and keynote talks and reviewed the technology leading articles for journals, such as *Scientific Reports-Nature*, the IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, the IEEE TRANSACTIONS ON RELIABILITY, IEEE ACCESS, the IEEE INTERNET OF THINGS, *Journal Big Data Research*, and *Human-centric Computing and Information Sciences*. He has authored/edited Conference Proceedings, Book Chapters, and two books published by *Springer*. He has been a part of various seminars, paper presentations, research article reviews, convener, and a session chair of the several conferences. He has displayed vast success in continuously acquiring new knowledge and applying innovative pedagogies. He has always aimed to be an effective Educator and have a Global Outlook. His current research interests include healthcare applications, secret image sharing scheme, digital image security, cryptography, the Internet of things, and optimization algorithms. He has served as the Chair (Program, Publicity, and Track) on several conferences and workshops. He has Guest-Edited several special issues at many journals published by Inderscience and MDPI. He has served as a Guest Editor and an Associate Editor in SCI and Scopus indexed journals, such as Elsevier, Springer, Wiley, and MDPI.

YIZHUO ZHANG is currently pursuing the master's degree with the College of Mathematics and Computer Science, Fuzhou University. His contributions have been published with IEEE ACCESS, WWW'2020, the IEEE ICCE-TW 2020, and so on. His research interests include deep learning and machine learning.

YIWEI LIU is currently pursuing the bachelor's degree with the College of Mathematics and Computer Science, Fuzhou University. Her contributions have been published with IEEE ACCESS, the IEEE ICCE-TW 2020, and so on. Her research interests include machine learning and intelligent transportation systems.

LING WU received the Ph.D. degree from the School of Economics and Management, Fuzhou University. She is currently a Lecturer with the College of Mathematics and Computer Science, Fuzhou University. Her contributions have been published with IEEE ACCESS, *Physica A*, and so on. Her recent research interests include social networks, grey systems, and machine learning.

CHI-HUA CHEN (Senior Member, IEEE) is currently a Distinguished Professor with Fuzhou University. He is also a Chair Professor with Dalian Maritime University, China. He has published over 300 academic articles and patents. His contributions have been published with the IEEE INTERNET OF THINGS JOURNAL, IEEE ACCESS, *IEICE Transactions*, WWW'2020, SIGIR 2020, and so on. His recent research interests include the Internet of Things, machine learning, deep learning, big data, cellular networks, and intelligent transportation systems. He serves as the Chair for several conferences, such as ICC 2020, APNOMS 2020, TrustCom 2020 Workshop, BIBM 2020 Workshop, and so on. He serves as an Editor for several international journals, such as IEEE ACCESS, *IEICE Transactions*, and so on.