

Received December 31, 2020, accepted January 18, 2021, date of publication February 2, 2021, date of current version April 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3056430

Discrimination of Diabetic Retinopathy From Optical Coherence Tomography Angiography Images Using Machine Learning Methods

ZHIPING LIU^{1,2}, CHEN WANG³, XIAODONG CAI³, HONG JIANG^{2,4}, AND JIANHUA WANG^{2,3}

¹Ophthalmic Center, The Second Affiliated Hospital, Guangzhou Medical University, Guangzhou 510260, China

²Department of Ophthalmology, Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL 33136, USA

³Department of Electrical and Computer Engineering, University of Miami, FL 33146, USA

⁴Department of Neurology, University of Miami Miller School of Medicine, Miami, FL 33136, USA

Corresponding author: Jianhua Wang (jwang3@med.miami.edu)

This work was supported in part by the National Institutes of Health (NIH) Center under Grant P30EY014801 and Grant NINDS1R01NS111115-01, and in part by the Research to Prevent Blindness (RPB). The work of Zhiping Liu was supported in part by the Guangzhou Municipal Science and Technology Project under Grant 201804010038, and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2020A1515010276.

ABSTRACT The goal was to discriminate between diabetic retinopathy (DR) and healthy controls (HC) by evaluating Optical coherence tomography angiography (OCTA) images from 3×3 mm scans with the assistance of different machine learning models. The OCTA angiography dataset of superficial vascular plexus (SVP), deep vascular plexus (DVP), and retinal vascular network (RVN) were acquired from 19 DR (38 eyes) patients and 25 HC (44 eyes). A discrete wavelet transform was applied to extract texture features from each image. Four machine learning models, including logistic regression (LR), logistic regression regularized with the elastic net penalty (LR-EN), support vector machine (SVM), and the gradient boosting tree named XGBoost, were used to classify wavelet features between groups. The area under the receiver operating characteristics curve (AUC), sensitivity, specificity, and diagnostic accuracy of the classifiers were obtained. The OCTA image dataset included 114 and 132 images from DR and HC subjects, respectively. LR-EN and LR using all three images, SVP, DVP, and RVN, provided the highest sensitivity of 0.84 and specificity of 0.80, the best diagnostic accuracy of 0.82, and an AUC of 0.83 and 0.84, respectively, which were slightly lower than that of LR using one image SVP (0.85) or two images DVP and SVP (0.85). The LR-EN and LR classification algorithms had the high sensitivity, specificity, and diagnostic accuracy in identifying DR, which may be promising in facilitating the early diagnosis of DR.

INDEX TERMS Diabetic retinopathy, machine learning, logistic regression, logistic regression regularized with the elastic net penalty, support vector machine.

I. INTRODUCTION

Diabetic retinopathy (DR) is one of the main causes of blindness among the working-age population [1]. Since it becomes incurable in its late stages, early diagnosis of DR is important. However, it is difficult to achieve early diagnosis because of the paucity of experienced retinal doctors and the increased number of DR patients [2]. Recently, several automated diagnosis systems have been developed to assist in the early diagnosis of DR, which are objective and reliable tools used

The associate editor coordinating the review of this manuscript and approving it for publication was Alberto Cano ¹.

to support clinical decision-making process without inter and intra-expert variability.

Artificial intelligence (AI) can analyze a large amount of data simultaneously and can correct automatically and learn continuously to improve upon its sensitivity and specificity as a diagnostic tool and predict disease progression in medicine [3]–[5]. An AI system is designed to mimic human brain perception for data processing and decisions making. Recently, there has been an increasing interest in using machine learning models for medical applications, which are branches of AI that can learn from data, identify patterns, and make decisions automatically [6].

In ophthalmic research, the application of AI systems has led to robust diagnostic accuracy for several ocular conditions such as DR [7], age-related macular degeneration (AMD) [8], cataract [9], glaucoma [10], and keratoconus [11].

Machine learning models using fundus images have been developed by researchers for retinopathy staging as well as identifying multiple ocular diseases. However, it is difficult to detect micro-vasculature alterations in different retinal layers and the area around the foveal avascular zone using fundus images. Additionally, a large and well-documented database of fundus images is needed to train and optimize convolutional neural networks. It is also extremely difficult to provide strong accuracy metrics because of the database variances from different imaging centers [7], [12].

A deep learning system (DLS) is a branch of machine learning using AI and representation learning methods to process large data and extract meaningful patterns that contribute to the excellent performance for discovering the intricate structures of high-dimensional data. Currently, DLS for the automatic classification and segmentation of optical coherence tomography angiography (OCTA) images in ophthalmology affords excellent results. In DLS, convolutional neural network (CNN), and multi-scaled encoder-decoder neural network (MEDnet) have been utilized to classify referable DR on optical coherence tomography angiography datasets with high predictive accuracy [13]–[18].

Abramoff and associates conducted a pivotal trial in which an autonomous AI-based diagnostic system detected DR, and that system has now been authorized by the Food and Drug Administration for use by health care providers to detect mild DR and diabetic macular edema [19]. Alam *et al.* presented the feasibility of a supervised machine learning method to detect DR with excellent diagnostic accuracy using 6×6 mm OCTA images from small datasets [20]. Recently, Sandhu *et al.* introduced a new automated system for detecting DR using OCTA. The system demonstrated a high degree of accuracy, sensitivity, and specificity in analyzing vessel density, vessel caliber, and FAZ area [21].

In the present study, we applied different machine learning models to discriminate between healthy individuals and DR by evaluating OCTA images from the 3×3 mm scans of superficial vascular plexus (SVP), deep vascular plexus (DVP), and retinal vascular network (RVN). We use sensitivity, specificity, and accuracy as metrics to compare the performance of different methods. Sensitivity is the ability of a method to correctly identify those with the disease (true positive), whereas specificity is the ability of the method to identify those without the disease (true negative) correctly. Accuracy is the proportion of true results, either true positive or true negative, identified by a method.

II. METHODS

A. STUDY DESIGN, SETTING, AND POPULATION

This study was approved by the institutional review board for human research of the University of Miami. All participants

were treated according to the Declaration of Helsinki. Written informed consent was obtained.

Nineteen patients with DR were recruited from Bascom Palmer Eye Institute, University of Miami, from August 2017 to June 2019. These patients underwent a complete fundus examination and were diagnosed by a retinal specialist (JT). Twenty-five age and sex-matched healthy control (HC) subjects were recruited.

B. OCTA, IMAGE SEGMENTATION, AND DATA ACQUISITION

The angiograph dataset of the 3×3 mm scans was acquired using the Optovue OCTA device (Optovue, Fremont, CA, USA), which consists of 70,000 Hz A-scans rate, an 840 nm central wavelength, a $5 \mu\text{m}$ resolution, and a $22 \mu\text{m}$ beamwidth. Each scan of retinal angiography included a raster scan with 304 (A-scan) \times 304 (B-scan). The OCTA image quality was quantified with the scan quality score metric provided in the Angiovue's software interface. Any OCTA image with a scan quality score of less than 5 was excluded. The SVP was segmented from the inner limited membrane layer (ILM) to the inner plexus layer (IPL); the DVP was segmented from the IPL to the outer plexus layer (OPL), and RVN was segmented from the ILM to the OPL.

C. DATA PRE-PROCESSING AND OCTA FEATURE EXTRACTION

The first step of applying a machine learning model to predict DR from images is to extract useful features from the images. It has been shown that wavelet transform can capture texture features of an image at multiple resolutions and that such features can be used to classify images with high accuracy [22], [23]. Wavelet transform has also been applied to fundus images for the diagnosis of glaucoma [24], [25].

We first applied the two-dimensional discrete wavelet transform (DWT) [26] to images to extract features in different frequency bands. Specifically, we employed the wavedec2 function in Matlab to implement the 2-dimensional DWT with the biorthogonal wavelet bior1.1. Of note, DWT had also been employed to extract features for the classification of fundus images of glaucoma [24], [27]. Since images were of size 304×304 pixels, we performed the 8-level ($\lfloor \log_2(304) = 8 \rfloor$) DWT. As shown in Figure 1, at the i th level ($i = 1, 3, \dots, 8$), we used biorthogonal wavelet bior1.1 to decompose the image in the LL_{i-1} band into four images in four frequency bands: HH_i , HL_i , LH_i , and LL_i , where LL_0 was the original image. Therefore, each image yielded 25 images of different resolutions with a total of 25 frequency bands. For each image, the energy in each of 25 frequency bands was calculated and standardized using the z-score, which generated 25 features that would be further used to classify images. Specifically, the image of the i th individual in the j th frequency band was represented as a $p \times q$ matrix D_{ij} . The energy of the band was $E_{ij} = \frac{1}{(pq)^2} \sum_{l=1}^p \sum_{k=1}^q D_{ij}^2(l, k)$, where $D_{ij}(l, k)$ was the entry of matrix D_{ij} on the l th row

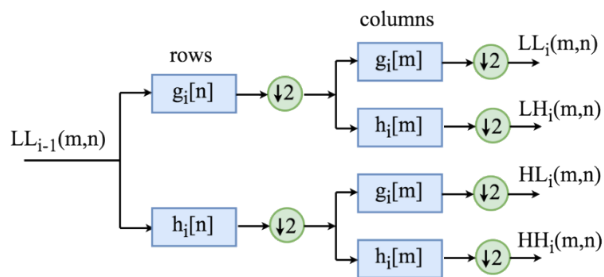


FIGURE 1. Block diagram of two-dimensional DWT. $g_i[n]$ and $h_i[n]$ are impulse responses of a low-pass wavelet filter and a high-pass wavelet filter, respectively. The output of each filter is down-sampled by a factor of 2.

and the k th column. Suppose that there were n individuals in the dataset, defined $\bar{E}_j = \frac{1}{n} \sum_{i=1}^n E_{ij}$ and $\sigma_j^2 = \frac{1}{n-1} \sum_{i=1}^n (E_{ij} - \bar{E}_j)^2$, then the z-score of E_{ij} is $z_{ij} = \frac{E_{ij} - \bar{E}_j}{\sigma_j}$. Finally, we represented each image of the i th individual as a 25×1 feature vector $x_i = [z_{i,1}, \dots, z_{i,25}]^T$, which was then used by a machine learning model for image classification. In our dataset, each eye has three images, including DVP, RVN, and SVP. We can use any one, two, or three images of all the three images to predict DR. If we use k ($k = 2$ or 3) images, we concatenate the feature vectors of these k images into a $25k \times 1$ vector as the input to a machine learning model. Of note, prior studies have shown that the energy distribution in different frequency bands provided discriminative power for image classification [22], [23].

D. DATA ANALYSIS

We employed four machine learning models to classify images based on wavelet features. These four models are logistic regression (LR), support vector machine (SVM), logistic regression regularized with the elastic net penalty (LR-EN), and the gradient boosting tree named XGBoost. LR and LR-EN were implemented with software glmnet [28]; SVM and XGBoost were implemented with the SVM function in R package e1071 [29] and the software XGBoost [30], respectively. Of note, Boosting trees are particularly powerful in supervised learning and have frequently won competitions [30], [31], sometimes outperforming deep neural networks. The mathematical formulation of LR and SVM can be found in a machine learning textbook [33], [34]; a detailed description of LR-EN and XGBoost is in [28] and [30], respectively. Next, we will describe these four methods briefly.

As mentioned earlier, the feature vector in the i th individual is denoted as x_i . Let $y_i = -1$ or 1 be the class label of the i th individual. The logistic regression model assumes the following probability $p(y_i|x_i) = 1/(1 + e^{-y_i(\beta_0 + x_i^T \beta)})$, where the vector β and the scalar β_0 contains model parameters. LR-EN finds (β_0, β) by minimizing the negative log-likelihood function of the data regularized by the elastic net penalty [24]. More specifically, the following optimization problem is

solved to find (β_0, β)

$$\min_{\beta_0, \beta} \log(1 + e^{-y_i(\beta_0 + x_i^T \beta)}) + \lambda(\alpha \|\beta\|_1 + (1 - \alpha) \|\beta\|_2) \tag{1}$$

where $\|\beta\|_k$, $k = 1, 2$, is the l_k -norm of β , and $\lambda > 0$ and $0 \leq \alpha \leq 1$ are two hyperparameters that can be determined with cross-validation (CV). If we set $\lambda = 0$ in (1), the solution of (1) gives the parameters of LR. Since the number of data samples is relatively small, and with the issue of quasi complete separation [35] in the datasets under consideration, the solution of (1) with $\lambda = 0$ is not stable. To overcome this problem, we implemented LR using (1) with $\alpha = 0$ and $\lambda > 0$. The hyperparameter λ was determined using cross validation [36].

SVM finds the decision boundary $\beta_0 + x^T \beta = 0$ by solving the following optimization problem

$$\min_{\beta_0, \beta} \frac{C}{n} \sum_{i=1}^n \max\{0, 1 - y_i(\beta_0 + x_i^T \beta)\} + \|\beta\|_2^2 \tag{2}$$

where $C > 0$ is a hyperparameter. The SVM in (2) finds a linear decision boundary in the feature space. We can also use the kernel trick to find a nonlinear decision boundary in the feature space. In this paper, we used the Gaussian kernel, which is defined as $K(x, x') = \exp(-\gamma \|x - x'\|_2^2)$, where $\gamma > 0$ is another hyperparameter of the model.

We used the nested CV procedure [37] to train the four models and estimate the classification errors. The nested CV procedure has two loops, both of which employed a leave-one-out (LOO) approach. The inner loop was used to tune the hyperparameters and train the model, while the outer loop was used to estimate the classification error. Since the samples used in the outer loop for error estimation were never used by the training process in the inner loop, the nested CV provided an unbiased estimate of the generalization error. First, one of the 82 samples was held out. The remaining 81 samples were used by a LOO procedure, which was referred to as the inner CV loop to be described in detail later, to train the model and select the optimal values of the hyperparameters. After the values of the hyperparameters were selected, the 81 samples were used to fit the model, which was then used to find the error on the one sample that was held out. This process, referred to as the outer loop, was repeated 82 times until all samples were held out and used to calculate the error. This error is the estimated classification error. Now, let us describe the inner CV loop. One of the 81 samples was left out, and the remaining 80 samples were used to train the model for a set of values within the hyperparameters, and the trained model was used to predict the label of the sample left out. This process was repeated 81 times until every sample was left out, and its label was predicted. The predicted labels of the 81 samples were used to calculate the CV error, and the value of the hyperparameter(s) that yielded the smallest CV error was selected as the optimal value [37].

We use sensitivity, specificity, and accuracy as performance measures. Sensitivity measures the proportion of true

TABLE 1. Sensitivity and specificity of four machine learning methods for predicting DR.

Images	Sensitivity				Specificity			
	LR-EN	LR	SVM	Xgb oost	LR-EN	LR	SVM	XGBo ost
DVP	0.61	0.58	0.55	0.68	0.82	0.77	0.66	0.75
RVN	0.61	0.61	0.58	0.61	0.84	0.86	0.77	0.75
SVP	0.66	0.66	0.63	0.68	0.82	0.84	0.75	0.66
DVP+RVN	0.66	0.68	0.71	0.55	0.77	0.8	0.75	0.7
DVP+SVP	0.76	0.76	0.68	0.68	0.77	0.8	0.8	0.75
RVN+SVP	0.71	0.68	0.58	0.71	0.82	0.84	0.73	0.77
DVP+RVN+SVP	0.84	0.84	0.63	0.61	0.8	0.8	0.8	0.77

LR-EN, logistic regression regularized with the elastic net penalty; LR, logistic regression; SVM, support vector machine; SVP, superficial vascular plexus; DVP, deep vascular plexus; RVN, retinal vascular network.

positives (TP) that are correctly identified with false negatives (FN), and is defined as $\text{sensitivity} = \text{TP} / (\text{TP} + \text{FN})$. Specificity measures the proportion of true negatives that are correctly identified with false positives (FP), and can be computed as $\text{specificity} = \text{TN} / (\text{TN} + \text{FP})$. Accuracy is the degree of closeness of measurements of a quantity to that quantity's true value and can be calculated as $\text{accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FN} + \text{TN} + \text{FP})$ [38].

III. RESULTS

The OCTA image dataset included 114 images from 38 eyes of 19 DR patients and 132 images from 44 eyes of 25 HC subjects. Tables 1 and 2 show the performance of the four machine learning models in predicting DR from one, two, or three images. As shown in Table 1, LR-EN and LR using the three images (DVP+RVN+SVP) offered the highest sensitivity (0.84) and specificity (0.80). As shown in Table 2, the AUCs of LR-EN and LR using the three images are 0.83 and 0.84, respectively, which are slightly lower than the AUCs of LR using one image SVP (0.85) or two images DVP and SVP (0.85), but are higher than or equal to the AUCs of LR-EN and LR using other images, and the AUCs of SVM and XGBoost. Table 3 shows that LR-EN and LR using the three images provided the best diagnostic accuracy (0.82) among all four classifiers.

Figure 2 depicted the receiver operating characteristic (ROC) curves and AUCs of LR-EN using different images. It was seen that the LR-EN using the three images offers almost the same performance as using one image, the SVP, and better performance than LR-EN using other images. Figure 3 showed the ROC curves and AUCs of LR-EN, LR, SVM, and XGBoost using the three images. It is clearly seen from Figure 2 that LR-EN and LR have almost the same performance but outperform SVM and XGBoost.

IV. DISCUSSION

We herein clarified the applicability of machine learning algorithms to predict diabetic retinopathy (DR) using

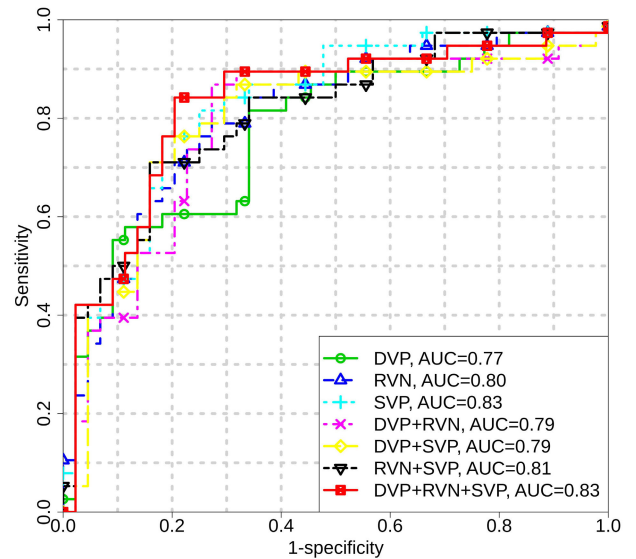


FIGURE 2. Receiver operating characteristic curve of LR-EN using one, two, or three images to predict DR.

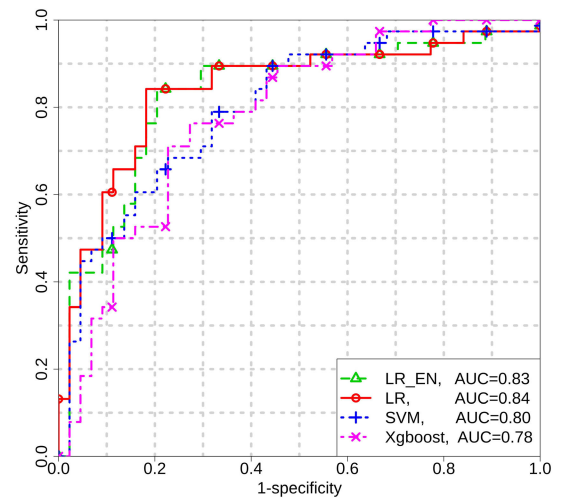


FIGURE 3. Receiver operating characteristic curve of four machine learning methods using three images (DVP+RVN+SVP) to predict DR. LR-EN and LR have almost the same performance, and they outperform SVM and XGBoost.

3 × 3 mm scans of optical coherence tomography angiography (OCTA) images. Subtle retinal microvascular alterations due to DR can be analyzed quantitatively based on OCTA images. The LR-EN and LR classifiers demonstrated the best diagnostic performance in all classification tasks in our current study. The LR-EN or LR algorithm would provide an effective strategy for clinicians to diagnose early stages of DR.

Generally, in our current dataset, using all three images improved the performance in terms of diagnostic accuracy and sensitivity when compared to the cases using only one or two images. However, using three images did not necessarily improve the performance in terms of the AUC when compared to the cases using the one image of SVP, or two images of DVP and SVP. This implies that the SVP image

TABLE 2. AUC of four machine learning methods for predicting DR.

Images	LR-EN	LR	SVM	XGBoost
DVP	0.77	0.77	0.76	0.75
RVN	0.80	0.82	0.80	0.75
SVP	0.83	0.85	0.84	0.74
DVP+RVN	0.79	0.82	0.82	0.73
DVP+SVP	0.79	0.85	0.84	0.75
RVN+SVP	0.81	0.84	0.73	0.74
DVP+RVN+SVP	0.83	0.84	0.80	0.78

AUC, area under the receiver operating characteristics curve.

TABLE 3. Diagnostic accuracy measured in four machine learning methods.

Images	LR-EN	LR	SVM	XGBoost
DVP	0.72	0.68	0.61	0.72
RVN	0.73	0.74	0.68	0.69
SVP	0.75	0.76	0.69	0.67
DVP+RVN	0.72	0.74	0.73	0.63
DVP+SVP	0.77	0.78	0.74	0.72
RVN+SVP	0.77	0.77	0.66	0.74
DVP+RVN+SVP	0.82	0.82	0.72	0.70

contains the most information that can help discriminate DR from normal tissue, and can explain why adding the other two images did not significantly change the performance of the classifier.

Sensitivity is an important criterion for any screening and diagnostic prediction system [15]. Studies have shown that the sensitivity of automatic DR screening ranges from 75% to 97.5%, and the accuracy is comparable [7], [12], [19], [20], [39], [40]. The 84% sensitivity of our system represented the capability to identify individual eyes with DR from a healthy control. Similarly, specificity is also an important factor because it will represent the capability of detecting subjects that do not require a referral to an ophthalmologist. Our study indicated that LR-EN and LR algorithm could achieve the best performances for maximum diagnostic accuracy in all classification tasks. As supported by the results from Table 3, we can observe that, in all performance metrics, the LR-EN and LR algorithms demonstrated better diagnostic proficiency than SVM and XGBoost algorithms.

Machine learning models for analyzing angiography images did not require alignment since the wavelet features were directly processed from the raw angiography images. The machine learning models, particularly LR-EN and LR, provided better performance than the traditional vessel density analysis (Dbox). The traditional vessel density analysis provided a sensitivity of 0.70 and specificity of 0.65 and used the vessel density of DVP (VDd), to discriminate DR from HC [41]. The improvement of machine learning models was mainly due to the following two factors. First, machine learning models exploited wavelet features of the angiography images of each eye, whereas the VDd analysis used the average density. Wavelet features can capture patterns in an

image at multiple resolutions, which provides more information and improves discriminative power. Second, the machine learning models were designed based on certain optimality criteria to discriminate images in the vector space of features, which can exploit the discriminative information appropriately [23], [24], [42]. XGBoost is a very efficient implementation of the gradient boosting tree [29]. Instead of fitting a tree to the gradient of the loss function at each iteration, as normally done by gradient boosting, XGBoost uses the second-order approximation of the loss function to build trees. The hyperparameters of XGBoost include the maximum depth of each tree, the number of trees, and the learning rate. In this work, the LR-EN and LR have almost the same performance, and they outperform SVM and XGBoost.

Although our findings indicated the LR-EN and LR classification algorithms had high sensitivity, specificity, and diagnostic accuracy in identifying DR, there are several limitations in this study. Firstly, the sample size is relatively small for each cohort. In future studies, we plan to include multiple imaging centers and a much more extensive OCTA database to test our AI screening algorithm for practical performance in retinal clinics. Additionally, we did not use retinal images to test the availability of our algorithm for detecting DR. There are many important components that should be included in the grading system for detecting DR using retinal images, such as microaneurysms, hemorrhages, hard and soft exudations, as well as blood vessel morphology [43]–[45]. Thirdly, we only analyzed the 3×3 mm OCTA images from mild DR patients. Patients with moderate and severe DR were not included. Since mild DR is hard to distinguish from normal individuals, the sensitivity and accuracy for detecting DR may have been found to be relatively lower. Future studies should enlarge the sample size and include moderate and severe DR patients. Finally, we did not use other data such as thickness maps of intra-retinal layers and retinal angiography, which, if incorporated into the machine learning process, may further improve the diagnostic accuracy and applicability.

V. CONCLUSION

In summary, in this evaluation of the 3×3 mm scans of OCTA angiography images from DR patients and healthy individuals, the LR-EN algorithm had high sensitivity and specificity in identifying DR, which may be promising in facilitating the early diagnosis of DR. Further large scale and multicenter studies are necessary to assess the applicability of LR-EN algorithm in DR and related eye diseases to improve vision outcomes.

ACKNOWLEDGMENT

All authors have no proprietary interest in any materials or methods.

REFERENCES

- [1] A. M. Hendrick, M. V. Gibson, and A. Kulshreshtha, "Diabetic retinopathy," *Prim Care*, vol. 42, no. 3, pp. 451–464, 2015.
- [2] D. S. W. Ting, G. C. M. Cheung, and T. Y. Wong, "Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges: A review," *Clin. Exp. Ophthalmol.*, vol. 44, no. 4, pp. 260–277, May 2016.

- [3] E. J. Topol, "High-performance medicine: The convergence of human and artificial intelligence," *Nature Med.*, vol. 25, no. 1, pp. 44–56, 2019.
- [4] J. He, S. L. Baxter, J. Xu, J. Xu, X. Zhou, and K. Zhang, "The practical implementation of artificial intelligence technologies in medicine," *Nature Med.*, vol. 25, no. 1, pp. 30–36, Jan. 2019.
- [5] P. Hamet and J. Tremblay, "Artificial intelligence in medicine," *Metabolism*, vol. 69, pp. S36–S40, Apr. 2017.
- [6] R. C. Deo, "Machine learning in medicine," *Circulation*, vol. 132, no. 20, pp. 1920–1930, 2015.
- [7] D. S. W. Ting et al., "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," *J. Amer. Med. Assoc.*, vol. 318, no. 22, pp. 2211–2223, 2017.
- [8] D. K. Hwang, C. C. Hsu, K. J. Chang, D. Chao, C. H. Sun, Y. C. Jheng, A. A. Yarmishyn, J. C. Wu, C. Y. Tsai, M. L. Wang, C. H. Peng, K. H. Chien, C. L. Kao, T. C. Lin, L. C. Woung, S. J. Chen, and S. H. Chiou, "Artificial intelligence-based decision-making for age-related macular degeneration," *Theranostics*, vol. 9, no. 1, pp. 232–245, 2019.
- [9] H. Lin et al., "Diagnostic efficacy and therapeutic decision-making capacity of an artificial intelligence platform for childhood cataracts in eye clinics: A multicentre randomized controlled trial," *EclinicalMedicine*, vol. 9, pp. 52–59, Mar. 2019.
- [10] A. A. Jammal, A. C. Thompson, E. B. Marittoni, S. I. Berchuck, C. N. Urata, T. Estrela, S. M. Wakil, V. P. Costa, and F. A. Medeiros, "Human versus machine: Comparing a deep learning algorithm to human gradings for detecting glaucoma on fundus photographs," *Amer. J. Ophthalmol.*, vol. 211, pp. 123–131, Mar. 2020.
- [11] S. D. Klyce, "The future of keratoconus screening with artificial intelligence," *Ophthalmology*, vol. 125, no. 12, pp. 1872–1873, Dec. 2018.
- [12] N. Asiri, M. Hussain, F. Al Adel, and N. Alzaidi, "Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey," *Artif. Intell. Med.*, vol. 99, Aug. 2019, Art. no. 101701.
- [13] M. Guo, M. Zhao, A. M. Y. Cheong, H. Dai, A. K. C. Lam, and Y. Zhou, "Automatic quantification of superficial foveal avascular zone in optical coherence tomography angiography implemented with deep learning," *Vis. Comput. Ind., Biomed., Art.*, vol. 2, no. 1, p. 21, Dec. 2019.
- [14] D. Le, M. Alam, C. K. Yao, J. I. Lim, Y.-T. Hsieh, R. V. P. Chan, D. Toslak, and X. Yao, "Transfer learning for automated OCTA detection of diabetic retinopathy," *Transl. Vis. Sci. Technol.*, vol. 9, no. 2, p. 35, Jul. 2020.
- [15] T. T. Hormel, H. Xiong, B. Wang, A. Camino, J. Wang, D. Huang, T. S. Hwang, and Y. Jia, "Development and validation of a deep learning algorithm for distinguishing the nonperfusion area from signal reduction artifacts on OCT angiography," *Biomed. Opt. Express*, vol. 10, no. 7, pp. 3257–3268, 2019.
- [16] M. Heisler, S. Karst, J. Lo, Z. Mammo, T. Yu, S. Warner, D. Maberley, M. F. Beg, E. V. Navajas, and M. V. Sarunic, "Ensemble deep learning for diabetic retinopathy detection using optical coherence tomography angiography," *Transl. Vis. Sci. Technol.*, vol. 9, no. 2, p. 20, Apr. 2020.
- [17] J. Lo, M. Heisler, V. Vanzan, S. Karst, I. Z. Matovinovic, S. Loncaric, E. V. Navajas, M. F. Beg, and M. V. Sarunic, "Microvasculature segmentation and intercapillary area quantification of the deep vascular complex using transfer learning," *Transl. Vis. Sci. Technol.*, vol. 9, no. 2, p. 38, Jul. 2020.
- [18] Y. Guo, A. Camino, J. Wang, D. Huang, T. S. Hwang, and Y. Jia, "MED-net, a neural network for automated detection of avascular area in OCT angiography," *Biomed. Opt. Express*, vol. 9, no. 11, pp. 5147–5158, 2018.
- [19] M. D. Abramoff, P. T. Lavin, M. Birch, N. Shah, and J. C. Folk, "Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices," *Yearbook Paediatric Endocrinol.*, vol. 1, p. 39, Sep. 2019.
- [20] M. Alam, D. Le, J. I. Lim, R. V. P. Chan, and X. Yao, "Supervised machine learning based multi-task artificial intelligence classification of retinopathies," *J. Clin. Med.*, vol. 8, no. 6, p. 872, Jun. 2019.
- [21] H. S. Sandhu, N. Eladawi, M. Elmogry, R. Keynton, O. Helmy, S. Schaal, and A. El-Baz, "Automated diabetic retinopathy detection using optical coherence tomography angiography: A pilot study," *Brit. J. Ophthalmol.*, vol. 102, no. 11, pp. 1564–1569, Nov. 2018.
- [22] T. Chang and C.-C.-J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Trans. Image Process.*, vol. 2, no. 4, pp. 429–441, Oct. 1993.
- [23] K. Huang and S. Aviyente, "Wavelet feature selection for image classification," *IEEE Trans. Image Process.*, vol. 17, no. 9, pp. 1709–1720, Sep. 2008.
- [24] S. Dua, U. R. Acharya, P. Chowriappa, and S. V. Sree, "Wavelet-based energy features for glaucomatous image classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 1, pp. 80–87, Jan. 2012.
- [25] S. Maheshwari, R. B. Pachori, and U. R. Acharya, "Automated diagnosis of glaucoma using empirical wavelet transform and correntropy features extracted from fundus images," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 3, pp. 803–813, May 2017.
- [26] D. Sundararajan, *Discrete Wavelet Transform: A Signal Processing Approach*, 2016th ed. Hoboken, NJ, USA: Wiley, 2020.
- [27] B. S. Kirar and D. K. Agrawal, "Computer aided diagnosis of glaucoma using discrete and empirical wavelet transform from fundus images," *IET Image Process.*, vol. 13, no. 1, pp. 73–82, Jan. 2019.
- [28] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *J. Stat. Softw.*, vol. 33, no. 1, pp. 1–22, 2010.
- [29] D. Meyer, "Support vector machines," in *The Interface to LIBSVM in Package*. 2019, p. e1071. [Online]. Available: [Online]. Available: <https://cran.r-project.org/web/packages/e1071/vignettes/svmdoc.pdf>
- [30] C. Tianqi and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 785–794.
- [31] C. Sutton, L. M. Ghiringhelli, T. Yamamoto, Y. Lysogorskiy, L. Blumenthal, T. Hammerschmidt, J. R. Golebiowski, X. Liu, A. Ziletti, and M. Scheffler, "Crowd-sourcing materials-science challenges with the NOMAD 2018 Kaggle competition," *NPJ Comput. Mater.*, vol. 5, no. 1, pp. 1–11, Dec. 2019.
- [32] T. W. Neller, "AI education matters: Lessons from a Kaggle click-through rate prediction competition," *AI Matters*, vol. 4, no. 3, pp. 5–7, 2018.
- [33] S. Ben-David and S. Shalev-Shwartz, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
- [34] J. H. Friedman, R. Tibshirani, and T. Hastie, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2009.
- [35] A. Albert and A. Anderson, *On the Existence of Maximum Likelihood Estimates in Logistic Regression Models*, 71st ed. 1984, pp. 1–10.
- [36] A. Albert and J. A. Anderson, "On the existence of maximum likelihood estimates in logistic regression models," *Biometrika*, vol. 71, no. 1, pp. 1–10, 1984.
- [37] S. Varma and R. Simon, "Bias in error estimation when using cross-validation for model selection," *BioMed Central*, vol. 7, no. 1, p. 91, 2006.
- [38] R. Parikh, A. Mathai, S. Parikh, S. G. Sekhar, and R. Thomas, "Understanding and using sensitivity, specificity and predictive values," *Indian J. Ophthalmol.*, vol. 56, no. 1, pp. 45–50, 2008.
- [39] S. K. Padhy, B. Takkar, R. Chawla, and A. Kumar, "Artificial intelligence in diabetic retinopathy: A natural step to the future," *Indian J. Ophthalmol.*, vol. 67, no. 7, pp. 1004–1009, 2019.
- [40] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega, and D. R. Webster, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *J. Amer. Med. Assoc.*, vol. 316, no. 22, pp. 2402–2410, 2016.
- [41] Z. Liu, J. Hong, J. Townsend, and J. Wang, "Retinal tissue perfusion reduction best discriminates early stage diabetic retinopathy in patients with type 2 diabetes mellitus," *Retina*, vol. 41, no. 3, pp. 546–554.
- [42] B. J. Erickson, P. Korfiatis, Z. Akkus, and T. L. Kline, "Machine learning for medical imaging," *Radiographics*, vol. 37, no. 2, pp. 505–515, 2017.
- [43] K. Akyol, B. Sen, and S. Bayir, "Automatic detection of optic disc in retinal image by using keypoint detection, texture analysis, and visual dictionary techniques," *Comput. Math Methods Med.*, vol. 2016, Mar. 2016, Art. no. 6814791.
- [44] E. Imani, H.-R. Pourreza, and T. Banaee, "Fully automated diabetic retinopathy screening using morphological component analysis," *Comput. Med. Imag. Graph.*, vol. 43, pp. 78–88, Jul. 2015.
- [45] J. Nayak, P. S. Bhat, and U. R. Acharya, "Automatic identification of diabetic maculopathy stages using fundus images," *J. Med. Eng. Technol.*, vol. 33, no. 2, pp. 119–129, 2009.

•••