A Deep Learning Grading Classification of Diabetic Retinopathy on Retinal Fundus Images with Bio-inspired Optimization

Radhakrishnan Ramesh
Department of Computer Application, Government Arts and Science College for Women, India
radha.esh@gmail.com (corresponding author)

Selvarajan Sathiamoorthy
Annamalai University PG Extension Centre, India
ks_sathia@yahoo.com

Received: 11 May 2023 | Revised: 25 May 2023 and 5 June 2023 | Accepted: 6 June 2023

 DOI: https://doi.org/10.48084/etasr.6033

ABSTRACT

Diabetic Retinopathy (DR) is considered the major cause of impaired vision for diabetic patients, particularly in developing counties. Treatment includes maintaining the patient's present grade of vision as the illness can be irreparable. Initial recognition of DR is highly important to effectively sustain the vision of the patients. The main problem in DR recognition is that the manual diagnosis procedure consumes time, effort, and money and also includes an ophthalmologist's analysis of retinal fundus imaging. Machine Learning (ML)-related medical image analysis is proven to be capable of evaluating retinal fundus images, and by using Deep Learning (DL) techniques. The current research presents an Automated DR detection method by utilizing the Glowworm Swarm Optimization (GSO) with Deep Learning (ADR-GSODL) approach on retinal fundus images. The main aim of the ADR-GSODL technique relies on the recognizing and classifying process of DR in retinal fundus images. To obtain this, the introduced ADR-GSODL method enforces Median Filtering (MF) as a pre-processing step. Besides, the ADR-GSODL technique utilizes the NASNetLarge method for deriving the GSO, and feature vectors are applied for parameter tuning. For the DR classification process, the Variational Autoencoder (VAE) technique is exploited. The supremacy of the ADR-GSODL approach was confirmed by a comparative simulation study.

Keywords-Diabetic Retinopathy (DR); DR screening; fundus images; deep learning; metaheuristics

I. INTRODUCTION

Diabetic Retinopathy (DR) is a complication of diabetes instigated by Diabetes Mellitus in which the glucose will block the blood vessels that feed the eye, leading to fluids or blood leakage and swelling that can cause serious eye injury [1]. The detrimental vision impairment due to DR befalls primarily if there is swelling in the retina. The estimated population that suffers from vision loss, whether severe or mild, due to trachoma, glaucoma, and DR, is 11 million [2]. To evade problems related to chronic diseases, namely diabetes, initial identification is crucial. Abnormal development of blood vessels in the retina has an effective consequence of DR, which may cause bleeding or scarring from retina and subsequent blindness [3]. Universally, DR accounts for nearly 2.6% of impaired vision cases. The period a person suffered from diabetics, high blood pressure readings, and high haemoglobin are regarded as the most high-risk elements linked with DR growth [4]. Regular screening is the most used diagnostic tool to ensure that DR is diagnosed at a primary phase. DR detection usually includes an examination by physician of retinal images for the appearance and shape of distinct kinds of lesions [5].

Retinal fundus image analysis is useful in medical processing [6]. Diseases such as DR, macular degeneration, hypertension, and atherosclerosis could modify the morphology of the arteries, generating modifications in their branching angle, diameter, and tortuosity [7]. Manual segmentation of retinal vascular ailments is a time-consuming and skill-demanding process. The initial diagnosis of DR plays an important role in the prevention of blindness. Henceforth, doctors prescribe annual retinal screening tests [8]. Remarkably, such retina scans are utilized for detecting diabetes, though this demands the judgment of ophthalmologists, which may take time. In the detection of DR, DL methods have illustrated better performance, with a high accuracy level that differentiates them from other methods [9]. Certainly, DL could not cover concealed components in images that clinical experts would never see. Due to their capacity in training and feature extractions in discriminating among many classes, Convolutional Neural Networks (CNNs) are the most employed DL method in the health care sector [10]. Often, the

use of the Transfer Learning (TL) method simplified the process of retraining Deep Neural Network (DNN) reliably and speed.

This research presents an Automated DR detection method using the DSO with DL (ADR-GSODL) approach on retinal fundus imaging. The main aim of the ADR-GSODL technique relies in the recognition and classification process of DR in retinal fundus imaging. To obtain this, the ADR-GSODL method enforces Median Filtering (MF) as a pre-processing step. Besides, the ADR-GSODL technique also utilizes the NASNetLarge method for deriving the GSO and feature vectors are applied in the parameter tuning process. For the DR classification process, the VAE technique is exploited. To illustrate the advanced performance of the ADR-GSODL approach, comparative simulation analysis is executed.

II. LITERATURE REVIEW

Authors in [11] introduced CAD techniques that use the DL results for image recognition. This method relies on a deep residual CNN to derive biased factors with no previous convolution images transforming to accentuate specific structures or to improve the image qualities. Authors in [12] developed the Quantum TL (QTL) algorithm to achieve DR diagnosis. QTL is a hybrid integration of quantum computing and traditional TL. Authors in [13] explored ways to utilize deeper TL for the diagnoses of DR through OCT images. The authors retrain the present DL model for these tasks and investigate how a retrained method could be enhanced. They demonstrate that utilizing an improved pre-training module as a feature extractor and training a traditional classification on this feature is an efficient method to identify DR through OCT images. Authors in [14] developed a novel two-phase architecture for the classification of automated DR. Initially, the authors used two different U-Net architectures for Blood Vessel (BV) and Optic Disc (OD) segmentation during the pre-processing. Next, the improved retinal imaging afterward the extraction of OD and BV are applied as the input of TL-oriented VGGNet architecture that implements DR diagnosis by recognizing retinal biomarkers namely Exudates (EX), Microaneurysm (MA), and Haemorrhages (HM). Authors in [15] introduced a learning-based model for the classification and segmentation of DR lesioning. The pre-training Xception module was applied for the deep factor extracting process in the segmenting stage. The derived feature was given to Deeplabv3 for semantic segmentation. For segmentation training, a study was conducted for the selection of optimum hyperparameters that the presented method resulted in the evaluation stage. The multiple classification method is proposed for extracting features through FC MatMul layers of pool-10 of the squeeze-net and efficient-net-b0. In [16], automated detection of severity of DR using CNN with a TL technique was developed to assist the diagnosis method. A comparison of distinct CNN frameworks like Inception-ResNet-v2 and ResNet was performed by the quadratic weighted kappa metric.

III. THE PROPOSED MODEL

In this research, we introduce the ADR-GSODL technique for DR classification on retinal fundus imaging. The ADR-GSODL technique intends to recognize and classify of DR in retinal fundus imaging. It follows a sequence of procedures, namely MF noise removing process, NASNetLarge feature extractor, GSO parameter tuning process, and VAE classification process. Figure 1 demonstrates the ADR-GSODL block diagram model.

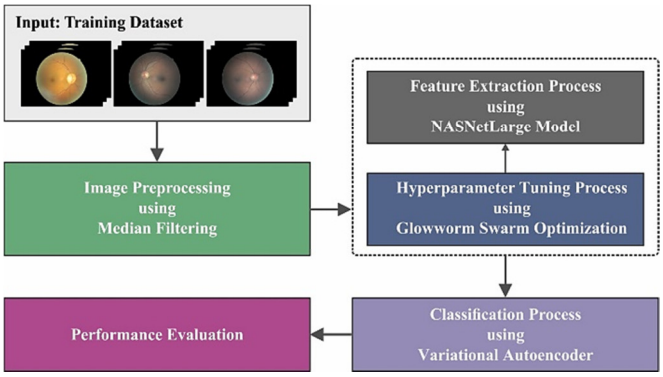

Fig. 1. Block diagram of the ADR-GSODL.

A. Feature Extraction

The ADR-GSODL technique utilizes NASNetLarge approach for the derivation of featuring vectors. The NASNetLarge networking contains encoding, decoding, and classifying layers [17]. There are 2 important variances related to Segnet which utilizes the pre-trained VGG16 infrastructure to encoder. The NasnetLarge-decoder net utilizes the primary 414 layers of NasnetLarge net (i.e. a highly trained network to ImageNet classification) as encoding to decompose the image. It employs the pre-trained weighted network and then retrains it with new data for appropriating NasnetLarge as the database can be different from ImageNet. Furthermore, the decryption is different and there is no pool index. The NASNetLarge net will produce comprehensive decoded datasets. The appropriately decoded upsample has its input feature mapping with max-pool layers. All blocks initiate with upsampling which would extend the ReLU, feature mapping, and convolutional layer. After that, a batch normalization layer is used. The primary decrypted data were closer to the final encrypted data generating a multichannel feature map, in a way analogous to Segnet.

B. Hyperparameter Tuning

In this work, the GSO algorithm is enforced for this process. In GSO, a Glowworm (GW) has a luciferin value and decisions are made on a local level [18]. The luciferin value relies upon its location and objective function. The superior the location of the GW, the brighter it is (i.e. has more luciferin) than its neighbors.

$$L_i(t+1) = (1-\rho)L_i(t) + \gamma F\big(p(t)\big) \quad (1)$$

For the GW i, the $L(t)$ characterizes its luciferin value. This value decay constant is represented as ρ and ranges from 0 to 1. The enhancement fraction of luciferin is denoted as γ, $F(p(t))$ provides the objective function at the existing location p_i of GW i. Then, the GW i will explore the adjacent area to find the neighbors with maximum luciferin values by employing the following rule:

$z \epsilon N_i(t)$ if $Distance_{iz} < rd_i(t)$ and $L_Z(t) > L_i(t)$ (2)

where z indicates the GW closer to the $i\text{-}th$ GW, $N(t)$ represents the set of neighboring GWs, $Distance_{iz}$indicates the Euclidean distance, $rd(t)$ provides the local decision range of the $i\text{-}th$ GW, and the luciferin levels of GWs z and i are denoted by $L(t)$ and $L(t)$, respectively.

The better neighboring GW from the neighboring set is chosen by evaluating the probability of each GW:

$$Probability_{iz} = \frac{L_Z(t) - L_i(t)}{\sum_{x \epsilon N_i} L_X(t) - L_i(t)} \quad (3)$$

The location of a GW is attuned according to the location of a better neighboring GW and is evaluated by:

$$p_i(t+1) = p_i(t) + s\frac{p_Z(t) - p_i(t)}{Distance_{iz}} \quad (4)$$

where $s > 0$ indicates the step size of the movement of a GW. The decision range $rd_i(t)$ is evaluated as follows:

$$rd_i(t+1) =$$
$$\min\{r_s, \max[0, rd_i(t) + \beta(n_r - |N_i(t)|)]\} \quad (5)$$

where r_s indicates the radial sensor range constant, β represents the model constant, and n_ζ restricts the neighborhood numbers.

C. VAE-based DR Detection

For DR classification process, the VAE model is utilized in this work. VAE belongs to the family of generative mechanisms and bears massive similarities to AE with respect to the design [19]. These mechanisms consist of an encoder (inference or recognition method) and a decoder that is called the generative mechanism. One more similarity among these two is that they try to recreate input datasets while learning from latent vectors. A difference among these two is that VAE's latent space is constant, because the encoder does not output an encoder vector of length n, 0 mean, and 1 Standard Deviation (SD). The mean vector aims to control the center position of the encoding input, while the SD vector aims to control the region where the encoding could differ. Because the encoding is randomly generated, the decoder also exposes itself to encoding variations. The VAE encoder will learn a map $Q_\theta(z|X)$ from the input x_i with covariance $o^2(x_i)$ and mean $\mu(x_i)$ vectors of the latent variable. VAE assumes data augmenting while considering that the latent variable z follows a uniform distribution (0, 1). The decoder should trial from the latent dispersion given as output through encoder because it no more results a latent depiction z as in AE. The VAE decoder would learn a map $P_\phi(|)$ from z latent illustration to the dispersion parameter of the training dataset X. For generating realtime samples for the decoder and to avoid overfitting, the log probability of input dataset $\log(p(X))$, should be maximized. A uniform dispersion has identity covariance matrices that denote that the covariance among the latent variables is 0 and they are not dependent. Therefore, the VAE learns the distribution of training dataset $p(X)$ with encoder $Q_\theta(z|X)$, decoder $P_\phi(X|z)$, and latent dispersion $p(z)$ and can be trained by means of the objective function (6):

$$\max_{\theta,\varphi} log(p(X)) \approx \max_{\theta,\varphi} E_{Q_\theta(z|X)}[log(P_\varphi(X|z))] - D_{KL}(Q_\theta(z|X)||p(z)) \quad (6)$$

whereas θ, ψ denote the parameters of the encoder and decoder, respectively and D_{KL} characterizes the KL divergence amongst two likelihood distributions.

IV. RESULTS AND DISCUSSION

The presented ADR-GSODL method is investigated using the DR database from Kaggle [20]. The database holds 35126 instances with 5-class labelling as represented in Table I. The presented modes are simulated by utilizing the Python 3.6.5 tool on an i5-8600k PC, with GeForce 1050Ti 4GB, 16 and 250 GB RAM and SSD, and 1TB HDD. The parameter setup is: learning rate: 0.01, epoch count: 50, batch size: 5, dropout rate: 0.5, and activation function: ReLU.

TABLE I. DATASET DETAILS

Label	DR class	Number of samples
0-DR	No	25810
1-DR	Mild	2443
2-DR	Moderate	5292
3-DR	Severe	873
4-DR	Proliferative	708
Total No. of Instances		**35126**

The confusion matrices of the ADR-GSODL approach are shown in Figure 2. In the 80% of the dataset (training –TR), the ADR-GSODL approach has classified 20434, 1878, 4202, 625, and 492 samples into DR of L0, L1, L2, L3 and L4 class, respectively. On 20% the dataset (testing-TS), the ADR-GSODL technique has classified 5148, 493, 996, 165, and 112 samples into DR of L0, L0, L1, L2, L3 and L4 class. In the same manner, on 70% TR, the ADR-GSODL technique categorized 17875, 1585, 3590, 590, and 475 samples into DR of L0, L1, L2, L3 and L4 class. Table II reports a comprehensive DR classifying of the results of the ADR-GSODL approach on 80 and 20% of TR and TS. On 80% TR, it is notified that the ADR-GSODL method has gained average $accu_y$ of 99.33%, $sens_y$ of 93.92%, $spec_y$ of 99.46%, F_{score} of 94.06%, and NPV of 99.30%. Besides, on 20% TS, the ADR-GSODL technique has obtained average $accu_y$ of 99.36%, $sens_y$ of 94.64%, $spec_y$ of 99.51%, F_{score} of 94.38%, and NPV of 99.29%.

TABLE II. DR CLASSIFYING OUTCOME OF THE ADR-GSODL FOR 80% AND 20% TR AND TS

Class	$Accu_y$	$Sens_y$	$Spec_y$	F_{score}	NPV
Training phase (80%)					
DR-0	99.04	99.13	98.80	99.34	97.62
DR-1	99.38	97.00	99.55	95.55	99.78
DR-2	99.21	98.20	99.39	97.43	99.68
DR-3	99.55	90.71	99.77	90.84	99.77
DR-4	99.48	84.54	99.80	87.16	99.67
Average	**99.33**	**93.92**	**99.46**	**94.06**	**99.30**
Testing phase (20%)					
DR-0	99.06	99.08	99.02	99.36	97.42
DR-1	99.40	97.24	99.57	95.91	99.78
DR-2	99.23	98.32	99.38	97.36	99.72
DR-3	99.56	89.67	99.82	91.41	99.72
DR-4	99.56	88.89	99.75	87.84	99.80
Average	**99.36**	**94.64**	**99.51**	**94.38**	**99.29**

Table III illustrates the comprehensive DR classifying output of the ADR-GSODL method on 70 and 30% of TR and TS, respectively. On the TR, the ADR-GSODL methodology acquired average $accu_y$ of 99.23%, $sens_y$ of 95.45%, $spec_y$ of 99.47%, F_{score} of 92.94%, and NPV of 99.15%. Besides, on the TS the ADR-GSODL method achieved average $accu_y$ of 99.26%, $sens_y$ of 95.57%, $spec_y$ of 99.50%, F_{score} of 92.51%, and NPV of 99.20%.

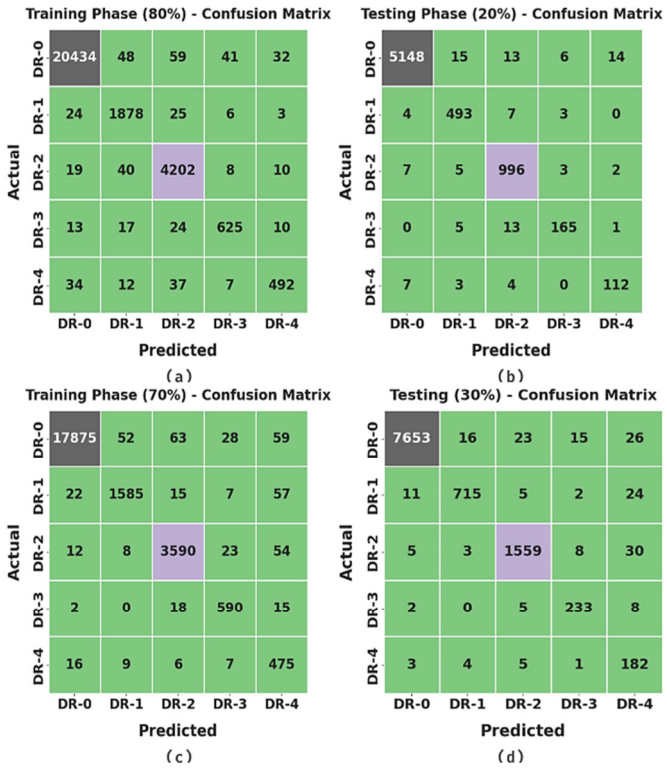

Fig. 2. Confusion matrix of ADR-GSODL method (a-b) 80:20 TR and TS and (c-d) 70:30 TR and TS data.

TABLE III. DR CLASSIFYING OUTCOME OF THE ADR-GSODL FOR 70% AND 30% TR AND TS

Class	$Accu_y$	$Sens_y$	$Spec_y$	F_{score}	NPV
Training phase (70%)					
DR-0	98.97	98.88	99.20	99.29	96.97
DR-1	99.31	94.01	99.70	94.91	99.56
DR-2	99.19	97.37	99.51	97.30	99.54
DR-3	99.59	94.40	99.73	92.19	99.85
DR-4	99.09	92.59	99.23	80.99	99.84
Average	**99.23**	**95.45**	**99.47**	**92.94**	**99.15**
Testing phase (30%)					
DR-0	99.04	98.97	99.25	99.34	97.21
DR-1	99.38	94.45	99.76	95.65	99.57
DR-2	99.20	97.13	99.57	97.38	99.49
DR-3	99.61	93.95	99.75	91.91	99.85
DR-4	99.04	93.33	99.15	78.28	99.87
Average	**99.26**	**95.57**	**99.50**	**92.51**	**99.20**

The Training and Validation Accuracy, TR_{acc} and VL_{acc}, respectively, reached by the ADR-GSODL approach under trial data are shown in Figure 3. The simulating values illustrated that the ADR-GSODL approach has attained greater TR_{acc} and VL_{acc} values. Seemingly the VL_{acc} is greater than TR_{acc}.

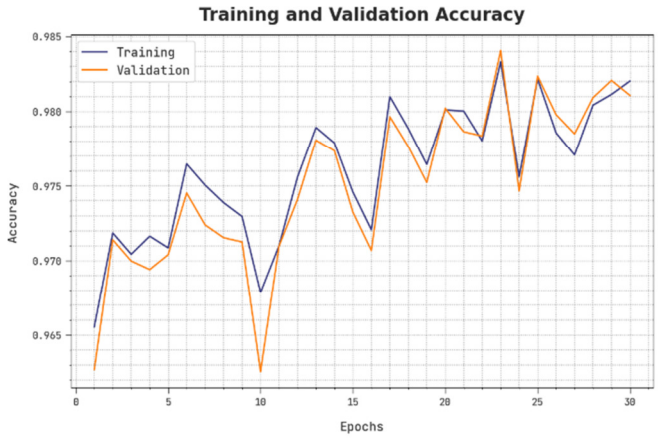

Fig. 3. TR_{acc} and VL_{acc} of the ADR-GSODL method.

To verify the enhanced performance of the ADR-GSODL method, a comprehensive relative study was conducted and its results are given in Table IV. It is noticed that the AlexNet model fails to achieve effectual outcomes The output specified that the ResNet-50 and ResNet-101 approaches have performed less satisfactory, in contrast with the VGG-16 and Inception v3 approaches. The ADR-GSODL approach achieved the optimum performance with $accu_y$ of 99.26%, $sens_y$ of 95.57%, and $spec_y$ of 99.50%. Thus, the ADR-GSODL model can be considered an effective tool for the DR grading process.

V. CONCLUSION

In this study, the novel ADR-GSODL model to classify DR on retinal fundus imaging is presented. The main aim of the ADR-GSODL model is in the classification and recognition of DR in retinal fundus images. To obtain this, the ADR-GSODL technique applied the MF approach as a pre-processing step. Also, the ADR-GSODL technique utilizes the NASNetLarge model to derive feature vectors, whereas GSO algorithm is applied for the tuning process. For DR classification, the VAE method is considered. To exhibit the enhanced ADR-GSODL performance, a widespread simulation analysis was performed and the comparison study assured the supremacy of the ADR-GSODL technique. In the future, the computation complexity of the proposed model can be examined. In addition, the proposed model can be tested on large scale datasets.

REFERENCES

[1] I. Kandel and M. Castelli, "Transfer Learning with Convolutional Neural Networks for Diabetic Retinopathy Image Classification. A Review," *Applied Sciences*, vol. 10, no. 6, Jan. 2020, Art. no. 2021, https://doi.org/10.3390/app10062021.

[2] C. Zhang, T. Lei, and P. Chen, "Diabetic Retinopathy Grading by a Source-Free Transfer Learning Approach," *Biomedical Signal Processing and Control*, vol. 73, Mar. 2022, Art. no. 103423, https://doi.org/10.1016/j.bspc.2021.103423.

[3] A. K. Gangwar and V. Ravi, "Diabetic Retinopathy Detection Using Transfer Learning and Deep Learning," in *Evolution in Computational Intelligence*, Singapore, 2021, pp. 679–689, https://doi.org/10.1007/978-981-15-5788-0_64.

[4] D. Le *et al.*, "Transfer Learning for Automated OCTA Detection of Diabetic Retinopathy," *Translational Vision Science & Technology*, vol. 9, no. 2, Jul. 2020, Art. no. 35, https://doi.org/10.1167/tvst.9.2.35.

[5] M. K. Jabbar, J. Yan, H. Xu, Z. Ur Rehman, and A. Jabbar, "Transfer Learning-Based Model for Diabetic Retinopathy Diagnosis Using Retinal Images," *Brain Sciences*, vol. 12, no. 5, Apr. 2022, Art. no. 535, https://doi.org/10.3390/brainsci12050535.

[6] N. E. M. Khalifa, M. Loey, M. H. N. Taha, and H. N. E. T. Mohamed, "Deep Transfer Learning Models for Medical Diabetic Retinopathy Detection," *Acta Informatica Medica*, vol. 27, no. 5, pp. 327–332, Dec. 2019, https://doi.org/10.5455/aim.2019.27.327-332.

[7] N. B. Thota and D. Umma Reddy, "Improving the Accuracy of Diabetic Retinopathy Severity Classification with Transfer Learning," in *2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS)*, Springfield, MA, USA, Dec. 2020, pp. 1003–1006, https://doi.org/10.1109/MWSCAS48704.2020.9184473.

[8] M. M. H. Milu, M. A. Rahman, M. A. Rashid, A. Kuwana, and H. Kobayashi, "Improvement of Classification Accuracy of Four-Class Voluntary-Imagery fNIRS Signals using Convolutional Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 13, no. 2, pp. 10425–10431, Apr. 2023, https://doi.org/10.48084/etasr.5703.

[9] D. Patil and S. Jadhav, "Road Segmentation in High-Resolution Images Using Deep Residual Networks," *Engineering, Technology & Applied Science Research*, vol. 12, no. 6, pp. 9654–9660, Dec. 2022, https://doi.org/10.48084/etasr.5247.

[10] S. Rani, Y. Chabrra, and K. Malik, "An Improved Denoising Algorithm for Removing Noise in Color Images," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8738–8744, Jun. 2022, https://doi.org/10.48084/etasr.4952.

[11] F. J. Martinez-Murcia, A. Ortiz, J. Ramírez, J. M. Górriz, and R. Cruz, "Deep residual transfer learning for automatic diagnosis and grading of diabetic retinopathy," *Neurocomputing*, vol. 452, pp. 424–434, Sep. 2021, https://doi.org/10.1016/j.neucom.2020.04.148.

[12] S. S, K. T, S. Bhattacharjee, D. Shahwar, and K. S. Sekhar Reddy, "Quantum Transfer Learning for Diagnosis of Diabetic Retinopathy," in *2022 International Conference on Innovative Trends in Information Technology (ICITIIT)*, Kottayam, India, Oct. 2022, https://doi.org/10.1109/ICITIIT54346.2022.9744184.

[13] K. T. Islam, S. Wijewickrema, and S. O'Leary, "Identifying Diabetic Retinopathy from OCT Images using Deep Transfer Learning with Artificial Neural Networks," in *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*, Cordoba, Spain, Jun. 2019, pp. 281–286, https://doi.org/10.1109/CBMS.2019.00066.

[14] A. Bilal, G. Sun, S. Mazhar, A. Imran, and J. Latif, "A Transfer Learning and U-Net-based automatic detection of diabetic retinopathy from fundus images," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 10, no. 6, pp. 663–674, Nov. 2022, https://doi.org/10.1080/21681163.2021.2021111.

[15] N. Shaukat, J. Amin, M. Sharif, F. Azam, S. Kadry, and S. Krishnamoorthy, "Three-Dimensional Semantic Segmentation of Diabetic Retinopathy Lesions and Grading Using Transfer Learning," *Journal of Personalized Medicine*, vol. 12, no. 9, Sep. 2022, Art. no. 1454, https://doi.org/10.3390/jpm12091454.

[16] A. S. Krishnan, D. Clive R., V. Bhat, P. B. Ramteke, and S. G. Koolagudi, "A Transfer Learning Approach for Diabetic Retinopathy Classification Using Deep Convolutional Neural Networks," in *2018 15th IEEE India Council International Conference (INDICON)*, Coimbatore, India, Sep. 2018, https://doi.org/10.1109/INDICON45594.2018.8987131.

[17] A. D. Algarni, N. Alturki, N. F. Soliman, S. Abdel-Khalek, and A. A. A. Mousa, "An Improved Bald Eagle Search Algorithm with Deep Learning Model for Forest Fire Detection Using Hyperspectral Remote Sensing Images," *Canadian Journal of Remote Sensing*, vol. 48, no. 5, pp. 609–620, Sep. 2022, https://doi.org/10.1080/07038992.2022.2077709.

[18] A. Khan, F. Aftab, and Z. Zhang, "Self-organization based clustering scheme for FANETs using Glowworm Swarm Optimization," *Physical Communication*, vol. 36, Oct. 2019, Art. no. 100769, https://doi.org/10.1016/j.phycom.2019.100769.

[19] J. Saldanha, S. Chakraborty, S. Patil, K. Kotecha, S. Kumar, and A. Nayyar, "Data augmentation using Variational Autoencoders for improvement of respiratory disease classification," *PLOS ONE*, vol. 17, no. 8, 2022, Art. no. e0266467, https://doi.org/10.1371/journal.pone.0266467.

[20] "Diabetic Retinopathy Detection," *Kaggle*, 2015. https://kaggle.com/competitions/diabetic-retinopathy-detection.