Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 9 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i09.49353

PAPER

Hybrid Classification Approach Utilizing DenseUNet+ for Diabetic Macular Edema Disorder Detection

Laith Abualigah[1–4](✉), Mohammad H. Almomani[5], Raed Abu Zitar[6], Mohammad Sh. Daoud[7], Hazem Migdady[8], Amer Al-Rahayfeh[9]

[1]Computer Science Department, Al al-Bayt University, Mafraq, Jordan

[2]Computer Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman, Jordan

[3]MEU Research Unit, Middle East University, Amman, Jordan

[4]Applied Science Research Center, Applied Science Private University, Amman, Jordan

[5]Department of Mathematics, Facility of Science, The Hashemite University, Zarqa, Jordan

[6]Sorbonne Center of Artificial Intelligence, Sorbonne University – Abu Dhabi, Abu Dhabi, UAE

[7]College of Engineering, Al Ain University, Abu Dhabi, UAE

[8]CSMIS Department, Oman College of Management and Technology, Barka, Oman

[9]Department of Computer Science, Al Hussein Bin Talal University, Maan, Jordan

aligah.2020@gmail.com

ABSTRACT

Diabetic macular edema (DME) poses a significant threat to vision. It is characterized by the enlargement of the macula due to the accumulation of plasma in the extracellular space of the retina. Detection of DME, crucial for timely intervention, traditionally relies on manual inspection of images, which is time-consuming and prone to human error. Leveraging advancements in computer-assisted diagnostics, this study proposes a novel approach utilizing the DenseUNet+ architecture tailored for precise segmentation across diverse image modalities. The proposed method integrates data from four modalities within dense block structures, followed by linear operations and concatenation to enhance feature representation. Evaluation using ResNet101V2 and DenseNet201 demonstrates superior performance, with accuracy exceeding 95% and 99%, respectively, showcasing their efficacy in screening retinal optical coherence tomography (OCT) images for DME. This research highlights the potential of deep learning techniques to enhance ophthalmologists' abilities to efficiently screen, diagnose, and manage DME, thus reducing the risk of irreversible vision loss.

KEYWORDS

diabetes mellitus (DM), diabetic macular edema (DME), ResNet, DenseNet

1 INTRODUCTION

Computerized diagnosis using imaging technology in medicine is of utmost importance in healthcare systems. The prevalence of diabetes mellitus (DM) has increased in modern times due to changes in social habits, work environments, nutrition, and daily routines. The World Health Organization (WHO) has raised concerns about the rapid increase in the diabetes epidemic attributed to lifestyle changes among individuals [1]. Diabetes induces metabolic and molecular changes in various parts of the body, leading to microvascular and structural issues. Diabetes can significantly impair eyesight. Ocular diseases such as diabetes-related retinopathy (DR), detached retinas (RD), cataracts, diabetes-related macular edema (DME), and glaucoma are

Abualigah, L., Almomani, M.H., Zitar, R.A., Daoud, M.Sh., Migdady, H., Al-Rahayfeh, A. (2024). Hybrid Classification Approach Utilizing DenseUNet+ for Diabetic Macular Edema Disorder Detection. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(9), pp. 113–129. https://doi.org/10.3991/ijoe.v20i09.49353

Article submitted 2024-02-26. Revision uploaded 2024-03-27. Final acceptance 2024-03-27.

caused by various factors. Around 135 million people worldwide suffer from visual difficulties caused by DR and DME. Diabetes patients with a 20-year or longer history are at a higher risk for DME, accounting for approximately 29% [2]. The percentage of people experiencing diabetes-related blindness is increasing over time due to the rapid expansion of the population and the daily rise in diabetic patients. The current population lacks sufficient access to equipment, resources, and healthcare specialists. As a result, there has never been a greater demand for a mass-screening automated diagnostic instrument capable of accurately diagnosing and classifying disease stages. To diagnose DME, the ophthalmologist will often capture a digital retinal image.

Diabetes is defined as having higher levels of glucose compared to the normal range. Diabetes may result in several visual complications, including diabetic retinopathy, diabetic macular edema, cataracts, and glaucoma. Diabetic retinopathy is a condition that affects the blood vessels in the retina and may result in the formation of microaneurysms, hemorrhages, as well as soft and hard exudates. DME, or diabetic macular edema, is the progressive phase of diabetic retinopathy that specifically affects the macula region. DME causes fluid to accumulate in the macula, leading to edema. The retina's central part, the macula, is responsible for clear vision [3]. DME, an eye condition that typically develops in patients already diagnosed with DR, can lead to vision impairment. Poor management of blood sugar levels can exacerbate medical complications, increasing the risk of blindness for individuals with DME. While DME can manifest at any stage of DR, it is more likely to occur in the later stages of the disease's progression. The condition involves the accumulation of fluid in the macula, leading to its enlargement. The macula, located at the center of the retina and responsible for central vision, can sustain damage, leading to the loss of central vision. Common symptoms of DME include blurred vision, double vision, the presence of floaters, and ultimately, blindness if left untreated. [4] The fundus image comprises several key structural features, including the optic disc (OD), macula area, and vascular network, as depicted in Figure 1. In contrast to the vascular network composed of blood vessels, the bright circular region represents the OD. The fovea, a dark area within the macula region, is highlighted as an approximate marker of the macula, depicted as a circular shape with green dots. There is a concentration of blood vessels in the orbital dilator (OD), which is where the optic nerve emerges from the eyeball and enters the brain.

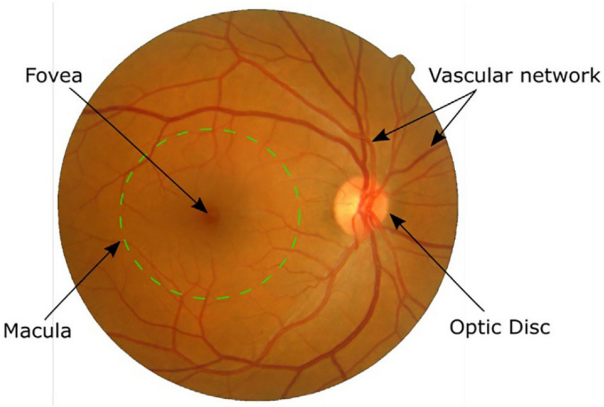

Fig. 1. The main anatomical elements of a fundus fluorescein angiography image

In color fundus imaging, the macula appears as a bright pinkish-yellow area that is spherical in shape. On the other hand, in red-free fundus imaging, the macula

appears as a region of high intensity. Found inside the main blood vessels, this area appears darker in color in fundus pictures and is densely populated with cones. It is also found in the arterial system. When it comes to central vision and color perception, the fovea, located in the middle of the macula, is an extremely important component. Individuals with diabetes are at a significant risk of developing exudates in this area due to the accumulation of leaky lipids from damaged capillaries, which can lead to diabetic maculopathy or diabetic macular edema.

There are several factors that contribute to the development of DME, including the accumulation of lipids in the retinal layer, the breakdown of the blood-retinal barrier, and the thickening of the retina leading to the formation of exudate in the macula [5]. Different types of DME are classified as clinically significant macular edema (CSME) or non-clinically significant macular edema (non-CSME) based on the geographic location, size, and quantity of exudates. This classification is represented in Figure 2. A retina is considered to be in good condition if it does not include any exudates (as seen in Figure 1) or if it has exudates that extend beyond the macular area. The exudates present in cases of non-CSME are situated away from the fovea, ensuring that the central vision remains unaffected. On the other hand, the condition known as CSME is characterized by the presence of exudate deposition on or near the fovea, which affects central vision. Figure 2 displays retinal images illustrating a healthy condition, non-CSME, and CSME. These photos feature the retina.

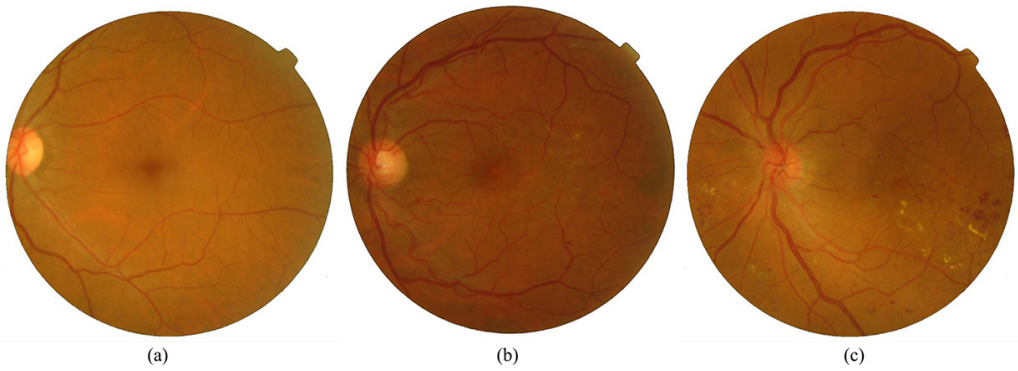

Fig. 2. Diabetic macular edema at different phases (a) Healthy retina, (b) Non-CSME, (c) CSME

Ophthalmologists diagnose and examine DME using fundoscopy. Ophthalmologists use fundus fluorescein angiography, or optical coherence tomography, to identify and analyze retinal diseases of various types and stages. Ophthalmologists typically examine the fundus or optical coherence tomography (OCT) images to detect any eye abnormalities. Nevertheless, detecting early-stage abnormalities solely through visual inspection of images is challenging. Hence, the development of a computer-aided diagnostic (CAD) tool capable of identifying early-stage retinal abnormalities would be advantageous in reducing the incidence of false negatives.

In recent years, the convergence of mobile technologies and healthcare has ushered in transformative advancements in medical diagnosis and management. Among various health challenges, DME stands out as a significant concern. It is characterized by the accumulation of leaked plasma in the macula, which can lead to vision impairment if left untreated. The traditional approach to DME detection relies heavily on manual inspection of retinal images, which is not only time-consuming but also subject to human error. With the proliferation of mobile technologies and the increasing demand for efficient healthcare solutions, there is a pressing need for innovative approaches to automate and streamline the detection and management

of DME. Leveraging the power of deep learning and mobile platforms, this research presents a novel hybrid classification approach based on the DenseUNet+ architecture for the precise identification of DME from retinal images.

By harnessing the capabilities of dense block structures and integrating data from multiple image modalities, the proposed method aims to enhance the accuracy and efficiency of DME detection. Furthermore, the utilization of mobile platforms facilitates seamless integration into clinical workflows, enabling timely diagnosis and intervention. This paper aims to contribute to the growing field of interactive mobile technologies in healthcare by highlighting the potential of deep learning techniques in revolutionizing the screening, diagnosis, and management of DME. Through rigorous evaluation and validation, we have demonstrated the efficacy of our approach in achieving high accuracy and reliability. This empowers healthcare professionals with a powerful tool to combat the vision-threatening complications of diabetes.

The primary objective of this study is to present a comprehensive and effective diagnostic approach for identifying and grading DME. This method aims to differentiate between normal patients and those with either non-center-involved diabetic macular edema (non-CSME) or CSME. Different considerations, such as the presence of other ocular structures resembling exudates, challenges in accurately segmenting the network of blood vessels, and the precise localization of the fovea, all contribute to the complexity of identifying exudates. Difficulties arise due to the presence of artifacts in the photos, variations in light levels ranging from low to high, and blurriness in the images. These issues are addressed by the CAD system, which aims to achieve better performance compared to previous approaches. The CAD system is designed for screening patients, making early diagnoses, and categorizing disease stages. The design simplicity facilitates installation in various environments, requiring only minimal technical skills from operators. It is possible to reduce the number of patients admitted to hospitals through screening. This, in turn, may decrease the workload of medical staff by identifying more complex cases. Through staff time, the reduction of future treatment expenditures, and the preservation of important healthcare resources, it is projected that implementing this automation in clinical settings would lead to a decrease in the costs associated with clinical care. The key contributions of this study include the identification of the blood vascular network, the accurate localization of the fovea, the segmentation of exudates (lesions), and the categorization of DME phases with significantly higher accuracy compared to existing models in use.

2 RELATED RESEARCH

This section discusses prior research on detecting OD, extracting blood vessels from the fovea, and segmenting exudates.

2.1 Detection of optic disc

When photographs of the retinal fundus are taken using a fundus camera, the digital processing of these images often includes localizing the OD and segmenting the borders of the fundus. It is essential to accurately segment the OD since differences in its shape, cupping size, cup-to-disc ratio, edge sharpness, swelling, notching, and color might be indicative of a variety of illnesses. In addition, the initial stages for the subsequent processing of fundus images involve the localization and segmentation

of the OD. When it comes to localizing other components of the retina, such as the fovea and the macula, the normal deviation (OD) acts as a reference point.

There is a close relationship between the features of the OD and the exudates, which highlights the importance of localizing the OD as the initial step in diagnosing DME. There is a hybrid method for OD detection that was developed, as mentioned in reference [6]. The identification of discs was first accomplished by using a clustering method called vessel transformation. After that, direction vectors were constructed based on the pattern of the blood vessel network to locate the OD. Phase-portrait analysis was then used to examine the convergence points of the vector field obtained from the vector field. In conclusion, decision model-based rules were used to maximize vessel transformation and phase portrait analysis, leading to an impressive average accuracy of 98.69%. Furthermore, for the purpose of identifying the best threshold value for OD localization, an additional study reported in [7] suggested a method. An accuracy of over 95% was achieved through the use of the elliptical fitting technique in this approach, which aimed to enhance the segmentation of the optical disc boundary.

[8] presents a hybrid method that combines neural networks (NN) with supervised learning approaches, such as support vector machines (SVM). This approach was presented in the field of artificial intelligence. The detection and separation of exudates in retinal images were achieved through the utilization of a pre-trained convolutional neural network (CNN) model, specifically ResNet-50, in conjunction with SVM. The method achieved an accuracy of 98%. In a similar manner, the approach proposed in [9] involved combining a CNN network with a decoding convolutional block to enhance the sensitivity from 94.8% to 97.48%. For the purpose of segmenting the OD and the optic cup, various CNN models were employed, resulting in an accuracy of 98% for the OD segmentation.

When it comes to computer-assisted diagnostic systems, the precise localization and segmentation of the OD in retinal fundus images are crucial processes essential for detecting disorders like DME. Within the context of subsequent investigations of retinal components, the OD serves as an essential reference point and plays a critical role in identifying anomalies that may indicate various ocular disorders. The hybrid methodologies and approaches mentioned in the cited publications offer promising outcomes in object detection and segmentation, achieving remarkable accuracies exceeding 95%. Through the utilization of NN-based methodologies, direction vectors from vessel networks, and clustering algorithms, these approaches showcase the effectiveness of employing contemporary computational techniques to enhance the precision and sensitivity of OD localization. Furthermore, the integration of CNNs and supervised learning techniques such as SVMs demonstrates the potential to achieve high accuracy in detecting and segmenting OD and associated features. This integration can facilitate early disease diagnosis and treatment planning. In general, the continuous advancement of digital processing techniques holds great promise for enhancing the effectiveness and efficiency of automated disease detection systems in the field of ophthalmology. This progress will ultimately lead to improved patient care and outcomes.

2.2 Extraction of blood vessels and foveal localization

The technique of retinal blood vessel segmentation is an essential method for identifying alterations in blood vessels and locating them in the retina. For the purpose of treating age-related macular degeneration, the retinal map is created

automatically and utilized. Within the center of the macula is where you will find the fovea. It is possible that damage to the macular region could result in vision loss or even blindness. Lesions located near the fovea can cause severe damage, and the distance between the lesions and the fovea is a clinically significant factor in determining the extent of the damage.

In the paper [10], it was suggested that a single CNN model could be utilized to segment both the fovea and the vascular network simultaneously. In order to identify each effective pixel in the fundus pictures, this CNN network with seven layers requires three input channels simultaneously. Approximately 92.68% of the segments were correctly segmented, according to the authors. Within the retinal fundus image, the blood vascular network is a prominent feature that constitutes a significant portion. An approach based on fuzzy rules was proposed in [11] for recognizing blood vessels. The first phase in this method is to perform morphological reconstruction on the binary image, and the second step is to utilize a gaussian mixture model (GMM) classifier to identify small vessel pixels. The fovea and the vascular network were segmented together in several of the experiments conducted. This is due to the fact that segmenting the blood vessel network in a fundus picture aids in localizing the fovea. A data-driven deep learning technique was used by [12] to determine the location of both the fovea and the optic disc.

2.3 Exudates segmentation

One of the most important steps in the diagnostic procedure for DME is the identification of the exudate. Exudates can appear anywhere in the fundus image and vary in size. For the purpose of ensuring a high level of accuracy in lesion diagnosis, it is necessary to identify all probable exudate locations to prevent false negatives. A review of previous studies on recognizing exudates from fundus photographs is included in this section.

Methods that integrate NN with supervised learning approaches, such as SVM, are described in [13]. The researchers utilized a pre-trained CNN model (ResNet-50) and an SVM to detect and classify exudates in retinal images. A supervised learning system that combines SVM and CNN approaches was proposed in [14] as a means of addressing the challenge of problematic exudate segmentation. The photo preparation approaches were used before the segmentation procedure. It was stated that the accuracy was 99.8% on a small dataset consisting of 89 photographs. Recently conducted studies utilize NN or a combination of supervised learning methodologies, as mentioned previously. Even though NN models have shown promising results with increasing accuracy, they have a number of limitations. These include their computational complexity, the requirement for large datasets, their dependence on data quality, and their sensitivity to variable initialization. An incremental learning strategy was proposed in [15], which utilizes both historical and contemporary segmentation models to improve segmentation results.

2.4 DME classification stage

When it comes to the classification of DME, researchers developed a deep learning model in the article [16]. Techniques such as image enhancement, a focus loss function, cosine annealing learning rate, and weighted random sampling were utilized to enhance the classification performance of the model. Despite the dataset

being relatively small and the classes not being evenly distributed, this was accomplished. Following the generation of class activation maps from fundus images, the system presents these maps in the form of heat maps to indicate potential lesion areas. The dependability of the training model is improved as a result of this, which is beneficial to ophthalmologists. The recent study demonstrated the potential of basic two-dimensional photographs through the application of deep learning. This has eliminated the need for specialists and costly three-dimensional imaging equipment, such as OCT. It was discovered that the model was able to effectively identify DME by using only a two-dimensional color fundus picture [17]. The suggested model also showed the capacity to identify intra- and sub-retinal fluid with accuracy percentages of 0.81 and 0.88, respectively. These capabilities were proven by the model. In the study [18], a deep learning network was trained on cropped pictures of increasing size around the fovea to predict OCT properties from fundus images. Three distinct CNNs were applied by the authors to accurately classify leakages, microaneurysms, and non-perfusion areas (NPA) in the images. These CNNs were DenseNet, ResNet50, and VGG16. During the NPA segmentation process, the U-Net architecture was utilized, offering the capability for both up-sampling and down-sampling. The generation of 64 × 64 feature maps was achieved by employing four max-pooling procedures during the downsampling process. Composition blocks that included batch normalization (BN), Re LU activations, and 3 × 3 convolutions were used before each down sampling operation. A total of four upsamplings were performed on the final feature maps to fully recover the original picture size. The objective of this strategy was to accurately designate NPA zones by utilizing the Attention U-Net architecture.

3 PROPOSED METHOD

"Pre-training" and "fine-tuning" are the two primary processes involved in the execution of transfer learning (TL). Six pre-trained ResNet models and three pre-trained DenseNet models were used in the aforementioned research project to categorize OCT images of individuals diagnosed with DME. In light of the fact that pre-trained models are utilized, the aim of this study is to enhance existing TL models. During the fine-turning process, the initial layers are frozen to preserve the learned characteristics.

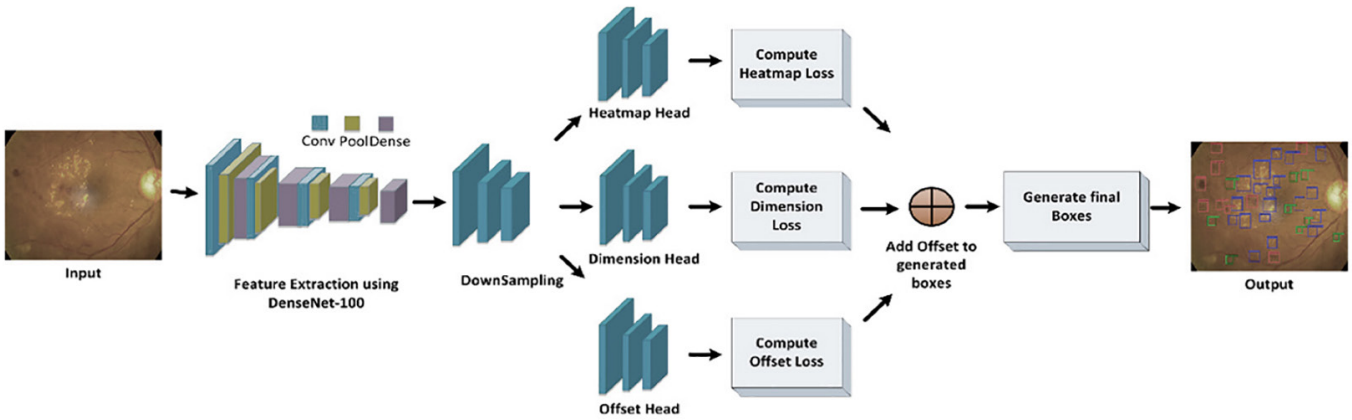

Fig. 3. Flow diagram of proposed technique

The FC layers that utilize SoftMax are the final layers incorporated in the TL model. In contrast to the FC layers, which utilize random parameters, the convolutional layers make use of discriminative filters that are trained based on data specific to a particular domain. It is possible that discriminative filters will be negatively affected if the gradient is allowed to propagate back through the network from random values. It is necessary to fine-tune and retrain the DL architecture to recognize multiple classes. During the fine-tuning process, a new FC layer is constructed and will be integrated after the design development process is finished. To perform fine-tuning, the last layers need to be removed, and a new FC layer needs to be added in their place. It is possible to prevent backpropagation from causing harm to discriminative filters by adjusting the weights of the convolution layers.

The pre-training and fine-tuning stages are the two primary procedures included in the suggested technique for applying TL in the research study. For the purpose of classifying OCT pictures of patients with DME, initially, six pre-trained ResNet models and three pre-trained DenseNet models are applied. The purpose of the research is to increase the accuracy of categorization by enhancing existing TL models with more advanced versions.

During fine-tuning, the early layers of the pre-trained models are frozen to retain the learned characteristics. This ensures that the network does not forget the general features it has learned in the past. For the purpose of retaining the discriminative filters that were learned from the initial domain, this step is absolutely necessary. In the suggested method, the final layers of the TL model consist of fully connected (FC) layers that utilize softmax activation. It is important to note that the convolution layers use discriminative filters that have been learned from the particular domain, while the FC layers initially contain random values. When gradients are allowed to flow through these random parameters during backpropagation, there is a possibility that the learned features in the convolution layers may be disrupted.

The fine-tuning process involves replacing the final FC layers with a new FC layer that will adjust to the specific classification task being carried out. This is done in order to alleviate the issue. By maintaining the weights of the convolution layers and only adjusting the parameters of the FC layers during training, the discriminative filters learned in the convolution layers are preserved. This ensures that the model can reliably categorize OCT images of DME patients. By using this technique, the need for fine-tuning and retraining is acknowledged to successfully detect various classes present in the OCT images. The objective of the proposed strategy is to enhance the performance of TL models in classifying DME images while retaining the valuable features acquired from the pre-trained models. This will be achieved by carefully regulating the changes to the network's parameters, particularly in the FC layers.

3.1 ResNet

In the previous iteration of CenterNet [19], ResNet was deployed to identify significant areas in images and for medical image analysis. Skip connections and identity approaches are utilized by ResNet to manage non-linear transformations. This allows for a direct gradient flow from the rear layers to the front layers by utilizing the identity function. An illustration of the construction of the ResNet-101 framework can be seen in Figure 4. It is important to note that the ResNet-101 model is characterized by a high number of parameters, which might lead to the vanishing gradient issue.

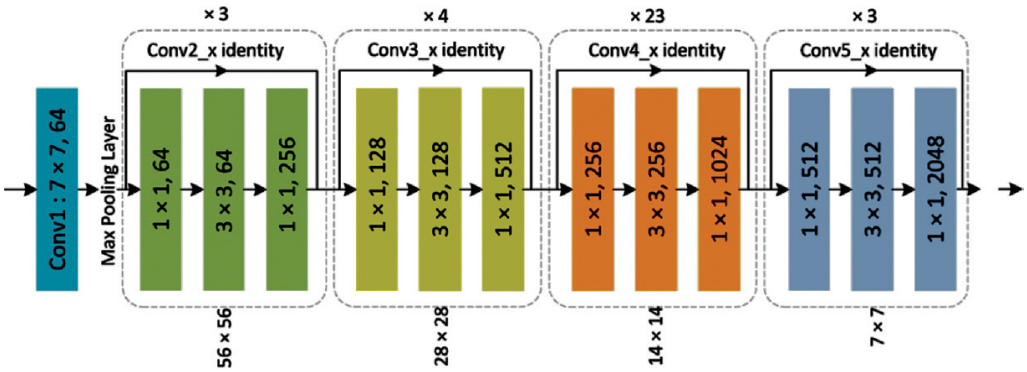

Fig. 4. Depicts the ResNet-101 framework's structure

A deep CNN with 101 layers, known as ResNet, was utilized in this study. The CNN consists of numerous layers that extract and convert input images into hierarchical feature maps. These feature maps may vary from basic elements such as edges and lines to more complex factors such as shapes and colors. It also contains layers, which integrate these characteristics and provide the final probability value of the class. Additionally, there are pooling layers that combine features from multiple levels into a single layer to reduce the dimensionality of the retrieved features. As the number of network layers increases, the algorithm can acquire a greater number of characteristics. Recent research has shown that the depth of a network enhances the accuracy of classification. As a consequence of this, ResNet enables achieving higher accuracy levels from significantly deeper networks compared to shallower networks when performing image categorization tasks.

ResNet, short for Residual Network, is a deep convolutional NN architecture that introduces residual connections to enable the effective training of very deep networks. Let's delve into the mathematical details of ResNet.

- Residual block

  The basic building block of ResNet is the residual block. It consists of two convolutional layers (typically 3×3 filters) with Rectified Linear Unit (ReLU) activations. The output of the second convolutional layer is added to the input of the block through a skip connection. Mathematically, the output of the residual block can be expressed as:

  $$\text{Output} = \text{ReLU}(\text{Batch Norm}(W2 \cdot \text{ReLU}(\text{Batch Norm}(W1 \cdot \text{Input}) + \text{Input})) + \text{Input})$$

  Here,

  1W and 2W represent the weights of the convolutional layers, while input denotes the input to the residual block.
- Residual connection

  The key innovation of ResNet is the inclusion of identity shortcut connections, also known as skip connections. These connections bypass one or more layers and directly add the input of a layer to the output of one or more subsequent layers. Mathematically, the residual connection can be represented as:

  $$\text{Output} = \text{ReLU}(\text{Batch Norm}(2 \cdot \text{ReLU}(\text{Batch Norm}(1 \cdot \text{Input})) + \text{Input})$$
  $$\text{Output} = \text{ReLU}(\text{Batch Norm}(W \cdot \text{ReLU}(\text{Batch Norm}(W1 \cdot \text{Input})) + \text{Input})$$

  Here, the addition operation directly adds the input to the output of the second convolutional layer.

- Residual learning

  Residual learning involves learning residual functions instead of the original underlying mappings. The residual function is defined as the difference between the output and input of a layer. Mathematically, the residual function

  F(x) of a residual block can be expressed as:

  $$F(x) = \text{ReLU}(\text{Batch Norm}(2\ \text{ReLU}(\text{Batch Norm}(1)) + F(x)$$
  $$= \text{ReLU}(\text{Batch Norm}(W2\ \text{ReLU}(\text{Batch Norm}(W1\ x)) + x)$$

  The output of the residual block is obtained by adding the residual function.

  F(x) to the input x: Output = F(x) + x
- Training and optimization

  During training, ResNet is optimized using gradient descent-based optimization algorithms such as stochastic gradient descent (SGD) or Adam. The weights of the network are updated iteratively to minimize a predefined loss function, typically categorical cross-entropy for classification tasks. The gradients are computed using backpropagation, and techniques such as weight decay and dropout may be employed to prevent overfitting.

  In your research, you can implement ResNet as a feature extractor or classifier for analyzing OCT images of patients with DME. Fine-tuning or training ResNet on your specific dataset may further enhance its performance for your particular classification task.

3.2 DenseNet

The DenseNet framework offers a solution to the issue of insufficient sequential location information for top-level features by showcasing sophisticated transformations. Because it enhances feature propagation and reusability, DenseNet is well-suited for recognizing DR and DME, and it also enables faster training [20, 21]. In addition to having the same number of layers as ResNet-101, DenseNet-100 consists of four modules that are closely interconnected. In comparison to the ResNet-101 model, the DenseNet-100 framework has fewer parameters, potentially offering computational benefits. Figure 5 illustrates that DenseNet relies significantly on the database, which presents the feature mappings (FMs) of the n–1 layer in the form of $N \times N \times M0$.

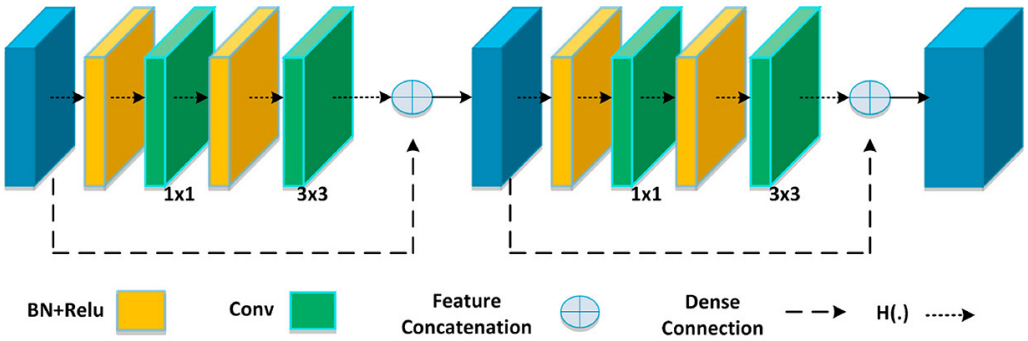

Fig. 5. Illustration of DenseNet

The algorithm makes use of various approaches, including BN, Rectified Linear Unit (ReLU) activation, a 1×1 convolution layer (ConvL) to decrease the total number of channels, and a 3×3 ConvL for key-point rearrangement. The dense links connect

the n–1 and, in layers, perform concatenation with the output of H. The relationships are indicated by the long-dashed arrows. In conclusion, the n + 1 layer produces sensors for the purpose of peer review (N Sensors, 2021, 21, x). 7 out of 17 3.3.1. Feature Extraction. In DenseNet-100, there are four modules tightly interconnected, with the same number of layers as in ResNet-101. DenseNet-100, on the other hand, has a smaller number of individual parameters. Creating dense connections between layers is achieved by DenseNet through the utilization of dense blocks. Each layer is connected to the preceding one using the feedforward approach of DenseNet. All of the layer feature maps have access to the gradient and loss functions, which leads to an improvement in the gradient flow across the network [22]. For the purpose of DME classification, pre-trained DenseNet models such as D121, D169, and D201 were optimized. For the purpose of multi-feature fusion approaches, deep learning techniques are used in the articles [23, 24].

4 RESULTS AND DISCUSSION

4.1 Dataset

To evaluate various TL models, DME and normal OCT images from the Mendeley dataset [25] were utilized. This dataset consists of 11,349 DME and 51,140 normal images for training and testing. In order to evaluate the performance of the best pre-trained model on a smaller OCT dataset, only 500 images were used in the experiment. The selected OCT images were divided into two sets: 400 for training and 100 for testing. The publicly available OCT dataset contains both DME and normal images with an initial individual size of 1024 × 496 pixels. To accommodate the selected input image size, OCT scans were resized to 224 × 224 pixels.

4.2 Results

In this investigation, the performance of six pre-trained CNN models, namely R101, R101V2, R152, D121, D169, and D201, was evaluated for feature extraction. These models were trained using 200 images of DME and 200 images of normal retinal OCT obtained from a public dataset. The training was conducted using Google Colab. Subsequently, the test set comprised 100 OCT images, with 50 classified as DME and the remaining 50 as normal.

Accuracy serves as a crucial metric for evaluating classification algorithms because it quantifies the number of correctly classified instances by the model. The pre-trained models underwent training four times with varying hyperparameter settings to determine the optimal configuration. Table 1 presents the various hyperparameter settings used in the experiment, while Table 2 shows the accuracy results for each model under six different hyperparameter configurations.

Table 1. Variable estimates for different sets

Settings	Parameter Values
Set 1	Optimizer-Adam, LR = 0.001, Epochs = 75
Set 2	Optimizer-Adam, LR = 0.0001, Epochs = 75
Set 3	Optimizer-SGD, LR = 0.001, Epochs = 75
Set 4	Optimizer-SGD, LR = 0.0001, Epochs = 75

Table 2. Accuracy values derived for different models and sets

Model	Set 1	Set 2	Set 3	Set 4
R101	79%	93%	72%	73%
R101V2	85%	69%	91%	95%
R152	55%	92%	90%	89%
D121	95%	91%	90%	93%
D169	82%	98%	81%	97%
D201	95%	99%	84%	69%

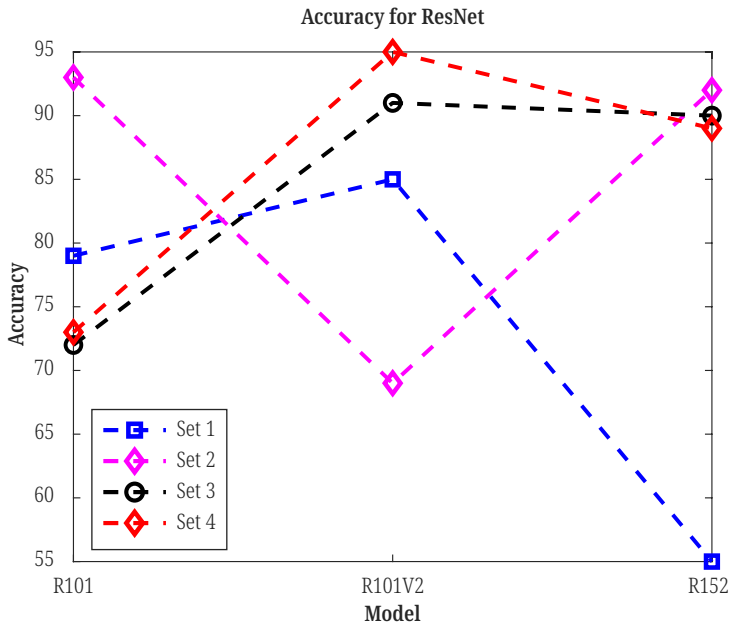

Fig. 6. Illustration of accuracy for ResNet

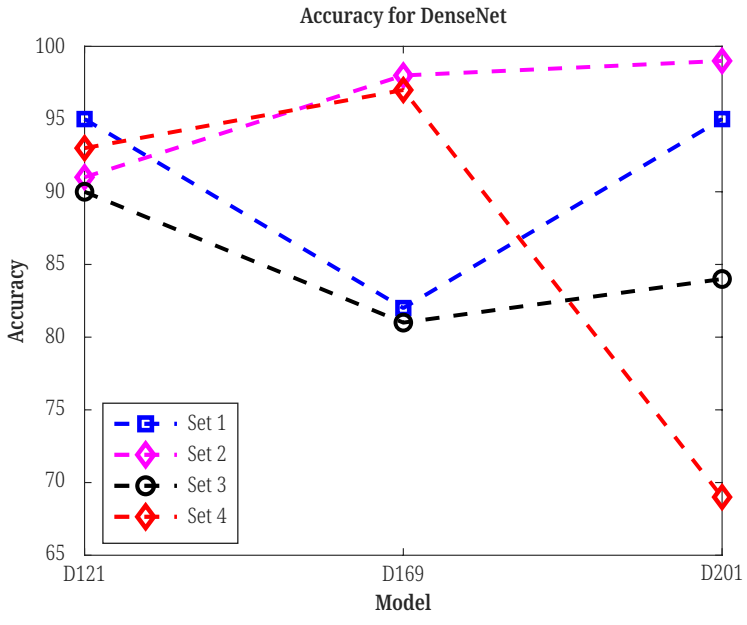

Fig. 7. Illustration of accuracy for DenseNet

The illustrations of the accuracy for ResNet and DenseNet are shown in Figures 6 and 7. Models R101 and R152 typically require longer training times, are susceptible to overfitting, and may not always be necessary. In experiments conducted with the OCT retinal dataset, satisfactory results were achieved using R101 by steadily reducing the learning rate (LR). Among the various settings, all versions of residual networks, except R101V2, performed better with settings S1 and S2. For R101V2, optimal results were obtained with RMSprop and a lower LR. However, the larger versions of ResNet implementations, such as R152, can also produce satisfactory results. Figure 8 depict the loss and accuracy learning curves for ResNet101V2 on Google Colab, respectively.

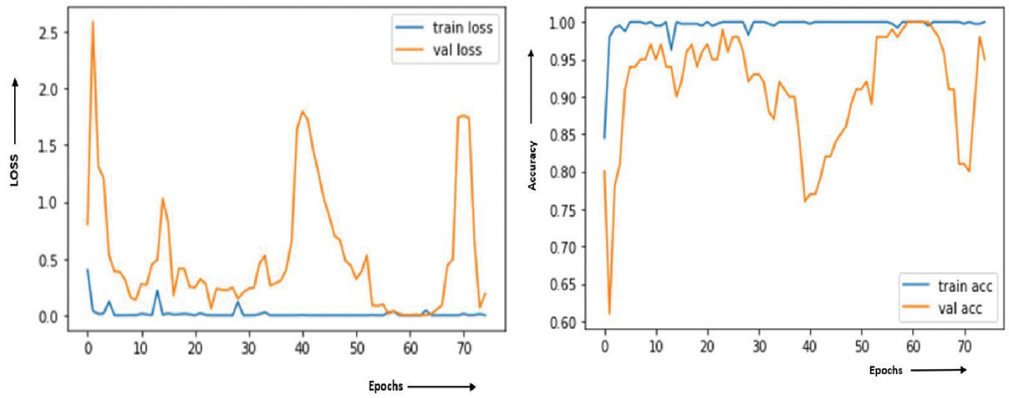

Fig. 8. Training and validation loss and accuracy for ResNet101V2

Unlike ResNet, DenseNet may benefit from increased network depth. DenseNet implementations D201, D169, and D121 demonstrated strong performance with settings S2 and S1, respectively. Upon scrutinizing the highest accuracy values achieved by each model, it becomes evident that Model D201 delivered exceptional results. DenseNet distinguishes itself by utilizing feature maps from preceding layers for additional computational input. Figure 9 illustrate the loss and accuracy learning curves for DenseNet201 on Google Colab, respectively.

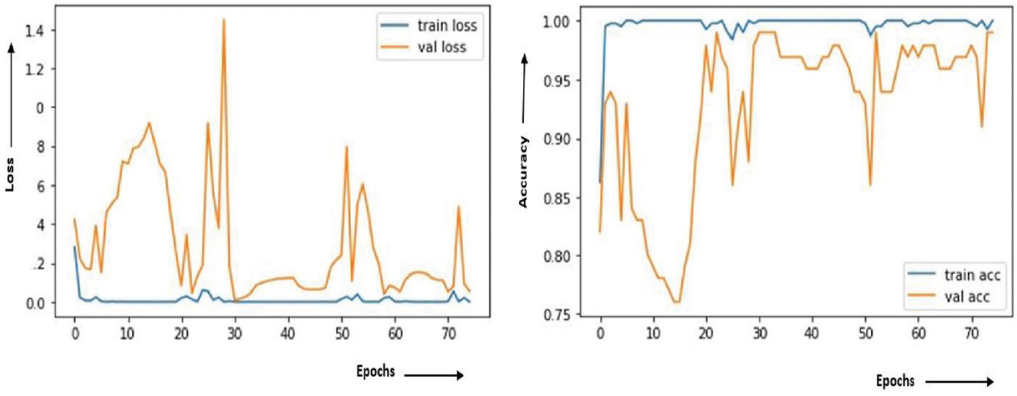

Fig. 9. Training and validation loss and accuracy for DenseNet201

The comprehensive analysis indicates that, for most models, settings S1 and S2 represent the optimal choices among the selected configurations. Figure 10 illustrates the accuracy results obtained for various settings that are considered ideal for DME detection using the pre-trained models.

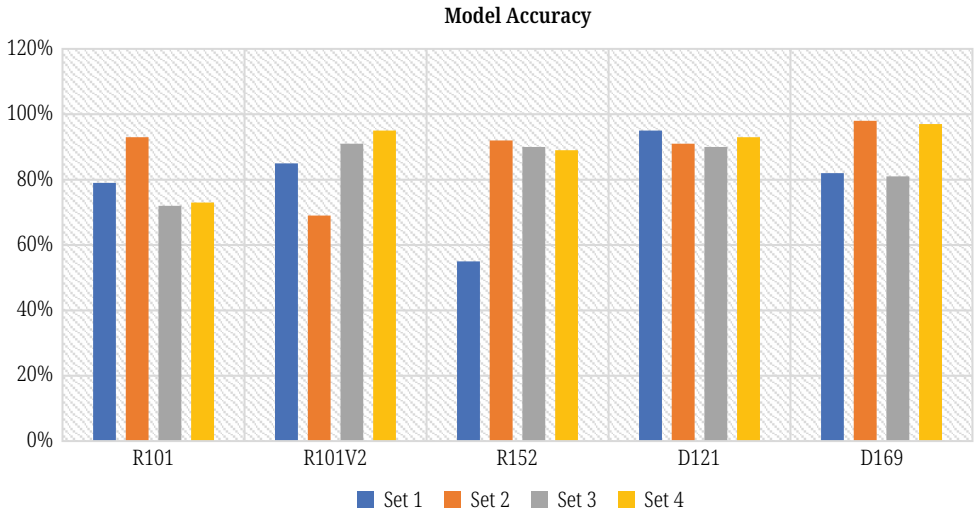

Fig. 10. Depicts the accuracy results

5 CONCLUSION

The algorithm makes use of a variety of approaches, including BN, ReLU activation, a 1 × 1 ConvL to decrease the total number of channels, and a 3 × 3 ConvL for key-point rearrangement. The dense links connect the n–1 and in layers, performing concatenation with the output of H. The relationships are indicated by the long-dashed arrows. In conclusion, the n + 1 layer produces N Sensors 2021, 21, and x for the purpose of peer review. 7 out of 17 (3.3.1). Feature Extraction. There are four modules that are tightly interconnected in DenseNet-100, and the number of layers is equivalent to ResNet-101. DenseNet-100, on the other hand, has a smaller number of individual parameters. Creating dense connections between layers is achieved by DenseNet through the utilization of dense blocks. Each layer is connected to the preceding one using the feedforward approach of DenseNet. All of the layer feature maps have access to the gradient and loss functions, which leads to an improvement in the gradient flow across the network [22]. For the purpose of DME classification, pre-trained DenseNet models D121, D169, and D201 were optimized. For the purpose of multi-feature fusion approaches, deep learning techniques are used in the articles [23, 24].

Building upon the efficient CAD system presented in this study for detecting DME, several promising future directions emerge. Firstly, there is a need to integrate multimodal data beyond OCT images, such as fundus photography and patient history records, to provide a more comprehensive understanding of diseases and enhance diagnostic accuracy. Secondly, the development of real-time CAD systems that can be directly integrated into medical imaging equipment would facilitate faster decision-making during patient examinations. Additionally, extending similar methodologies to detect and classify other diseases, employing advanced data augmentation techniques, and synthesizing realistic medical images could enhance the CAD system's applicability and robustness. Clinical validation studies and collaboration with medical professionals are crucial for ensuring regulatory compliance and seamless integration into clinical workflows. Continual model improvement, personalized medicine approaches, and enhancing the interpretability and explicability of CAD models are essential for advancing the field and improving healthcare outcomes globally.

6 COMPLIANCE WITH ETHICAL STANDARDS

6.1 Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

6.2 Ethical approval

This paper does not contain any studies with human participants or animals performed by any of the authors.

6.3 Informed consent

Informed consent was obtained from all individual participants included in the study.

6.4 Data availability statements

Data is available from the authors upon reasonable request.

7 ACKNOWLEDGMENT

Not Applicable

7.1 Funding

Not Applicable

7.2 Author contributions

All authors read and approved the final paper.

8 REFERENCES

[1] H. King, R. E. Aubert, and W. H. Herman, "Global burden of diabetes, 1995–2025: Prevalence, numerical estimates, and projections," *Diabetes Care,* vol. 21, pp. 1414–1431, 1998. https://doi.org/10.2337/diacare.21.9.1414

[2] R. Klein, B. E. Klein, S. E. Moss, M. D. Davis, and D. L. DeMets, "The Wisconsin epidemiologic study of diabetic retinopathy : Iv. diabetic macular edema," *Ophthalmology,* vol. 91, no. 12, pp. 1464–1474, 1984. https://doi.org/10.1016/S0161-6420(84)34102-1

[3] U. R. Acharya, M. R. K. Mookiah, J. E. Koh, J. H. Tan, S. V. Bhandary, A. K. Rao, Y. Hagiwara, C. K. Chua, and A. Laude, "Automated diabetic macular edema (DME) grading system using DWT, DCT Features and maculopathy index," *Comput. Biol. Med.*, vol. 84, pp. 59–68, 2017. https://doi.org/10.1016/j.compbiomed.2017.03.016

[4] T. Nazir, A. Irtaza, A. Javed, H. Malik, D. Hussain, and R. A. Naqvi, "Retinal image analysis for diabetes-based eye disease detection using deep learning," *Appl. Sci.,* vol. 10, no. 18, p. 6185, 2020. https://doi.org/10.3390/app10186185

[5] D. A. Antonetti, E. Lieth, A. J. Barber, and T. W. Gardner, "Molecular mechanisms of vascular permeability in diabetic retinopathy," *Seminars in Ophthalmology*, vol. 14, pp. 240–248, 1999. https://doi.org/10.3109/08820539909069543

[6] N. Muangnak, P. Aimmanee, and S. Makhanov, "Automatic optic disk detection in retinal images using hybrid vessel phase portrait analysis," *Med. Biol. Eng. Comput.*, vol. 56, pp. 583–598, 2018. https://doi.org/10.1007/s11517-017-1705-z

[7] A. S. Abdullah, Y. E. Özok, and J. Rahebi, "A novel method for retinal optic disc detection using bat meta-heuristic algorithm," *Med. Biol. Eng. Comput.*, vol. 56, pp. 2015–2024, 2018. https://doi.org/10.1007/s11517-018-1840-1

[8] L. Wang, H. Liu, Y. Lu, H. Chen, J. Zhang, and J. Pu, "A coarse-to-fine deep learning framework for optic disc segmentation in fundus images," *Biomed. Signal Process. Control,* vol. 51, pp. 82–89, 2019. https://doi.org/10.1016/j.bspc.2019.01.022

[9] H. Veena, A. Muruganandham, and T. S. Kumaran, "A novel optic disc and optic cup segmentation technique to diagnose glaucoma using deep learning convolutional neural network over retinal fundus images," *J. King Saud Univ.-Comput. Informat. Sci.,* vol. 34, no. 8, pp. 6187–6198, 2022. https://doi.org/10.1016/j.jksuci.2021.02.003

[10] R. Kamble, M. Kokare, G. Deshmukh, F. A. Hussin, and F. Mériaudeau, "Localization of optic disc and fovea in retinal images using intensity-based line scanning analysis," *Comput. Biol. Med.,* vol. 87, pp. 382–396, 2017. https://doi.org/10.1016/j.compbiomed.2017.04.016

[11] F. Orujov, R. Maskeliunas, R. Damaševičius, and W. Wei, "Fuzzy based image edge detection algorithm for blood vessel detection in retinal images," *Appl. Soft Comput.,* vol. 94, p. 106452, 2020. https://doi.org/10.1016/j.asoc.2020.106452

[12] B. Al-Bander, W. Al-Nuaimy, B. M. Williams, and Y. Zheng, "Multiscale sequential convolutional neural networks for simultaneous detection of fovea and optic disc," *Biomed. Signal Process. Control,* vol. 40, pp. 91–101, 2018. https://doi.org/10.1016/j.bspc.2017.09.008

[13] P. Khojasteh, L. A. P. Júnior, T. Carvalho, E. Rezende, B. Aliahmad, J. P. Papa, and D. K. Kumar, "Exudate detection in fundus images using deeply-learnable features," *Comput. Biol. Med.*, vol. 104, pp. 62–69, 2019. https://doi.org/10.1016/j.compbiomed.2018.10.031

[14] N. Theera-Umpon, I. Poonkasem, S. Auephanwiriyakul, and D. Patikulsila, "Hard exudate detection in retinal fundus images using supervised learning," *Neural Comput. Appl.,* vol. 32, pp. 13079–13096, 2019. https://doi.org/10.1007/s00521-019-04402-7

[15] W. He *et al.*, "Incremental learning for exudate and hemorrhage segmentation on fundus images," *Infr. Fusion,* vol. 73, pp. 157–164, 2021. https://doi.org/10.1016/j.inffus.2021.02.017

[16] T. Wu, L. Liu, T. Zhang, and X. Wu, "Deep learning-based risk classification and auxiliary diagnosis of macular edema," *Intell. -Based Med.*, vol. 6, p. 100053, 2022. https://doi.org/10.1016/j.ibmed.2022.100053

[17] A. V. Varadarajan *et al.*, "Predicting optical coherence tomography-derived diabetic macular edema grades from fundus photographs using deep learning," *Nat. Commun.*, vol. 11, no. 130, 2020. https://doi.org/10.1038/s41467-019-13922-8

[18] K. Jin *et al.*, "Automatic detection of non-perfusion areas in diabetic macular edema from fundus fluorescein angiography for decision making using deep learning," *Sci. Rep.,* vol. 10, pp. 1–7, 2020. https://doi.org/10.1038/s41598-020-71622-6

[19] Y. Wang, H. Li, P. Jia, G. Zhang, T. Wang, and X. Xiaoyun Hao, "Multi-scale DenseNets-based aircraft detection from remote sensing images," *Sensors,* vol. 19, p. 5270, 2019. https://doi.org/10.3390/s19235270

[20] S. Zeng and Y. Huang, "A hybrid-pipelined architecture for FPGA-based binaryweight DenseNet with high performance-efficiency," in *Proceedings of the 2020 IEEE High Performance Extreme Computing Conference (HPEC)*, Waltham, MA, USA, 2020, pp. 1–5. https://doi.org/10.1109/HPEC43674.2020.9286185

[21] S. Albahli, T. Nazir, A. Irtaza, and A. Javed, "Recognition and detection of diabetic retinopathy using DenseNet-65 based faster-RCNN," *Comput. Mater. Contin.*, vol. 67, no. 2, pp. 1333–1351, 2021. https://doi.org/10.32604/cmc.2021.014691

[22] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2017, pp. 2261–2269. https://doi.org/10.1109/CVPR.2017.243

[23] B. Lalitha, V. Madhurima, Ch. Nandakrishna, J. Satish Babu, J. N. Chandra Sekhar, and P. Venkat Reddy, "Data augmentation based cross-lingual multi-speaker TTS using DL with sentiment analysis," *ACM Transactions on Asian and Low-Resource Language Information Processing*, 2023.

[24] J. Kumar Raja, N. Praveena, K. Madhavi, J. N. Chandra Sekhar, G. Kumari, D. Bullarao, and M. Venkata Rao, "Deep learning based multilingual speech synthesis using multi feature fusion methods," *ACM Transactions on Asian and Low-Resource Language Information Processing*, 2023.

[25] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, and J. Dong, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018. https://doi.org/10.1016/j.cell.2018.02.010

9 AUTHORS

Laith Abualigah, Computer Science Department, Al al-Bayt University, Mafraq 25113, Jordan; Computer Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman 19328, Jordan; Artificial Intelligence and Sensing Technologies (AIST) Research Center, University of Tabuk, Tabuk 71491, Saudi Arabia; MEU Research Unit, Middle East University, Amman 11831, Jordan; School of Engineering and Technology, Sunway University Malaysia, Petaling Jaya 27500, Malaysia; Applied Science Research Center, Applied Science Private University, Amman 11931, Jordan (E-mail: aligah.2020@gmail.com).

Mohammad H. Almomani, Department of Mathematics, Faculty of Science, The Hashemite University, P.O. Box 330127, Zarqa, 13133, Jordan.

Raed Abu Zitar, Sorbonne Center of Artificial Intelligence, Sorbonne University-Abu Dhabi, Abu Dhabi, UAE.

Mohammad Sh. Daoud, College of Engineering, Al Ain University, 112612 Abu Dhabi, UAE.

Hazem Migdady, CSMIS Department, Oman College of Management and Technology, 320 Barka, Oman.

Amer Al-Rahayfeh, Department of Computer Science, Al Hussein Bin Talal University, Maan, Jordan.