IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 28, 2020, accepted January 9, 2021, date of publication January 18, 2021, date of current version January 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3052477

DiaNet: A Deep Learning Based Architecture to Diagnose Diabetes Using Retinal Images Only

MOHAMMAD TARIQUL ISLAM[1], HAMADA R. H. AL-ABSI[2], ESSAM A. RUAGH[3], AND TANVIR ALAM[2]

[1]Computer Science Department, Southern Connecticut State University, New Haven, CT 06515, USA
[2]College of Science and Engineering, Hamad Bin Khalifa University, Doha 34110, Qatar
[3]Magrabi Eye, Dental and Ear Centre, Magrabi Hospital, Doha 23293, Qatar

Corresponding author: Tanvir Alam (talam@hbku.edu.qa)

This open access publication of this article was supported by the Qatar National Library (QNL), Doha, Qatar. The study was in part supported by the Qatar Biobank (QBB) under Grant QF-QBB-RES-ACC-0236.

ABSTRACT Diabetes is one of the leading fatal diseases globally, putting a huge burden on the global healthcare system. Early diagnosis of diabetes is hence, of utmost importance and could save many lives. However, current techniques to determine whether a person has diabetes or has the risk of developing diabetes are primarily reliant upon clinical biomarkers. In this article, we propose a novel deep learning architecture to predict if a person has diabetes or not from a photograph of his/her retina. Using a relatively small-sized dataset, we develop a multi-stage convolutional neural network (CNN)-based model DiaNet that can reach an accuracy level of over 84% on this task, and in doing so, successfully identifies the regions on the retina images that contribute to its decision-making process, as corroborated by the medical experts in the field. This is the first study that highlights the distinguishing capability of the retinal images for diabetes patients in the Qatari population to the best of our knowledge. Comparing the performance of DiaNet against the existing clinical data-based machine learning models, we conclude that the retinal images contain sufficient information to distinguish the Qatari diabetes cohort from the control group. In addition, our study reveals that retinal images may contain prognosis markers for diabetes and other comorbidities like hypertension and ischemic heart disease. The results led us to believe that the inclusion of retinal images into the clinical setup for the diagnosis of diabetes is warranted in the near future.

INDEX TERMS Convolutional neural network, deep learning, diabetes, machine learning, Qatar, Qatar Biobank (QBB), retina.

I. INTRODUCTION

Diabetes mellitus or diabetes is considered as a collection of metabolic conditions that can predominantly be described by hyperglycemia rising from the deficiency in insulin discharge [1]. The prolonged hyperglycemia of diabetes is correlated with long-term impairment and collapse of heart, kidneys, and microvascular circulation of the retina [2]. Among the diabetic individuals in the USA, almost 30% of them have the tendency of growing diabetic retinopathy (DR), a common complication for diabetic patients which may lead to blindness [3], [4]. Diabetes may adversely affect the vascular system of the retina causing structural change of it [2]. As the changes in vascular structure in retina can provide visual cues for diabetes, most of the clinical guidelines recommended annual retinal screen for the diabetic patients through retinal fundus images or dilated eye examinations [5], [6]. Alternatively, these retinal images could be used to detect diabetes as well, but it requires subjective judgement from the ophthalmologist, and it might be time consuming as well. The human oriented subjective judgement could be avoided if we could implement the automation of retinal image-based diabetes diagnosis in clinical setup. Such automation could alleviate the workload of the ophthalmologist as well as screen a large number of patients objectively within a short amount of time [7]. Though there are multiple studies [8], [9] that aim at detecting diabetic retinopathy from retinal images, none addressed the task of detecting diabetes using retinal images from a holistic point of view. In fact, a significant number of studies focused on

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

the diagnosis of diabetes mainly based on clinical markers e.g., HbA1c, Glucose [10]. Therefore, there is an emerging need for novel, cost-effective, non-invasive and fast diabetes screening solutions that could be implemented easily.

Diabetes in Qatar is the third leading cause of death in its population [11] and one of the topmost five factors in healthcare that affects its economy [12]. The estimated direct and indirect cost for all noncommunicable diseases, including diabetes, in Qatar were $36.2 billion in 2013 [12]. So, early detection of diabetes is of key importance for Qatar to prevent the spread of this disease. There exist many studies, based on Qatar, for the diagnosis and the potential risk factors of diabetes. Shi and Abou-Samra showed that magnesium deficiency is a very common symptom in the Qatari diabetic cohort [13]. Nazeemudeen *et al.* showed that the Qatari diabetic cohort is following the food and drink guidelines, but they need to improve the level of physical activities in their life [14]. Awad *et al.* also showed that physical activity could be considered as an intervention to reduce the diabetes incidence in Qatar [15]. Previously, we developed a machine learning (ML) model which considered 237 clinical measurements based on a Qatar Biobank (QBB) diabetic cohort to predict diabetes groups with over 78% accuracy [10]. Ullah *et al.* considered different clinical measurements and biochemical markers from QBB to develop ML model which discovered magnesium, chloride, insulin as potential risk factors for diabetes in a Qatari diabetes cohort [16]. There exist many studies ([7], [17]–[25]), though not based on Qatar, which focused on DR detection and grading (e.g., ''mild'', ''moderate'', ''severe'', ''proliferate'') using deep learning-based techniques. Though DR is the leading cause of vision loss for the diabetic patients, visual impairment might be caused by other retinal and non-retinal problems, like cataract, glaucoma, macular degeneration, ischaemic optic neuropathy etc. [26]. So, the inclusion of retinal images in diabetes diagnosis will add more insights for better treatment plan of the diabetic patients. But not a single study exists, as per the best of our knowledge, on the diagnosis of diabetes as a disease from a holistic view based on retinal images in Qatar. To fill this gap, we propose DiaNet, a Convolutional Neural Network (CNN)-based approach, to solve the problem of estimating if a person has diabetes given only a photograph of his/her retina—specifically, RGB fundus photography. It is important to emphasize that our goal is more general, i.e., differentiating between diabetic and non-diabetic retinal images, the former of which may include both DR and non-DR cases. While existing studies (mentioned above) mainly focused on detecting DR grading, our study is unique in the sense that it aims to predict diabetes from retinal images irrespective of the existence of DR.

We formulate the problem as a supervised learning task, specifically, a classification problem. This entailed estimating the conditional probability distribution P(D|I, w) of the label D indicating whether a person has diabetes or not given the input I, the retinal photography of the person and w, the parameters of the probability distribution. These parameters are estimated using a data-driven approach by iteratively minimizing a loss function L(Y, Y') parameterized on the actual label Y, and the estimated label Y' computed using the current estimation of the parameters. The proposed computational workflow consists of multi-stage fine-tuning of a neural network using multiple retinal image datasets which yields a superior performance compared to that of a network trained only the target dataset, which can be attributed to its smaller size. Our contributions in this paper can be summarized as follows:

1) We proposed a novel method to predict whether a person has diabetes or not from an image of his/her retina and introduced a multi-stage CNN-based model, DiaNet for the purpose.
2) We introduced a novel dataset for diabetes detection containing retinal images from 492 control and diabetic patients from Qatar.
3) We performed an extensive set of experiments to show that a small dataset is sufficient for diabetes detection from retinal images in order to reach a reasonably high accuracy using the proposed approach.

II. MATERIALS AND METHODS

In this section, we discuss in detail the process involved in the collection, curation, and pre-processing of the dataset as it pertains to this work and the development of our proposed solution for the problem at hand. Our proposed approach uses two datasets to achieve state-of-the-art in detecting diabetes from retinal images. The larger of these datasets contains retinal images from patients with different stages of DR, while the other, smaller dataset, contains retinal images from diabetic patients and a control group. We describe the smaller dataset first since it is more closely related to the goal of our work and is also one of the contributions of this paper.

A. THE QBB RETINA-IMAGE DATASET

This study was conducted under the regulation of the Ministry of Public Health, Qatar. All procedures were approved by the Institutional Review Board (IRB) of Hamad Medical Corporation, Qatar and only de-identified images were collected from QBB.The dataset consists of retinal images from a diabetes cohort of size 246 and a control group of size 246. The medical practitioners interviewed all the participants at QBB to collect their medical and family history, lifestyle, and their habitual factors. Then both the diabetes and the control groups were determined with the help of QBB medical practitioners and nurses. The diabetes group was free from self-reported diabetic status or HbA1c% >= 6.5. The control group was free from diabetes, obesity, and cardiovascular disease. All the subjects from the cohort were Qatari nationals. The cohort was evenly distributed based on male and female gender (i.e. 50% each) for both groups. Topcon TRC-NW6S retinal camera was used at QBB to capture the ''microscopic'' features of the optic nerve and macula from the participants. The details of the data collection protocol

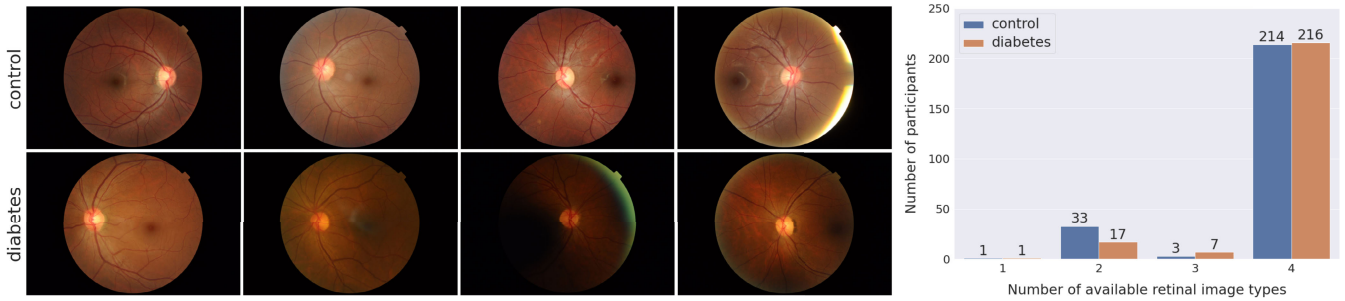

FIGURE 1. A set of retinal images from the QBB dataset is shown in (a), while (b) shows the number of participants with each of the four types of image availability configuration: (left, right) × (macula-centered, disc-centered). Most participants had both macula-centered and disc-centered images from both eyes.

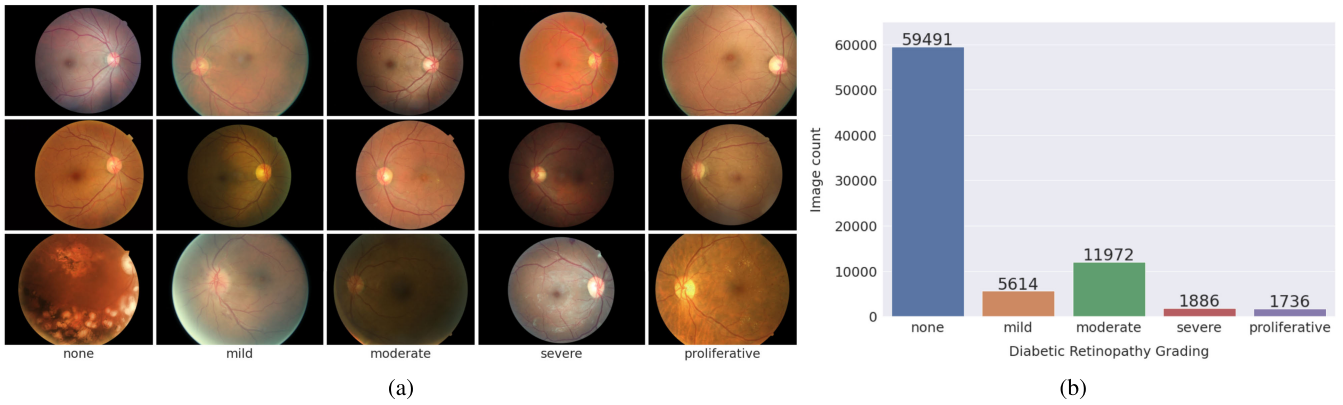

FIGURE 2. Few examples from the EyePACS dataset. Columns in montage (a) show three randomly selected retinal images from each of the five DR gradings. The bar chart in (b) shows the distribution of image counts for the five different DR gradings in the dataset.

can be found in [27], [28]. It is important to emphasize that in QBB, DR was not evaluated or graded at all for the participants. So, the participants from QBB may or may not have DR. The dataset is not publicly available in accordance with the Qatar Biobank data-sharing policy.

The QBB dataset doesn't contain the same number of images for each participant – there were at most two types of images from the left and right eyes: (a) macula-centered images, and (b) disc-centered images. There were 1852 images in total from 492 participants. The mode of the original image sizes was 3696 × 2448 pixels. Figure 1 shows some randomly selected images from the QBB dataset, and the patient counts for each type of retinal image count in the dataset.

B. THE EyePACS DR DATASET

We also used existing retinal images from the EyePACS dataset [29] to train models. The EyePACS dataset contained over 80,000 retinal images and the corresponding labels indicating one of the five different gradings of diabetic retinopathy: "none", "mild", "moderate", "severe", and "proliferative" based on the severity of diabetic retinopathy. Figure 2 shows some samples from the dataset and the number of images in each category.

C. IMAGE PRE-PROCESSING AND AUGMENTATION

We applied multiple pre-processing and data augmentation techniques on our dataset to increase the robustness of our approach. We adopted and slightly modified the pre-processing steps used by the winning solution [30] of the Kaggle Diabetic Retinopathy Detection competition for DR staging considering the EyePACS dataset. Specifically, the following steps were performed in the pre-processing stage. We first extracted the circular region from each image, then resized it so that the radius of the retina is 300 pixels. In the next step, we cropped the outer 10% to eliminate the border-noise followed by a subtraction of the local mean from a 4 × 4 pixel neighborhood, and finally, placed the cropped retina in dark background inside a square-shaped image with tight borders. This sequence of pre-processing steps transformed each image in our dataset from a varying sized image to a 570 × 570 image with a black background so that all the pre-processed images are aligned with each other and have a similar background. For data augmentation, we applied random horizontal flip, and random brightness and contrast perturbation. We performed the same pre-processing and augmentation steps on retinal images from both the QBB and the EyePACS dataset. Figure 3 shows a few example preprocessed images from the QBB dataset.

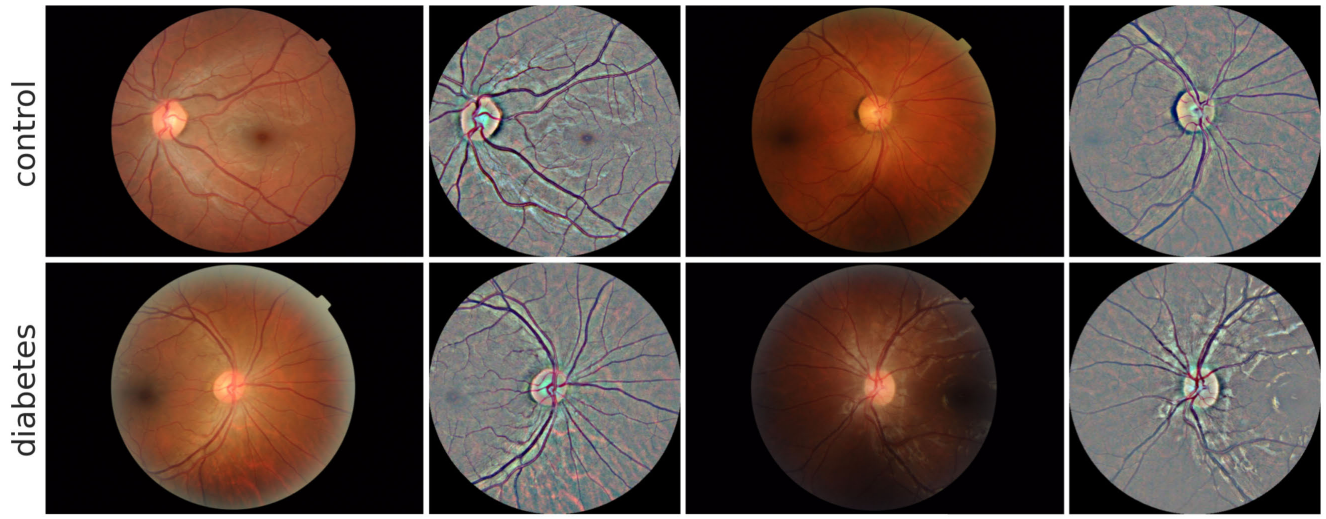

FIGURE 3. A few example images from the QBB dataset before and after pre-processing. The first and third columns show the raw images before pre-processing, while the columns next to them (second and fourth) show the corresponding images after pre-processing.

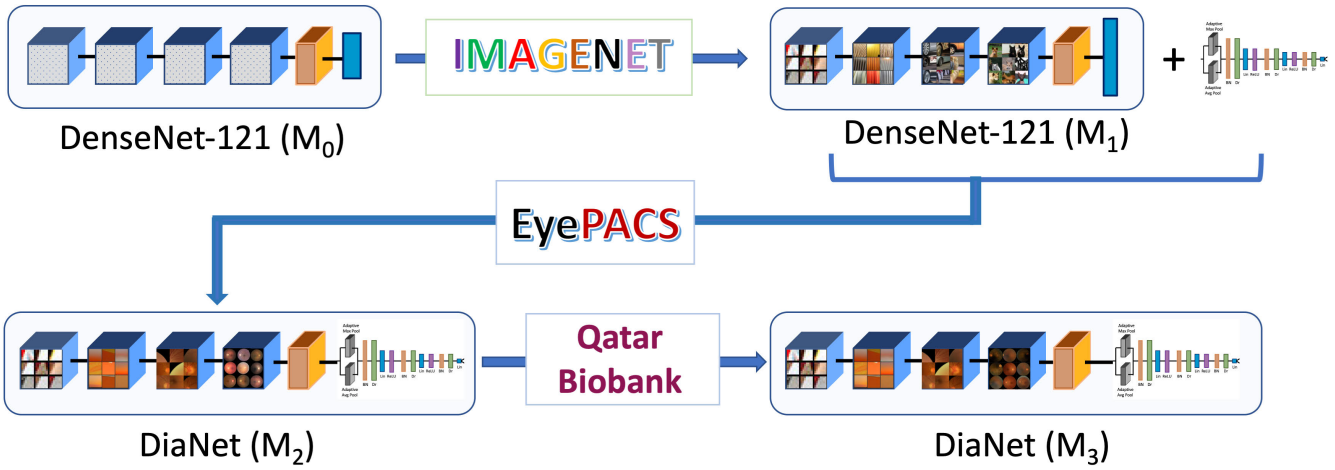

FIGURE 4. A high-level view of the multi-stage fine-tuning approach for training DiaNet. The top left model (M_0) is a randomly initialized DenseNet-121. Training M_0 on the ImageNet dataset yields M_1, which is capable of image classification. Adding a few extra layers and fine-tuning it in on the EyePACS retinal image dataset gives us M_2; a model with retinal image understanding capability. Finally, fine-tuning M_2 on the QBB dataset results in our proposed model (M_3) for diabetes detection from retinal images.

D. DEVELOPMENT OF DiaNet CONSIDERING EyePACS AND QBB DATASET

1) PROPOSED APPROACH

To develop DiaNet, we start with a CNN model M1 that has already been pre-trained on the ImageNet [31] dataset. We then augment the network with a few additional layers before the final layer to increase its ability of understanding more complex patterns from the data. In order to imbue this augmented network with retinal image understanding capabilities, we first fine-tune it on a dataset for the DR detection task, which is defined as the binary task of identifying whether a person has DR or not. At the end of this fine-tuning stage, we have a model M2 that is capable of distinguishing between DR and non-DR retinal images with a reasonably high accuracy. However, our goal is more general, i.e. differentiating between diabetic and non-diabetic retinal images, the former of which includes both DR and non-DR cases. To this end, we take the model M2 and further fine-tune it on the diabetes dataset related to our target task that contains retinal images from non-diabetic subjects and diabetic subjects with and without DR. This gives us M3, a model that can identify diabetic and non-diabetic patients from their retinal images. Our proposed model, DiaNet, used DenseNet-121 [32] as the base CNN. Figure 4 shows an overview of this multi-stage fine-tuning approach.

2) NETWORK ARCHITECTURE

Our proposed approach is based on the popular image-classification CNN called Dense Convolutional Network, known as DenseNet [32]. This architecture allows training

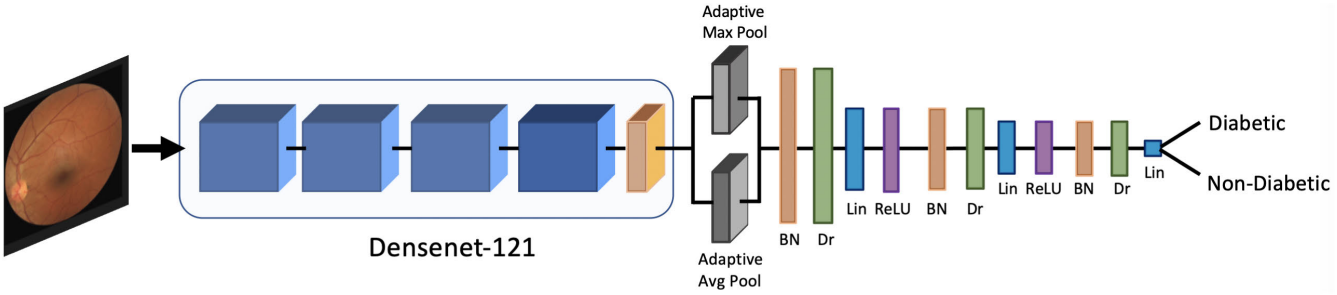

FIGURE 5. The proposed architecture of DiaNet consists of a DenseNet-121 backbone and a few additional layers consisting of a pair of pooling layers followed by three composite layers each consisting of primarily a sequence of batch normalization (BN), dropout (Dr), linear (Lin), and a ReLU activation layer. The final layer contains a single neuron indicating the predicted label (diabetic/non-diabetic). Detailed configurations of these layers can be found in Table 1.

models that are much deeper than AlexNet [33] or VGG [34], and differs from ResNet [35] in the manner in which the feature maps from the previous layers are combined. In ResNets, feature maps are combined downstream using the addition operation, which means feature maps are summed element-by-element along the channel (feature map) dimension, while in DenseNet, they are concatenated, along the same dimension. This allows a layer in DenseNets to have access to the features from all the previous layers, which aids in improving its performance as decisions can be made based on features extracted from the input image at different scales. We used the 121-layer variant of DenseNet which consists of four dense blocks due to its superior performance. In spite of the network being 121-layer deep, overfitting was not an issue due to (i) the dense connections themselves acting as regularizers [32] and (ii) counter-intuitively, the network having less parameters compared to a non-dense CNN [32]. The vanilla DenseNet-121 network trained on ImageNet has 1,000 neurons in the final layer, corresponding to the number of labels in the dataset. Since diabetes detection is a binary task, we needed to change the final layer into a 2-neuron layer. Additionally, we also found that adding a concatenation of global average pooling layer and a global max pooling layer followed by a series of batch normalization, dropout, linear, and ReLU layers before the final linear layer boosted the performance by over 2%. This additional layer configuration is inspired by the approach used by fastai [36] for fine tuning its models. Table 1 shows the output shape and number of parameters for each added layer. We used the method introduced in [37] to initialize these layers. Figure 5 shows a detailed diagram of our proposed network architecture; we call it DiaNet.

E. DiaNet AND ITS VARIANTS

To evaluate the performance of our proposed approach and establish baselines we developed several other models, which comprises of end-to-end CNNs and CNNs combined with Gradient Boosting Machines (GBMs). Specifically, we developed two end-to-end models: DiaNet, and a variant of it, DiaNet_Res50 by replacing its backbone (DenseNet-121) with a ResNet50 instance pre-trained on ImageNet. For the combined approach, we take the activations from the penultimate layer of a CNN and use it as the input to the GBM, which acts as the classifier. We used XGBoost (XGB) as the GBM implementation. This results in two more classification models: DiaNet+XGB, and DiaNet_Res50+XGB. Having established the four candidate models, namely: DiaNet, DiaNet_Res50, DiaNet+XGB, and DiaNet_Res50+XGB, we fine-tuned these in multiple stages. In the first stage, we fine-tuned these four models using the EyePACS dataset and in the second stage on the QBB dataset.

TABLE 1. Details of the layers added at the end of Densenet-121 in DiaNet. The first two pooling layers were concatenated into a single input of size 2048 for the next layer.

Layer Name	Output Size	Number of Parameters
AdaptiveMaxPool2D	[1024, 1, 1]	0
AdaptiveAvgPool2D	[1024, 1, 1]	0
Flatten + Concat	[2048]	0
BatchNorm1D	[2048]	4096
Dropout	[2048]	0
Linear	[512]	1,049,088
ReLU	[512]	0
BatchNorm1D	[512]	1024
Dropout	[512]	0
Linear	[256]	131,328
ReLU	[256]	0
BatchNorm1D	[256]	512
Dropout	[256]	0
Linear	[2]	512

F. DEVELOPMENT OF QBBNet AND ITS VARIANTS CONSIDERING ONLY THE QBB DATASET

Unlike DiaNet, we considered only the QBB dataset to develop QBBNet. Similar to DiaNet, QBBNet considered DenseNet-121 and ResNet50 for configuring end-to-end CNNs and their variants with an XGB as the classifier. We then fine-tuned the models on the QBB dataset only. This results in a total of four additional experimental configurations for QBBNet: QBBNet (with a DiaNet-like architecture), QBBNet_Res50 (ResNet50 backbone instead of a DenseNet121), QBBNet+XGB (XGB classifier on extracted features from QBBNet), and

QBBNet_Res50+XGB (ResNet-50 backbone and XGB as the classifier on extracted features from QBBNet).

G. EXPERIMENT SETUP

Model selection, generalization, and performance estimation in the fine-tuning stage on the QBB dataset were carried out using nested cross-validation [38] to prevent data leakage. We used a 5-fold setup in both inner and outer folds. To obtain even more consistent results, we conducted each such experiment n times (n=5) and computed the arithmetic mean of the representative metrics for reporting. Due to the enormous size of the EyePACS dataset, only a single train-validation-test split in the fine-tuning stage using this dataset was sufficient to ensure that these sets have similar distributions.

The data was pre-processed and augmented using the same approach in multi-stage tuning for DiaNet and single-stage tuning for QBBNet. The Dropout layers we added in our networks used 0.4 as the dropout probability. For minimizing the loss, we used the AdamW optimizer with a One-Cycle Learning Rate Schedular [39].

For multi-stage fine-tuning using the EyePACS and the QBB dataset, models were first trained on the former. To this end, we binarized the labels so that non-DR cases and the DR cases (''mild'', ''moderate'', ''severe'', and ''proliferative'') are the negative and positive classes, respectively. Since the dataset remained severely imbalanced even after this re-labeling, we used class weights inversely proportional to the number of examples in each class. Using a batch size 64 and a maximum one-cycle learning rate of 3e-2, we fine-tuned the models for 20 epochs, which led to convergence of the training and validation loss curves. We found using L2 regularization was not necessary in this stage to prevent overfitting. After this fine-tuning stage, which completed in approximately 8 hours for DiaNet and 5 hours for DiaNet_Res50, we fine-tuned the models on the QBB dataset. Unlike the previous stage, we found that L2 regularization (with the regularization constant set to 0.02) was needed to prevent overfitting, which is expected given the small size of the dataset. We fine-tuned each model with a batch size of 32 for: (a) 30 epochs keeping everything but the added layers frozen, using a maximum one-cycle learning rate of 1e-4, then (b) 2o epochs unfreezing the entire network and training using discriminative learning rates with the following range: (1e-5 and 1e-7). DiaNet and DiaNet_Res50 took 3 and 1.5 hours to complete training over these 50 epochs, respectively. Since QBB dataset is balanced, there was no need to employ any technique to compensate for any imbalance, such as class weights that were used in the earlier stage.

For the single-stage fine-tuning experiments conducted solely on the QBB dataset, we found that using L2 regularization (with the regularization constant set to 0.01) prevented overfitting. The rest of the parameter values were identical to those used in the latter stage of our multi-stage fine-tuning experiment.

H. HARDWARE AND SOFTWARE SETUP

We conducted the experiments on a workstation with the following compute configuration: CPU: Core i7 8700K, main memory: 64 GB DDR4, GPU: Nvidia Titan X (Pascal). Python 3.7.4 was used as the base language, and fastai [36] v1.4 was used along with Pytorch [40] as the deep learning frameworks. We also used Pandas, Numpy, Scikit-learn, and Matplotlib python packages in our experiments for data preprocessing, manipulation, cross-validation, etc.

III. RESULTS

In this section, we present the outcomes of our experiments and analyze the findings. We compare the performance of our proposed approach quantitatively against other candidate methods and show its superiority in predicting the onset of diabetes among the test subjects. For quantitative performance reporting, we used mean accuracy, sensitivity, specificity, precision, F1 score and AUC ROC. Since QBB dataset is balanced, using the mean accuracy as the measure of model assessment suffices.

As an overview of the performances of the methods experimented on, Table 2 shows the corresponding accuracies and other evaluation metrics. DiaNet fine-tuned on the EyePACS and the QBB dataset performed best with an 84.47% mean accuracy. In comparison, QBBNet, which was trained only on the smaller, QBB dataset, yields 79.02% accuracy. It could be mentioned that the base architecture of DiaNet (Densenet-121) achieved only 80% accuracy on the EyePACS test dataset. This shows that, even though the EyePACS dataset is targeted to a task (DR staging) different from ours (Diabetes detection), the network benefitted from the first-stage fine-tuning as it helped the network understand from a larger dataset what retina images look like and transfer that knowledge to our task. This essentially proves that our proposed approach is utilizing the immensely effective transfer learning technique, which entails first training a neural network on a large dataset to train it with domain knowledge, then fine-tuning it on a dataset aimed at a slightly different, but related task.

From Table 2, we also observed the highest F1 score of 84.71 and 79.47 from DiaNet and QBBNet, respectively. We achieved the highest accuracy of 84.47 and 80.65 from DiaNet and QBBNet, respectively. Figure 6 highlights the ROC curves for both QBBNet and DiaNet. The highest AUC for the DiaNet and QBBNet was 84.4 and 83.10, respectively. Figure 6 highlights the ROC curve for the proposed models (DiaNet and QBBNet).

IV. DISCUSSIONS

A. WHY DID DiaNet INCORPORATE IMAGES FROM EyePACS

In this article, we proposed DiaNet, a deep learning-based approach for estimating the presence of diabetes in a test subject from his/her retinal images. Specifically, we used a CNN based architecture that takes a retinal image as input,

TABLE 2. Performance of different candidate model variants for DiaNet and QBBNet. DiaNet and its variants were trained on both EyePACS and QBB, whereas QBBNet and its variants were trained on only the QBB dataset.

Metric	Accuracy	Precision	Sensitivity/Recall	Specificity	F1 Score	AUC ROC
QBBNet_Res50+ XGB	80.65	79.27	83.15	78.14	81.16	80.64
DiaNet_Res50 + XGB	82.01	80.41	84.78	79.23	82.53	82.08
QBBNet + XGB	76.56	76.63	76.63	76.5	76.63	76.56
DiaNet + XGB	83.92	81.09	82.58	79.23	84.67	83.91
QBBNet_Res50	80.38	80.94	76.63	79.15	79.66	80.39
DiaNet_Res50	83.1	81.77	85.32	80.87	83.51	83.1
QBBNet	79.02	78.01	80.97	77.04	79.47	79.01
DiaNet	**84.47**	**83.59**	**85.86**	**83.06**	**84.71**	**84.46**

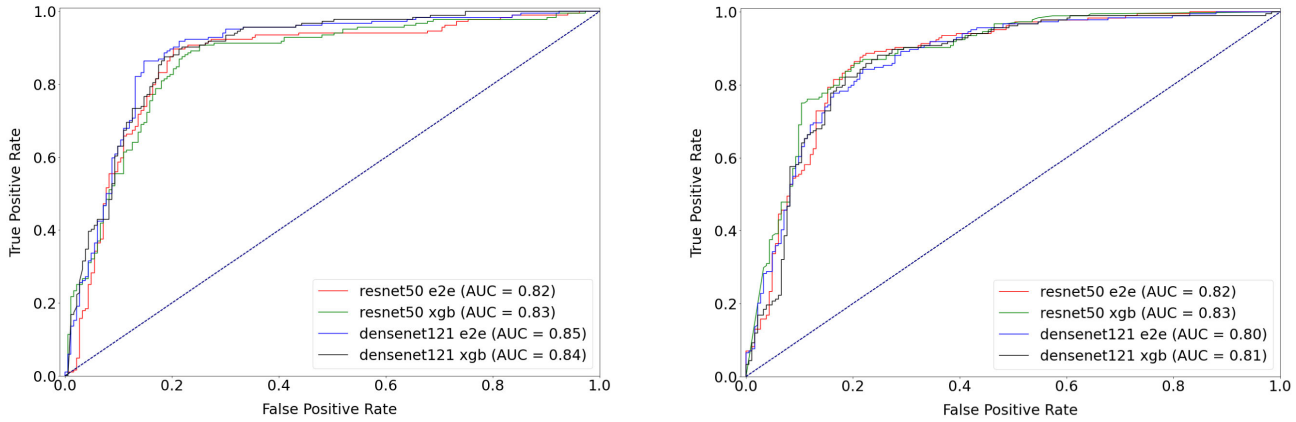

FIGURE 6. ROC plots for the classifiers. Figure 6 (a) shows the model performances of DiaNet, while Figure 6 (b) shows the model performances of QBBNet.

and outputs a probability distribution over the possible labels; i.e. control and diabetes. We applied transfer learning in multiple stages to achieve state-of-the-art performance in the task. The primary motivation behind this multi-stage fine-tuning approach was limited size of the available dataset (from QBB) for the target task. Since many deep learning-based models are data-bound and result in suboptimal performance when trained on a smaller dataset, we augmented the model's understanding of retinal images by incorporating a larger dataset of retinal images, but labeled for a different, but related task – DR detection. Our experiments showed that the proposed multi-stage fine-tuning approach which integrated retinal image from EyePACS and QBB results in an improvement of the model's performance compared to the same when trained on only the QBB retinal image dataset (Table 2).

B. CLASS ACTIVATION MAPS (CAM) HIGHLIGHTING THE REGION OF INTEREST FOR DiaNet

Figure 7 shows some retinal images and the overlaid heatmaps from the diabetes and the control group. The color-coded regions on the input images represent the influence on the predictions across the image at various degrees. Only the images with high-confidence predictions (probability larger than 0.80) from DiaNet were selected for Figure 7. The bottom row in Figure 7 highlights some retinal images from the diabetes group. Interestingly, in all the images the overlaid heatmap was mainly focusing on the central retinal area (located between the optic disc area and the macular region) where characteristics of DR may develop. The images (d), (e), and (f) shows a widespread microaneurysm, which are tiny bulges protruding from the walls of the smaller vessels and are the earliest clinically visible changes of DR. A retinal bleeding in a form of dot and blot shaped intraretinal hemorrhages that is a sign of diabetic retinopathy. All three images also show intraretinal hemorrhages and exudation (a distinct yellow-white intra retinal deposits of lipids and protein in the extracellular space due to leaky abnormal retinal capillaries) that can be found in the eyes of people affected with DR. In addition, the image in Fig 7(f) is showing tortuous retinal blood vessels with arteriovenous nicking and venous dilation which is associated with systemic diseases such as hypertension, diabetes and ischemic heart disease, thus bolstering our claim that our model is capable of identifying retinal characteristics of diabetes in general. In the control group images (Figure 7.a-c), the absence of the said characteristics in the focused regions (as indicated by the heatmaps) indicates that these retinas are free from the symptoms of diabetes.

C. COMPARISON AGAINST EXISTING ML BASED MODELS USING QBB CLINICAL DATASET

It is important to emphasize that we could not compare the performance of the proposed model against any model due to the lack of existence of any study which considered

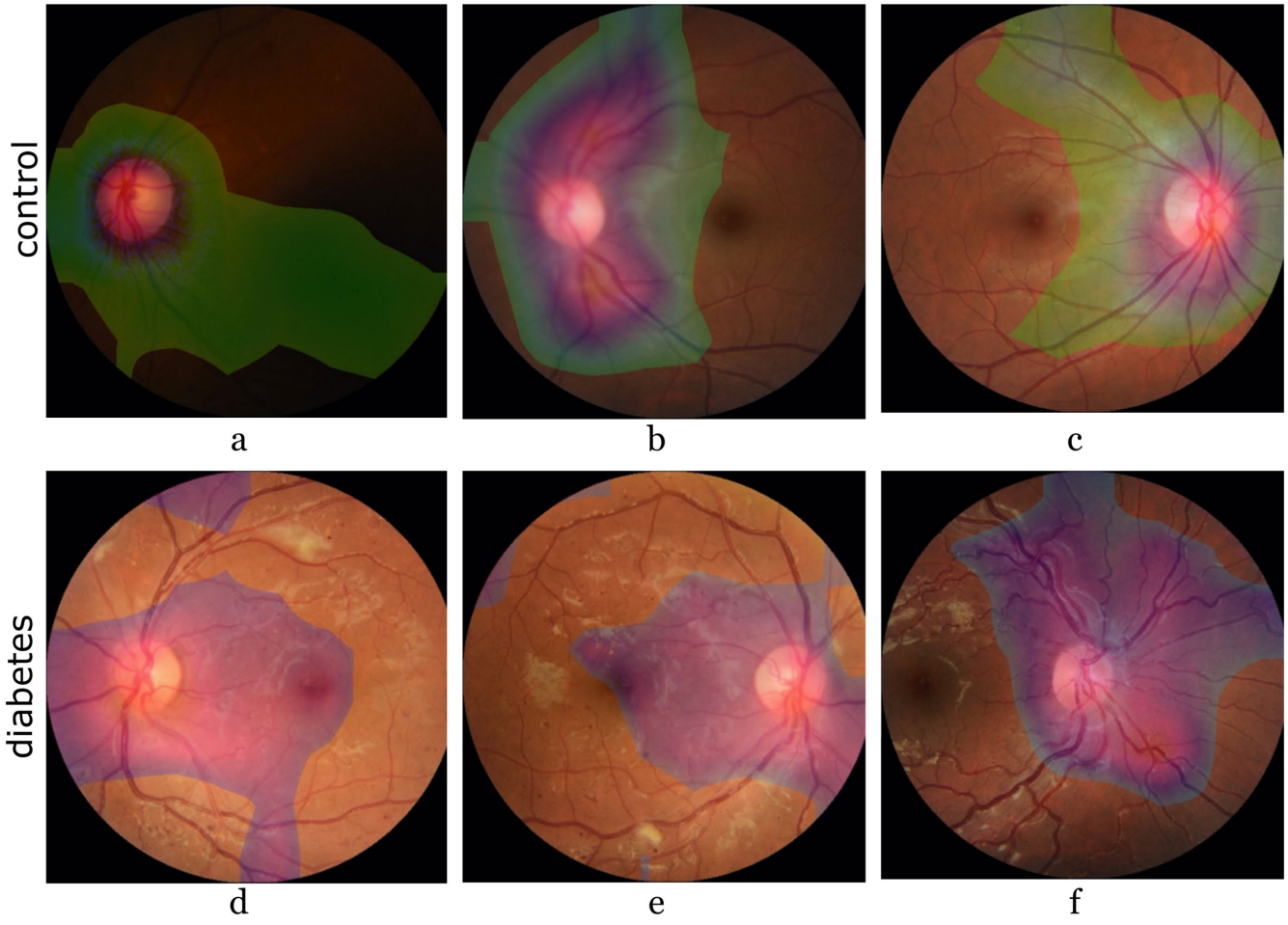

FIGURE 7. A few Retinal images overlaid with heatmaps. (top: the Control group, bottom: the Diabetes group) The composite retinal heatmap was generated by combining images from the test set with corresponding class activation map (CAM). Purple regions indicate greater influence on the prediction decision while the yellow-greenish regions correspond to relatively lesser impact on the same.

QBB retinal images to distinguish diabetes group from control group. But, previously, we considered 237 clinical measurements, covering bioimpedance, spirometry, biomarkers, anthropometric measurements and lifestyle related variables from QBB diabetes cohort to detect diabetes patients from the control group with nearly 78% accuracy [10]. And in this study, we achieved an even higher level of accuracy (over 84% and 79% accuracy for DiaNet and QBBNet, respectively) to predict the diabetes group considering retinal images only. This indicates that the proposed model DiaNet as well as QBBNet outperformed the previous model for distinguishing the diabetes group from the control group. This also entails that retinal images contain sufficient information, like other clinical measurements, to distinguish the diabetes group and this could be considered in clinical setup for the diagnosis of diabetes as well.

D. PRACTICAL APPLICATIONS OF THE PROPOSED MODEL

Recently, the International Diabetes Federation (IDF) and the World Health Organization (WHO) have prioritized low-cost screening of digital retinal photography by non-physicians and remote grading using mobile healthcare services [41], [42]. DiaNet, the solution we propose, if implemented in a clinical setup, will revolutionize the eye care system considering the cost-effectiveness for low- and middle-income countries. For high-income countries our solution will reduce the workload of the physicians as well as help to implement mass level diabetes screening within a shorter period of time.

V. LIMITATIONS

DiaNet was able to distinguish diabetic patients from the control group at over 84% accuracy, which means the proposed model was able to work correctly for four persons out of five persons. Ideally, we would expect near-perfect accuracy, but considering the limited number of retinal images we had, our model can be further improved by incorporating more images. Though the achieved accuracy seems low compared to the other existing deep learning-based model to predict different grading of DR, we would like to emphasize that the goal of this study was not to predict DR, rather distinguish diabetes from the control group. And to date, there exists no

study which focused on the diagnosis of diabetes considering retinal images; therefore, our study stands out as unique as well.

VI. CONCLUSION

In this work, we proposed DiaNet, a novel deep learning-based model to distinguish diabetes from the control group using QBB retinal photography. Our model, based on retinal images, achieved over 84% accuracy to classify the diabetes group from the control group achieving a higher level of accuracy than what we achieved using clinical dataset only, for the same purpose. As per our best knowledge, our study is the first, which predicts diabetes considering only retinal images. So, we believe the retinal images can be used in clinical setup to diagnose diabetes and incorporate more images the model could perform at a higher level than what we achieved.

REFERENCES

[1] D. Mellitus, "Diagnosis and classification of diabetes mellitus," *Diabetes Care*, vol. 37, no. 1, pp. S81–S90, 2014. [Online]. Available: https://care.diabetesjournals.org/content/37/Supplement_1/S81

[2] E. S. Shin, C. M. Sorenson, and N. Sheibani, "Diabetes and retinal vascular dysfunction," *J. Ophthalmic Vis. Res.*, vol. 9, no. 3, pp. 362–373, Sep. 2014.

[3] X. Zhang, J. B. Saaddine, C.-F. Chou, M. F. Cotch, Y. J. Cheng, L. S. Geiss, E. W. Gregg, A. L. Albright, B. E. K. Klein, and R. Klein, "Prevalence of diabetic retinopathy in the United States, 2005-2008," *JAMA*, vol. 304, no. 6, pp. 649–656, Aug. 2010.

[4] J. Yau, S. Rogers, and R. Kawasaki, "Global prevalence and major risk factors of diabetic retinopathy," *Diabetes Care*, vol. 35, no. 3, pp. 556–564, 2012.

[5] S. D. Solomon, E. Chew, E. J. Duh, L. Sobrin, J. K. Sun, B. L. VanderBeek, C. C. Wykoff, and T. W. Gardner, "Diabetic retinopathy: A position statement by the American diabetes association," *Diabetes Care*, vol. 40, no. 3, pp. 412–418, Mar. 2017. [Online]. Available: https://care.diabetesjournals.org/content/40/3/412

[6] 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," *JAMA*, vol. 316, no. 22, p. 2402, Dec. 2016. [Online]. Available: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2016.17216

[7] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, Jul. 2017.

[8] M. M. Islam, H.-C. Yang, T. N. Poly, W.-S. Jian, and Y.-C. (Jack) Li, "Deep learning algorithms for detection of diabetic retinopathy in retinal fundus photographs: A systematic review and meta-analysis," *Comput. Methods Programs Biomed.*, vol. 191, Jul. 2020, Art. no. 105320. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0169260719311010

[9] A. Bora, S. Balasubramanian, B. Babenko, S. Virmani, S. Venugopalan, A. Mitani, G. de Oliveira Marinho, J. Cuadros, P. Ruamviboonsuk, G. S. Corrado, L. Peng, D. R. Webster, A. V. Varadarajan, N. Hammel, Y. Liu, and P. Bavishi, "Predicting the risk of developing diabetic retinopathy using deep learning," *Lancet Digit. Health*, vol. 3, no. 1, pp. e10–e19, Jan. 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2589750020302508

[10] S. Musleh, T. Alam, A. Bouzerdoum, S. B. Belhaouari, and H. Baali, "Identification of potential risk factors of diabetes for the qatari population," in *Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT)*, Feb. 2020, pp. 243–246.

[11] *Diabetes Mellitus in Qatar*. Accessed: Aug. 3, 2020. [Online]. Available: https://www.worldlifeexpectancy.com/qatar-diabetes-mellitus

[12] S. K. Al-Kaabi and A. Atherton, "Impact of noncommunicable diseases in the State of Qatar," *ClinicoEconomics Outcomes Res., CEOR*, vol. 7, pp. 377–385, Jul. 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4494184/

[13] Z. Shi and A. B. Abou-Samra, "Association of low serum magnesium with diabetes and hypertension: Findings from qatar biobank study," *Diabetes Res. Clin. Pract.*, vol. 158, Dec. 2019, Art. no. 107903.

[14] A. Nazeemudeen, H. R. H. Al-Absi, M. A. Refaee, M. Househ, Z. Shah, and T. Alam, "Understanding the Food Habits and Physical Activities of Diabetes Cohort in Qatar," *Stud. Health Technol. Informat.*, vol. 272, pp. 453–456, Jun. 2020.

[15] A. Bener and A. Al-Hamaq, "Predictions burden of diabetes and economics cost: Contributing risk factors of changing disease prevalence and its pandemic impact to qatar," *Experim. Clin. Endocrinol. Diabetes*, vol. 124, no. 08, pp. 504–511, Mar. 2016.

[16] E. Ullah, R. Mall, R. Rawi, N. Moustaid-Moussa, A. A. Butt, and H. Bensmail, "Harnessing qatar biobank to understand type 2 diabetes and obesity in adult qataris from the first qatar biobank project," *J. Transl. Med.*, vol. 16, no. 1, p. 99, 2018.

[17] M. Ashikur, M. Arifur, and J. Ahmed, "Automated detection of diabetic retinopathy using deep residual learning," *Int. J. Comput. Appl.*, vol. 177, no. 42, pp. 25–32, Mar. 2020. [Online]. Available: https://www.ijcaonline.org/archives/volume177/number42/31185-2020919927

[18] W. Zhang, J. Zhong, S. Yang, Z. Gao, J. Hu, Y. Chen, and Z. Yi, "Automated identification and grading system of diabetic retinopathy using deep neural networks," *Knowl.-Based Syst.*, vol. 175, pp. 12–25, Jul. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0950705119301303

[19] S. B. Hathwar and G. Srinivasa, "Automated grading of diabetic retinopathy in retinal fundus images using deep learning," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Sep. 2019, pp. 73–77.

[20] W. Pratumgul and W. Sa-ngiamvibool. (2019). *Classification and Identification of Diabetic Retinopathy Severity Stages in Thai Patients Using Deep Learning Based Convolution Neural Networks*. [Online]. Available: https://www.semanticscholar.org/paper/Classification-and-Identification%-of-Diabetic-in-Pratumgul-Sa-ngiamvibool/49b0aff4c1d6cee0d106071f84ee11f4bcd8f%eb3

[21] C. González-Gonzalo, V. Sánchez-Gutiérrez, P. Hernández-Martínez, I. Contreras, Y. T. Lechanteur, A. Domanian, B. Ginneken, and C. I. Sánchez, "Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration," *Acta Ophthalmologica*, vol. 98, no. 4, pp. 368–377, Jun. 2020.

[22] R. AlSaad, S. Boughorbel, and S. Al-Maadeed, "Automated classification of diabetic retinopathy severity: A deep learning approach," in *Qatar Foundation Annual Research Conference Proceedings*. Doha, Qatar: Hamad Bin Khalifa Univ. Press, 2018. [Online]. Available: https://www.qscience.com/content/papers/10.5339/qfarc.2018.HBPD1007

[23] H. R, A. V, N. J, and V. G, "A detailed study on diagnosis and prediction of diabetic retinopathy using current machine learning and deep learning techniques," *Int. J. Psychosocial Rehabil.*, vol. 23, no. 1, pp. 412–417, Feb. 2019.

[24] L. Qiao, Y. Zhu, and H. Zhou, "Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms," *IEEE Access*, vol. 8, pp. 104292–104302, 2020.

[25] B. Tymchenko, P. Marchenko, and D. Spodarets, "Deep learning approach to diabetic retinopathy detection," in *Proc. 9th Int. Conf. Pattern Recognit. Appl. Methods*, 2020, pp. 501–509.

[26] A. Khan, I. N. Petropoulos, G. Ponirakis, and R. A. Malik, "Visual complications in diabetes mellitus: Beyond retinopathy," *Diabetic Med.*, vol. 34, no. 4, pp. 478–484, Apr. 2017.

[27] A. Al Thani, E. Fthenou, S. Paparrodopoulos, A. Al Marri, Z. Shi, F. Qafoud, and N. Afifi, "Qatar biobank cohort study: Study design and first results," *Amer. J. Epidemiology*, vol. 188, no. 8, pp. 1420–1433, Aug. 2019.

[28] H. Al Kuwari, A. Al Thani, A. Al Marri, A. Al Kaabi, H. Abderrahim, N. Afifi, F. Qafoud, Q. Chan, I. Tzoulaki, P. Downey, H. Ward, N. Murphy, E. Riboli, and P. Elliott, "The qatar biobank: Background and methods," *BMC Public Health*, vol. 15, no. 1, p. 1208, Dec. 2015, doi: 10.1186/s12889-015-2522-7.

[29] J. Cuadros and G. Bresnick, "EyePACS: An adaptable telemedicine system for diabetic retinopathy screening," *J. Diabetes Sci. Technol.*, vol. 3, no. 3, pp. 509–516, May 2009.

[30] B. Graham, "Kaggle diabetic retinopathy detection competition report," Univ. Warwick, Coventry, U.K., Tech. Rep., 2015.
[31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
[32] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.
[33] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, Stateline, NV, USA, vol. 25, Dec. 2012, pp. 1097–1105.
[34] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," in *Proc. 3rd IAPR Asian Conf. Pattern Recognit. (ACPR)*, Nov. 2015, pp. 730–734.
[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[36] J. Howard. (2018). *Fastai*. [Online]. Available: https://github.com/fastai/fastai
[37] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.
[38] G. C. Cawley and N. L. Talbot, "On over-fitting in model selection and subsequent selection bias in performance evaluation," *J. Mach. Learn. Res.*, vol. 11, pp. 2079–2107, Jul. 2010. [Online]. Available: http://jmlr.org/papers/v11/cawley10a.html
[39] L. N. Smith, "A disciplined approach to neural network hyper-parameters: Part 1—Learning rate, batch size, momentum, and weight decay," 2018, *arXiv:1803.09820*. [Online]. Available: https://arxiv.org/abs/1803.09820
[40] A. Paszke, S. Gross, F. Massa, and A. Lerer, "Pytorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2019, pp. 8024–8035. [Online]. Available: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-pe%rformance-deep-learning-library.pdf
[41] M. W. M. Wintergerst, D. K. Mishra, L. Hartmann, P. Shah, V. K. Konana, P. Sagar, M. Berger, K. Murali, F. G. Holz, M. P. Shanmugam, and R. P. Finger, "Diabetic retinopathy screening using smartphone-based fundus imaging in india," *Ophthalmology*, vol. 127, no. 11, pp. 1529–1538, Nov. 2020.
[42] C. Sabanayagam, W. Yip, D. S. W. Ting, G. Tan, and T. Y. Wong, "Ten emerging trends in the epidemiology of diabetic retinopathy," *Ophthalmic Epidemiol.*, vol. 23, no. 4, pp. 209–222, Jul. 2016.

MOHAMMAD TARIQUL ISLAM is currently an Assistant Professor with the Computer Science Department, Southern Connecticut State University. He has published at notable peer-reviewed conferences and journals, such as the International Conference on Image Processing, the International Conference on Bioinformatics and Biomedicine, *Computer Vision and Pattern Recognition*, *International Journal on Image and Video Processing*, and so on. His primary area of research interests include computer vision, deep learning, and applied bioinformatics. He was a recipient of several grants on the application of deep learning in computer vision and has supervised multiple graduate students in their pursuit of the master's degree.

HAMADA R. H. AL-ABSI received the Bachelor of Technology degree (Hons.) in information and communication technology and the Master of Science degree in information technology from Universiti Teknologi PETRONAS, Malaysia, in 2009 and 2010, respectively. He is currently pursuing the Ph.D. degree with Hamad Bin Khalifa University, Qatar. His research interest includes machine learning applications to diseases detection and diagnosis. He is a member of the Australian Computer Society (ACS) and a Certified Professional by ACS.

ESSAM A. RUAGH received the ophthalmology training from U.K. and became a Fellow of the Royal College of Ophthalmologists, London. Then, he received the advanced fellowship training in vitreoretinal surgery and medical retina working in units with high volume of retinal procedures from U.K. He is currently a Consultant Ophthalmologist, a Vitreoretinal Specialist, and a Lead Clinician for Retina Service with the Magrabi Eye, Dental and Ear Centre, Magrabi Hospital, Qatar. He has a special interest in the treatment of diabetic retinopathy and the application of artificial intelligence in common diseases in ophthalmology. He has delivered over 40 presentations at national and international conferences for opthalmologist.

TANVIR ALAM is currently an Assistant Professor with the College of Science and Engineering, Hamad Bin Khalifa University. Among his notable research works are on the transcription regulation of non-coding RNAs and their roles in different diseases. His research work also centered around the application of artificial intelligence (AI) on the diagnosis and prognosis of communicable and non-communicable diseases. He is a member of FANTOM Consortium. He also served as a reviewer for number of international conferences and reputed journals.