ISSN: 0067-2904

Binary Classification of Diabetic Retinopathy Using CNN Architecture

Ali Hassan Khudaier, Abdulkareem Merhej Radhi
Department of Computer Science, Al-Nahrain University, Baghdad, Iraq

Received: 14/8/2022 Accepted: 28/2/2023 Published: 29/2/2024

Abstract

Diabetes mellitus (DM), a chronic, clinically heterogeneous condition, is becoming increasingly common all over the world. Insulin deficiency, resistance to insulin's actions on the body's surface, or both may lead to pancreatic beta-cell degeneration. Diabetes makes people more prone to its consequences, the most prevalent of which is diabetic retinopathy (DR). Deep learning algorithms surpass traditional classification models for illness diagnosis on photos of medical problems. Deep transfer learning models for medical DR detection were evaluated using the APTOS 2019 dataset. Deep transfer learning algorithms for medical diabetic retinopathy (DR) detection are being evaluated. MobileNet Convolutional Neural Networks (CNN) architecture is used to detect the DR in binary class classification tasks, which leverages pre-trained weights collected during the training process using the ImageNet database. Cohen Kappa, F1 score, recall, accuracy, and precision are some of the performance indicators used. According to the data, the given model is the most effective in terms of accuracy and training time for handling our challenges. Overall, MobileNet is a good pick. The following metrics were found to be accurate: 0.9455, precise: 0.94651, recall: 0.9455, F1 score: 0.94556, and Cohen Kappa score: 0.89083. This method might aid medical personnel in the early detection of diabetic retinopathy.

Keywords Gaussian filter, Deep learning, Diabetic retinopathy, MobileNet, and Diabetes mellitus.

التصنيف الثنائي لاعتلال الشبكية السكري باستعمال *CNN*

علي حسن خضير، عبد الكريم مرهج راضي
قسم علوم الحاسوب، جامعة النهرين، بغداد، العراق

الخلاصة

مرض السكري (DM) ، وهو حالة مزمنة وغير متجانسة سريريًا ، يتزايد بشكل متكرر في جميع أنحاء العالم. قد يؤدي نقص الأنسولين أو مقاومة الأنسولين على سطح الجسم أو كليهما إلى تنكس خلايا بيتا في البنكرياس. مرض السكري يجعل الناس أكثر عرضة لعواقبه ، وأكثرها انتشارًا هو اعتلال الشبكية السكري (DR). تتجاوز خوارزميات التعلم العميق نماذج التصنيف التقليدية لتشخيص الأمراض في صور المشكلات الطبية. تم تقييم نماذج التعلم العميق للنقل لاكتشاف DR الطبي باستعمال مجموعة بيانات APTOS 2019. يتم تقييم خوارزميات التعلم العميق للنقل للكشف عن DR الطبي. لاكتشاف DR في مهام تصنيف الفئة الثنائية ، استعملنا بنية MobileNet CNN ، والتي تستفيد من الأوزان المدربة مسبقًا التي تم جمعها أثناء عملية التدريب باستعمال

*Email: programer.1422014@gmail.com

قاعدة بيانات ImageNet. Cohen Kappa, F1 score, recall, accuracy, and precision هي بعض مؤشرات تقييم الأداء التي استعملت. وفقًا للبيانات ، يعد النموذج المحدد هو الأكثر فعالية من حيث الدقة ووقت التدريب للتعامل مع تحدياتنا. بشكل عام ، يعد MobileNet اختيارًا جيدًا. تم العثور على المقاييس التالية لتكون accurate 0.9455, precise: 0.94651, recall 0.9455, F1 score 0.94556, and Cohen Kappa score: 0.89083. قد تساعد هذه الطريقة العاملين في المجال الطبي في الكشف المبكر عن اعتلال الشبكية السكري.

1. Introduction

Diabetes mellitus (DM) [1], a chronic metabolic illness with a variety of clinical manifestations, was expected to affect 126.6 million people in 2010, with that figure expected to rise to 191 million by 2030 [2], and its prevalence has been steadily increasing globally [3]. 366 million people were predicted to have DM in 2011, and 552 million were projected to have it by 2040 [4]. The global prevalence of DR accounts for more than 2.6% of blindness [5]. Chronic hyperglycemia causes pancreatic beta-cell failure, which can be caused by a decrease in insulin production, resistance to insulin's peripheral effects, or both [6]. People with DM are more prone to many short- and long-term consequences because metabolic imbalances can impair numerous organ systems and result in the development of incapacitating and life-threatening health issues. Microvascular problems (such as nephropathy, retinal degeneration, and neuropathy) are the most common of these consequences [7].

Diabetic retinopathy, or DR, occurs when high blood sugar levels damage the blood vessels within the retina [8]. Issues like blindness are typically the outcome of DR. Non-proliferative DR (NPDR) and proliferative DR (PDR) are the two basic phases of DR (PDR). Microaneurysms (MA), blood vessels, hemorrhage, and exudates are the main lesions that are evaluated for grading [9]. DR at its advanced stage with neovascularization is referred to as PDR, while DR without any of these lesions is referred to as NPDR. The retina, as shown in Figure 1, has many lesions [10]. The walls of the small blood arteries get swollen, a condition known as MA.

One theory holds that, among those who already have DR, these little bulging MAs are the first visible symptoms of the disease. They appear as little red flecks on the retina [11]. It becomes worse as the condition gets worse. Another retinal condition brought on by DR is retinal hemorrhage. High blood pressure and blockage of the retinal veins are other causes of bleeding (HM). When they're young, they look a lot like MAs. Exudates arise in the retina when blood from ruptured capillaries mixes with protein and lipid leftovers to create yellow flickers. The DR classifications that are used most frequently include PDR, mild NPDR, moderate NPDR, severe NPDR, and no DR. The "4-2-1" criterion is used to identify severe NDPR. A significant component is intra-retinal microvascular abnormalities (IRMA) [12]. A retinal blood vessel's aberrant branching or expansion is referred to as IRMA. Table 1 shows the grading criteria for DR. Recovery from DR is difficult in its later stages. Total blindness will be the outcome. Reduce its global spread to the greatest extent possible [8].

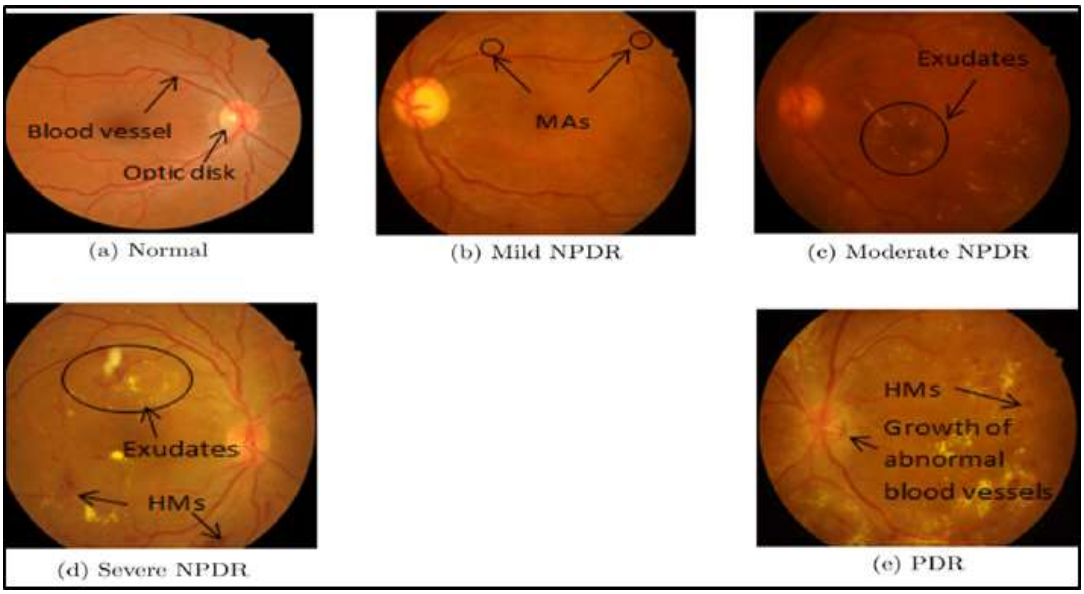

Figure 1: Fundus photos at different stages of DR where (a: No DR, b: Mild NPDR, c: Moderate NPDR, d: Severe NPDR, e:PDR).

Table 1: Diabetic retinopathy Grading

Grading level	Identification of lesions
No DR	There were no anomalies found.
Mild NPDR	MAs were only seen.
Moderate NPDR	A limited number of MAs, mild yet severe, with or without cotton-wool patches, venous beading, or presence of IRMA that is less than the 4-2-1 recommendations
Severe NPDR	Any of the 4-2-1 There are rules in place: Dot blot hemorrhage in all four quadrants, at least two quadrants with venous beading, and at least one quadrant with significant IRMA

Detecting DR and assessing its severity remain difficult challenges, despite the development of early detection tools. Many ways have been created to detect the problem in its early stages in order to fix it. The procedures must be trustworthy, precise, and cost-effective. The aforementioned anomalies are frequently used by ophthalmologists to detect DR and estimate its severity. When the lesions initially occur, they may be difficult to notice and differentiate from normal retinal conditions. The difficulty in removing each participant's retinal lesions complicates DR identification.

To acquire the most reliable characteristics from the fundus image that can discriminate between normal and DR images, very excellent feature extraction algorithms are required for effective automated detection. In this study, an efficient automated feature learning method for feature extraction from binary retinal fundus pictures was investigated. The detection process would be made simpler if all DR-related characteristics could be automatically obtained. In this aspect, artificial intelligence will be useful. A quick and effective method for categorizing binary DR from fundus images is offered. The rest of this essay is structured as follows: The associated work in DR image categorization from the first section is included in the second section. Section 3 describes the material and the intended procedure. Commentary and experimental analysis are found in Section 4. Conclusions are offered in Section 5.

2. Related Works

Given the significance of computer-aided design (CAD) systems in the early diagnosis of DR, there has been much study on the use of fundus photos to identify DR [13]. Both deep learning (DL)-based DR and conventional detection techniques can be used in this research. The preprocessing stage of most DR detection techniques involves noise reduction and image enhancement. After segmenting the fundus image, retinal elements and other types of abnormal lesions are removed. To train machine learning models for categorizing fundus photographs into diseased images and normal ones, attributes from the segmented lesions are collected last. DL-based DR detection techniques, on the other hand, use various deep convolutional neural network (DCNN) architectures to distinguish between pictures that are healthy and those that are ill. Using the back-propagation approach to train on a set of labeled training pictures, a DCNN is a feed-forward neural network that finds the most effective features for classification. An overview of existing convolutional neural network (CNN)-based DR diagnostic approaches is offered.

In [14], five categories were used to investigate how to identify retinopathy. Features extracted from photographs were used as input into the SVM classifier to capture the contours and diversity of shapes. Its specificity is 88 percent and its accuracy is 82 percent, respectively. The SVM classifier receives as input from the images an estimate of the locations of hard exudates, hemorrhages, blood vessels, and micro-aneurysms. 300 retinal pictures of patients with mild NPDR, moderate NPDR, severe NPDR, PDR, and normal instances were analyzed. The National University Hospital in Singapore supplied these statistics. With this method, it was discovered that the sensitivity and specificity were, respectively, 82 and 86 percent, with an accuracy of 85.9 percent.

In [15], researchers have compiled a database of fundus pictures of DR patients that have been annotated with the appropriate therapy. In order to quantify the severity of DR in fundus pictures, they used this data set to train deep convolutional neural network models. In studies, performance on a four-degree classification test improved to 88.72% accuracy, leading to cloud-based model deployment and pilot DR diagnostic service provision at many hospitals. Clinical assessment shows 91.8 percent ophthalmologist agreement; this demonstrates the success.

In [16], researchers describe a novel method (for assessing the DR severity stages) based on a binary tree-based multiclass classifier of convolutional neural networks (CNN). Our suggestion was to base the ensemble's construction on a set of values for the confusion matrix between any two DR phases. The recommended approach achieved an overall accuracy of 83 percent when evaluating the DR phases (i.e., severity levels) using fundus images from the publicly accessible EyePACS dataset, which is a significant improvement over existing state-of-the-art models.

In [17], researchers use transfer learning techniques, which maintain the more universal properties of a basic architecture that has already been taught while adapting to a novel dataset. To choose the optimal basic architecture and hyperparameters, they develop a custom heuristic equation and ranking system for assessment metrics. Images from the EYEPACS diabetic retinopathy dataset were imported. The Xception base architecture, Adam optimizer, and mean squared error loss function exceed the competition with 90% accuracy, 94% sensitivity, and 86% specificity.

In [18], CNNs were well employed for DR staging on retinal fundus pictures, with a

validation sensitivity of 95%. They used pre-trained GoogLeNet and AlexNet models and a contrast-limited adaptive histogram equalization strategy for preprocessing and transfer learning. As a result, they were able to achieve test accuracies on 2-ary, 3-ary, and 4-ary classification models of 74.5 percent, 68.8 percent, and 57.2 percent, respectively.

In [19], an automated DR disclosure mechanism has been developed using deep CNN. It is based on 35,000 images from an extensive collection. Images are scaled down in this project to 448x448 pixels. They also used a range of data augmentation methods. The authors' final accuracy results were 81 percent for class 0 and 88 percent for class 1.

In [11], this paper describes a unique CNN approach for extracting features from retinal fundus pictures in order to improve classification performance. In the proposed system, the CNN output features are employed as input for several machine learning classifiers. Using pictures from the generic IDRiD, MESSIDOR, and KAGGLE datasets, the model is tested using multiple classifiers (Support Vector Machine, AdaBoost, Naive Bayes, Random Forest, and J48). The specificity, precision, recall, false-positive rate (FPR), Kappa-score, and accuracy scores for each classifier are compared to determine the classifier's effectiveness. The assessment findings show that the suggested feature extraction approach, in conjunction with the J48 classifier, beats all other classifiers on the MESSIDOR, IDRiD, and KAGGLE datasets, with an average accuracy of 99.89% for binary classification and 99.59% for multiclass classification. Furthermore, the average Kappa-score (K-score) for the J48 classifier is 0.994 for binary classification and 0.994 for multi-class classification.

3. Materials And Methods

A. Proposed Approach

The primary layers and secondary layers of CNN are separated into two categories. Convolutional neural networks (CNNs) primarily use convolutional, activation, pooling, flattening, and dense layers. Secondary layers are additional layers that may be added to CNNs to increase their generalizability and resistance to overfitting. Dropout layers, batch regularization layers, and normalization layers are among them. A CNN structure is shown in Figure 2. Deep convolutional layers may be built in many ways. InceptionResNetV2, VGG16, InceptionV3, DenseNet169, DenseNet121, MobileNetV2, VGG19, ResNet50, and ResNet101 are all examples of mobile networks.

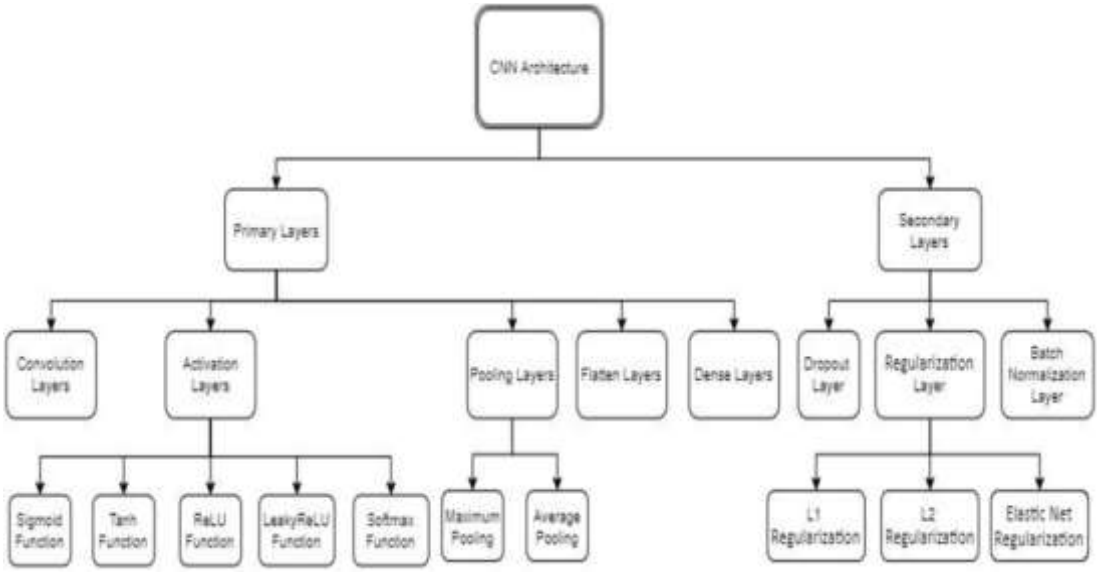

Figure 2: Convolutional neural network (CNN) structure [20].

- **EfficientNet**: Mingxing Tan and Quoc V. Le [21] investigate ConvNet scaling in depth and find that appropriately balancing network width, depth, and resolution is a crucial yet overlooked component that stops us from enhancing accuracy and efficiency. They propose a simple but incredibly efficient compound scaling technique to overcome this challenge, which allows us to quickly scale up a baseline ConvNet to any required resource constraints while maintaining model efficiency. A mobile-size EfficientNet model may be scaled up to attain the most cutting-edge accuracy by a factor often shown in this paper. Thanks to the implementation of this compound scaling technique, ImageNet and five frequently used transfer learning datasets require fewer parameters and FLOPS. A multi-can model may be used to detect DR and extract significant properties from retinal fundus images shown in this paper.
-

To learn from people who have DR, deep learning techniques are used. APTOS, a publicly available dataset, was used throughout the whole experiment. The retinal fundus of the eyes is depicted in 3662 color photographs in this collection. The images are available in various resolutions, such as 640*480 pixels, 1050*1050 pixels, 2416*1736 pixels, and so on. Before inter-network transmission, all photographs in our experiment will be shrunk to 224*224. Each image's diagnosis was based on medical expert determinations, which included determining the existence of DR and assigning a retinopathy grade. Because a binary categorization system is used, only two classes are considered. A performance comparison with the InceptionResNetV2, VGG16, InceptionV3, DenseNet169, DenseNet121, MobileNetV2, VGG19, ResNet50, EfficientNet, and ResNet101 designs is included. The chosen architectures' network structures are as follows:

- **MobileNet:** is a collection of depth-wise separable convolution layers created by Google's research team. With a limited set of hyperparameters, the MobileNet architecture is capable of achieving high accuracy rates. Cross-channel and spatial correlations are mapped using depth-separable convolution layers in the feature maps of the input pictures. A depthwise and a pointwise convolution make up the two fundamental parts of a depthwise separable convolution. Each input feature map is subjected to a single spatial filter during depthwise convolution, while a cross-channel pattern filter is used during pointwise convolution (1 * 1). While separable convolutional layers independently record spatial and cross-channel patterns, standard convolutional layers gather both concurrently [22].
-
- **InceptionResNetV2:** Neural network using convolutions over a million images from the ImageNet collection served as the training data for Inception-ResNet-v2. The 164-layer network can classify images into 1,000 distinct object categories, including the keyboard, mouse, pencil, and a wide range of animals. As a result, the network has acquired knowledge of several rich feature representations for a range of images. A list of predicted class probabilities is produced by the network using a 299 × 299 image as input [23].
-
- **VGG16:** An artificial neural network known as “ConvNet” employs convolutional neural networks. A convolutional neural network consists of an input layer, an output layer, and several hidden layers. One of the top current computer vision models is the CNN (convolutional neural network), known as VGG16. The model's developers studied the networks and improved the depth using an architecture with incredibly small (33) convolution filters, outperforming prior-art configurations considerably. There are now 138 trainable parameters as a result of their expansion in depth to 16–19 weight layers [24].

- InceptionV3: The Inception module, which contains 42 layers, was suggested by Szegedy. The third iteration of the Inception module for Google Brain comprises 159 levels and is referred to as InceptionV3. The Inception module's primary objective is to learn multi-scale representations while reducing computational expense and parameter count by combining small and big kernels [25].

- DenseNet169, DenseNet121: One of the models in the DenseNet family created for picture classification is the denseness-169 model. The dense net's major distinction (the 121 model) is a measure of the model's size and accuracy. Compared to the denseness-169, the size of the former is slightly bigger (about 55 MB), and 121 is around 31 MB in size. The writers changed them from their original Torch training into Caffe* format. On the ImageNet picture database, all of the DenseNet models have undergone pretraining. If you want to learn more about this model family, look at the repository. The sole picture in the model input is a blob that is 1x3x224x224 in BGR order. As shown, subtract the BGR mean values: Before uploading the picture blob to the network, [103.94, 116.78, 123.68]. The values also need to be divided by 0.017. The standard object classifier output for the 1000 alternative classifications that match those in the ImageNet database matches the model output for denseness-169 [26].

- MobileNet-v2: is a deep convolutional neural network with 53 layers**.** An untrained version of the network may be imported from the ImageNet database [27]. This has a training set of more than a million images. The network can group images into 1,000 distinct types of objects, including keyboards, mice, pencils, and many different species of animals. The network has acquired a range of rich feature representations for a variety of images. The input image size for the network is 224 by 224 pixels [28]

- VGG19: A VGG model variant with 19 layers (16 convolution layers, 3 fully connected layers, 5 MaxPool layers, and 1 SoftMax layer). They include several variations of VGG, such as VGG11 and VGG16. The VGG19 FLOP count is 19.6 billion [29].

- ResNet50: With the deep residual learning network, the idea of a residual block is first presented (ResNet). Residual blocks aim to connect the first block's input to the second block's output. The residual block can use this approach to learn the residual function and avoid parameter explosion. The deep residual learning network introduces the notion of a residual block (ResNet). Residual blocks aim to connect the first block's input to the second block's output. The residual block can use this approach to learn the residual function and avoid parameter explosion [30].

- ResNet101: is a deep convolutional neural network with 101 layers. From the ImageNet database, where the network has been trained on more than a million images, you may import a pre-trained version of the network. The network can group images into 1,000 distinct types of objects, including keyboards, mice, pencils, and many different species of animals. The network has acquired a range of rich feature representations for a variety of images. The input image size for the network is 224 by 224 pixels [31]. Figure 3 shows the proposed system block diagram.

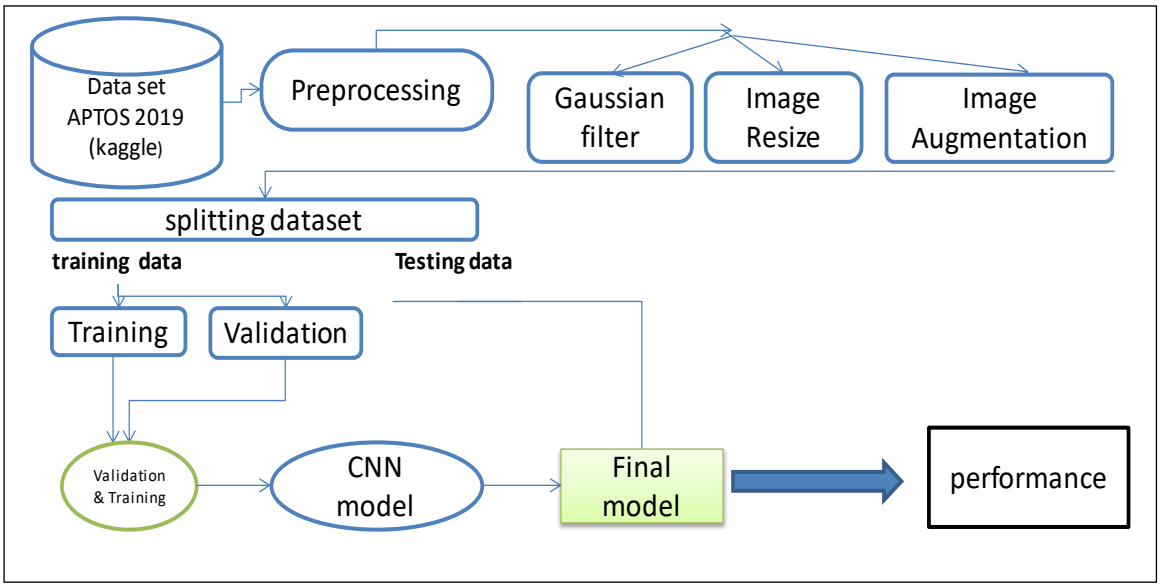

Figure 3: Proposed System block diagram

B. Overview Of The Dataset

The Asia Pacific Tele-Ophthalmology Society's APTOS 2019 diabetic retinopathy (https://www.kaggle.com/c/aptos2019-blindness-detection/) categorization challenge provided the data for this study, which may be obtained here [25]. The construction of machine learning models that can automatically scan fundus images for the early detection of DR in remote areas where medical screening is difficult and time-consuming is the goal of this project. The collection includes 3662 retinal pictures that were captured utilizing fundus photography at several clinics by India's Aravind Eye Hospital under various imaging circumstances. These classifications apply to the fundus pictures in this dataset: no DR (class 0), mild DR (class 1), moderate DR (class 2), severe DR (class 3), and proliferative DR (class 4). Table (2) illustrates the APTOS dataset's severely skewed class distribution. Figure 4 shows the number of pictures in each category.

Table 2 : the APTOS dataset's class distribution and the percentage of each class in the dataset.

Class	Label	Count
0	No_DR	**1805**
1	Mild_DR	**370**
2	Moderate_DR	**999**
3	Severe_DR	**193**
4	Proliferative_DR	**295**

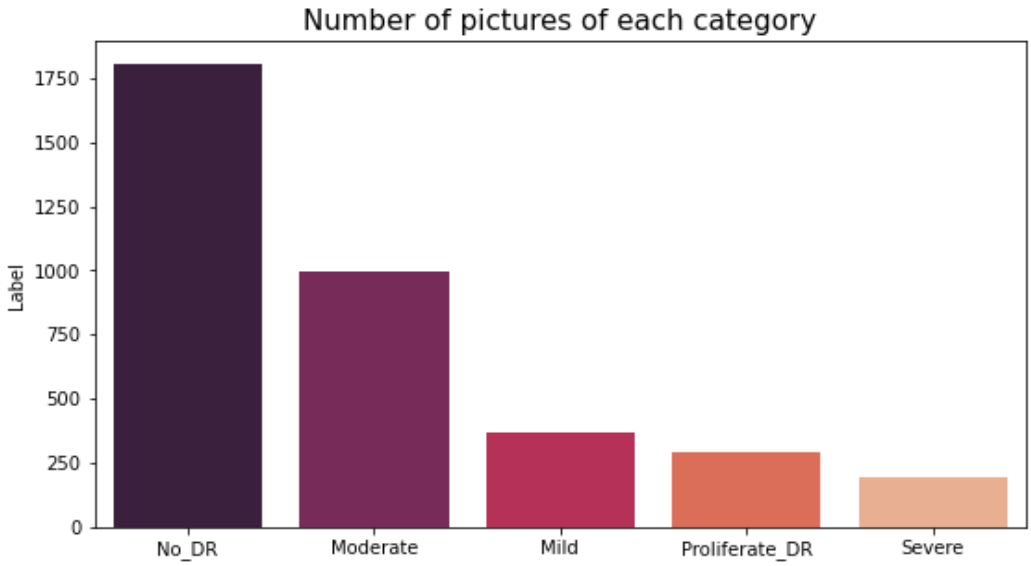

Figure 4: Number of pictures in each category

C. Preprocessing of Data

- Images were captured utilizing a variety of imaging techniques, lighting setups, and clinic locations. The pictures in the collection, therefore, differ in size, brightness, and even focus. Preprocessing methods were employed to eliminate contradictions and enhance critical properties. The Gaussian filter is one of the most important techniques in image processing. It is a low-pass filter that operates in the frequency domain. Since the filter passes low frequencies and noise has a high frequency, it has been the subject of much research in image processing and computer vision. Despite being considered the best filter in many ways, it has a variety of issues. A Gaussian filter for noise suppression is used to smooth out the noise while still distorting the signal. A Gaussian filter for noise suppression is used to smooth out the noise while still distorting the signal.

Edge position edges, displacements, and phantom edges all disappear when edge detection with a Gaussian filter as a preprocessor is used. A Gaussian filter performs better than a monopolization filter at separate frequencies. This task is simplified by having access to an online windowed source filter. It may be used on a 256-by-256-pixel image on average as long as it is clear. The filters and transform algorithms reduce noise while maintaining smoother, more defined edges in the photos. Performance decreases as noise variation rises. With the use of smoothing components, you may use filters to eliminate noise from photos. To prevent the image from degrading, it is crucial to reduce the high-frequency input; otherwise, the low-frequency input would have no impact. Figure 5 shows the difference between original fundus photos and processed images.

- **Resizing:** The proportions and sizes of the photos also vary. The pictures in the dataset range in width and height. However, they must be the same size to be comparable. As a result, a photo resizing process is now necessary. Because EfficientNetB1 is utilized, each image is enlarged to 224 * 224, and then bicubic interpolation is used to drop each image from each retinal circle at the exact center of the picture, from the final resolution's center.

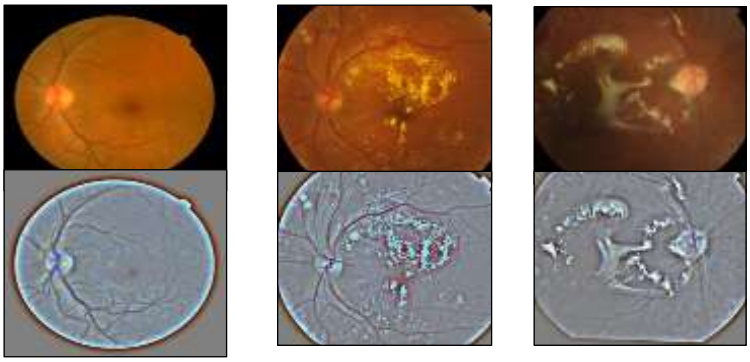

Figure 5 : Top row: original fundus photos; bottom row: processed images (bottom row).

- **Data Augmentation:**

Effective training of a neural network requires sufficient training data, yet in most neural network applications, this need is rarely satisfied. The lack of data is becoming more and more troublesome in medical imaging applications due to the cost of annotation and the imbalance in the occurrence of disorders. To alleviate and effectively utilize the data and data shortages that are already accessible in our experiment, the Augmentator software package is employed, and certain data augmentation techniques must be used [32]. The following techniques were used in particular to supplement our data: The image is rotated at random between 25 and 25 degrees. Change the picture randomly. Data augmentation is a frequently used tactic for increasing the quantity and variety of labeled training sets. It works by utilizing input modifications that retain matching output labels. To avoid overfitting in deep learning models, image augmentations are a common implicit regularization technique used in computer vision and are frequently employed to boost performance. The possibilities for basic photo transformations are often limited to flipping, rotating, scaling, and cropping, even though the majority of deep learning frameworks do include them. Additionally, several existing image augmentation libraries operate at different speeds while processing pictures. Table (3) shows the number of images in each row before and after data augmentation.

Table 3 : the number of images in each row before and after data augmentation.

Class no	Label	Before augmentation	After augmentation
0	No DR	(1500)	1500
1	Mild DR	(370)	1500
2	Moderate DR	(999)	1500
3	Severe DR	(193)	1525
4	Proliferative DR	(295)	1500

D. Performance Metrics

Additional performance matrices in this study must be examined to assess the effectiveness of the models that were given. Accuracy, pinpoint accuracy, recall, and F1 score are the most widely used performance metrics in the field of deep learning, and they are supplied by equations (1) through (5).

- **Accuracy** is the ratio of properly categorized events to all occurrences [33].

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN) \quad (1).$$

- **Sensitivity** or Recall: measures how frequently good occurrences are categorized [34].

$$\text{Recall}=TP/(TP+FN) \quad (2).$$

- **Precision** is defined as the ratio of positively recognized samples that were successfully counted to those that were mistakenly counted [35].

$$\text{Precision}=TP/(TP+FP) \quad (3).$$

- **F1-Score**: The arithmetic average of recall and accuracy is known as the F1-score.

$$\text{F1}=2\times(Precision\times Sensitivity\ )/(Precision+Sensitivity) \quad (4).$$

- **Cohen Kappa Score** is a typical statistic for assessing the degree of agreement between two raters. It can also be used to evaluate the efficacy of a categorization model [36].

$$Cohen\ Kappa=p0-pe/1-pe \quad (5).$$

Where the total number of positive samples (TP), the total number of negative samples (TN), the total number of positive samples (FP), and the total number of erroneous samples (FN) are all included. (pe) indicates the level of agreement between model predictions and actual class values as if they occurred randomly. (p0) indicates the model's overall accuracy.

E. Confusion Matrix

A N x N matrix, where N is the number of target classes, is used to evaluate the efficacy of a classification model. The figure below shows the confusion matrix. A matrix that contrasts the predicted values of the machine-learning model with the actual goal values [37].

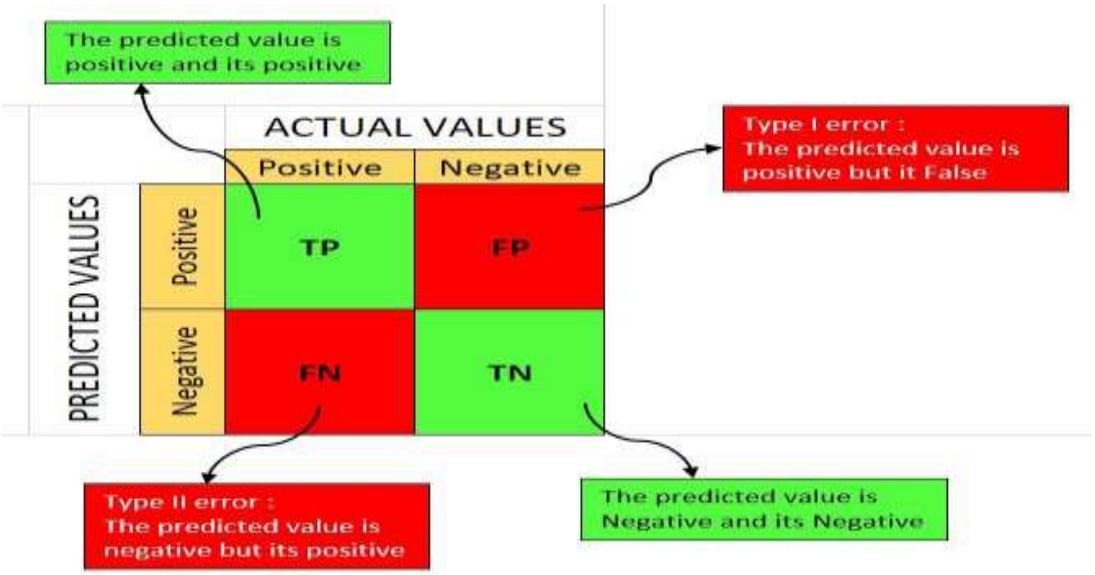

Figure 6 : The four cases in Confusion Matrix

4. Experimentation And Result

A. *Setup for The Experiment*

The following describes our data separation policy: Eighty percent of the photos from each class are submitted to the training set, the validation set, and the test set, respectively. The images have been reduced in size to 224 by 224 pixels. Avg-pooling reduces the number of parameters that must be learned, which lowers the computational cost and gives the internal representation some basic translation invariance. Adam optimizer used the ReLU activation

function for completely connected layers. Eleven convolutional network models were produced with a batch size of 32. The models are trained for a single iteration. Utilizing the Keras package as a backend and TensorFlow as a frontend, our study runs on an Nvidia GeForce GTX 1080 Ti GPU with 8 GB RAM.

B. Results and Discussions

Following validation based on results, all models are tested on 752 additional photographs, and 6020 photographs are obtained using 753 test images. The training cycle for the models was one. Models are compared using training time (sec), validation time (sec), and training accuracy. The outcomes they got are summarized in Table 4. As demonstrated in Table 4, the model MobileNet surpasses all analogs and achieves the maximum degree of accuracy for five classification criteria by combining data from intermediate convolutional layers. Table 4 demonstrates that the recommended model offered excellent training accuracy and high validation accuracy with little training time (sec).

Table 4 : Convolutional network was built. The models are trained for one iteration. The comparison is made based on accuracy and training time.

#	Model	train_accuracy	val_accuracy	Training time (sec)
0	MobileNet	0.7299	0.7660	29.37
1	InceptionResNetV2	0.6908	0.7386	151.38
2	VGG16	0.6490	0.7386	261.54
3	InceptionV3	0.7073	0.7325	62.62
4	DenseNet169	0.7299	0.7234	138.20
5	DenseNet121	0.7161	0.7204	112.59
6	MobileNetV2	0.7212	0.7204	35.49
7	VGG19	0.6224	0.7143	321.74
8	ResNet50	0.6365	0.6900	107.54
9	ResNet101	0.6237	0.6869	187.14
10	EfficientNetB1	0.4966	0.4468	86.20

Performance of InceptionResNetV2, VGG16, InceptionV3, DenseNet169, DenseNet121, MobileNetV2, VGG19, ResNet50 EfficientNetB1, and ResNet101 are evaluated and compared to the model MobileNet to validate the performance of the recommended design as shown in Table 4.

Our model increases InceptionResNetV2 accuracy by 3% and decreases training time by 122.01 seconds, VGG16 accuracy by 8% and decreases training time by 232.16 seconds, InceptionV3 accuracy by 2% and decreases training time by 33.25 seconds, DenseNet169 accuracy by 0% and decreases training time by 108.83 seconds, and DenseNet121 accuracy by 1% and decreases training time by 83.22 seconds. The MobileNetV2 design outperforms the InceptionResNetV2, VGG16, InceptionV3, DenseNet169, DenseNet121, VGG19, ResNet50, EfficientNetB1, and ResNet101 designs in terms of training accuracy, validation accuracy, and training time (sec). In terms of training accuracy, validation accuracy, and training time, the MobileNetV2 design surpasses the InceptionResNetV2, VGG16, InceptionV3, DenseNet169, DenseNet121, VGG19, ResNet50, EfficientNetB1, and ResNet101 designs (sec).

The statistics show that, in terms of accuracy and training time, the MobileNet model is the most effective model for resolving our problems. Figure 7 displays the accuracy of MobileNet.

As shown in Figure 7, the accuracy increased gradually, and the accuracy of the test deteriorated at the beginning and then increased. This proves that there is no overfitting.

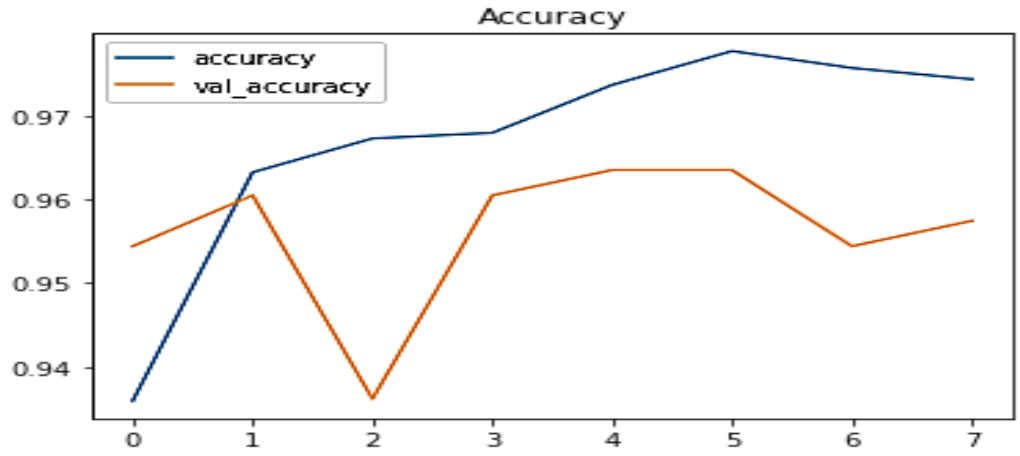

Figure 7: Accuracy chart for MobileNet

The Adam optimizer and ReLU activation function will be utilized with this model, along with 10 epochs and 32 batch sizes. Results show the recall value is 0.9455, the accuracy value is 0.9455, the precision value is 0.94651, the F1 score is 0.94556, and the Cohen Kappa value is 0.889083 when utilizing MobileNetV2 with 10 epochs, 32 batch sizes, the Adam optimizer, and the ReLU activation function. The Cohen Kappa score is 0.80275, while the accuracy, precision, recall, F1 score, and accuracy scores are each 0.90188. Figure 8 shows the MobileNet confusion matrix, whereas Figure 9 depicts the MobileNetV2 confusion matrix.

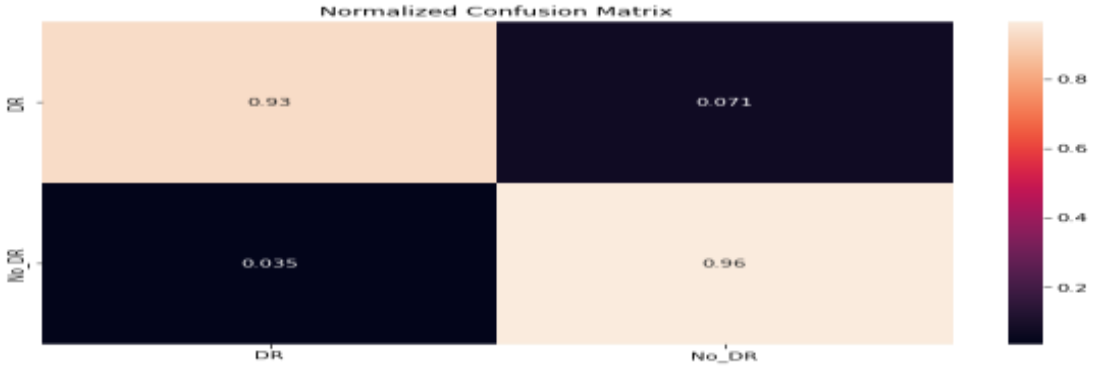

Figure 8 : Depicts the MobileNet confusion matrix

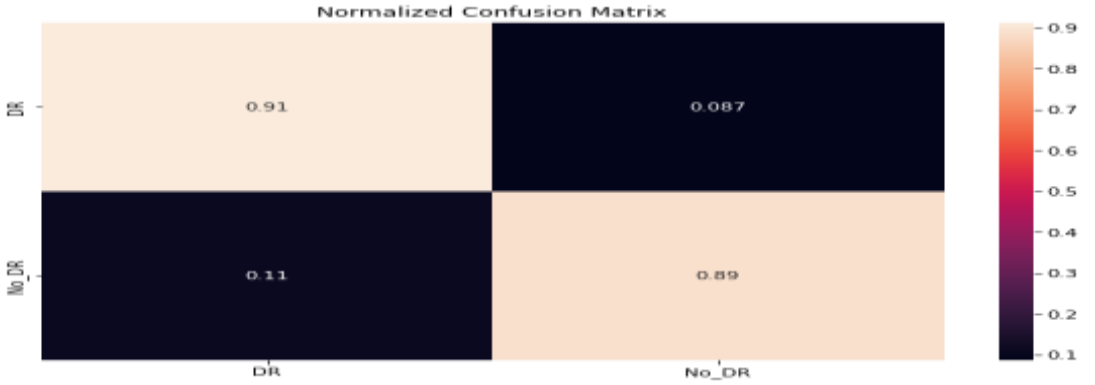

Figure 9 : Depicts the MobileNetV2 confusion matrix

5. Conclusion And Future Work

Diabetic retinopathy is the medical term for diabetes that causes eye damage. This condition could show symptoms or just cause minor visual issues, but it might ultimately result in blindness. The possibility of detecting DR early has grown thanks to the development of artificial intelligence (AI) and deep learning (DL) models, among other computer technologies. Early identification will improve the odds of recovery and reduce the danger of vision loss for patients. One of the most crucial image processing techniques, the Gaussian filter, is used in this work to enhance the quality of fundus images. To cut down on training time, deep transfer learning models for medical DR detection were looked at. The study examines how effectively binary categorization is handled by CNN architectures. There were 11 trained convolutional network models. The models for one iteration are built based on their precision and training length. As a deep transfer learning model, MobileNet was compared in this study to InceptionResNetV2, VGG16, InceptionV3, DenseNet169, DenseNet121, MobileNetV2, VGG19, ResNet50m, EfficientNetB1, and ResNet101. According to the data, the MobileNet model is the most effective model for solving our issues in terms of accuracy and training time. The MobileNet model got the greatest score overall. This model, with 10 epochs and 32 batch sizes, used the Adam optimizer and the ReLU activation function. The findings revealed that the F1 score was 0.94556, the Cohen Kappa score was 0.89083, the accuracy was 0.9455, the precision was 0.94651, and the recall was 0.9455.

A plan to test our methodology in the future using datasets from conditions that affect the eyes similarly, including hypertensive retinopathy, maculopathy, and macular degeneration, can be considered as future work. A model that would run on cloud services was also created to get around memory and processing constraints. Additionally, the fundamental IOT method utilizing DL will be described.

Reference

[1] H. S. Al-Musawi, M. Al-Lami, and A. H. Al-Saadi, "Assessment of Glycemic Control, Renal Function, and Oxidative Stress Parameters in Type 2 Diabetes MellitusPatients," *Iraqi Journal of Science*, pp. 4628-4638, 20 Feb 2021, doi: 10.24996/ijs.2021.62.12.4.

[2] M. H. Yingfeng Zheng, and Nathan Congdon, "The worldwide epidemic of diabetic retinopathy," *Indian J Ophthalmol*, vol. 60, pp. 429-431, Sep-Oct 2020 2020, doi: 10.4103/0301-4738.100542.

[3] L. M. Chen, D. J. Zimmet, P. Z., "The worldwide epidemiology of type 2 diabetes mellitus--present and future perspectives," *Nat Rev Endocrinol*, vol. 8, no. 4, pp. 228-36, Nov 8 2011, doi: 10.1038/nrendo.2011.183.

[4] L. W. Guariguata, D. R. Hambleton, I.Beagley, J.Linnenkamp, U.Shaw, J. E., "Global estimates of diabetes prevalence for 2013 and projections for 2035," *Diabetes Res Clin Pract*, vol. 103, no. 2, pp. 137-49, Feb 2014, doi: 10.1016/j.diabres.2013.11.002.

[5] R. R. A. S. Bourne, Gretchen A.White, Richard A. Smith, Jennifer L. Flaxman, Seth R. Price, HollyJonas, Jost B. Keeffe, Jill Leasher, Janet Naidoo, Kovin Pesudovs, Konrad Resnikoff, Serge Taylor, Hugh R., "Causes of vision loss worldwide, 1990–2010: a systematic analysis," *The Lancet Global Health*, vol. 1, no. 6, pp. e339-e349, dec 2013 2013, doi: 10.1016/s2214-109x(13)70113-x.

[6] M. E. Okur, I. D. Karantas, and P. I. Siafaka, "Diabetes mellitus: A review on pathophysiology, current status of oral medications and future perspectives," *Acta Pharm. Sci.*, Vol 55, No 1, 2017.

[7] M. A. Lotfy, J. Kalasz, H. Singh, J. deghate, E., "Chronic Complications of Diabetes Mellitus: A Mini Review," *Curr Diabetes Rev*, vol. 13, no. 1, pp. 3-10, 17 oct 2017, doi: 10.2174/1573399812666151016101622.

[8] R. C. Ankita Guptaa, "Diabetic Retinopathy: Present and Past," presented at the International Conference on Computational Intelligence and Data Science (ICCIDS 2018), 11 feb, 2018.

[9] M. D. G. Abramoff, M. K. Sonka, M., "Retinal imaging and image analysis," *IEEE Rev Biomed Eng*, vol. 3, pp. 169-208, 1 Jan 2010, doi: 10.1109/RBME.2010.2084567.

[10] N. W. Cheung, J. J. Klein, R. Couper, D. J. Sharrett, A. R. Wong, T. Y., "Diabetic retinopathy and

the risk of coronary heart disease: the Atherosclerosis Risk in Communities Study," *Diabetes Care*, vol. 30, no. 7, pp. 1742-6, Jul 2007, doi: 10.2337/dc07-0264.

[11] G. G. S, Varun P.Palanisamy, P., "A lightweight CNN for Diabetic Retinopathy classification from fundus images," *Biomedical Signal Processing and Control*, vol. 62, 2020, doi: 10.1016/j.bspc.2020.102115.

[12] S. H. K. Kassani, Peyman Hosseinzadeh Khazaeinezhad, Reza Wesolowski, Michal J. Schneider, Kevin A. Deters, Ralph, "Diabetic Retinopathy Classification Using a Modified Xception Architecture," presented at the IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2019.

[13] I. M. Qureshi, Jun Abbas, Qaisar, "Recent Development on Detection Methods for the Diagnosis of Diabetic Retinopathy," *Symmetry*, vol. 11, no. 6, pp. 1-34, 3 June 2019, doi: 10.3390/sym11060749.

[14] U. R. C. Acharya, C. K. Ng, E. Y. Yu, W. Chee, C., "Application of higher order spectra for the identification of diabetes retinopathy stages," *J Med Syst*, vol. 32, no. 6, pp. 481-8, Dec 2018, doi: 10.1007/s10916-008-9154-8.

[15] Z. L. Gao, Jie Guo, Jixiang Chen, Yuanyuan Yi, Zhang Zhong, Jie, "Diagnosis of Diabetic Retinopathy Using Deep Neural Networks," *IEEE Access*, vol. 7, pp. 3360-3370, 2019, doi: 10.1109/access.2018.2888639.

[16] M. M. Adly, "On the Grading of Diabetic Retinopathies using a Binary-Tree-based Multiclass Classifier of CNNs," *IJCSIS*, 2019.

[17] R. Jammula, "Optimal Transfer Learning Model for Binary Classification of Funduscopic Images Through Simple HeuristicsRohit Jammula," *arXiv1*, 2018.

[18] S. B. Shivashish Naithani, Dhiraj Kumar., "Automated Detection of Diabetic Retinopathy using Deep Learning," *IRJET*, vol. 06, pp. 2945-2947, Apr 2019 2019, doi: 2395-0072.

[19] D. S. Doshi, Aniket Sidhpura, Deep Gharpure, Prachi, "Diabetic retinopathy detection using deep convolutional neural networks," presented at the International Conference on Computing, Analytics and Security Trends (CAST), 13 Dec, 2016.

[20] I. K. a. M. Castell, "Transfer Learning with Convolutional Neural Networks for Diabetic Retinopathy Image Classification. A Review," *Medical Artificial Intelligence*, vol. 6, pp. 1-10, 16 March 2020 2020, doi: https://doi.org/10.3390/app10062021.

[21] Q. V. L. Mingxing Tan "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *arXiv*, pp. 1-11, 11 Sep 2020, doi: 1905.11946v5

[22] A. G. H. M. Z. B. C. D. Kalenichenko., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv*, pp. 1-9, 17 Apr 2017, doi: 1704.04861v1.

[23] A. B. D. Nicholas Dominic And T. W. C. , Bens Pardamean, , "Transfer Learning Using Inception-Resnet-V2 Model To The Augmented Neuroimages Data For Autism Spectrum Disorder Classification.," *SCIK Publishing Corporation*, pp. 1-21, February 16, 2021 2021, doi: 10.28919.

[24] J. M. Zhong , Ling Liu, And Dong-Yang Zhou, "Crack Detection of Concrete Pavement With Cross-Entropy Loss Function and Improved VGG16 Network Model," *IEEE*, pp. 1-10, March 27, 2020 2020, doi: 10.1109.

[25] kaggle. " Asia Pacific Tele-Ophthalmology Society (APTOS) 2019 Blindness Detection " https://www.kaggle.com/c/aptos2019-blindness-detection/ (accessed.

[26] W. G. Hao Ran, Jihang Mi, Zhenwei Zhao, "Fine-grained recognition of Chinese food image based on DenseNet with attention mechanism," *International Conference on Graphics and Image Processing (ICGIP 2020)*, vol. 11720, pp. 1-10, 27 January 2021 2021, doi: 10.1117/12.2589449.

[27] P. Prof. Li Fei-Fei, Prof. Jia Deng, Prof. Olga Russakovsky, Prof. Alex Berg, U Prof. Kai Li,, " ImageNet.," ed. Stanford University, 2021.

[28] V. C. Aurelia Michelea, Diaz D. Santika, "MobileNet Convolutional Neural Networks and Support Vector Machines for Palmprint Recognition," *International Conference on Computer Science and Computational Intelligence*, pp. 1-9, 12–13 September 2019.

[29] J. W. Jian Xiao, Shaozhong Cao and Bilong Li, "Application of a Novel and Improved VGG-19 Network in the Detection of Workers Wearing Masks," *Journal of Physics: Conference Series* pp. 1-9, 11 Oct 2020, Art no. 1518, doi: 10.1088/1742-6596/1518/1/012041.

[30] K. Z. He, Xiangyu Ren, Shaoqing Sun, Jian, "Deep Residual Learning for Image Recognition," presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 23

Feb, 2016.

[31] M.-V. N. Quoc-Huy Trinh, "Custom Deep Neural Network for 3D Covid Chest CT-scan Classification," *arXiv*, pp. 1-3, 3 Jul 2021, doi: 10.48550/arXiv.2107.01456.

[32] M. D. Bloice, "Augmentor: An Image Augmentation Library for Machine Learning," *arXiv*, pp. 1-5, 11 Aug 2017, doi: 1708.04680v1.

[33] M. M. F. Farag, Mariam Abdel-Hamid, Amr T., "Automatic Severity Classification of Diabetic Retinopathy Based on DenseNet and Convolutional Block Attention Module," *IEEE Access*, vol. 10, pp. 38299-38308, 6 April 2022, doi: 10.1109/access.2022.3165193.

[34] E. S. D. Nasser, Faten Abd Ali, "Diagnosis and Classification of Type II Diabetes based on Multilayer Neural Network," *Iraqi Journal of Science*, pp. 3744-3758, 14 April 2021, doi: 10.24996/ijs.2021.62.10.33.

[35] M. M. S. Rahma, Aymen Dawood, "Heart Disease Classification–Based on the Best Machine Learning Model," *Iraqi Journal of Science*, pp. 3966-3976, 30 Sep 2022, doi: 10.24996/ijs.2022.63.9.28.

[36] M. J. W. DAVIDE CHICCO, AND GIUSEPPE JURMAN "Cohen's Kappa: what it is, when to use it, how to avoid pitfalls," *IEEE Access*, vol. 9, pp. 78368-78381, June 3 2021, doi: 10.1109/ACCESS.2021.3084050.

[37] M. M. Pawel Trajdos, "Bayes metaclassifier and Soft-confusion-matrix classifier in the task of multi-label classification.," *arXiv*, pp. 1-34, 25 Jan 2019, doi: 10.48550/arXiv.1901.08827