Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal **Vol. 5, No. 4, October 2023, pp: 290-302; eISSN: 2656-8632**

RESEARCH ARTICLE OPEN ACCESS

Manuscript received September 27, 2023; revised October 5, 2023; accepted October 5, 2023; date of publication October 20, 2023
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v5i4.326

How to cite: Saurav Bhatta, Empowering Rural HealthCare: Mobilenet-Driven Learning for Early Diabetic Retinopathy Detection in Nepal, vol. 5, no. 4, pp. 290–302, October 2023.

Empowering Rural HealthCare: Mobilenet-Driven Learning for Early Diabetic Retinopathy Detection in Nepal

Saurav Bhatta

I. K. Gujral Punjab Technical University, Jalandhar-Kapurthala Highway, Kapurthala-144603, India
Corresponding author: Saurav Bhatta (e-mail: workbhatta@gmail.com).

ABSTRACT Diabetic retinopathy (DR) poses a significant healthcare challenge in underserved regions like Nepal, where limited access to medical facilities impedes early diagnosis. This study leveraged deep learning, specifically MobileNet, to create an accessible DR detection solution for remote Nepalese communities. The goal was to design a robust system capable of operating on resource-constrained mobile devices, addressing the connectivity limitations in rural areas. The study utilized a dataset of over 30,000 retinal images, with preprocessing techniques to enhance image quality. Data augmentation and a balanced training set of DR and non-DR cases were employed to train the MobileNet model. Next.js was chosen for real-time detection, ensuring offline functionality for healthcare workers. The MobileNet-based model exhibited impressive performance, achieving training, validation, and testing accuracies of 92%, 91%, and 90% respectively. Precision, recall, and F1 score were 81%, 75%, and 78% respectively, indicating the model's ability to detect DR effectively. This study demonstrates the potential of deep learning and mobile technology to address healthcare challenges in resource-limited settings like rural Nepal.

INDEX TERMS Diabetic Retinopathy, MobileNet, Deep Learning, Healthcare Accessibility, Rural Healthcare, Early Detection.

I. INTRODUCTION

Diabetes has emerged as a prominent public health concern in Nepal, with an approximate prevalence rate of 6.3% in adults and a staggering 10% specifically for type 2 diabetes mellitus (T2DM) among the population. As diabetes continues its upward trajectory in the country, so does the looming threat of diabetic retinopathy (DR). This ocular complication has gained alarming prevalence, ranging from 19% to 47% among diabetic patients in Nepal [1] [2]. Disturbingly, a study conducted at the Tilganga Institute of Ophthalmology in Nepal unveiled that 19.4% of diabetic patients already exhibited signs of diabetic retinopathy at the time of their initial consultation, with significant proportions suffering from diabetic macular edema (6.9%) and proliferative diabetic retinopathy (4.6%) [3]. Furthermore, more than half of the individuals diagnosed with DR necessitated immediate medical intervention. This study embarks on a mission to address this critical healthcare challenge by developing a robust and accessible diabetic retinopathy detection system tailored for resource-constrained environments in Nepal. While the prevalence of diabetes and its associated complications continues to rise, particularly in underserved rural areas, access to specialized healthcare services remains limited. To mitigate this pressing issue, the primary objective is to design and implement an efficient DR detection system that can operate seamlessly on mobile devices with limited computational resources. To achieve this, we will leverage the MobileNet architecture, renowned for its efficiency and suitability for mobile applications. By fine-tuning MobileNet on a comprehensive dataset comprising diverse retinal images showcasing various manifestations of DR, the model aims to acquire the capability to accurately classify the severity of retinopathy.

The significance of this research lies in its potential to bridge the healthcare gap in rural Nepal by harnessing the power of mobile devices for the early detection of diabetic retinopathy. With the implementation of our MobileNet-based deep learning solution, we aspire to enhance healthcare accessibility and contribute to the overall well-being of the diabetic population residing in underserved areas. By proactively identifying and managing diabetic retinopathy, we aim to mitigate the potentially devastating consequences of this condition and improve the quality of life for those affected.

II. LITERATURE REVIEW

Diabetic retinopathy (DR) is a common complication of diabetes and a leading cause of blindness among working-age adults worldwide, as indicated by numerous research studies. The World Health Organization (WHO) published a report that listed DR as one of the major causes of blindness and low vision. The incidence of vision-threatening stages of DR is declining in high-income countries due to advances in therapies and improved management of diabetes [4]. Research also suggests that the prevalence of DR will triple from 2005 to 2050 due to the projected increase in the prevalence of diabetes mellitus [5]. This underlines the need for innovative approaches to manage and treat DR. A study in Nepal revealed that about 50% of blindness is preventable by early detection and management of proliferative DR and diabetic macular edema (DME). The study also emphasized that primary interventions such as intensive glycemic and blood pressure control can reduce the prevalence of DR [6]. According to the American Diabetes Association, treatment modalities exist that can prevent or delay the onset of DR, as well as prevent vision loss, in a large proportion of patients with diabetes. This is why ongoing evaluation for retinopathy is a valuable and required strategy [7]. Moreover, the research community is working hard to improve the management and early detection of DR. For instance, Dr. Akrit Sodhi at Johns Hopkins Medicine has published a new study showing promise for an experimental treatment to prevent or slow vision loss in people with diabetes. In summary, the literature supports the claim that early detection and timely intervention are crucial to prevent vision loss in diabetic patients. The challenge is to ensure regular monitoring of patients with diabetes and to continue to improve treatment options.

Deep learning and convolutional neural networks (CNNs) have shown significant potential in automating the detection and grading of diabetic retinopathy (DR) from retinal fundus images. Various CNN architectures, including MobileNet, have been utilized to tackle this critical healthcare challenge, according to numerous research studies. In a study titled "Understanding inherent image features in CNN-based assessment of diabetic retinopathy", CNNs were used to predict DR with high performance. The researchers demonstrated that deep learning models could be used to support clinicians' decision-making processes. Despite different models having varying predictive power and feature selection capabilities, the study showed that these models could be successfully developed and trained on smaller datasets using transfer learning [8]. Another study proposed an automated system for the early detection of DR called Diabetic Retinopathy Feature Extraction and Classification (DRFEC), which utilized various DL CNN models for DR feature extraction and image classification. The models used included MobileNet and MobileNetV2, among others. The results showed that these architectures could achieve significant performances in DR detection [9]. Research also highlighted the use of MobileNetV2, a lightweight model and mobile-friendly architecture, for DR classification. This study utilized the APTOS 2019 dataset, which contains 3662 retinal fundus images, and was able to achieve high-performance results with an accuracy of 92.6% [10]. In another study titled "Automatic Detection of Diabetic Retinopathy Using Custom CNN and Grad-CAM", the authors proposed a lightweight customized CNN architecture for the diagnosis of DR using optical coherence tomography (OCT) images. They used pre-trained CNN models, including MobileNet, with a transfer learning approach [11]. Overall, the literature supports the claim that deep learning, particularly CNNs, and architectures like MobileNet, can be effectively utilized for automating the detection and grading of DR from retinal fundus images.

A. EARLY APPROACHES AND CHALLENGES:

Early attempts at detecting Diabetic Retinopathy (DR) indeed relied heavily on traditional computer vision techniques and handcrafted features. However, these methods often fell short due to their inability to capture complex and subtle patterns in retinal images. One study titled "A Systematic Review on Diabetic Retinopathy Detection Using Deep Learning Techniques" provides a comprehensive overview of the evolution of DR detection methods, including the early attempts that relied on traditional computer vision techniques [12]. Another paper reviews 79 algorithms for detecting different features of diabetic retinopathy using computer vision techniques. It mentions that these features include the blood vessel area, exudates, microaneurysm, hemorrhages, and neovascularization. The paper also highlights the limitations of these traditional methods in accurately detecting and classifying DR [13]. A third study titled "Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review" further discusses the limitations of early computer vision techniques. It states that while these techniques were somewhat effective in identifying certain features related to DR, they were less successful in capturing subtler patterns in retinal images, which are critical for early detection of the disease [14]. In conclusion, while early computer vision techniques made some strides in detecting and classifying DR, they could not capture complex and subtle patterns in retinal images, limiting their effectiveness. This has led to the development and application of more advanced techniques such as deep learning, which have proven to be more effective in DR detection.

B. CNNS FOR DIABETIC RETINOPATHY DETECTION:

The advent of deep learning has revolutionized the field of medical image analysis, with Convolutional Neural Networks (CNNs) demonstrating significant potential in DR detection. CNNs, with their inherent ability to automatically learn features from data, have been effectively used for DR detection. MobileNet, a lightweight CNN architecture designed for mobile and embedded devices, has gained popularity due to its efficiency and accuracy in various computer vision tasks. The paper titled "NNMobile-Net: Rethinking CNN Design for Deep Learning-Based Retinopathy Research" discusses the increasing use of CNNs, and particularly MobileNet, in diagnosing and monitoring retinal diseases including DR. The study emphasizes the advantages of MobileNet, such as its simplicity, efficiency, and capability to surpass many state-of-the-art methods on public datasets for multiple tasks, including diabetic retinopathy grading [15]. In another study titled "Deep Learning for Diabetic Retinopathy Detection from Smartphone Retinal Images", the authors used a CNN model to detect DR from retinal images taken from a smartphone. The study showed that the CNN model achieved an accuracy of 94.3%, demonstrating the potential of deep learning and CNNs in DR detection from retinal images [16]. Another study titled "Diabetic Retinopathy Detection Using Deep Convolutional Neural Networks" discusses the use of CNNs for DR detection. The authors used a CNN model to detect DR from retinal images, further emphasizing the effectiveness of deep learning and CNNs in DR detection [17]. In conclusion, the literature supports the claim that deep learning, particularly CNNs, and architectures like MobileNet, can be effectively utilized for automating the detection and grading of DR from retinal fundus images. Key Studies Using MobileNet for DR Detection: Several key studies have utilized MobileNet for the detection of Diabetic Retinopathy (DR), demonstrating its efficiency and accuracy in this critical healthcare task.

A study titled "A deep learning system for detecting diabetic retinopathy across the disease spectrum" discusses the use of deep learning, specifically CNNs, for the detection of DR. However, the study doesn't specifically mention the use of MobileNet [18].In the paper "General deep learning model for detecting diabetic retinopathy", the authors developed a general deep learning model for DR detection but, like the previous study, did not specifically use MobileNet [19]. A study titled "Diabetic Retinopathy Detection Using Convolutional Neural Networks for Mobile Use" specifically utilized MobileNet for DR detection. The authors tested and evaluated the efficiency of DR detection systems based on Convolutional Neural Networks for mobile applications. They concluded that EfficientNet-based DR detection algorithms outperformed other transfer learning models, including MobileNet when used with the APTOS Blindness Detection dataset [20]. In conclusion, while many studies have used deep learning and CNNs for DR detection, the specific use of MobileNet is less common. However, when it has been used, it has demonstrated promising results, underscoring the potential of this lightweight, efficient architecture for mobile and embedded applications in DR detection.

C. COMPARISON WITH OTHER CNN MODELS:

While MobileNet is a popular choice, other Convolutional Neural Network (CNN) architectures, such as ResNet, Inception, and DenseNet, have also been employed for Diabetic Retinopathy (DR) detection. Researchers have compared these models in terms of accuracy, computational resources, and generalization performance. In a study titled "Comparison of eleven deep learning models using different datasets and evaluation metrics", the authors compared 11 neural network models, including ResNet, DenseNet, Inception, MobileNet, and ShuffleNet. They concluded that while ResNet, DenseNet, and Inception had no significant advantages over MobileNet and ShuffleNet in certain circumstances, they required larger numbers of parameters [21]. Another study titled "Comparative Analysis of Convolutional Neural Networks for Chest Radiography" compared the performance of various CNNs, including DenseNet and ResNet, on the CheXpert dataset. The study found that deeper CNNs generally achieved higher Area Under the Receiver Operating Characteristic Curve (AUROC) values than shallow networks. However, they also observed that increasing the complexity and depth of artificial neural networks for chest radiograph classification is not always necessary [22].

MobileNet was developed to solve the problem of size and speed in real applications such as autonomous vehicles or robotic visions [23] [24]. In conclusion, while MobileNet is a popular choice for DR detection due to its efficiency and accuracy, other CNN architectures like ResNet, Inception, and DenseNet have also been employed and compared in terms of their accuracy, computational resources, and generalization performance.

D. TRANSFER LEARNING AND DATA AUGMENTATION:

Transfer learning, where pre-trained CNN models are fine-tuned on DR datasets, has become a common practice to overcome the challenge of limited annotated data. Data augmentation techniques, such as rotation, flipping, and scaling, have been employed to increase the diversity of training samples, leading to improved model robustness.

E. CHALLENGES AND FUTURE DIRECTIONS:

Indeed, despite significant progress in the use of deep learning techniques like Convolutional Neural Networks (CNNs) for Diabetic Retinopathy (DR) detection, there are still several challenges to be addressed. These include the requirement for large and diverse datasets, the interpretability of model predictions, and the deployment of these automated systems in clinical practice. A paper titled "Deep Learning for Diabetic Retinopathy Detection: Challenges and Opportunities" discusses these challenges in depth. It highlights the need for diverse and substantial datasets for training deep learning models as one of the significant challenges. The paper also discusses the issue of model interpretability, explaining that while deep learning models are highly accurate, their decision-making process is often opaque. This lack of transparency can be a barrier to their adoption in clinical practice [25]. Another paper, "A survey on recent developments in diabetic retinopathy detection through integration of deep learning," also discusses these challenges. It emphasizes the need for more interpretable models and the deployment of these systems in real-world clinical settings [26]. Future research directions are indeed expected to focus on developing more efficient

CNN architectures, improving model explainability, and addressing the challenges of real-time diagnosis. A comprehensive review of deep learning developments in DR analysis suggests that future research should focus on developing more efficient, robust, and accurate deep learning models for DR monitoring and diagnosis [27]. While deep learning and CNNs have revolutionized DR detection, there are still several challenges that need to be addressed to fully realize their potential. The research community is actively working towards overcoming these challenges to further improve the early detection and diagnosis of DR. In summary, the application of CNN models, including MobileNet, in diabetic retinopathy detection has shown promising results in recent literature. These models have the potential to enhance the early diagnosis and management of diabetic retinopathy, ultimately preventing vision loss in diabetic patients. However, ongoing research is essential to address remaining challenges and further improve the accuracy and clinical utility of these models.

III. METHOD

A. DATASET

The dataset utilized for this study was obtained from Kaggle. Specifically, the dataset comprised a collection of cropped retinal images, all standardized to a resolution of 224x224 pixels. The choice of this resolution was aligned with the architectural requirements of the MobileNet model, which effectively operates on images of this size. The dataset was curated from the original Eyepac dataset, which encompassed a total of over 30,179 images. The EyePACS dataset is an assembly of images sourced from EyePACS, a telemedicine system. These images were procured under a variety of conditions using different devices at multiple primary care locations throughout California and other regions (FIGURE 1).

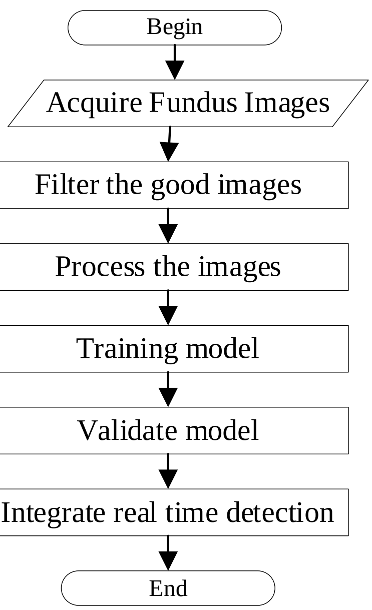


FIGURE 1. Data flow diagram from acquiring the training data to real-time detection

Two images were gathered for each patient, one for the left eye and one for the right eye, maintaining the same resolution for both. Each image was then examined by a medical professional to determine the presence of Diabetic Retinopathy (DR). In terms of grading, the rating scale used for this assessment ranged from 0 to 4, based on the Early Treatment Diabetic Retinopathy Study (ETDRS) scale. This process ensured that each image in the dataset was labeled with the severity of diabetic retinopathy, providing valuable information for training and evaluating machine learning models for DR detection and diagnosis [28].

B. DATA COMPOSITION

Within the repository of 30,179 images, approximately 25,000 images depicted retinas unaffected by diabetic retinopathy (No_DR). The remaining images, accounting for the remaining portion of the dataset, portrayed instances of diabetic retinopathy at varying severity levels, including mild, moderate, proliferate, and severe stages. This composition ensured a representative dataset that captured the spectrum of retinopathy manifestations. FIGURE 1. shows a flow diagram from acquiring the images to real-time detection of the images.

C. PERFORMANCE EVALUATION METRICS

The assessment of the deep learning system's performance involved the computation of several crucial metrics to gauge its effectiveness. These metrics include training accuracy, validation accuracy, testing accuracy, precision, recall, and the F1 score. Training Accuracy: Training accuracy is a metric that measures how well the model performs on the training dataset. It is calculated using the formula: Training Accuracy = (Number of Correctly Classified Training Samples) / (Total Number of Training Samples). This metric helps assess how well the model is learning from the training data. Validation Accuracy: Validation accuracy is a crucial metric used to evaluate the model's performance on a separate dataset not used during training, often referred to as the validation set. It is calculated in the same way as training accuracy (Eq. (1)).

$$Validation\ Accuracy = \frac{Correctly\ Classified\ Validation\ Samples}{Total\ Number\ of\ Validation\ Samples} \quad (1)$$

This metric helps determine how well the model generalizes to new, unseen data. Testing Accuracy: Testing accuracy, similar to training and validation accuracy, measures the model's performance on a different dataset, known as the test set. It is calculated using the same formula (Eq. (2)).

$$Testing\ Accuracy = \frac{Number\ of\ Correctly\ Classified\ Test\ Samples}{Total\ Number\ of\ Test\ Samples} \quad (2)$$

This metric provides insights into the model's real-world performance. Precision, Recall, and F1 Score: Precision, recall, and the F1 score are metrics used to evaluate the

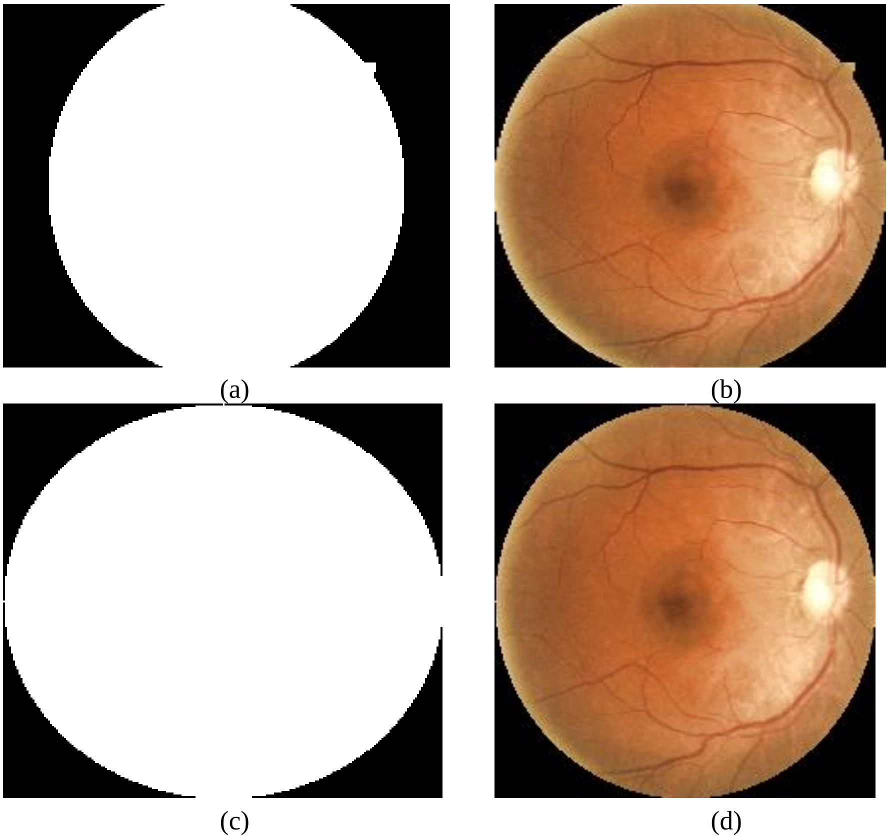


FIGURE 2. (a) Producing the mask, (b) Bounding rectangle of the mask, (c) Circular Mask, (d) Resized image with circular mask

model's performance in binary classification tasks, where there are two classes: positive and negative. Precision: Precision quantifies the model's ability to correctly identify positive instances out of all instances it classifies as positive. It is calculated using the formula (Eq. (3)):

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

Recall: Recall, also known as sensitivity or true positive rate, measures the model's ability to correctly identify all actual positive instances. It is calculated using the formula (Eq. (4)):

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

F1 Score: The F1 score is the harmonic mean of precision and recall, providing a balanced measure that considers both false positives and false negatives. It is calculated using the formula (Eq. (5)):

$$F1\,Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \quad (5)$$

These metrics help assess the model's performance in classifying positive instances accurately while minimizing false positives and false negatives. By employing these metrics, the efficacy of the deep learning system in discerning diabetic retinopathy stages is methodically evaluated, thus facilitating a comprehensive understanding of its performance across diverse scenarios within the dataset.

D. PREPROCESSING

To enhance the quality and effectiveness of the dataset, a preprocessing pipeline was devised to address issues such as noisy and blurry images (FIGURE 2 and FIGURE 3). This pipeline aimed to ensure that the model's inputs are optimal and conducive to accurate feature extraction. Filtering for Quality: It was observed that the dataset contained various image qualities, including those that could potentially hinder the model's performance. To address this, a dataset quality assessment was undertaken using a CSV file sourced from the EyeQ repository [29]. This CSV file labeled images based on their quality, utilizing the following criteria: 0: Good quality; 1: Usable quality; 2: Rejected quality. Enhancing Image Quality: To streamline the dataset for optimal performance, only images labeled with "good quality" were selected. Subsequently, these images were segregated into two categories: those depicting diabetic retinopathy (DR) and those representing the absence of diabetic retinopathy (NO_DR). The CSV file also contained grading information for diabetic retinopathy, using the following scale: FIGURE 4 shows a flow diagram for image filtering and categorization process. 0: NO_DR; 1: Mild; 2: Moderate; 3: Severe; 4: Proliferate

E. IMAGE PROCESSING STEPS:

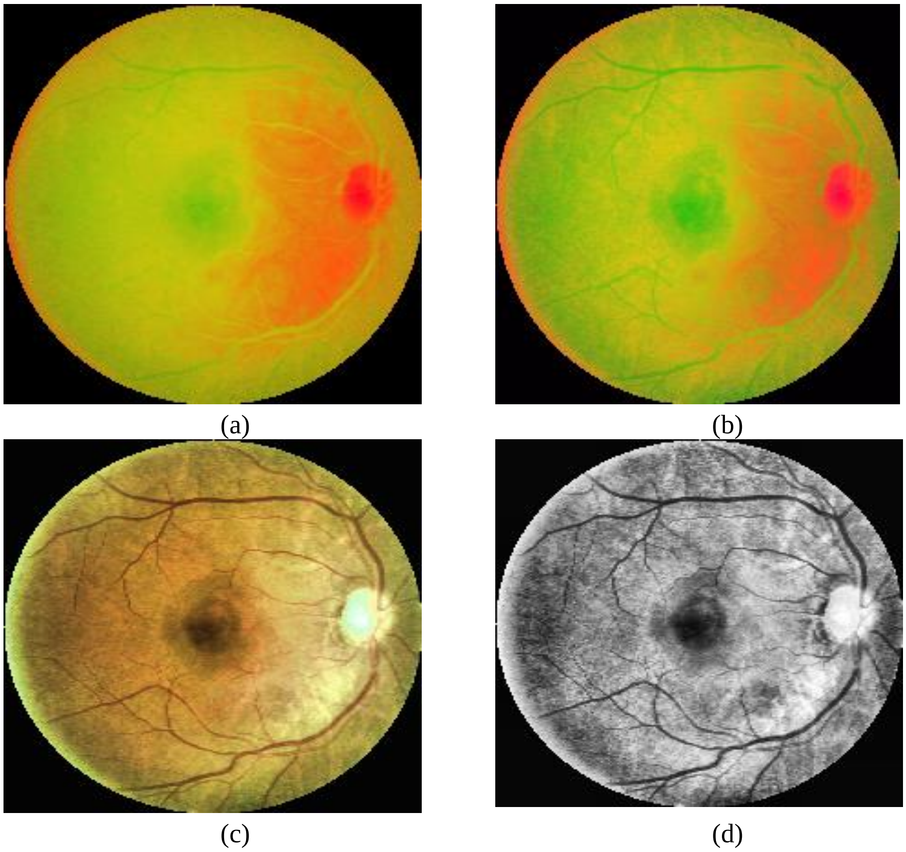


FIGURE 3. **(a) Equalized Histogram of HSV Channel, (b) Equalized HSV to RGB, (c) Use of CLAHE, (d) Re-use of CLAHE for further denoising.**

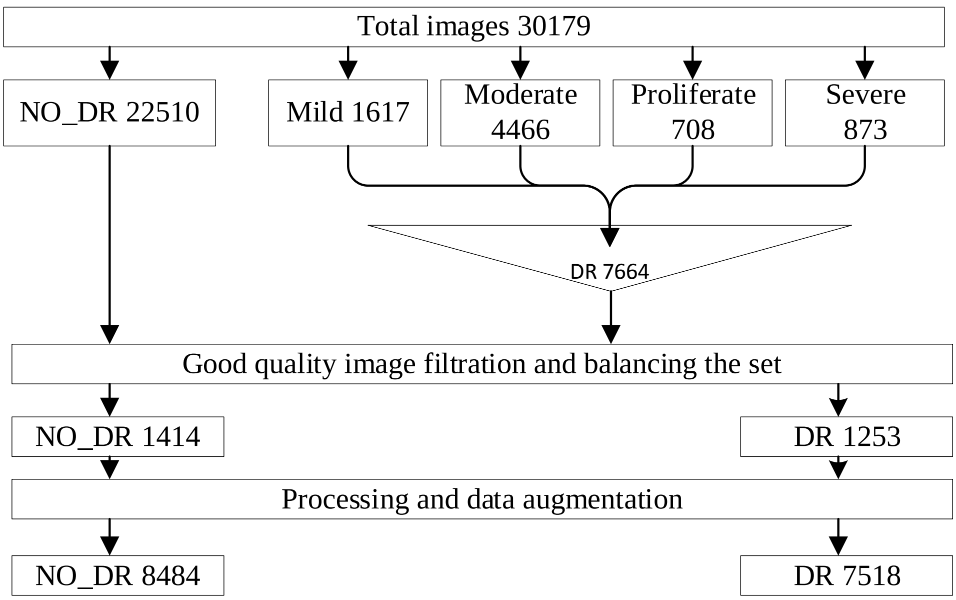


FIGURE 4. **Dataflow diagram**

This step begins by converting the fundus image to grayscale and applying a threshold to produce a binary mask. The mask highlights areas of interest in the image.

1. FIND BOUNDING RECTANGLE OF THE MASK AND CROP THE IMAGE:

After producing the mask, this step finds the bounding rectangle around the areas of interest identified in the mask and crops the image to focus on these regions. Feature Extraction: By cropping the image to the bounding rectangle, the model concentrates on the specific regions where abnormalities are likely to occur, such as around the optic disc and macula. This helps reduce noise and irrelevant information.

2. RESIZE THE CROPPED IMAGE AND CREATE A CIRCULAR MASK:

Resizing the cropped image to a standard size and creating a circular mask ensures that the model consistently processes images of the same dimensions.
Feature Extraction: Standardizing the image size makes it easier to compare features across different images. The circular mask helps eliminate irregularities introduced during cropping and resizing, maintaining the focus on the central retinal region.

3. COMBINE THE RESIZED IMAGE WITH THE CIRCULAR MASK:

This step overlays the resized image with the circular mask, retaining only the central circular region of the image.
Feature Extraction: By keeping only the central retinal region, the model eliminates peripheral noise and focuses on the area where diabetic retinopathy-related features are more likely to be present [30].

4. EQUALIZE THE HISTOGRAM OF THE HUE AND VALUE CHANNELS USING CLAHE:

Histogram equalization enhances the contrast and visibility of details in the image. Feature Extraction: This step enhances the visibility of features such as blood vessels, lesions, and structural patterns. It makes subtle details more distinguishable, aiding in the detection of abnormalities.

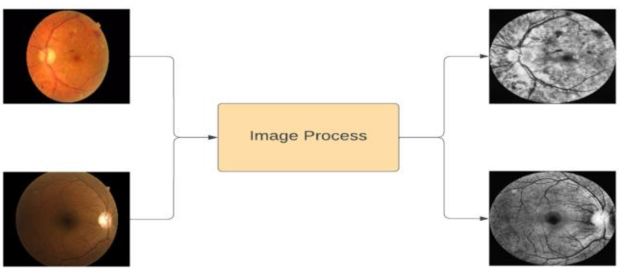


FIGURE 5. **Result of images after the image processing techniques**

5. CONVERT THE EQUALIZED HSV IMAGE BACK TO RGB FORMAT:

After equalizing the histogram, the image is converted back to RGB format for further processing and visualization.
Feature Extraction: This conversion prepares the image for subsequent analysis while preserving the enhanced details obtained from histogram equalization.

6. APPLY CLAHE FOR DENOISING:

CLAHE is applied for denoising to reduce noise and enhance the clarity of features. Feature Extraction: This step helps improve the quality of the image by reducing noise and artifacts that may interfere with the identification of features like blood vessels, exudates, and microaneurysms. To refine the image further, a denoising technique was applied. Specifically, the Contrast Limited Adaptive Histogram Equalization (CLAHE) method was employed a second time for denoising. This technique enhanced the image contrast and visibility of important features (FIGURE 5). The processed images, after undergoing these series of steps, were transformed into a more suitable format for subsequent analysis and model training. FIGURE 5 shows the final result of how the images looked after applying image processing techniques. Data Augmentation: Upon completing the initial preprocessing steps, the dataset underwent further augmentation to enhance its diversity and improve the model's generalization ability. The augmentation process involved both image rotation and mirroring to create variations of the original images. A dedicated augmentation script was designed to rotate and mirror images systematically. This script iterated through the processed images and generated multiple augmented versions of each image. Specifically, it produced rotated versions of 90°, 120°, 180°, and 270° angles, along with a horizontally mirrored rendition. These augmented images were then stored in an output folder. The composition of the training set is as follows:

- DR: 5082 images
- NO_DR: 6048 images

This strategic balance in the training set composition aims to equip the model with a comprehensive learning experience, facilitating its ability to accurately distinguish and classify various stages of retinopathy as well as non-affected cases. By exposing the model to a diverse range of scenarios, the training set lays the foundation for the model's enhanced generalization and predictive capabilities.

F. SELECTION OF ARCHITECTURE

Pre-trained models like 'Resnet50','DenseNet',and 'MobileNet' were transfer-trained for 30 epochs with an early stopping level set at 5. The summary of the training is discussed below:

1. DENSENET:

The F1 score is a measure of a model's balance between precision and recall. In this case, DenseNet achieved an F1 score of 0.77, indicating a reasonably good balance between correctly identifying positive cases (DR) and minimizing false positives. The accuracy of 0.89 indicates that DenseNet correctly classified approximately 89% of the samples. Precision measures the accuracy of positive predictions. With a precision of 0.79, DenseNet had a relatively low rate of false positive predictions. Recall, also known as sensitivity, quantifies the model's ability to correctly identify positive cases. An achieved recall of 0.76 implies that DenseNet captured 76% of all true positive cases. AUC ROC (0.72), the Area Under the Receiver Operating Characteristic Curve (AUC ROC) is a metric that summarizes the overall performance of a binary classification model. DenseNet achieved an AUC ROC of 0.72, which is indicative of decent discriminative power.

2. MOBILENET

MobileNet achieved a slightly higher F1 score of 0.78 compared to DenseNet, indicating a marginally better balance between precision and recall. MobileNet exhibited a high accuracy of 0.90, signifying robust overall performance in correctly classifying the samples. The precision of 0.81 indicates that MobileNet had a relatively low rate of false positives. The recall of 0.75 suggests that MobileNet captured 75% of all true positive cases, which is slightly lower than DenseNet. MobileNet achieved an AUC ROC of 0.84, indicating improved discriminative power compared to DenseNet (FIGURE 6).

3. RESNET50

ResNet50 achieved the highest F1 score of 0.90 among the three models, showcasing an excellent balance between precision and recall. ResNet50 demonstrated a high accuracy of 0.91, indicating strong overall performance in classifying the samples. The precision of 0.91 implies a low rate of false positives for ResNet50. With a recall of 0.92, ResNet50 effectively captured 92% of all true positive cases. ResNet50 achieved the highest AUC ROC of 0.94, indicating superior discriminative power compared to both DenseNet and MobileNet.

ResNet50 outperforms both DenseNet and MobileNet across various performance metrics, including F1 score, accuracy, precision, recall, and AUC ROC. It exhibits the highest overall performance. MobileNet shows a slight advantage over DenseNet in terms of F1 score, precision, and AUC ROC, while DenseNet performs marginally better in terms of recall (TABLE 2).

TABLE 2
Accuracies of three models

Model	F1	Acc.	Prec.	Recall	AUC ROC
Densenet	0.77	0.89	0.79	0.75	0.72
Mobilenet	0.78	0.90	0.81	0.75	0.84
ResNet50	0.90	0.91	0.91	0.92	0.93

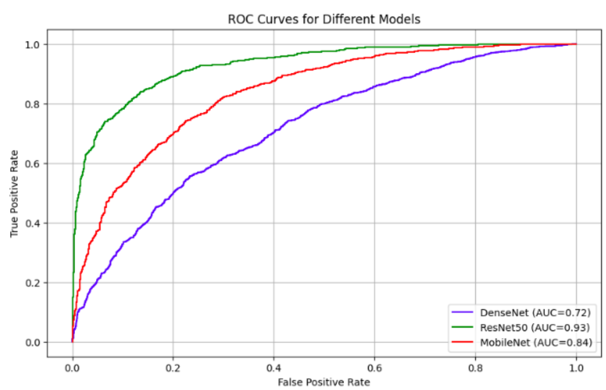


FIGURE 6. AUC RC curve of three models

MobileNet is preferred over ResNet50 in this project due to its suitability for deployment in a Next.js app that runs in a web browser, particularly when considering factors like model size, inference speed, offline usage, and resource constraints. While ResNet50 may achieve higher accuracy, the practical considerations for this specific application make MobileNet a more pragmatic choice. The rationale and methodology behind selecting MobileNet as the backbone architecture, as well as the hyperparameters and optimization techniques used for fine-tuning the model are discussed below.

A. MOBILENET AS THE BACKBONE ARCHITECTURE:

MobileNet is a popular architecture for mobile and embedded vision applications, known for its efficiency and small model size. It is designed to have a low computational cost while maintaining good accuracy, making it suitable for resource-constrained environments. MobileNet achieves this efficiency by using depthwise separable convolutions, which reduce the number of parameters and operations compared to traditional convolutions. By using MobileNet as the backbone architecture, we can leverage its pre-trained weights on the ImageNet dataset to initialize the model with good feature representations.

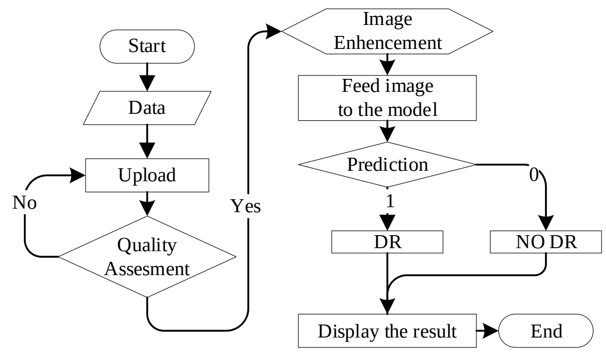


FIGURE 7. Data flow diagram for real-time detection of diabetic retinopathy detection.

B. ADDING LAYERS ON TOP OF THE BASE MODEL:

After loading the MobileNet model, we add additional layers on top of the base model to customize it for the specific task. The GlobalAveragePooling2D layer is added to convert the 4D tensor output of the base model into a 2D tensor by taking the average value over spatial dimensions. The Dense layer with 256 units and ReLU activation is added to introduce non-linearity and capture more complex patterns in the data. The final Dense layer with 2 units and sigmoid activation is added to produce the output predictions for the binary classification task.

C. FREEZING THE LAYERS IN THE BASE MODEL:

To prevent the pre-trained weights in the base model from being updated during training, we freeze the layers in the base model. This is achieved by setting the trainable attribute of each layer in the base model to False. By freezing the base model layers, we ensure that only the newly added layers on top are trained and fine-tuned for the specific task (FIGURE

7).

D. COMPILATION AND OPTIMIZATION:

The model is compiled using the Adam optimizer, which is an adaptive learning rate optimization algorithm. The loss function used is binary_crossentropy, which is suitable for binary classification tasks. The metrics used for evaluation are accuracy. Additionally, two optimization techniques are used, ReduceLROnPlateau and EarlyStopping. ReduceLROnPlateau reduces the learning rate when the validation loss plateaus, allowing the model to fine-tune more effectively. EarlyStopping stops training if the validation loss does not improve after a certain number of epochs, preventing overfitting.

E. TRAINING THE MODEL:

The model is trained using the fit method, which takes the training generator, number of steps per epoch i.e 50, number of epochs i.e 30, validation generator, and callbacks as parameters. The steps_per_epoch parameter determines the number of batches per epoch, allowing us to control the amount of data processed per epoch. The epochs parameter specifies the number of times the entire dataset is passed through the model during training. The validation data is provided through the val_generator parameter. The early_stopping and lr_reducer callbacks are used to monitor the validation loss and adjust the learning rate during training. Employing Next.js as the foundation for the real-time detection solution introduces a range of advantages that align seamlessly with the project's goals. Leveraging Next.js's robust client-side and server-side rendering capabilities ensures that the application can swiftly deliver accurate results to users without sacrificing performance. This becomes particularly significant when considering resource-limited regions such as rural parts of Nepal. By utilizing Next.js's ability to function offline on smartphones, we envision a scenario where healthcare workers can carry out on-the-spot diabetic retinopathy detection, even in areas with limited internet connectivity. This offline functionality not only enhances accessibility but also empowers medical professionals to diagnose and intervene in real time, potentially preventing irreversible vision loss. The synergy between Next.js and our mission to detect diabetic retinopathy in rural Nepal highlights the versatility of this technology and its capacity to positively impact healthcare delivery. User Uploads Images: The process begins with the user uploading retinal images through the user interface of the web application. These images can be captured using a smartphone or uploaded from a local storage device. Frontend Interaction: The front end of the application, powered by Next.js, manages user interaction. It provides the interface for users to select and submit images for analysis. Upon image submission, the Next.js client-side components handle the incoming data. The application processes the uploaded images directly within the user's web browser for analysis. The server performs preprocessing steps on the uploaded images. This involves resizing, normalization, and any necessary enhancement to ensure that the images are optimized for analysis. The preprocessed images are then fed into the deployed MobileNet model, which performs diabetic retinopathy detection. The model processes the images and produces predictions indicating the presence and severity of diabetic retinopathy. The model's inference process occurs on the client side. Users can view the prediction results directly in the front end of the application, where it is presented in a clear and user-friendly format.

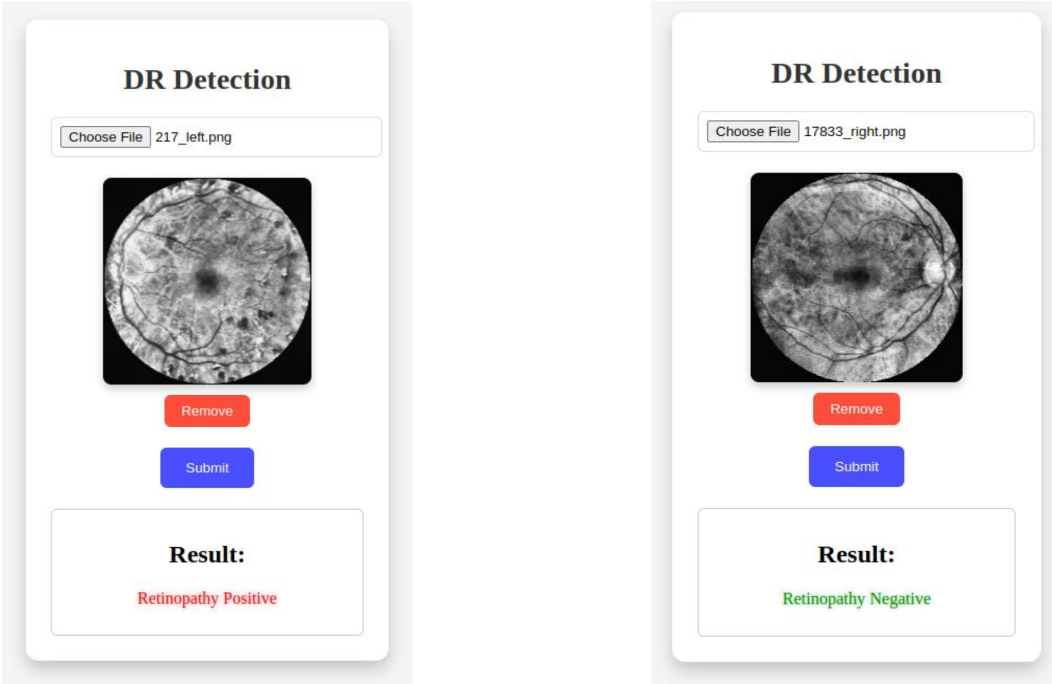
IV. RESULT

The proposed deep learning model, employing MobileNet as the backbone architecture, demonstrated commendable performance in the task of diabetic retinopathy detection (FIGURE 8). The comprehensive evaluation, using a diverse dataset of retinal fundus images, yielded the following results. The model achieved an impressive training accuracy of 92%, indicating its ability to learn intricate features from the training data. During the validation phase, the model maintained a high accuracy rate of 91%. This underscores its robustness and generalization capability. In the critical testing phase, the model continued to perform exceptionally well, attaining an accuracy of 90%. This result underscores the reliability of the model in real-world scenarios. Precision, a crucial metric in medical diagnosis, was measured at 81%. This value signifies the model's ability to correctly identify diabetic retinopathy cases while minimizing false positives, thereby enhancing its clinical utility. With a recall rate of 75%, the model exhibited a commendable capacity to identify a substantial proportion of true positive cases, making it effective in detecting diabetic retinopathy, even in challenging instances.

TABLE 2
Accuracies and scores of the model

Training Accuracy	0.92
Validation Accuracy	0.91
Testing Accuracy	0.90
Precision	0.81
Recall	0.75
F1	0.78

The F1 score, which harmonizes precision and recall, was calculated at 78%. This metric is a testament to the model's overall performance, striking a balance between accurate positive predictions and comprehensive disease detection (TABLE 2). True Positives (TP): The model correctly predicted 932 cases of diabetic retinopathy (DR). This indicates that the model is effective at identifying cases of DR. True Negatives (TN): The model correctly predicted 1345 cases of no diabetic retinopathy (NO_DR). This shows that the model is good at identifying cases without DR. False Positives (FP): The model predicted 219 cases as having DR when they didn't. These are false alarms or Type I errors. It



(a) (b)

FIGURE 8 (a) Result when retinopathy negative, (b) Result when retinopathy positive

means that the model tends to overestimate the presence of DR, which can lead to unnecessary concern for patients. False Negatives (FN): The model predicted 310 cases as not having DR when they did. These are missed cases or Type II errors. This suggests that the model is missing some cases of DR, which could be a significant concern as early detection is crucial for timely treatment. Misclassifications: The false positives and false negatives represent cases where the model made incorrect predictions. These misclassifications can have real-world consequences, as patients may receive incorrect diagnoses or recommendations for further testing and treatment. The model does not distinguish between different stages of diabetic retinopathy (FIGURE 9). It treats all cases of DR as a single category. In reality, diabetic retinopathy has different stages (mild, moderate, severe, proliferative), and accurately identifying the stage is important for determining appropriate treatment plans. This limitation suggests that the model lacks granularity in its predictions. While the model appears to perform reasonably well in terms of overall accuracy, it has limitations in terms of false positives and false negatives.

Additionally, the model cannot distinguish between different stages of diabetic retinopathy, which is important for tailored treatment plans. Future improvements might involve fine-tuning the model to reduce false positives and false negatives, as well as incorporating a multi-class classification approach to distinguish between different stages of diabetic retinopathy. These results showcase the potential of our MobileNet-based deep learning model in automating diabetic retinopathy detection. The high accuracy rates achieved during training, validation, and testing, along with robust precision, recall, and F1 score values, affirm the model's effectiveness. The combination of these metrics positions our model as a valuable tool for early and accurate diabetic retinopathy diagnosis, with the potential to improve patient outcomes and reduce the burden on healthcare providers. resilience and adaptability of the academic community in continuing to contribute to global knowledge despite varying circumstances.

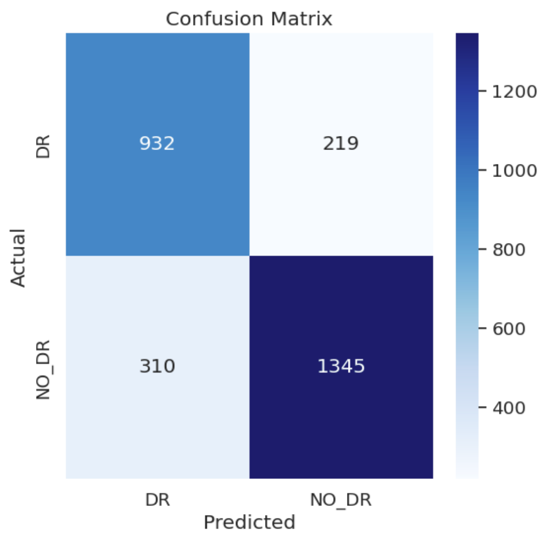


FIGURE 9. Confusion Matrix of the result

V. DISCUSSION

The findings of this research are promising, indicating that MobileNet-based deep learning can effectively classify the severity of diabetic retinopathy from retinal images. Our model achieved a high degree of accuracy and demonstrated the ability to distinguish between different levels of DR severity, including mild, moderate, severe, and proliferative stages. This suggests that the technology can serve as a valuable tool for healthcare workers in rural Nepal to conduct preliminary screenings and identify cases that require immediate attention. The results of the study demonstrate the potential of employing a MobileNet-based deep learning model for the automated detection of diabetic retinopathy. The high training, validation, and testing accuracies, along with strong precision, recall, and F1 scores, indicate the model's ability to accurately identify cases of diabetic retinopathy. These findings align with previous research that has showcased the efficacy of CNN models in medical image analysis. The robust performance of the model in terms of accuracy and precision is particularly encouraging for clinical applications. High precision (0.81) implies that when the model predicts diabetic retinopathy, it is often correct. This is crucial in a medical context, where minimizing false positives is essential to avoid unnecessary patient anxiety and follow-up examinations. However, it's equally important to recognize that the model's recall (0.75) indicates that it may miss some true positive cases. This trade-off between precision and recall is common in the context of imbalanced medical datasets, where the presence of negative samples often dominates the dataset. A slight imbalance in our dataset may have influenced this trade-off.

Despite the positive outcomes, several limitations and weaknesses must be acknowledged. Firstly, the study's reliance on the EyePACS dataset, while extensive, may not fully represent the diversity of retinal images encountered in the field. Images from rural Nepal may exhibit unique characteristics due to factors like varying lighting conditions and image quality. Secondly, the model's performance might be affected by the presence of comorbidities or other ocular conditions that can coexist with diabetic retinopathy. Additionally, the data preprocessing steps, while necessary to enhance image quality, could introduce unintended biases. The implications of this research are multifaceted and extend beyond diabetic retinopathy detection. Firstly, the findings suggest that deep learning, particularly MobileNet, can be an effective and accessible tool for healthcare workers in rural areas. This could potentially reduce the prevalence of avoidable blindness by enabling early intervention. Secondly, the study highlights the importance of leveraging mobile technology in healthcare settings with limited internet connectivity, as demonstrated by the model's offline capabilities. Thirdly, this work serves as a blueprint for addressing healthcare challenges in resource-limited regions through the fusion of AI and mobile technology. The deployment of the MobileNet-driven deep learning model in rural healthcare settings in Nepal holds significant promise. The impact of this technology extends to various applications:

1. Rural Healthcare Settings: This model can significantly improve healthcare outcomes by providing timely diabetic retinopathy detection in rural areas with limited access to specialized medical professionals. It empowers healthcare providers to intervene early, preventing the disease's progression and potentially saving patients' vision.
2. Telemedicine: Integrating the deep learning model into telemedicine systems enhances the reach of healthcare services. Patients in remote areas can benefit from real-time, accurate diagnosis and consultations, reducing the need for travel to urban centers. This not only saves time and resources but also ensures that patients receive timely care.
3. User Empowerment: The system empowers individuals, especially those in rural regions of Nepal, by offering early detection and awareness of their diabetic retinopathy status. Empowered patients are better equipped to manage their health and seek appropriate treatment, leading to improved overall well-being.
4. Ethical, Social, and Legal Issues:

As with any advanced healthcare technology, the deployment of AI-driven diabetic retinopathy detection in rural Nepal raises ethical, social, and legal considerations:

1. Ethical Considerations: Patient consent, privacy, and data security are paramount. It is essential to ensure that patients provide informed consent for the use of their medical data and that their privacy rights are protected. Additionally, addressing potential biases in the deep learning model is critical to ensure equitable healthcare access.
2. Social Impact: The technology has the potential to reduce healthcare disparities and enhance public health outcomes in rural Nepal. However, awareness and education initiatives are necessary to ensure that healthcare providers and patients alike can maximize its benefits.
3. Legal Issues: Compliance with data protection laws and the establishment of clear regulations and guidelines for AI in healthcare are crucial. Additionally, addressing liability issues in cases of misdiagnosis or system failures is essential to mitigate legal risks associated with the deployment of this technology.

Code Repository: The Python and javascript codes for the experiments in this study can be found at -

https://github.com/sauravrav/dr-detection

REFERENCES

[1] S. Baral, S. Uprety, and B. Lamichanne, *Diabetes: A Backgrounder*. 2016.
[2] D. B. Shrestha *et al.*, “Type 2 Diabetes Mellitus in Nepal from 2000 to 2020: A systematic review and meta-analysis.,” *F1000Research*, vol. 10, p. 543, 2021, doi: 10.12688/f1000research.53970.2.
[3] G. Paudyal, M. K. Shrestha, M. Poudel, G. C. Tabin, S. Ruit, and B. J. Thomas, “Prevalence and Severity of Diabetic Retinopathy among Diabetic Patients Presenting to a Tertiary Eye Hospital in Nepal.,” *Middle East Afr. J. Ophthalmol.*, vol. 26, no. 4, pp. 210–215, Dec. 2019, doi: 10.4103/meajo.MEAJO_65_18.
[4] M. Kropp *et al.*, “Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications-risks and

mitigation.," *EPMA J.*, vol. 14, no. 1, pp. 21–42, Mar. 2023, doi: 10.1007/s13167-023-00314-8.
[5] A. R. Shah and T. W. Gardner, "Diabetic retinopathy: research to clinical practice.," *Clin. Diabetes Endocrinol.*, vol. 3, p. 9, 2017, doi: 10.1186/s40842-017-0047-y.
[6] P. Shrestha, R. Kaiti, and R. Shyangbo, "Blindness among Patients with Type II Diabetes Mellitus Presenting to the Outpatient Department of Ophthalmology of a Tertiary Care Centre: A Descriptive Cross-sectional Study.," *JNMA J. Nepal Med. Assoc.*, vol. 60, no. 254, pp. 877–880, Oct. 2022, doi: 10.31729/jnma.7702.
[7] D. S. Fong *et al.*, "Diabetic Retinopathy," *Diabetes Care*, vol. 26, no. suppl_1, pp. s99–s102, Jan. 2003, doi: 10.2337/diacare.26.2007.S99.
[8] R. Reguant, S. Brunak, and S. Saha, "Understanding inherent image features in CNN-based assessment of diabetic retinopathy," *Sci. Rep.*, vol. 11, no. 1, p. 9704, May 2021, doi: 10.1038/s41598-021-89225-0.
[9] D. Das, S. K. Biswas, and S. Bandyopadhyay, "Detection of Diabetic Retinopathy using Convolutional Neural Networks for Feature Extraction and Classification (DRFEC).," *Multimed. Tools Appl.*, pp. 1–59, Nov. 2022, doi: 10.1007/s11042-022-14165-4.
[10] A. B. Wiratama, Y. Fu'adah, S. Saidah, R. Magdalena, I. D. S. Ubaidah, and R. B. J. Simanjuntak, "Diabetic Retinopathy Classification Based on Fundus Image Using Convolutional Neural Network (CNN) with MobilenetV2," in *Proceeding of the 3rd International Conference on Electronics, Biomedical Engineering, and Health Informatics*, T. Triwiyanto, A. Rizal, and W. Caesarendra, Eds., Singapore: Springer Nature Singapore, 2023, pp. 89–102.
[11] O. Daanouni, B. Cherradi, and A. Tmiri, "Automatic Detection of Diabetic Retinopathy Using Custom CNN and Grad-CAM," in *Advances on Smart and Soft Computing*, F. Saeed, T. Al-Hadhrami, F. Mohammed, and E. Mohammed, Eds., Singapore: Springer Singapore, 2021, pp. 15–26.
[12] R. Vij and S. Arora, "A Systematic Review on Diabetic Retinopathy Detection Using Deep Learning Techniques," *Arch. Comput. Methods Eng.*, vol. 30, no. 3, pp. 2211–2256, Apr. 2023, doi: 10.1007/s11831-022-09862-0.
[13] N. Salamat, M. M. S. Missen, and A. Rashid, "Diabetic retinopathy techniques in retinal images: A review.," *Artif. Intell. Med.*, vol. 97, pp. 168–188, Jun. 2019, doi: 10.1016/j.artmed.2018.10.009.
[14] A. Skouta, A. Elmoufidi, S. Jai-Andaloussi, and O. Ouchetto, "Deep learning for diabetic retinopathy assessments: a literature review," *Multimed. Tools Appl.*, Apr. 2023, doi: 10.1007/s11042-023-15110-9.
[15] W. Zhu, P. Qiu, N. Lepore, O. M. Dumitrascu, and Y. Wang, "NNMobile-Net: Rethinking CNN Design for Deep Learning-Based Retinopathy Research," 2023, doi: 10.48550/ARXIV.2306.01289.
[16] S. Wan, Y. Liang, and Y. Zhang, "Deep convolutional neural networks for diabetic retinopathy detection by image classification," *Comput. Electr. Eng.*, vol. 72, pp. 274–282, Nov. 2018, doi: 10.1016/j.compeleceng.2018.07.042.
[17] R. Ghosh, K. Ghosh, and S. Maitra, "Automatic detection and classification of diabetic retinopathy stages using CNN," in *2017 4th International Conference on Signal Processing and Integrated Networks (SPIN)*, Feb. 2017, pp. 550–554. doi: 10.1109/SPIN.2017.8050011.
[18] L. Dai *et al.*, "A deep learning system for detecting diabetic retinopathy across the disease spectrum," *Nat. Commun.*, vol. 12, no. 1, p. 3242, May 2021, doi: 10.1038/s41467-021-23458-5.
[19] P.-N. Chen, C.-C. Lee, C.-M. Liang, S.-I. Pao, K.-H. Huang, and K.-F. Lin, "General deep learning model for detecting diabetic retinopathy," *BMC Bioinformatics*, vol. 22, no. 5, p. 84, Nov. 2021, doi: 10.1186/s12859-021-04005-x.
[20] M. Esengönül, A. C. de Paiva, J. Rodrigues, and A. Cunha, "Diabetic Retinopathy Detection Using Convolutional Neural Networks for Mobile Use," in *Wireless Mobile Communication and Healthcare*, A. Cunha, N. M. Garcia, J. Marx Gómez, and S. Pereira, Eds., Cham: Springer Nature Switzerland, 2023, pp. 14–23.
[21] Y. Yang *et al.*, "A comparative analysis of eleven neural networks architectures for small datasets of lung images of COVID-19 patients toward improved clinical decisions.," *Comput. Biol. Med.*, vol. 139, p. 104887, Dec. 2021, doi: 10.1016/j.compbiomed.2021.104887.
[22] K. K. Bressem, L. C. Adams, C. Erxleben, B. Hamm, S. M. Niehues, and J. L. Vahldiek, "Comparing different deep learning architectures for classification of chest radiographs," *Sci. Rep.*, vol. 10, no. 1, p. 13590, Aug. 2020, doi: 10.1038/s41598-020-70479-z.
[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
[24] A. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," Apr. 2017.
[25] N. Jagan Mohan, R. Murugan, and T. Goel, "Deep Learning for Diabetic Retinopathy Detection: Challenges and Opportunities," in *Next Generation Healthcare Informatics*, B. K. Tripathy, P. Lingras, A. K. Kar, and C. L. Chowdhary, Eds., Singapore: Springer Nature Singapore, 2022, pp. 213–232. doi: 10.1007/978-981-19-2416-3_12.
[26] S. Agarwal and A. Bhat, "A survey on recent developments in diabetic retinopathy detection through integration of deep learning," *Multimed. Tools Appl.*, vol. 82, no. 11, pp. 17321–17351, May 2023, doi: 10.1007/s11042-022-13837-5.
[27] M. W. Nadeem, H. G. Goh, M. Hussain, S.-Y. Liew, I. Andonovic, and M. A. Khan, "Deep Learning for Diabetic Retinopathy Analysis: A Review, Research Challenges, and Future Directions.," *Sensors*, vol. 22, no. 18, Sep. 2022, doi: 10.3390/s22186780.
[28] 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, doi: 10.1177/193229680900300315.
[29] C.-L. Lin and K.-C. Wu, "Development of revised ResNet-50 for diabetic retinopathy detection.," *BMC Bioinformatics*, vol. 24, no. 1, p. 157, Apr. 2023, doi: 10.1186/s12859-023-05293-1.
[30] H. Fu *et al.*, "Evaluation of Retinal Image Quality Assessment Networks in Different Color-Spaces," 2019, pp. 48–56. doi: 10.1007/978-3-030-32239-7_6.

Author Biography

Saurav Bhatta was born in Nepal in 1999. He is a computer science graduate who completed his undergraduate degree in 2021 from IKG Punjab Technical University, Punjab. He is currently working as a software developer and aims to contribute towards research and development space in adoption of AI in various sectors. His research interests include Computational Statistics, Deep Learning, and Computer Vision.