CC BY

Strategies for Detecting Diabetic Retinopathy with Deep Learning and Image Processing

Dr. Jitendra Sheetlani[1]*, Sapana Desai[2]

[1]*Professor of Computer Application, Medicaps University, Indore
[2]Research Scholar, SSSUTMS, Sehore

***Corresponding Author:** Dr. Jitendra Sheetlani
*Professor of Computer Application, Medicaps University, Indore

Abstract:
Diabetic retinopathy is one of the most severe complications of diabetes, potentially leading to complete blindness if left untreated. Early detection is essential for effective treatment, but it presents significant challenges. Diagnosing the stage of diabetic retinopathy is particularly difficult and requires skilled interpretation of fundus images. Simplifying this detection process could greatly benefit millions. Diabetic retinopathy primarily affect working individuals with diabetes. It often involves extensive time spent analyzing fundus images after a patient's visit to the ophthalmologist. Our project aims to streamline this process, allowing doctors to care for more patients by speeding up result analysis and minimizing the risk of misdiagnosis, thus supporting ophthalmologists in their work. Diabetic retinopathy predominantly affects working-age individuals with diabetes. Diagnosing this condition typically requires significant time spent processing fundus images after each patient visit to the ophthalmologist. Our project aims to streamline this process, enabling doctors to see more patients due to faster result processing. Additionally, it seeks to assist ophthalmologists in preventing misdiagnoses.

Keywords: Deep learning, diabetic , retinopathy,Deep convolutional neural , network, multi- Target learning, ordinal , Regression.

1 INTRODUCTION

Diabetic retinopathy progresses through four stages: Mild, Moderate, Severe, and Proliferative. Neglecting regular specialist visits or proper care can lead to visual impairment, potentially causing lifelong blindness. The three main components of the human retina—vessels, optic disc, and fovea—are crucial for applications such as retinal image selection, illumination correction, and pathology recognition within the retina. Identifying these essential structures manually is time-consuming and relies heavily on the user's expertise.

Many diabetic patients are surprised to discover they have been living with diabetic retinopathy, a degenerative eye disease, for years without realizing it. Often, patients do not seek specialist care until they start experiencing vision loss. Symptoms and side effects of diabetic retinopathy, such as blurred vision, retinal clouding, eye itching and swelling, and occasional eye pain, typically appear once the disease has progressed. If you notice any of these symptoms, it is crucial to seek treatment from an experienced ophthalmologist promptly. Additionally, the side effects of diabetic retinopathy can change as the disease advances. Even in the absence of visual problems, changes in the retina observed during an eye exam can indicate diabetic retinopathy. These changes include retinal swelling, leaking blood vessels, scar tissue, or abnormal deposits on the retina. Therefore, it is important for patients to monitor their condition to prevent the disease from reaching more severe stages, which can have serious consequences.

The uneven lighting is associated with the acquisition technique, and the excessively high level of complexity arises from the procurement procedure and the process by which various vessels are distinguished from the background. Consequently, we aim to develop an automated vein identification tool that can accurately delineate the retinal veins quickly. Various methods and strategies have been explored using fragment retinal images. The algorithms in this field are categorized into three types: window-based, classifier-based, and tracking-based techniques.

Window-based approaches, such as edge detection, search for a match between a predefined model and the surrounding pixel window at each pixel. The passage of a vessel in a retinal image was validated by a Gaussian-shaped curve and then identified using rotated matched filters. The vertebrate retina is a light-sensitive tissue lining the inner surface of the eye. The eye's optics create an image of the visible world on the retina, functioning similarly to film in a camera. Light hitting the retina initiates a series of chemical and electrical events that eventually lead to nerve impulses. These impulses are sent to various visual centers of the brain through the fibers of the optic nerve.

When bright light is shone on the eye, it contracts as a response, known as the pupillary light reflex, an important test of brainstem function. Additionally, when a person observes something interesting, their pupil dilates. The circular iris sphincter muscle is where the oculomotor nerve, specifically the parasympathetic component originating from the Edinger-Westphal nucleus, terminates. When this muscle contracts, it reduces the pupil's range of motion.

Diabetic retinopathy is a degenerative eye disorder and one of the most serious complications of diabetes. While there are several symptoms and side effects of diabetic retinopathy that patients should be aware of, a diagnosis must be made by an experienced ophthalmologist. Although all diabetic individuals are at risk of developing diabetic retinopathy, not all will experience vision loss. There are a few key signs and symptoms to look for regarding diabetic retinopathy.

1. RELATED WORK

With advancements in the field, numerous studies and research have been conducted on the automated diagnosis of diabetic retinopathy using various advanced and increasingly precise techniques. The following papers have been thoroughly studied and analyzed to develop our project. These papers, published in esteemed journals, were instrumental in enhancing our understanding of diabetic retinopathy.

Many studies have focused on the early detection of diabetic retinopathy using various techniques. Initially, researchers used classical computer vision and machine learning methods. For instance, Priya et al. (2012) developed a method using color fundus photographs and SVM, achieving 98% sensitivity, 96% specificity, and 97.6% accuracy on 250 images. Others, like Conde et al. (2012), applied PCA and models like decision trees and k-NN, achieving 73.4% accuracy with 151 images.The Extreme Learning Machine (ELM) model shows superior performance compared to other models in detecting exudates in retinal images, as evidenced by trial results.

Furthermore, many researchers have utilized data mining techniques, particularly automated screening of these images, to assist specialists in accurately identifying patient conditions. This underscores the importance of using appropriate image processing and data mining methods to ensure the accurate classification of normal and abnormal retina images, thereby reducing the workload for ophthalmologists.

2. PROPOSED SYSTEM

The proposed design utilizes various image processing techniques and machine learning methods to effectively segment fundus input images, detailed in the upcoming subsections. Our system offers several advantages: it features an intuitive Graphical User Interface (GUI) accessible to anyone, ensures faster processing times compared to similar systems, and is cost-effective, costing less than manual strategies currently in use. Its user-friendly design improves usability over previous predictive models. Moreover, it is more reliable due to training on a robust dataset directly sourced from ophthalmologists (DIATREB).

The image data used in this research was taken from several datasets. We used an open dataset from Kag-gle Diabetic Retinopathy. Furthermore, we utilized two additional, smaller datasets: the Indian Diabetic Retinopathy Image Dataset (IDRiD) (Sahasrabuddhe and Meriaudeau, 2018), which provided 413 fundus photographs, and the MESSIDOR (Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology) dataset (Decencire et al., 2014). We specifically chose the version of the MESSIDOR dataset that had been regraded to standard grading by a panel of ophthalmologists, as the original MESSIDOR dataset's grading differed from that of other datasets (Google Brain, 2018). The evaluation was conducted on the APTOS2019 Blindness Detection dataset (APTOS, 2019), as we only had access to the training portion.

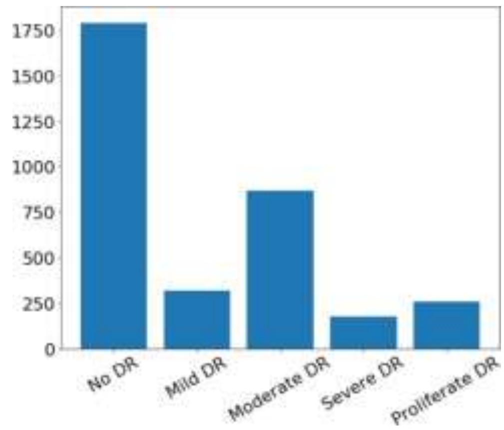

Figure 1: Classes distribution in APTOS.

The primary input of the proposed system will be a fundus image obtained following respective screening tests. The fundus refers to the inner part of the retina.The user needs to open the interface and subsequently click on the "upload image" button to upload the image.

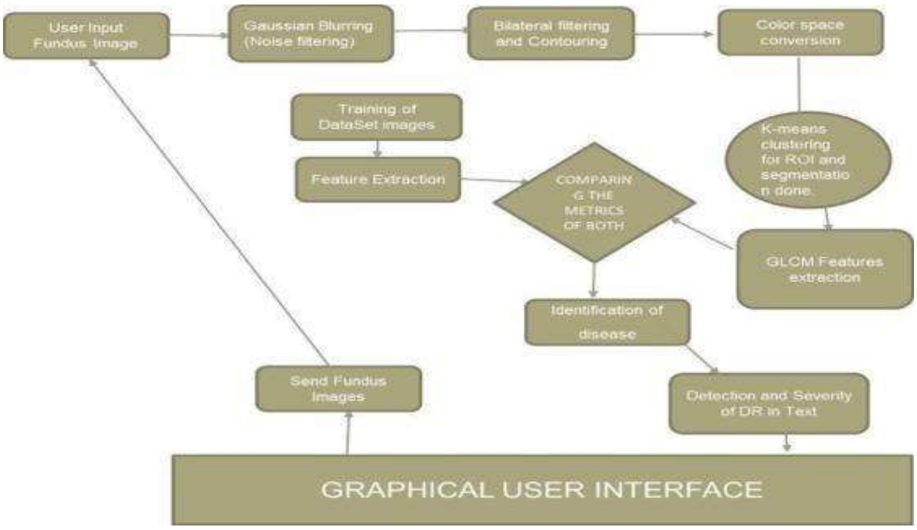

Fig. 2. Architecture Diagram

Gaussian filtering is a technique used to blur input images, effectively reducing noise presence. Unlike a uniform low-pass filter, the Gaussian filter applies a two-dimensional Gaussian function specifically tailored for image processing. A sample Gaussian filter is illustrated below.

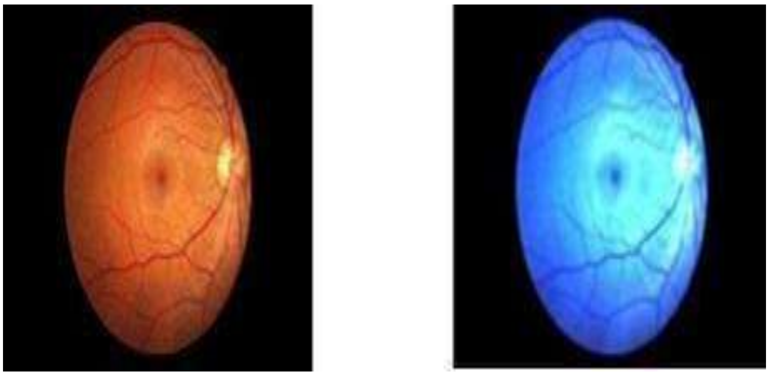

Fig. 3. Gaussian Filtering Implementation

Bilateral filtering is a smoothing filter used in image processing to preserve edges while reducing noise. It replaces each pixel's intensity with a weighted average of intensities from neighboring pixels. These weights are often based on a Gaussian distribution and consider both the Euclidean distance between pixels and their radiometric differences (such as color intensity or depth). This approach effectively maintains sharp edges in the image.

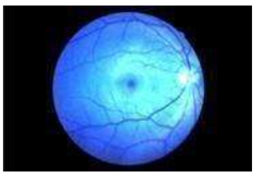

Fig. 4. Bilateral Filtered Image

Converting a color image into a grayscale image while preserving its distinct characteristics is a complex process. Traditional methods of conversion can often lead to loss of contrast, sharpness, shadow detail, and overall structure from the original color image. To address these issues, a new algorithm has been proposed. This algorithm applies RGB assumption, reduces and expands chrominance and luminance values to convert the color image into grayscale. Experimental grayscale images generated using this algorithm demonstrate that it effectively preserves the unique attributes of the original color image, such as sharpness, contrast, image structure, and shadow details.

Fig. 4. A Gray Scale converted image

After converting the image to grayscale, it undergoes contouring or edge detection, commonly known as edge detection. This process identifies and removes unnecessary parts of the image that are not required for training. Following this, the image undergoes blood vessel segmentation to differentiate and identify hemorrhages or swollen blood vessels.

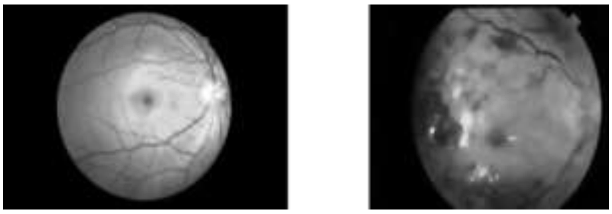

Fig. 5.Segmentation implementation

The image below illustrates the basic principles of K-means clustering.

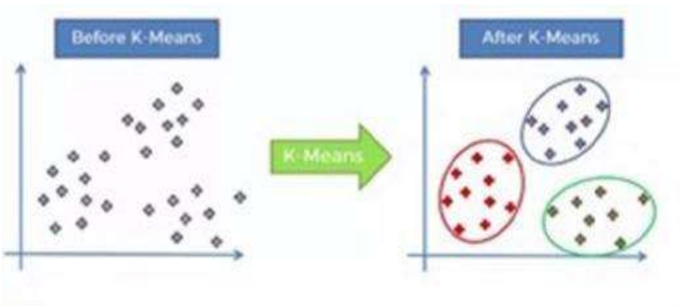

Fig.6. K-means Clustering Principle

4. Methodology

The problem of diabetic retinopathy detection can be approached from multiple perspectives: as a classification problem, a regression problem, and an ordinal regression problem (Ananth and Kleinbaum, 1997). This versatility arises because the stages of the disease occur sequentially.

The methodology simplifies existing approaches. The user interface allows users to upload a fundus image, which is then processed by the model. Initially, Gaussian blurring is applied to remove image noise. Subsequently, bilateral filtering and contouring are performed. The image undergoes color space conversion to create a grayscale fundus image. Finally, k-means clustering is used to identify regions of interest in the image.

After segmentation, the resulting Region of Interest (ROI) undergoes GLCM feature extraction to form a classification feature vector. This vector includes seven features derived from segmenting retinal structures and analyzing entropy, contrast, and homogeneity. The training process is illustrated in the image below.

This method calculates the grayscale intensity of each pixel within a structure to obtain measurements. It utilizes images from an available dataset for this process. Segmentation is applied to these images, and the measurements from both the predefined dataset and user input are compared to validate the input image. Recommendations are then provided based on this validation assessment.

We train all heads and the feature extractor jointly in order to reduce training time. We keep the linear regression model frozen until the post-training stage.

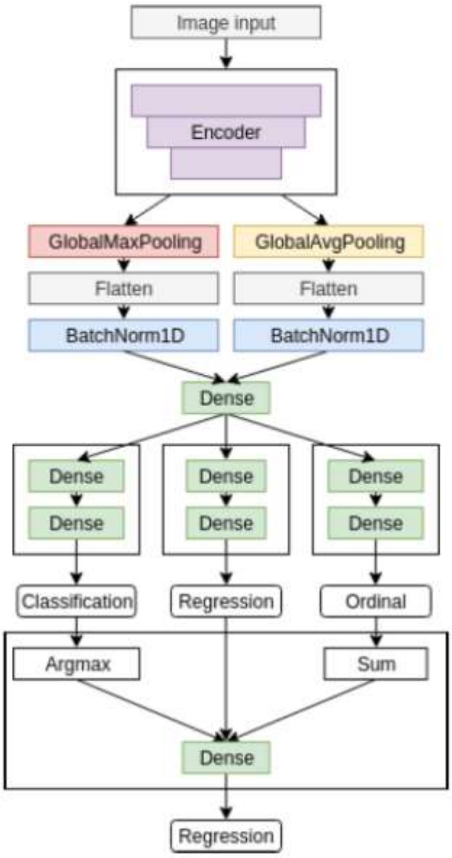

Figure 6: Three-head CNN structure

We employ a multi-stage training process, varying settings and datasets at each stage. To address diverse labeling strategies across datasets, we opted to pretrained our CNNs using the largest dataset from 2015. Transfer learning is suitable because the fundamental characteristics of diabetic retinopathy are consistent across different datasets and individuals.

Additionally, datasets often employ various imaging technologies. By integrating this knowledge into the model, we reduce sensitivity to device-specific noise, thereby improving its ability to generalize and highlight intrinsic features.

The feature extractor starts with pretrained weights from ImageNet, while the heads begin with random weights (He et al., 2015). Training involves 20 epochs on the 2015 dataset using minibatch stochastic gradient descent (SGD) and a cosine-annealing learning rate schedule (Loshchilov and Hutter, 2016).

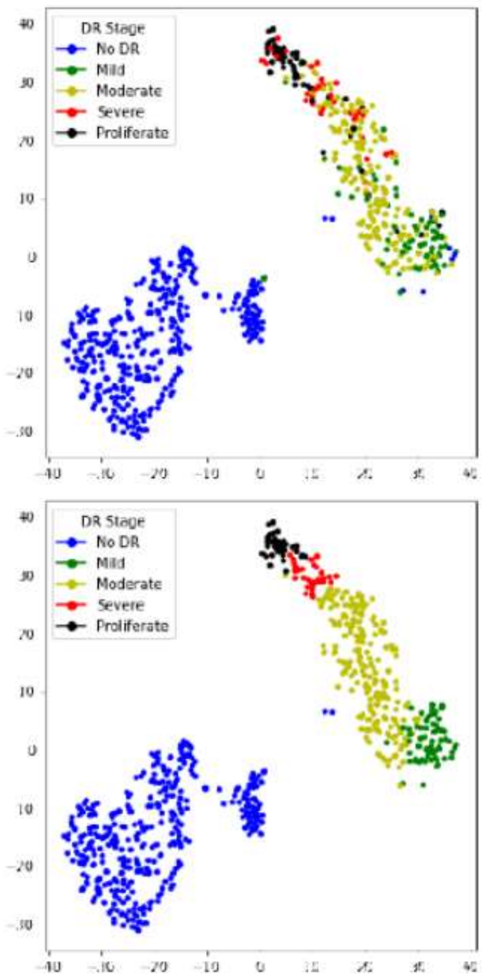

Figure 7:Feature embedding with T-SNE.

Initial weights for every head were set to 1⁄3 and then trained for five epochs to minimize mean squared error function. We used Catalyst framework (Kolesnikov, 2018)based on PyTorch (Paszke et al., 2017) with GPU support.

The final stage involves displaying the results on the graphical user interface after completing all processing steps. The interface provides textual feedback indicating whether the person has diabetic retinopathy. If the condition is detected, the system also suggests symptoms and indicates the severity of diabetic retinopathy on the GUI. This allows the patient to understand their current health status regarding the ailment.

5. RESULTS

The project outcomes include the application of image processing techniques to the input fundus image, which can optionally be disabled. In the graphical user interface displayed in Figure below, the patient's input image is shown. Figures [9,10] illustrate the Haralick metrics generated by the Gray Level Co-occurrence Matrix for each image in the training dataset, which are later compared with those of the input image. The final figure shows the result, indicating that the patient is diagnosed with proliferative diabetic retinopathy. It also provides next steps for the patient to follow.

Our testing phase was divided into two segments: local testing and Kaggle testing. Locally, we determined that the ensembling method was the most effective, and we validated this on Kaggle's validation and test datasets.

On a local dataset comprising 736 images, resembling with TTA showed slightly lower performance compared to ensemble without it. However, on the testing dataset of 13,000 images, ensemble with TTA demonstrated improved generalization ability on unseen images.

6. INTERPRETATION

In medical applications, accurately interpreting models' predictions is crucial. While good performance on validation datasets helps select the best-trained model for production, it may not suffice for real-world applications.
SHAP (Shapley Additive exPlanations) (Lundberg and Lee, 2017) enables the visualization of features contributing to disease stage assessment. It integrates multiple prior methods and offers a consistent and locally accurate method for attributing feature importance.

Using SHAP ensures that the model effectively learns relevant features during training and correctly utilizes them during inference. Moreover, in uncertain cases, visualizing salient features can aid physicians in focusing on regions where features are most pronounced.

When designing a graphical user interface for diabetic retinopathy detection, the primary goal is to allow users/patients to upload fundus images for reliable diagnosis with faster processing times. Future enhancements may involve integrating advanced algorithms such as convolutional neural networks to improve image classification accuracy. Additionally, features like a helpline or user manual should be included in the interface to support users who may not be familiar with technological advancements and application usage.

7. CONCLUSION

In this study, we propose a multistage transfer learning approach and an automated method for detecting diabetic retinopathy stages from single fundus photographs. Our approach utilizes an ensemble of three CNN architectures (EfficientNet-B4, EfficientNet-B5, SE-ResNeXt50) with transfer learning as the final solution. Experimental results demonstrate that our method achieves consistently high performance, even with variability in metrics. The key advantage lies in enhancing generalization and reducing variance through ensemble learning, leveraging pretrained models on extensive datasets and fine-tuning on specific datasets. Future research could extend this method by applying SHAP across the entire ensemble rather than individual networks and improving hyper parameter optimization. Additionally, exploring pretrained encoders for related ocular conditions could further enhance the approach. Investigating meta-learning with these models is also possible, but it requires separate and thorough research.

REFERENCES

1. A. Buslaev, A. Parinov, E. K. V. I. I. and Kalinin, A. A.(2018). Albumentations: fast and flexible image aug-mentations.*ArXive-prints*.
2. APTOS (2019).APTOS 2019 blindness detection.Ac-cessed:2019-10-20.
3. Ananth, C. V. and Kleinbaum, D. G. (1997).Regressionmodels for ordinal responses: a review of methods andapplications. *International Journal of Epidemiology*,26(6):1323–1333.
4. Madhukeerthana, Y. Bevish Jinila, "A review on rough set theory in medical images", Research Journal of Pharmaceutical, Biological and Chemical Sciences,Vol.7, Issue 1, pp.815-822, ISSN : 0975- 8585, 2016 [11] Gladence, L. Mary, and T. Ravi. "Heart Disease Prediction and Treatment Suggestion." RESEARCH JOURNAL OF PHARMACEUTICAL BIOLOGICAL AND CHEMICAL SCIENCES 7, no. 2 (2016): 1274- 1279.
5. Hagos, M. T. and Kant, S. (2019). Transfer learning baseddetection of diabetic retinopathy from small dataset.*CoRR*,abs/1905.07203.
6. Harry Pratt, FransCoenen, D. M. B. S. P. H. Y. Z. (2016).Convolutional neural networks for diabetic retinopa-thy.
7. A.M.Mendonca and A.Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction,"IEEE Trans. Med. Imag., vol. 25, no. 9, pp. 1200–1213, Sep. 2009. Conde, P., dela Calleja, J., Medina, M., and Benitez Ruiz, A. B. (2012). Application of machine learning to clas-sifydiabeticretinopathy.
8. Decencire, E., Zhang, X., Cazuguel, G., Lay,B.,Cochener,B.,Trone,C.,Gain,P.,Ordonez,R.,Massin.
9. P., Erginay, A., Charton, B., and Klein, J.C.(2014).Feedbackonapubliclydistributeddatabase: themessidor database.*Image Analysis & Stereology*,33(3):231–234.
10. T.H. Kauppi and T.Kälviäinen, "Simple and robust optic disc localization using colour de correlated templates," in Proc. 10th Int. Conf. Advanced Concepts for Intell. Vision Syst.,. Berlin, Germany: Springer Verlag, vol 7, pp. 719–729, Apr 2008.
11. EyePACs (2015). Diabetic retinopathy detection.Accessed:2019-10-20