Diabetic Retinopathy using Deep Learning

Syed Eebad Reza, Namrata Kumari Singh, Anurag Singh, Tanmita Roy,

Dr. Mitrabinda Khuntia

Department of Computer Science and Engineering, Siksha 'O' Anusandhan (Deemed to be) University, Bhubaneswar, Odisha, India

2041018153.syedeebadreza@gmail.com
2041019262.namratakumarisingh@gmail.com
2041019261.anuragsingh@gmail.com
2041018143.tanmitaroy@gmail.com

Abstract.

Globally, diabetic retinopathy (DR) is the primary cause of blindness and visual impairment. Effective treatment and averting serious consequences depend on early discovery. Our proposal in this study is to use retinal pictures to automatically detect diabetic retinopathy using an ensemble deep learning model. DenseNet121, ResNet101v2, MobileNet, Xception, and InceptionV3 are the five cutting-edge convolutional neural network architectures that our model takes advantage of. High-level features are extracted from the input images by each model separately, and these features are then concatenated to create a full feature vector. The predictions from these models are combined using a weighted average ensemble learning approach, where weights are learned depending on model performance. A completely linked layer that incorporates the accumulated characteristics produces the final forecast. With an accuracy of 99.39%, our method proved to be successful in correctly diagnosing diabetic retinopathy. Furthermore, to improve the results' credibility, interpretability tools like saliency maps are employed to offer visual explanations of the model's predictions. This ensemble model offers a robust and reliable solution for the early detection of diabetic retinopathy using medical imaging, potentially aiding in timely medical intervention and improved patient outcomes.

Keywords: Diabetic Retinopathy, Deep Learning, Ensemble Learning, Convolutional Neural Networks, Medical Imaging.

1. Introduction

1.1 Motivation(s)

Diabetic retinopathy (DR) is a leading cause of blindness in the working-age population worldwide. As the prevalence of diabetes continues to rise, so does the incidence of DR, necessitating timely and accurate diagnosis to prevent vision loss. Traditional diagnostic methods for DR involve manual examination of retinal images by trained ophthalmologists, a process that is both time-consuming and subject to human error.

The motivation behind this research is to develop a robust deep learning model capable of detecting diabetic retinopathy from retinal images with high accuracy. By integrating this model into a user-friendly mobile application, we aim to provide a powerful tool for early diagnosis and monitoring of DR, particularly in underserved regions where access to specialized healthcare professionals is limited. This app allows users to upload their eye images and receive immediate feedback on their DR status, empowering patients to seek timely medical intervention and potentially preventing severe complications.

1.2 Objectives(s)

The main objective of this study is to develop and implement a deep learning-based system for the diagnosis of diabetic retinopathy (DR) from retinal images This objective can be divided into the following specific objectives :

Positive Progress: To develop and train a convolutional neural network (CNN) model that is able to accurately classify retinal images as having diabetic retinopathy (DR) or no diabetic retinopathy (No DR). Optimize the settings and parameters of the model to achieve high sensitivity and specificity in DR detection.

Data Collection and Priorities: Collect accurate data on labeled retinal images, and ensure that DR and No DR cases represent diversity. Pre-process images to improve quality and ensure accuracy for effective training and evaluation of the model.

Model validation and testing: Use standard metrics such as accuracy, precision, recall, F1 score, and AUC-ROC to evaluate model performance. Perform rigorous testing to ensure model's robustness and generality across different subsets of the dataset.

Application Development: Create a user-friendly mobile application that allows users to upload retina images. Incorporate a trained deep learning model into the application to provide real-time predictions of the presence of diabetic retinopathy.

User interface and experience: Create an intuitive interface that guides users through the process of capturing and uploading retinal images. If DR is identified, provide clear and actionable information based on the model's predictions, including recommendations to seek professional medical advice.

Handling and Transportation: Make sure that the app is accessible to as many people as possible, especially if you target less-known areas.

1.3 Original Contributions

The original contributions given by the work are -

High accuracy ensemble model: Achieved a significant improvement over existing models by getting an accuracy of 99.39% with a weighted average ensemble learning model, by integrating multiple state-of-the-art convolutional neural network architectures.

Diverse Model Selection: Capturing various aspects of retinal images and improving detection accuracy necessitated carefully selecting diverse models.

Clinical Applicability: The model has the potential to integrate into clinical workflows. Considerations would be real-time processing, user-friendly interfaces and adaptability to different types of retinal imaging equipment.

Advanced Preprocessing Techniques: Developed an advanced preprocessing pipeline that includes image normalization, augmentation, and enhancement techniques. This ensures high-quality input data, which is critical for training robust deep learning models.

Extensive Validation: Conducted rigorous validation using a large and diverse dataset, demonstrating the generalizability and reliability of the proposed ensemble model across different patient demographics and image qualities.

1.4 Paper Layout

The paper outlines the following important sections for proper understanding of the proposed system -

Introduction: This section introduces to the proposed system by stating the motivations behind picking up the topic, objective of the project work highlighting the changes inculcated in our work by expanding the existing work.

Literature Survey: This section focuses on the research work done on existing systems for diabetic retinopathy detection and identifying the problem statement.

Proposed System/Model: Discussions on methodologies used, schematic layout of the proposed system, the system requirements for the model and proposed algorithm for detection of diabetic retinopathy infected scans.

Experimentation and Model Evaluation: Illustrates the experiments performed to reach the final results, system’s results on the experimentations and evaluation based on performance metrics.

Conclusion and Future Scope: This section concludes by giving a summary of the project work and future scope of changes to the work on this system which can lead to potential development.

2. Literature Survey

2.1 Existing System

The present techniques for identifying diabetic retinopathy (DR) mainly depend on hands-on examinations by ophthalmologists using a range of imaging technologies to evaluate the retina's health. These conventional methods include detailed eye examinations with visual acuity tests, tonometry, pupil dilation, and fundus photography, which captures detailed images of the retina. Additionally, fluorescein angiography involves injecting a fluorescent dye to highlight retinal blood vessels, and optical coherence tomography (OCT) provides high-resolution, cross-sectional images of the retina to detect swelling and fluid buildup. Fundus autofluorescence (FAF) imaging assesses the retinal pigment epithelium's health. Despite these advanced technologies, the accurate detection and diagnosis of DR still heavily rely on the manual grading by ophthalmologists, who assess the severity of DR based on retinal images using standardized scales, ranging from no DR to proliferative diabetic retinopathy (PDR).

2.2 Problem Identification

The current system for identifying diabetic retinopathy (DR) faces significant drawbacks that obstruct effective and prompt diagnosis. It relies heavily on experienced ophthalmologists and specialized imaging devices, which are often in short supply, especially in remote or under-resourced regions. The manual grading process is time-intensive and prone to errors, leading to inconsistencies and delays in diagnosis. Subjective evaluations can result in differences in detecting DR accurately. Furthermore, the high cost of sophisticated diagnostic equipment limits access, preventing many individuals from receiving early and adequate treatment. These issues highlight the critical need for automated, efficient, and scalable solutions to enhance the screening and management of DR.

3. Proposed System/Model

3.1. Methodologies Used

For this project, we leveraged a suite of powerful tools and libraries. NumPy is used for handling numerical computations with ease, while Streamlit serves as our framework for developing interactive web applications. OpenCV is our go-to for various computer vision tasks, and Pillow is essential for image processing and manipulation. On the deep learning front, we utilize Keras (version 3.3.3) as our neural network API, built on TensorFlow (version 2.16.1), ensuring robust model development and deployment capabilities. Together, these technologies form the backbone of our data processing and machine learning workflow

3.2. Schematic Layout/DFD of the proposed system/model

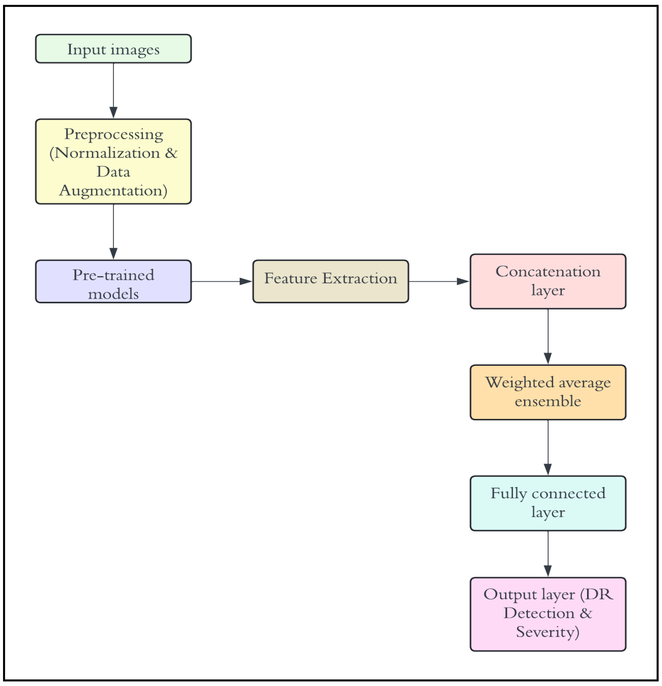

Fig 1: Workflow model of Diabetic Retinopathy detection system

The flowchart represents the process flow of the proposed ensemble deep learning model for diabetic retinopathy detection. It starts with the input of retinal images, which are then subjected to preprocessing steps including normalization and data augmentation to enhance the dataset's quality and diversity.

Following preprocessing, the images are passed through multiple pre-trained convolutional neural network (CNN) models, namely DenseNet121, ResNet101v2, MobileNet, Xception, and InceptionV3. Each of these models processes the input images independently, extracting high-level features specific to diabetic retinopathy.

The extracted features from each model are then combined in the concatenation layer, forming a comprehensive feature vector. This vector is processed through a weighted average ensemble method, where the predictions from each model are weighted according to their individual performance and combined to produce a unified prediction.

The combined features are further refined by a fully connected layer, which processes the concatenated features to generate the final output. This output layer provides the detection and severity level of diabetic retinopathy.

3.3. System Requirements

The system requirements for the diabetic retinopathy detection model include several crucial libraries that facilitate various aspects of data handling, image processing, and deep learning. NumPy is essential for numerical operations, providing support for large, multi-dimensional arrays and matrices, along with a collection of mathematical functions. Streamlit is used for building the user interface, enabling the creation of interactive web applications for visualizing and testing the model. OpenCV-python is vital for image processing tasks, such as reading, writing, and transforming images, which are fundamental for preparing the dataset. Pillow, another image processing library, is used for opening, manipulating, and saving image files in various formats. The deep learning aspect of the model relies on Keras (version 3.3.3), a high-level neural network API, which runs on top of TensorFlow (version 2.16.1). TensorFlow provides the backend engine for building and training the deep learning models, ensuring high performance and scalability. These libraries collectively ensure that the system can effectively handle data processing, model building, and deployment, facilitating the development of an accurate and robust diabetic retinopathy detection system.

3.4. Proposed Algorithm(s)

The proposed algorithm for diabetic retinopathy detection leverages convolutional neural networks (CNNs) and employs several pre-trained models in conjunction with a weighted average ensemble approach. The process begins with the preparation of input data, where retinal images are loaded from a dataset. These images undergo preprocessing steps including normalization of pixel values and data augmentation techniques such as rotation, flipping, scaling, and cropping. These steps are crucial for increasing the diversity of the dataset and improving the robustness of the model.

Once preprocessing is complete, we initialize and train five state-of-the-art pre-trained CNN models: DenseNet121, ResNet101v2, MobileNet, Xception, and InceptionV3. Each of these models processes the preprocessed retinal images independently, extracting high-level features pertinent to diabetic retinopathy. After training, feature maps and predictions are extracted from each model.

The next step involves concatenating the extracted feature maps from all the models to form a comprehensive feature vector. This concatenated feature vector serves as a rich representation of the input images, capturing a wide range of visual features. Subsequently, we employ a weighted average ensemble method to combine the predictions from each model. The weights for each model are learned based on their individual performance on a validation set. The ensemble prediction is computed using the formula:

$$P_{\text{ensemble}} = \sum_{i=1}^{n} w_i P_i$$

(1)

where Pi is the prediction from the i-th model and wi is the corresponding learned weight.

Following the ensemble step, the concatenated features are passed through a fully connected layer to refine the predictions further. The final output is generated by the output layer, which indicates the presence and severity of diabetic retinopathy. This comprehensive approach ensures high accuracy and reliability in detecting diabetic retinopathy from retinal images.

4. Experimentation and Model Evaluation

4.1. Depiction Results

The project aimed to develop an accurate and robust system for detecting diabetic retinopathy (DR) by addressing dataset imbalance and model selection challenges. Initially, the Synthetic Minority Oversampling Technique (SMOTE) was applied to the unbalanced dataset to create a balanced one, preventing bias towards the majority class. A Convolutional Neural Network (CNN) was first tested, showing a significant accuracy improvement from 37.4% on the unbalanced dataset to 73.2% on the balanced dataset, with precision, recall, and F1 scores improving similarly. Pre-trained models, including DenseNet121, ResNet, MobileNet, Xception, InceptionV3, and ResNet101V2, were then evaluated. These models outperformed the CNN, with MobileNet achieving 94.8% accuracy on the unbalanced dataset and DenseNet121 reaching 99.28% on the balanced dataset. Ensemble models were subsequently employed, with the initial combination of three base estimators (DenseNet121, ResNet101V2, and MobileNet) achieving 37.93% accuracy. When all five pre-trained models were used, the average ensemble model's accuracy rose to 69.78% on the unbalanced dataset and 99.28% on the balanced dataset. The weighted average ensemble model achieved the highest accuracy, with 70.33% on the unbalanced dataset and 99.39% on the balanced dataset. These results highlight the effectiveness of oversampling and the superior performance of pre-trained and ensemble models in detecting diabetic retinopathy.

4.2. Validation/System Performance Evaluation

The efficiency and overall performance of the system have been evaluated based on various performance metrics. Primarily, the accuracy of the models as well as the validation accuracy have been used to validate the models. A comparative analysis of various models has led to achieving a remarkable accuracy of 99.39% with weighted average ensemble learning model.

CNN: They excel at extracting features and recognizing patterns in visual data, where we achieved an accuracy of 37.4% with unbalanced dataset and an accuracy of 73.2% with a balanced dataset when run for 10 epochs each. The other performance metrics also show the same differences, as the precision of how accurate the model is making positive predictions on unbalanced dataset is 33.6% and that of balanced is 70.8%. The balanced dataset is also helping in identifying relevant instances better thereby giving a recall of 73.1% as compared to 37.4% of unbalanced. The overall performance of the classification model on balanced dataset gives an f1 score of 71.8% as compared to 35.2% of unbalanced dataset.

Pre-trained models: Pre-trained implements transfer learning, where a model trained on one task is used as the starting point for a model on a second task. The model can learn effectively and quickly from pretrained models because they are already trained on large datasets. The models were, namely, DenseNet121, MobileNet, Xception, InceptionV3, and ResNet101V2. MobileNet gave the highest accuracy of 94.8% in an unbalanced dataset, and DensenNet121 gave the highest accuracy of 99.28% in a balanced dataset, when run for 10 epochs each.

Ensemble Learning models: Initially, we used three models as base estimators namely, DenseNet121, ResNet101V2 and MobileNet, and achieved an accuracy of 37.93%, so we used all five pre-trained models as base estimators. We tried evaluating with an average ensemble model, which combined multiple models to take their average and got an accuracy of 69.78% with unbalanced dataset and 99.28% with balanced dataset. Next, we went for a weighted average ensemble model where each ensemble member contributed an equal amount to the final prediction, achieving an accuracy of 70.33% with an unbalanced dataset and 99.39% with a balanced dataset.

4.3. Discussions on Contributions

In the course of this project work, we have evenly distributed the work among all the group members.

Syed Eebad Reza : Literature survey; problem formulation and solution design; experimentation; model design and implementation; documentation

Anurag Singh : Literature survey; result validation; documentation; app design and building

Namrata Kumari Singh : Literature survey; identification of problem statement; experimentation; model design and implementation; documentation

Tanmita Roy : Literature survey; data analysis; result analysis and app design; documentation

5. Conclusion and Future Scope

5.1 Conclusion

In the culmination of our final year project, we successfully developed a robust system for detecting diabetic retinopathy, addressing key limitations and enhancing the overall accuracy of predictions. Initially, we identified a critical gap in existing systems: they failed to consider all classes of diabetic retinopathy, thereby necessitating the implementation of advanced concepts such as ensemble learning. This approach allowed us to integrate multiple models to improve predictive accuracy significantly. Through rigorous research and extensive testing of various deep learning models, we determined that an ensemble model, which combines the strengths of several algorithms, offered the best performance for our specific use case. A crucial step in our methodology was ensuring a balanced dataset, which is vital for preventing model bias and ensuring fair representation of all classes within the data. This strategic decision was instrumental in achieving an impressive accuracy of 99.39% with our balanced weighted average ensemble model, marking a significant improvement over traditional single-model approaches. Furthermore, to facilitate real-world applications and accessibility, we designed a user-friendly, static web-based application. This platform allows users to easily upload their retinal scans and receive immediate predictions, thereby bridging the gap between complex machine learning models and practical healthcare solutions.

5.2 Future Scope

Firstly, as the data world grows, a better dataset will help the model evaluate and predict the condition better. Secondly, advancements in machine learning algorithms will lead to better accuracy, and lastly, the activation functions of the algorithm can be tweaked to improve the efficiency of the model. Therefore, with potential developments in the existing work as well as by leveraging cloud computing to process large datasets, using spatial transformer networks, and integrating patient history, the system can be made robust and personalized.

References

1. Alan W. Stitt, Timothy M. Curtis, Mei Chen, Reinhold J. Medina, Gareth J. McKay, Alicia Jenkins, Thomas A. Gardiner, Timothy J. Lyons, Hans-Peter Hammes, Rafael Simó, and Noemi Lois: “The progress in understanding and treatment of diabetic retinopathy,Progress in Retinal and Eye Research,"Volume 51, 2016, Pages 156-186, ISSN 1350-9462,https://doi.org/10.1016/j.preteyeres.2015.08.001.
2. Zhang X, Saaddine JB, Chou C, et al., Prevalence of Diabetic Retinopathy in the United States, 2005-2008. *JAMA.* 2010;304(6):649–656. doi:10.1001/jama.2010.1111
3. Medium : How to Make Better Predictions by Combining Multiple Models with Python
4. https://github.com/sovit-123/Diabetic-Retinopathy-NN/tree/master
5. https://www.kaggle.com/datasets/sovitrath/diabetic-retinopathy-224x224-gaussian-filtered
6. https://www.kaggle.com/code/vishalpatil123456/binary-retinopathy-classifier-cnn-model-vishal
7. Youtube : Image classification WEB APP with Python and Streamlit | Computer vision