Amber Sautner

UCID - As4393

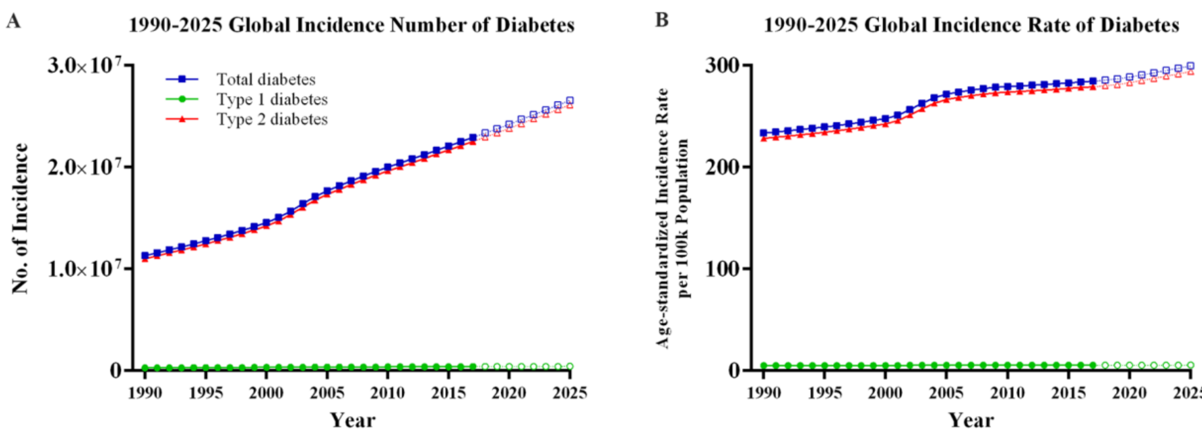
As4393@njit.edu

NJIT-ID 31609002

Table of Contents

Introduction

Diabetes mellitus is a global disease affecting millions each year, leading to reduced life expectancy, long-term disabilities, and in severe cases mortality. Despite years of research and the endless dedication of doctors, the incidence rate of this disease is projected to rise to over 570.6 million cases by 2025 [Lin et al, 2019]. It was observed that China, India, United States, Indonesia, and Mexico had over 222 million cases in 2017, with China in the lead. Depicted below, the global incidence rate of diabetes has been increasing each year, with Type 2 far surpassing Type 1 [Lin et al, 2019].



The graph clearly illustrates the worsening trend of diabetes each year, which in turn leads to a myriad of associated health complications. This report explores a revolutionary way to detect diabetic retinopathy, a severe eye disease that can result from long-term untreated diabetes.

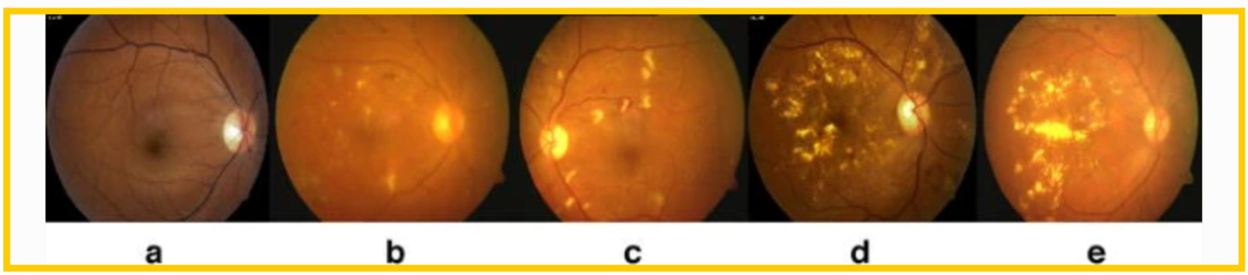
Overview

Diabetic retinopathy (DR) is a progressive eye disease caused by damaged blood vessels in the retina that leads to swelling and blurred vision. If left untreated, abnormal blood vessels can grow on the surface of the retina, leading to permanent vision loss and blindness. Since the disease progresses slowly, the symptoms may not be noticeable until it reaches an advanced stage. This leaves patients little time to receive proper treatment and often results in irreversible damage. There are various challenges that DR patients face, such as a shortage of worldwide ophthalmologists, limited availability of effective treatment, etc. When given effective treatment in time, clinical trials have reported decreased visual disability up to 90% [Das, 2016]. This demonstrates how important early detection is in preventing blindness, which is why the following proposed model is vital. It is important to note that this disease is more common in areas where citizens lack proper healthcare to prevent or treat DR, such as mainland China [Zhong et al, 2018]. Accurate and timely detection is critical in providing optimal care before it reaches a severe stage with permanent damage.

In this report, we will be analyzing the paper, *An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification* by H. Jiang, K. Yang, M. Gao, D. Zhang, H. Ma, and W. Qian [Jiang et al, 2019]. In their studies, the researchers propose a novel approach to DR detection by combining the strengths of three well-trained deep learning algorithms into one interpretable ensemble deep learning model. This paper shows significant advancement in the field of diabetic retinopathy classification by introducing a powerful method to accurately diagnosis DR for patients to acquire treatment. Throughout this report, we will investigate other methods for DR disease classification for comparison, analyze the proposed model by Jiang et al, and conclude with a personal discussion on the findings and what the future of DR patients looks like.

NPDR vs PDR

Before we delve into the methods used in DR detection, it is important to understand how DR is classified within these models. DR can be divided into two cases: non-proliferative (NPDR) and proliferative (PDR). The former represents the earlier stages of the disease, the latter is the severe stage. In NPDR, there are three levels: mild, moderate, and severe where each level is identified by unique symptoms [Nayak et al, 2008]. For example, severe NPDR images will show lesions of the retina and venous bleeding, whilst PDR will show retinal detachment, vitreous hemorrhage, etc. The following image is a visual example of what each stage of DR looks like [Nayak et al, 2008].



Retinal fundus images of **a** normal **b** mild **c** moderate **d** severe **e** PDR

This distinction between stages helps to properly classify how severe the disease has become, and which treatment is best to apply. With this knowledge, modern disease classification algorithms are trained on the images to find DR in its early stages, NPDR. If caught here, patients may be able to avoid surgical intervention, conduct lifestyle changes or obtain quick treatment.

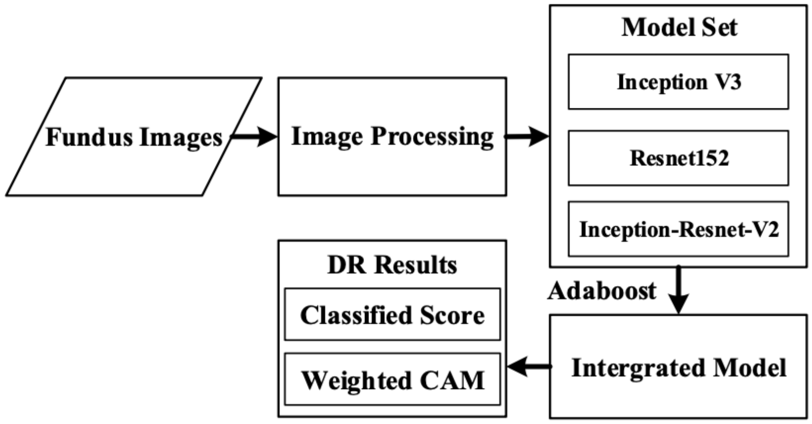
Lesion vs Image Method

DR classification models are structured on two main ideas: lesion-based or image-based [Jiang et al, 2019]. The lesion-based model works with “pixel-level annotations” to analyze the lesions in the retina. However, this approach is seen as time-consuming since researchers need to analyze each lesion in the retina, which required extensive computational resources and limits its effectiveness in large-scale applications. On the other hand, image-based algorithms are deemed more efficient by instead training the model through image labeling. This proved to be a faster, more accurate method in disease classification [Das, 2019]. Some of the main challenges that limit these two models are class imbalance, limited availability of annotated datasets, and lack of interpretability. Misunderstanding of deep learning models restricts its usefulness in real-world scenarios and leads to inefficiency and distrust [Varoquaux & Cheplygina, 2022]. We will analyze how the researchers overcame these obstacles in their proposed model.

Due to the large dataset acquired of fundus images and the desire for a faster computation time, the researchers chose an image-based method to build their model off. In the following section, we will examine how the researchers used the Adaboost method to integrate 3 deep learning models and Class Activation Maps to increase interpretability.

Proposed Model

The researchers structured their model's architecture into 5 sections: fundus images, image processing, model set, Adaboost, and DR results. In the following section, we will analyze each of these processes.



Fundus Images

In *An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification*, the researchers wanted to highlight the importance of DR detection in places that lack ophthalmologists and easily accessible healthcare. As mentioned before, China has one of the highest diabetic rates in the world. In a 2012 study, researchers in China studied the likelihood of DR rates in the general population vs diabetic citizens. The results found that the prevalence of any citizen with either NPDR or PDR in the general population was 1.3%, with a confidence interval of 0.5% to 3.2%, whilst the prevalence of NPDR alone was 1.1%, with a confidence interval of 0.6% to 2.1%. The prevalence of PDR alone was 0.1%, with a confidence interval of 0.1% to 0.3%. Among diabetic citizen, the researchers saw that the rate of having any sort of DR was high at 23%, with a confidence interval of 17.8% to 29.2%. In addition, the rate of diabetic citizens with solely NPDR was 19.1%, with a confidence interval of 13.6% to 26.3%, while the prevalence of PDR alone was 2.8%, with a confidence interval of 1.9% to 4.2%. These findings show that diabetic Chinese citizens are at a much higher risk of developing diabetic retinopathy compared to the general population [Zhong et al, 2018]. With this knowledge in mind, the authors of this paper decided to focus on DR in China in order to help address this issue. In collaboration with Beijing Tongren Eye Center, they were able to construct an extensive fundus images dataset of over 30K fundus images from patients ranging from mild to severe. The patients with mild eye conditions were labeled 'negative', at 16,661 samples. The patients with moderate or severe NPDR or PDR were labeled 'positive', at 13,583 samples. With this data, the data began to be pre-processed to prepare the data for the ensemble learning model.

Image Processing

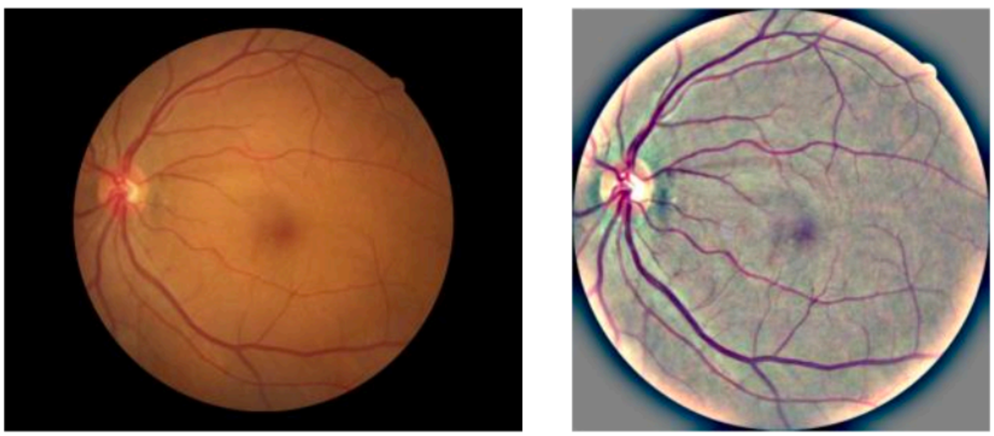
The first step in preparing the data was putting the images through various operations that included "image clipping, resizing, enhancement and augmentation" [Jiang et al, 2019]. To normalize the fundus images and increase diversity of the training data, the researchers designed their own image augmentation module that randomly transformed the images into various forms, through translation, rotation, mirroring, inclining, brightness, contrast, and sharpness. This image augmentation module contained eight different types of transformation, which helped reduce bias through randomization and improve the quality of fundus images. For the image augmentation algorithm, the researchers used a weighted summation of the original image and the effect of a Gaussian filter with a 17x17 kernel. The parameters α, β, and λ were set to 4, -4, and 128. The algorithm is shown below.

$$I_{gaussian} = I_{norm} \otimes G(x, y)$$

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$I_{enhance} = \alpha \cdot I_{norm} + \beta \cdot I_{gaussian} + \lambda$$

The performance of the image pre-processing is depicted in the following image. We can see how in image (b), the photo provides more information than the original photo (a). We conclude that the image augmentation module was successful in enhancing the fundus images details.



(a) **(b)**

Model Set

With the data now pre-processed, the researchers selected three algorithms for image classification: Inception V3, Resnet152, and Inception-Resnet-V2. These are three well-known popular classification algorithms that were chosen by the researchers. Before training the models, the images were again resized to 512 x 512 pixels to be the input map of each model using the Adam optimization algorithm with fixed learning rate of .001.

Adaboost Integrated Model

Although Inception V3, Resnet152, and Inception-Resnet-V2 are popular classification algorithms, they each have their own limitations. As a result, the researchers decided to utilize the Adaboost algorithm to combine each individual models' strengths. This formed a stronger classifier that decreased the bias of each individual algorithm. The ensemble deep learning model, also known as Adaboost algorithm, that combined the results of Inception V3, Resnet152, and Inception-Resnet-V2 worked in three main steps: distribution initialization, iterative learning, and model combination.

1. Distribution initialization

   The process of initializing the samples (x_i, y_i). The researchers labeled i=1,...m, x_i as the feature vector and y_i as the label to x_i. This step is important in improving the performance of the algorithm because it initializes the weights with appropriate values. The distribution of the samples is initialized as:

$$D_1(i) = 1/m$$

2. Iterative Learning

   The process of constantly updating the model's parameters as it is exposed to new data. This step helps the algorithm adapt and improve its accuracy as it receives new information, thus making it more effective in real-world applications. In this paper, the maximum iterative numbers for each model were all set to one million.

For t=1,…,T: select hypothesis model h_t with weighted error:

$$\varepsilon_t = P_{i \sim D_t}\left[h_t(x_i) \neq y_i\right] \tag{5}$$

The weight of each h_t is computed as:

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right) \tag{6}$$

The distribution of the samples is updated as:

$$D_{t+1}(i) = \frac{D_t(i)\exp(-\alpha_t y_i h_t(x_i))}{Z_t} \tag{7}$$

where Z_t represents the normalization factor.

3. Model combination

The process of combining the predictions of multiple machine learning models to make a final prediction, which helps improve the accuracy and help mitigate overfitting. When the weighted error is convergent and almost unchanged, the final integrated model is obtained. In the paper, the following formula was used with T=3 and the weights of the three models as $\alpha 1$=1.39, $\alpha 2$=0.536, $\alpha 3$=0.151, respectively.

$$H(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$$

Class Activation Maps

One of the key features of a successful ensemble deep learning model is its interpretability. There would be no point in creating a DR detection model if healthcare providers were unable to understand and implement it in clinical settings to improve patients' health. In this paper, the researchers utilized Class Activation Maps (CAM) to address this issue. Each model produced an individual CAM that was weighted by $\alpha i(i=1,2,3)$. Furthermore, a CAM was created for the integrated model, Adaboost. In the next section, we will discuss what the CAM's results demonstrated.

Results

The experimental results presented in *An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification* demonstrate the success of the proposed model in accurately classifying diabetic retinopathy. In the conducted experiments, the collected fundus images went through three rounds of rigorous annotation and grading by a group of trained ophthalmologists [Jiang et al, 2019]. All images were held to the international standard of DR grading category and any image of low quality was rejected to uphold accuracy. The following figure demonstrates the performance of the paper's results through the receiver operating characteristic (ROC) curve. As shown below, each individual model as well as the Adaboost integrated model performed well in terms of ROC.

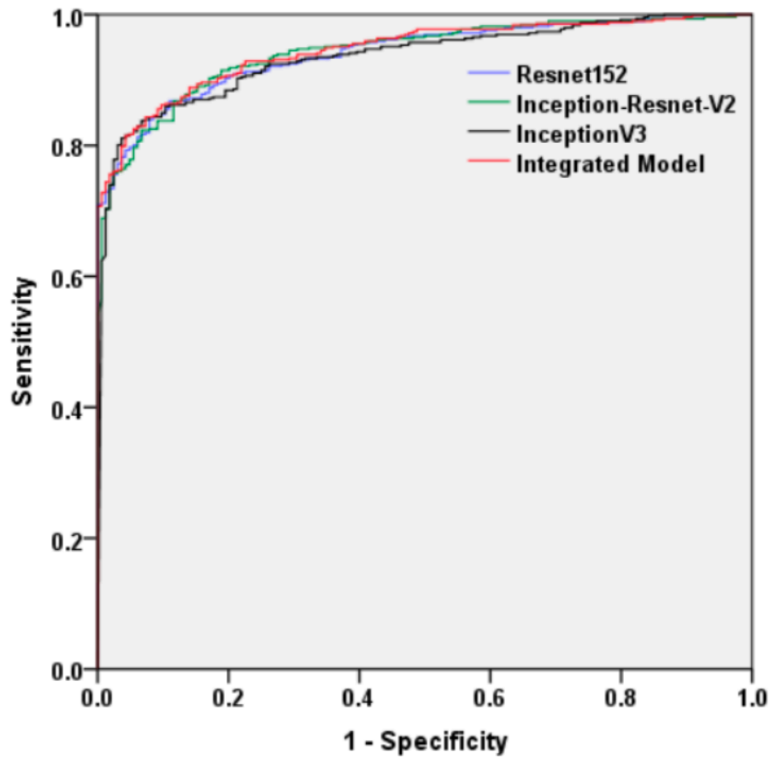


Figure 3. The performance comparison of ROC curves.

The specific performance indexes are indicated in the table below. One can see that the integrated ensemble model yields the highest accuracy at 88.21% over the other methods.

TABLE I. PERFORMANCE INDEXS OF THREE INDIVIDUAL MODELS AND THE ITEGRATED MODEL

Model Type	Sensitivity	Specificity	Accuracy	AUC
Inception V3	84.35%	91.46%	87.91%	0.935
Resnet152	84.76%	89.63%	87.20%	0.940
Inception-Resnet-V2	83.94%	88.41%	86.18%	0.943
Integrated Model	85.57%	90.85%	88.21%	0.946

These two charts support the researcher's conclusion that the combination of three deep learning models outperforms an individual method in terms of accuracy. In the final discussion, there will be a deeper analysis of these results.

The results of the CAM of the individual models can be seen below.

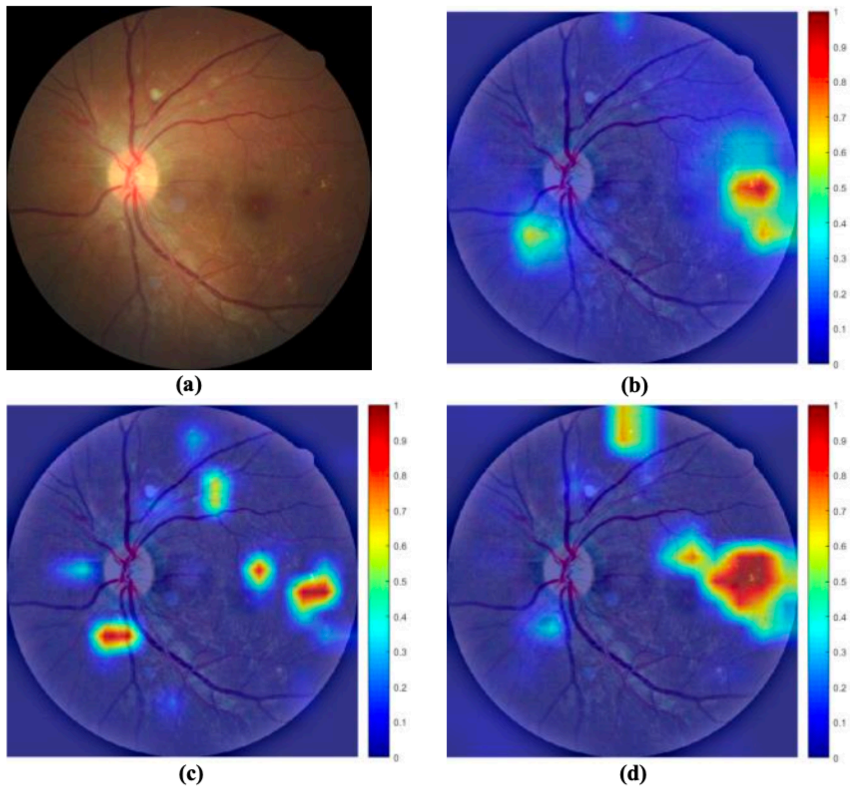


Figure 4. (a) The original fundus image. (b) The CAM of Inception V3. (c) The CAM of Resnet152. (d) The CAM of Inception-Resnet-V2.

This photo shows us that each CAM of the deep learning models show different results, which may be confusing to healthcare providers. Thus, an integrated CAM combining all three would increase interpretability because it would be a single image with the highest accuracy. Rather than analyzing multiple images, doctors could be quicker in their diagnosis and provide proper treatment. The following is a CAM of the integrated model.

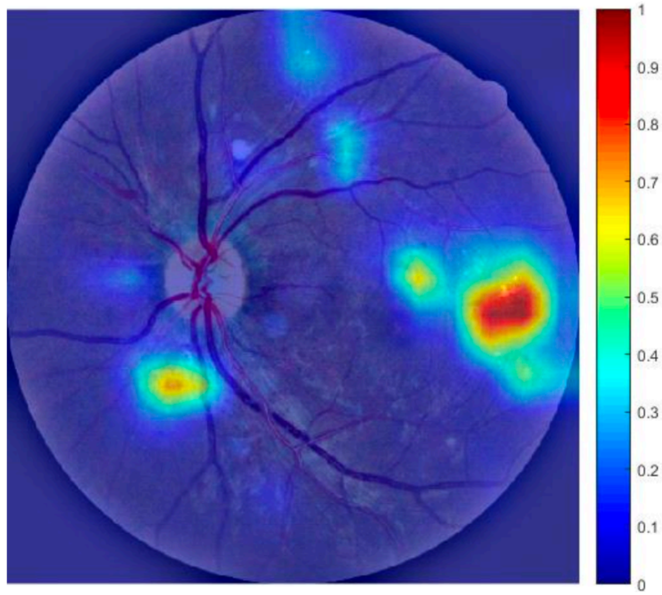


Figure 5. The CAM of integrated model.

In summary, before even applying the deep learning model the researchers designed an image augmentation module to increase diversity. Then, an integrated model was constructed using Inception V3, Resnet152, Inception-Resnet-V2 that reduced bias and obtained the highest accuracy possible. The results showed that the proposed model outperformed all the existing approaches.

Discussion

As the number of diabetic patients continues to rise each year, the importance in developing accurate detection and classification methods becomes increasingly crucial. Diabetic retinopathy is one of the many underlying side effects of long-term untreated diabetes that has serious, irreversible side effects when left untreated. Due to the lack of available treatment options and ophthalmologists, advanced detection systems and preventative measures are needed more than ever. In this section, we will analyze the results, implications, limitations, and conclusion.

Analysis

Jiang et el's combined algorithm was able to outperform in almost every category, proving its capabilities in DR detection. It achieved the highest accuracy, boasting an impressive 88.21% rate over its competitors. In addition, it yielded the highest Area under the Curve (AUC) at .946, which shows it had the highest success rate of predicting between positive and negative fundus images. It was also the second highest in specificity at 90.85%, which is the rate it was able to predict negative fundus images. Finally, the integrated model was able to predict the highest sensitivity rate at 85.57%, which is a measure of accurately predicted positive cases. Therefore, we see that the only area that this integrated model fell mildly short in was behind Inception V3 in "specificity", which was only by .61%. These results demonstrate how successful this model is and shows us the power of ensemble deep learning models.

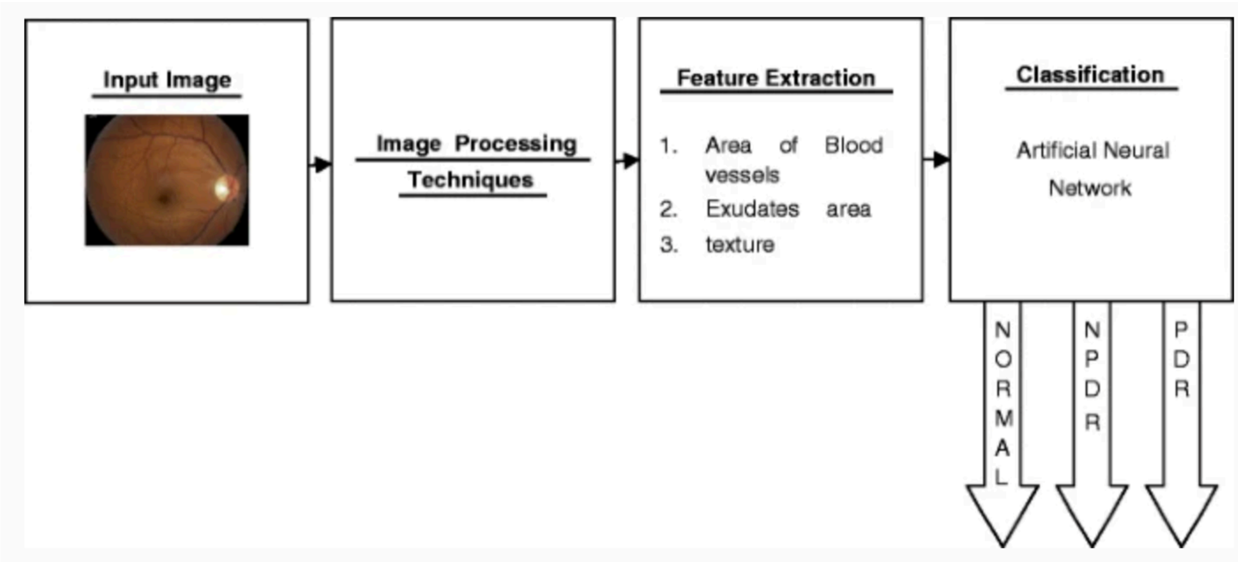
It is also important to note how the researchers utilized CAMs to increase interpretability. This allows healthcare providers to use the algorithm in real-life applications and trust in the results. With the quick advancement of technology in the past few decades and the increasing distrust around deep learning models, otherwise known as "black box" algorithms, this step was extremely important. Not only does it provide important insight for healthcare professionals to better understand the fundus images, but it is also an outline for other researchers to build interpretable ensemble deep learning models for disease classification. Thus, we see the importance in creating interpretable models in clinical use and this paper demonstrates a way for other researchers to incorporate CAMs in healthcare.

Limitations

Despite the promising results presented in the paper, *An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification*, there are still some limitations to the proposed model that should be noted. The purpose of this algorithm is to be able to apply it in a real-life clinical setting. However, the researchers required a large dataset of fundus images for training, to be specific over 30K fundus images. This large of a dataset is not practical or readily available in many clinical settings. Furthermore, real life clinical settings have confounding factors that could influence the results and impact the model, thus reducing its accuracy. In addition, this data was only tested on Chinese patients. There must be additional tests run on other datasets with varying medical histories or backgrounds [Varoquaux & Cheplygina, 2022].

Other than a clinical perspective, the proposed model could improve its interpretability. Since this model is being applied in clinical settings and affecting human health, it is vital that the results are clear to understand and applicable in real-life scenarios. One method that could be utilized is Grad-CAM, which would generate a heatmap to highlights the regions of an input fundus image that are important in classification. While CAM use global average pooling to convert feature maps, Grad-CAM uses the gradients of the predicted class with respect to the feature maps. This allows Grad-CAM to be applied to any convolutional neural network architecture, not just those with global average pooling layers [Selvaraju et al, 2017].

It would also be helpful to compare this model to other disease classification models that have been used in DR detection. In the below example, an Artificial Neural Network is fed objective measurements to automatically detect the blood vessels, hard exudates, and texture [Nayak et al, 2008]. Although these two techniques differ, the researchers could learn from other models to improve their own.



Conclusion

Early detection of DR is crucial for effective treatment and preservation of vision. For patients that cannot afford surgery or expensive treatment, early detection is life changing. The proposed model combines the strengths of three different algorithms to produce an aggregated accuracy that is both interpretable and applicable to real-life. This algorithm could be applied to regions that lack ophthalmologists or extensive treatment options, since it detects it in the early stages and preventative care could begin. Furthermore, this model is a gateway to other retina-related disease classification algorithms. Although there are some limitations and drawbacks, the promising results of this study demonstrate the potential of ensemble deep learning models for improving the accuracy and effectiveness of diabetic retinopathy disease classification.

References

Das, A. (2016). Diabetic Retinopathy: Battling the Global Epidemic. Investigative Ophthalmology & Visual Science, 57(15), 6669–6682. https://doi.org/10.1167/iovs.16-21031

Jiang, H., Yang, K., Gao, M., Zhang, D., Ma, H., & Qian, W. (2019). An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 2045-2048. https://doi.org/10.1109/EMBC.2019.8857160

Lin, X., Xu, Y., Pan, X., Xu, J., Ding, Y., & Sun, X. (2020). Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. Scientific Reports, 10, 14790. https://doi.org/10.1038/s41598-020-71908-9

Nayak, J., Bhat, P. S., Acharya, R., Lim, C. M., & Kagathi, M. (2008). Automated identification of diabetic retinopathy stages using digital fundus images. *Journal of medical systems*, *32*(2), 107–115. https://doi.org/10.1007/s10916-007-9113-9

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 618-626. https://doi.org/10.1109/ICCV.2017.74

Varoquaux, G., & Cheplygina, V. (2022). Machine learning for medical imaging: methodological

failures and recommendations for the future. npj Digital Medicine, 5, 48.
https://doi.org/10.1038/s41746-022-00592-y

Zhong, Y., Wu, J., Yue, S., Zhang, G., Liu, L., & Chen, L. (2018). Burden of diabetic retinopathy
in mainland China: Protocol for an updated systematic review and meta-analysis of prevalence and risk factors to identify prevention policies. Medicine, 97(50), e13678.
https://doi.org/10.1097/MD.0000000000013678