Deep Learning for Eye Disease Detection with Confidence Estimation and Explainable AI

ADEYINKA MAYOWA-MAJARO
University of Hull

Abstract- ***This research made progress on not only identifying valid deep learning models that can detect various eye diseases but also making the diagnosis process easier for physicians. This study focuses on three eye diseases, Cataract, Glaucoma and Diabetic Retinopathy. These three diseases are the leading causes of eye disorders which have resulted in irreversible visual impairment. Also, these three diseases are at different locations of the eye, the detection of their symptoms and analysis of the fundus images have set up different challenges. Various models, including Convolutional Neural Network (CNN), transfer learning architectures such as VGG16, InceptionV3, ResNet152V2, InceptionResNetV2, and DenseNet201 were used for Eye disease detection from patients' retinal images. DenseNet201 excelled, achieving a notable accuracy of 86.14%, a precision of 90.52%, a recall of 82.58%, and an AUC of 97.95%. To enhance model robustness, Monte Carlo (MC) dropout was integrated to estimate uncertainty levels, revealing the model's favourable performance under uncertainty. Furthermore, Local Interpretable Model-agnostic Explanations (LIME) was employed to elucidate the model's decision-making process, providing insights into how predictions were derived. This comprehensive approach showcases the efficacy of combining diverse models, leveraging transfer learning, and employing uncertainty estimation and explainability techniques for accurate eye disease detection.***

I. INTRODUCTION

Eye diseases, known as the global leading cause of vision loss, have affected millions of people (Bourne et al.2021). These conditions can affect any part of the eye, including the cornea, iris, lens, retina, and optic nerve. Cataracts, glaucoma, age-related macular degeneration (AMD), and diabetic retinopathy are the most common eye diseases that have posed a huge threat to people's life quality (Chalakkal et al., 2020). Cataracts occur as a clouding of the lens as a result of a build-up of proteins in the eye. This leads to blurred vision (Fekrat et al., 2021). Glaucoma is a disease of the optic nerve and involves a characteristic pattern of progressive damage to the nerve that transmits visual information to the brain. This can mean patchy losses of vision that are not noticed until late in the disease (Mélik et al.2020). Age-related macular degeneration is a painless eye condition that leads to the gradual loss of central vision (Thier & Holmberg, 2022). Diabetic retinopathy is a complication of diabetes, caused by high blood sugar levels damaging the back of the eye. It can cause blindness if left undiagnosed and untreated. This is because it affects the retina and the retinal blood vessels. There are two types of diabetic retinopathy: background retinopathy and proliferative retinopathy. Background retinopathy can occur at any stage and often does not affect sight. Proliferative retinopathy is when background retinopathy advances and the blood vessels in the retina start to become blocked. The retina is the part of the eye that converts light coming through the eye into electrical signals (Kropp et al.2023). It is reported that approximately 75% of total blindness can be avoided by early detection and treatment (Wong & Sabanayagam, 2020). Many of these cases can be prevented by early detection and treatment (Allison et al., 2020). As the current trend in the medical field suggests the convergence of artificial intelligence and medical science, the combination of AI and computer vision in the field of ophthalmology is a very promising development. This paper focuses on the use of deep learning algorithms for the detection of eye diseases. This study will explore the use of confidence estimation techniques combined with explainable artificial intelligence models to provide insightful information to medical professionals about the reliability of computer-aided diagnosis.

II. LITERATURE REVIEW

A good number of research has been done in this field by several researchers using different machine learning and deep learning algorithms and diagnostic methods. Ramanathan et al. (2021) explored Eye Disease (ED) detection using Machine Learning (ML) algorithms including Logistic Regression (LR), Random Forests (RF), Gradient Boosting (GB) and Support Vector Machines (SVM). Their findings emphasized the potential of ML in automating ED detection, achieving a notable accuracy rate of 90% in classifying various EDs but overlooked the quantification of prediction uncertainty. Rahul Pahuja et al. (2022) utilised ML architectures effectively. Their study's strength lies in the use of ML and Deep Learning (DL) architectures for an automated diagnosis of ED by creating CNN and SVM models to classify ED. The study showcased the effectiveness of ML models in achieving 87.08% and 87.5% accuracies respectively. However, like the rest of the works mentioned here, it falls short in discussing uncertainty estimation and interpretability, thereby leaving room for improvement in ensuring reliable and transparent decision-making. Arunkumar, (2021) focused on ML models for the detection of human eye disease. The author utilised Neural Networks (NNET), RF, K-Nerest Neighbor (KNN) and SVM models, achieving accuracies of 75.32%, 63.63%, 64.50% and 57.07% respectively. However, the study lacks explicit discussion on confidence estimation and explainability, leaving a gap in understanding the model's decision-making process. Nouf et al. (2022) applied various ML model classifiers including SVM, KNN, Naıve Bayes (NB), Multi-layer perceptron (MLP), Decision Tree (DT) and RF as well as Deep Learning (DL) models such as CNN based on Resnet152 model on the Ocular Disease Intelligent Recognition dataset to detect human eye infections of Glaucoma disease. With this study, the RF and MLP classifiers achieved the highest accuracy of 77% in comparison to the other ML classifiers while the deep learning model (CNN model: Resnet152) attained an accuracy of 84% for the same task and dataset. Zahraa et al. (2023) presented an ML-based method for targeted ocular detection using the Ocular Disease Intelligent Recognition (ODIR) dataset. The study utilized ML models such as NB, DT, RF and KNN on both binary and multiclass classification tasks. Amongst other models, NB achieved the highest accuracy (75%) in binary classification while NB achieved the highest accuracy (88%) for multiclass classification. Rodr et al.(2022) customised a new dataset with 20 classes which they called the MuReD (Multi-label Retinal Diseases) dataset, this was customised from the publicly available datasets to have varieties of eye diseases to predict. The C-Tran architecture was selected as the classification model, this model achieved 90% accuracy. While successful in customization, it also leaves a gap in understanding model reliability and interpretability.

III. OBJECTIVES

All the works discussed earlier contributed substantially to the development of ML algorithms for various eye disease detection. Although the achieved levels of accuracy were high (75%-90%) across the range of the outlined studies, none of these works explicitly address confidence estimation or employ explainable AI techniques. This is important because it enhances reliability, trust, and transparency. The work presented herein aims to extend upon these foundations by introducing confidence estimation methods like Monte Carlo (MC) dropout and explainable AI methods like Interpretable Model-Agnostic Explanations (LIME). These additions seek to provide a measure of confidence in predictions and offer insights into features influencing the model's decisions, crucial for adoption in clinical settings.

The research objectives of the present study are to:

1. Develop ML models for various eye disease detection that are accurate and reliable for diagnosing various eye diseases.
2. Integrate uncertainty estimation techniques, such as MC dropout, to quantify prediction uncertainty.
3. Incorporate explainable AI methods, specifically LIME, to provide insights into the features influencing the model's decisions.

IV. MATERIALS AND METHODS

Figure 1 shows the process taken to achieve automatic ED detection. Starting from detecting and classifying ED into different classes to implementing model confidence estimation to quantify model predictions, and finally introducing model explainability to explain

how the model has made its prediction. In this study, all experiments were programmed and implemented using Python 3.11.7 with Keras and Tensorflow (Tensorflow,2024). Anaconda Jupiter notebook was used as the software environment of the experiment and models were run on an Intel(R) Core(TM) i5-1145G7 @ 2.60GHz, 2611 Mhz, 4 Core(s), and 8 Logical Processor(s) with 32GB RAM.

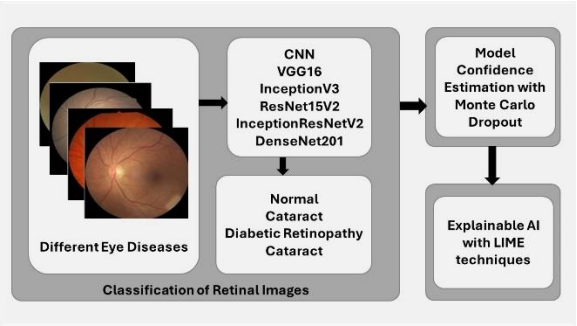


Figure 1: Process Schema for Automatic DR Detection

V. DATA COLLECTION AND PREPROCESSING

An eye disease dataset sourced from Kaggle (Eye-disease (kaggle.com)) was used in this study. This dataset comprises 4,217 colour images (Figure 2) categorised into four groups: Cataract (1,038 images), DR (1,098 images), Glaucoma (1,007 images), and Healthy (1,074 images). To ensure all images have a similar size, images were resized to 128x128 pixels, retaining the aspect ratio. For easier computation, all matrix representations of images were further flattened into vectors. Each value within the vector was normalized to between 0 and 1, this step is crucial to mitigate the effect of different pixel values, as the model could be sensitive to the magnitude of input. Also, a data augmentation technique is applied. Through using affine transformation in the Tensorflow library, the numbers of each category are increased to 1100. Horizontal flip, rotation, zooming, shifting, etc., were used to artificially increase the size of the images. Specifically, an extra number of the data is created by applying these linear transformations. Moreover, both horizontal and vertical flipping were used to further increase the images in the dataset. This step is essential to balance each category of the dataset, alleviate overfitting and expand the diversity of the dataset, hence improving the generalization capability in the later model training phase.

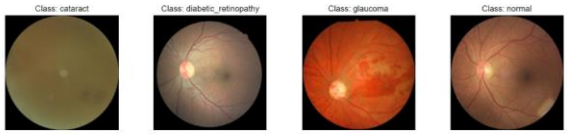


Figure 2: Representative samples of different Eye diseases

VI. MODEL TRAINING

Upon data preprocessing, models were trained for eye disease detection. Baseline CNN and transfer learning models including models VGG16, InceptionV3, ResNet152V2, InceptionResNetV2 and DenseNet201 were employed. Models were trained on 80% of the dataset and tested on 20%. Considering the dataset size (Table 1) the batch size was set to 16, this helps the models generalize better on unseen data. Models underwent training for a maximum of 10 epochs, with a callback patience of 2. This monitored the model's accuracy and stopped the training process when the model stopped improving, it also prevented overfitting and saved training time. Optimisers such as Adam (Adaptive Moment Estimation), RMSprop (Root Mean Square prop) and SGD (Stochastic Gradient Descent) were tested with Learning Rates (LR) of 0.001, 0.0005, and 0.0001 on the baseline CNN model. The best-performing optimiser and LR were then applied to the transfer learning models. Meanwhile, categorical crossentropy was used as the loss function with softmax activation to measure the difference between the predicted output and the true output. The designed CNN for image classification consists of three convolutional layers followed by max-pooling for down-sampling. The architecture incorporates Rectified Linear Unit (ReLU) activation functions to introduce non-linearity. The final layers include flattening to transform feature maps into a 1D vector, a dense layer with ReLU activation, and an output layer for class prediction. The transfer learning models (VGG16, Inceptionv3, ResNet152V2, InceptionResNetV2 and DenseNet201) were loaded with frozen weights and layers. Additional layers, including dropout, batch normalization, and dense layers, were added to fine-tune the model for a specific classification task.

Table 1: Hyperparameter Configurations.

Parameters	Multiclass Classification
Batch Size	32
Dataset Size	4,400 images
Epoch Number	10
Learning Rate (LR)	0.001, 0.0005, 0.0001
Optimizer	Adam, RMSprop, SGD
Activation Function	*softmax*
Loss Function	Categorical Crossentropy

VII. EVALUATION METRICS

All ML models were evaluated based on Accuracy, Precision, Recall and AUC given from equations 1-3 below. Accuracy is used to show how often the model is correct overall, Precision shows how many of the predicted positive cases are truly positive, Recall of a model shows how many of the actual positive cases the model correctly identifies and AUC measures how well the model distinguishes between positive and negative cases.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \text{ x } 100\% \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \text{ x } 100\% \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \text{ x } 100\% \quad (3)$$

VIII. CONFIDENCE ESTIMATION INTEGRATION

Confidence Estimation plays a crucial role in comprehending the model's behaviour, empowering users to make more informed decisions and bolstering the overall reliability of ML systems. The Monte-Carlo Dropout method is utilized to obtain the confidence score from the machine learning model's prediction. Monte Carlo Dropout is a technique that involves randomly dropping out neurons during the training process and performing multiple forward passes to yield a distribution of predictions (Sun et al.2023). This distribution provides not only the class label but also the associated confidence scores. By visualizing the confidence levels of the model's predictions for different regions in an eye fundus image, medical professionals can easily identify high-risk regions. A threshold can be set based on the confidence level, and the corresponding regions in the image can be classified into different categories based on their confidence levels. This information provides valuable insights for medical professionals in identifying high-risk regions more effectively. During testing, the model was iteratively evaluated with dropout enabled, providing insights into its level of confidence or uncertainty for individual predictions.

IX. EXPLAINABLE AI TECHNIQUES IMPLEMENTATION

Explainable AI is crucial in domains like healthcare to explain how the model arrived at the prediction made. LIME (Local Interpretable Model-Agnostic Explanations) is a widely used explainable AI technique in various domains, including credit scoring, healthcare, and remote sensing (Silva et al.2023). LIME aims to explain the reasoning behind the output predictions by attributing the change in the output to different input features.
In this study, LIME was leveraged to provide explanations for the outputs of the trained DenseNet201 model. LIME selects a specific instance for explanation and generates a perturbed dataset by introducing slight variations to its features. After passing the perturbed dataset through the original model, predictions are obtained for each sample. A simple interpretable model is then trained on the perturbed dataset, mapping input features to model predictions. The resulting locally fitted model provides heatmap explanations for the original model's prediction on the selected instance, highlighting the importance of each feature (Allgaier et al., 2023). The heatmap of these explanations is then visualized as seen in Figures 5 and 6 to aid users in understanding the model's decision-making process.

X. RESULT AND DISCUSSION

- Performance Evaluation of ML Models

Table 2 presents the performance metrics of several deep learning models trained on eye disease datasets, including CNN, VGG16, InceptionV3, ResNet50V2, InceptionResNetV2, and DenseNet201. The evaluation metrics include accuracy, precision, recall, and area under the receiver operating characteristic curve (AUC). These metrics provide insights into the models' ability to correctly classify instances,

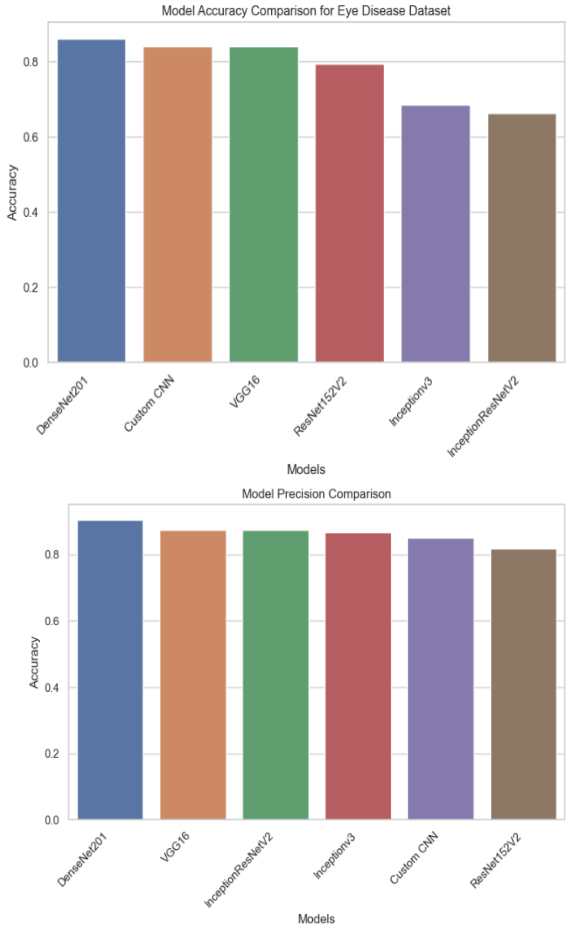
distinguish between classes, and handle class imbalances. DenseNet201 achieved the highest accuracy of 86.14%, followed closely by CNN and VGG16 with 84.12%. InceptionV3 had the lowest accuracy of 68.48%, indicating that it misclassified a significant portion of instances. DenseNet201 also exhibited the highest precision of 90.52%, indicating its ability to minimize false positives. VGG16 followed closely with a precision of 87.39%. InceptionV3 exhibited respectable precision (86.68%), despite its lower accuracy, suggesting that when it predicts a positive outcome, it is usually correct. CNN and DenseNet201 demonstrated the highest recall (>82%), indicating its ability to capture a large proportion of positive instances. However, InceptionV3 exhibited the lowest recall (51.66%), indicating its weakness in identifying true positives. DenseNet201 achieved the highest AUC of 97.95%, indicating its strong discriminative ability across various thresholds. CNN, VGG16, and ResNet50V2 also demonstrated high AUC values (>95%), indicating excellent overall performance. InceptionV3 and InceptionResNetV2 had relatively lower AUC values (90.42% and 89.39%, respectively), indicating weaker discrimination ability compared to other models. Based on the performance metrics, DenseNet201 emerges as the top-performing model in terms of accuracy, precision, recall, and AUC. VGG16 also performs well across these metrics, closely following DenseNet201. In contrast, InceptionV3 and InceptionResNetV2 show comparatively lower performance, particularly in recall and AUC.

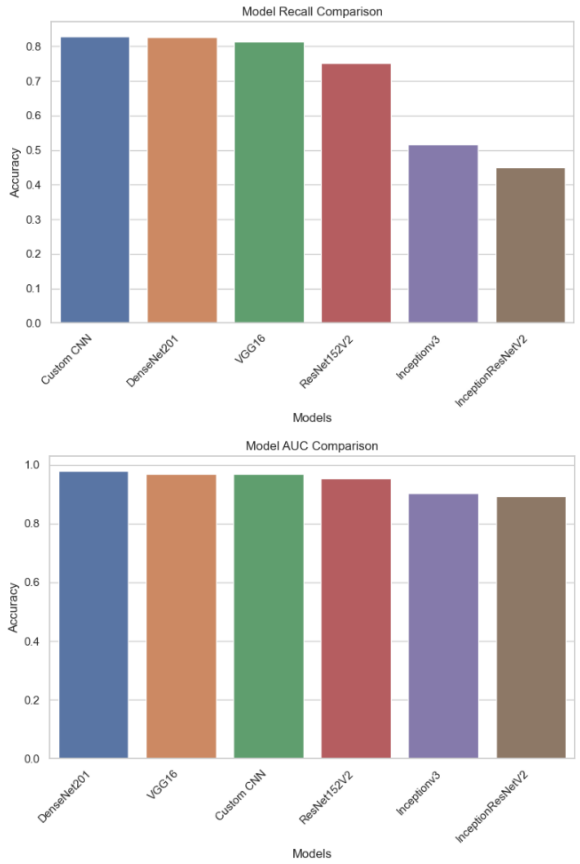
DenseNet201 outperforms other models as evident in Figure 3. Its robust performance is attributed to its features which include a 201-layer deep model that distinguishes itself through its unique dense connectivity pattern, where each layer directly receives input from all preceding layers. This architecture facilitates efficient feature reuse, allowing the model to capture intricate patterns effectively. Additionally, DenseNet201 incorporates batch normalization, pooling, and separable convolution layers, enhancing computational efficiency and minimizing memory requirements (Wang et.al 2020).

Table 2: Performance metrics by six models for multiclass classification of different Eye Diseases

Model	Accuracy (%)	Precision (%)	Recall (%)	AUC (%)
CNN	84.12	85.16	82.94	96.89
VGG16	84.12	87.39	81.28	96.89
InceptionV3	68.48	86.68	51.66	90.42
ResNet152V2	79.50	81.83	75.24	95.23
InceptionResNetV2	66.23	87.33	44.91	89.39
DenseNet201	86.14	90.52	82.58	97.95

Figure 3: Comparison of different classification models in terms of Accuracy, Precision, Recall and AUC





- Confidence Estimation Accuracy Analysis

Utilizing MC Dropout, the model's uncertainty levels were assessed. The calibration curve, which shows the relationship between the confidence score outputted by a model and the actual accuracy of those predictions, was utilized to inspect the level of confidence estimated by the model. Figure 4 visually represents the evaluation of uncertainty predictions in the classification model (DenseNet201) on ten image samples. Notably, nodes 3, 5, and 7 show deviations, indicating instances of under-confidence, while the model is generally confident in predicting other nodes. To further validate the confidence level of the model, 5 random samples were chosen from the validation set of the dataset and a bar chart was plotted for each selected sample showing both predicted and actual class (Figure 5). All 5 samples demonstrated true prediction of the actual class, which validates that the model is highly confident in its prediction.

Figure 4: Visualization depicting the evaluation of uncertainty predictions of the classification model (DenseNet201) on ten image samples. Each node's position corresponds to the level of uncertainty. The closer the node is to the perfectly calibrated line, the higher the confidence level.

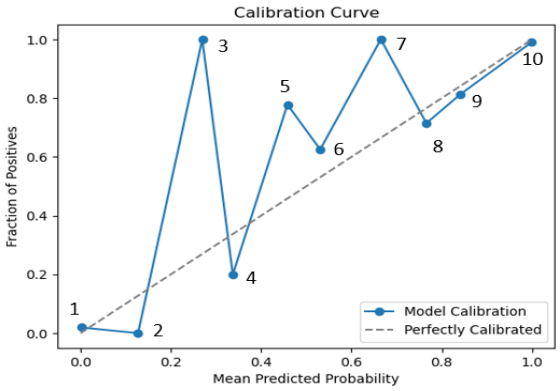
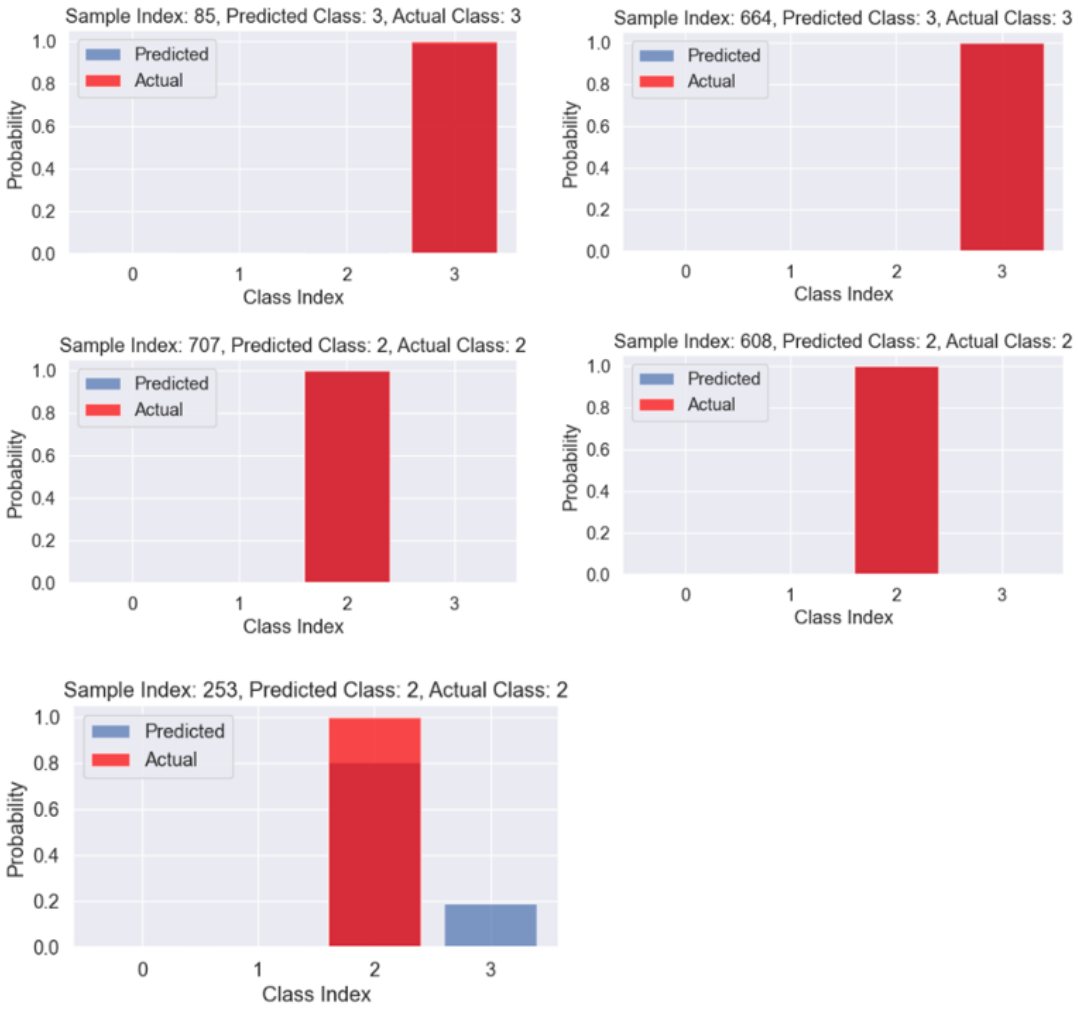


Figure 5: *Visualization depicting the evaluation of uncertainty predictions of the classification model (DenseNet201) on five image samples.*



- Interpretability of Explainable AI Technique

Preserving consistent effectiveness, the model's accuracies and AUCs hovered at approximately 86% and 97% respectively across all evaluated categories, adhering to its fundamental principles. The result of the LIME explanation is visualized into four images (Figure 6), the predicted class from the dataset, the positive-only LIME Explanation (POLE), the Positive and Negative LIME Explanation (PNLE) and the LIME Heatmap Explanation (LHE). In the Positive Only LIME Explanation (POLE), the focus is on highlighting the parts of the retinal image that the model considers important for making its prediction. POLE identifies specific features or patterns in the retinal image that strongly suggest the presence of the condition the model is predicting. In addition to highlighting the positive aspects, the negative ones are also considered in PNLE. It's like looking at both the light and shadow in a photograph. The positive explanation still shows us where the model sees signs of the condition, but also looks at areas that might contradict that diagnosis. This gives a more balanced understanding of how the model is interpreting the image and making its prediction. The heatmap helps to visualize the overall influence of different parts of the image on the model's decision. The intensity of colours (red and blue) in Figure 7 shows how strongly certain areas contribute to the prediction. The darker the color, the stronger the influence. In this heatmap, areas highlighted in blue indicate strong positive influence, suggesting they strongly support the model's prediction. Conversely, areas in red indicate a strong negative influence, suggesting they might go against

the prediction. Lighter shades of red and blue show areas of lesser influence, providing a gradient of importance across the image. Looking at these different types of explanations together shows a clearer picture of how the model analyzed the retinal image and made its prediction. These insights can help us validate the model's predictions, identify areas for improvement, and ultimately build trust in its ability to assist in diagnosing conditions based on retinal images.

Figure 6: LIME explanation depicting the retinal images with corresponding heatmaps generated by LIME

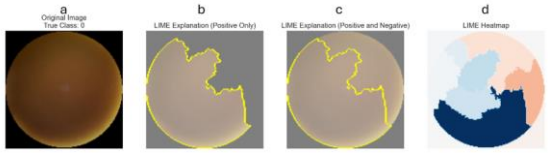
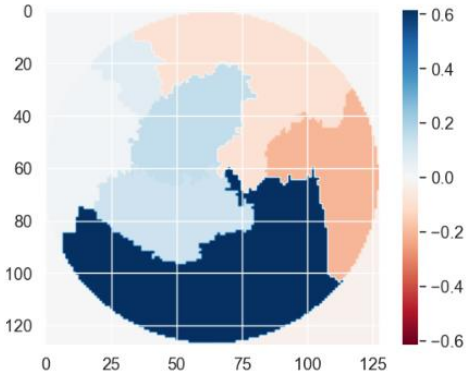


Figure 7: Intensity of colors for the LIME explanation



CONCLUSION

Early detection of various eye diseases, including glaucoma, diabetic retinopathy, and cataracts, is crucial for preventing vision loss. Machine learning models offer promise in revolutionizing eye disease detection, enabling timely intervention and personalized treatment. Numerous research groups explored ED detection using ML methods, however, some gaps were still present. This study addressed the gaps by presenting ML models that can detect and classify different eye diseases with a measure of confidence and explanations in predictions. This study applied six classification models, namely CNN, VGG16, InceptionV3, ResNet152V2, InceptionResNetV2 and. DenseNet201 model excelled in these tasks with an accuracy of 86.14% and AUC of 97.95%. Additionally, MC dropout was integrated to estimate the level of uncertainty. This study further used heatmaps under the LIME explainable AI paradigm for scientific validation. These outcomes underscore the versatility of the selected models, showcasing high confidence in predictions and their adaptability to different classification scenarios. The robust performance highlights the potential of these models for diverse medical imaging applications. Overall, this project addresses the broader spectrum of model behaviour. By enabling the model to quantify its uncertainty, the project advances beyond black-box predictions, creating a foundation for AI-assisted clinical decisions that are informed by reliability metrics.

FUTURE WORK

The model interpretability can be enhanced by exploring advanced interpretability methods like SHAP (Shapley Additive explanations) or attention mechanisms for more accurate explanations. Additionally, It is recommended to focus on utilizing diverse and large datasets containing high-resolution images. This approach can enhance the model's ability to learn robust features and generalize effectively to unseen data. By leveraging high-resolution images, finer details and nuances can be captured, leading to improved feature representation and discriminative pattern learning.

REFERENCES

[1] Bourne, R., Steinmetz, J.D., Flaxman, S., Briant, P.S., Taylor, H.R., Resnikoff, S., Casson, R.J., Abdoli, A., Abu-Gharbieh, E., Afshin, A. and Ahmadieh, H., 2021. Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. The Lancet global health, 9(2), pp.e130-e143.thelancet.com

[2] Chalakkal, R. J., Abdulla, W. H., & Hong, S. C., 2020. Fundus retinal image analyses for screening and diagnosing diabetic retinopathy, macular edema, and glaucoma disorders. Diabetes and fundus OCT.HTML

[3] Fekrat, S., Glaser, T. S., & Feng, H. L., 2021. All about Your Eyes, revised and updated.dukeupress.edu

[4] Mélik Parsadaniantz, S., Réaux-le Goazigo, A., Sapienza, A., Habas, C. and Baudouin, C., 2020. Glaucoma: a degenerative optic neuropathy related to neuroinflammation?. Cells, 9(3), p.535.mdpi.com

[5] Thier, A. & Holmberg, C., 2022. The patients' view: age-related macular degeneration and its effects–a meta-synthesis. Disability and Rehabilitation.HTML

[6] Kropp, M., Golubnitschaja, O., Mazurakova, A., Koklesova, L., Sargheini, N., Vo, T.T.K.S., de Clerck, E., Polivka Jr, J., Potuznik, P., Polivka, J. and Stetkarova, I., 2023. Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications—Risks and mitigation. EPMA Journal, 14(1), pp.21-42.springer.com

[7] Wong, T. Y. & Sabanayagam, C., 2020. Strategies to tackle the global burden of diabetic retinopathy: from epidemiology to artificial intelligence. Ophthalmologica.karger.com

[8] Allison, K., Patel, D., & Alabi, O., 2020. Epidemiology of glaucoma: the past, present, and predictions for the future. Cureus.cureus.com

[9] Ramanathan, Gauri, et al. "Eye Disease Detection Using Machine Learning." *IEEE Xplore*, 1 Oct. 2021, ieeexplore.ieee.org/abstract/document/9587740.

[10] Rahul Pahuja, et al. "A Dynamic Approach of Eye Disease Classification Using Deep Learning and Machine Learning Model." *Lecture Notes on Data Engineering and Communications Technologies*, 1 Jan. 2022, pp. 719–736, https://doi.org/10.1007/978-981-16-6289-8_59. Accessed 6 Feb. 2024.

[11] Arunkumar, (2021) Arunkumar, K. *MACHINE LEARNING MODELS for the DETECTION of HUMAN EYE DISEASE*.

[12] Nouf B., Amal A., Ashwaq A. & Raouia M., 2022. Automatic Eye Disease Detection Using Machine Learning and Deep Learning Models.

[13] Zahraa N.A. & Abbas M AI-Bakry, 2023. Diagnose eye diseases using various features extraction approaches and machine learning algorithms.

[14] M. A. Rodr´ıguez, H. AlMarzouqi, & P. Liatsis, 2022. Multi-label Retinal Disease Classification Using Transformers.

[15] TensorFlow. (n.d.). *Module: tf.keras | TensorFlow Core v2.4.1*. [online] Available at: https://www.tensorflow.org/api_docs/python/tf/keras [Accessed 25 Feb. 2024].

[16] Tensorflow Library accessed at: https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/apply_affine_transform.

[17] Sun, T., Yin, B. and Bohté, S., 2023, September. Efficient Uncertainty Estimation in Spiking Neural Networks via MC-dropout. In International Conference on Artificial Neural Networks (pp. 393-406). Cham: Springer Nature Switzerland. [PDF]

[18] Silva, R.M., Sbrana, A., de Castro, P.A. and Soma, N.Y., 2023. Developing and Assessing a Human-Understandable Metric for Evaluating Local Interpretable Model-Agnostic Explanations. International Journal of Intelligent Engineering & Systems, 16(4). inass.org

[19] Allgaier, J., Mulansky, L., Draelos, R.L. and Pryss, R. (2023). How does the model make predictions? A systematic literature review on the explainability power of machine learning in healthcare. *Artificial Intelligence in Medicine*, [online] 143, p.102616. doi:https://doi.org/10.1016/j.artmed.2023.102616.

[20] Wang, S.-H. and Zhang, Y.-D. (2020). DenseNet-201-Based Deep Neural Network with Composite Learning Factor and Precomputation for Multiple Sclerosis Classification. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 16(2s), pp.1–19. doi: https://doi.org/10.1145/3341095.