ANSAN (Amrita Nethra Samrakshan System)

A Mobile Application to Assist Hospitals for Glaucoma Screening

Ashwin Narayanan S
Department of Computer Science and Engineering
Amrita School of Computing Coimbatore
Amrita Vishwa Vidyapeetham
Tamil Nadu, India
cb.en.u4cse21008@cb.students.amrita.edu

A.V. Viswa
Department of Computer Science and Engineering
Amrita School of Computing Coimbatore
Amrita Vishwa Vidyapeetham
Tamil Nadu, India
cb.en.u4cse21168@cb.students.amrita.edu

Ashwin R Sharma
Department of Computer Science and Engineering
Amrita School of Computing Coimbatore
Amrita Vishwa Vidyapeetham
Tamil Nadu, India
cb.en.u4cse21108@cb.students.amrita.edu

Harinandan N
Department of Computer Science and Engineering
Amrita School of Computing Coimbatore
Amrita Vishwa Vidyapeetham
Tamil Nadu, India
cb.en.u4cse21114@cb.students.amrita.edu

Murugesh JLP
Department of Computer Science and Engineering
Amrita School of Computing Coimbatore
Amrita Vishwa Vidyapeetham
Tamil Nadu, India
cb.en.u4cse21137@cb.students.amrita.edu

K Vaisakhkrishnan
Department of Computer Science and Engineering
Amrita School of Computing Coimbatore
Amrita Vishwa Vidyapeetham
Tamil Nadu, India
cb.en.u4cse20069@cb.students.amrita.edu

Selvanayaki Kolandapalayam Shanmugam
Department of Mathematics and Computer Science
Ashland University
USA
skolanda@ashland.edu

Somasundaram K*
Department of Mathematics
Amrita School of Physical Sciences
Amrita Vishwa Vidyapeetham
Tamil Nadu, India
s_sundaram@cb.amrita.edu

***Abstract*—In India, around 12 Million people suffer from glaucoma, and 1.5 million are blind due to it. However, more than 75% of glaucoma cases are undiagnosed. Millions of individuals worldwide experience blindness due to the unfortunate reality that they were unaware of their glaucoma diagnosis. As per the World Health Organization (WHO), glaucoma is among the primary causes of blindness and visual impairment, particularly in high-income nations. This motivated us to develop and implement a system advantageous to both healthcare facilities and potential patients (i.e., the general public). We posit that this system will substantially mitigate the disease's severity, which will be discussed in this paper. Individuals can self-register for the application, or frontline workers, doctors, or hospital management personnel can add them. Upon registration, their medical history is collected, fundus images of their eyes are captured and uploaded, and an immediate prediction of the presence or absence of glaucoma is provided. A report of the findings is shared with the hospital's doctors, among other actions. Both the patient and the doctor can review the reports. Depending on the severity, doctors can add comments to the reports, which patients can view and respond to. There were apps like the Yanbao app, but it only works on iOS, and our app ANSAN works on both Android and iOS platforms and has more personalized features in comparison such as offering scalable data storage options using both relational and NoSQL databases. ANSAN can be a real-world application to help doctors, potential patients, and the world in reducing cases where glaucoma is the cause of blindness through early detection.**

***Index Terms*—Glaucoma, Deep Learning, Comparative Study, Telemedicine, Disease Detection, Computer-Aided Diagnosis, Healthcare Accessibility**

I. Introduction

We have only one pair of eyes through which we experience the beautiful world around us. Sight allows us to navigate our environment, connect with others, and engage with the world in a multitude of ways. Imagine losing this precious gift. Imagine a world plunged into darkness, replaced by a constant blur or an impenetrable veil. The emotions associated with vision loss can be profound. Fear of the unknown, frustration with daily tasks, and a deep sense of isolation are just some of the feelings people with vision loss experience as suggested by Docia L. Demmin [6] in their paper about the same. Now imagine losing this vision purely because of an individual unaware of the disease he might have.

Glaucoma involves a progressive loss of retinal ganglion cells (RGC) and characteristic changes in neuroretinal rim tissue in the optic nerve head (ONH), which are accompanied by visual field (VF) construction. There are several types of glaucoma constituting a group of eye diseases that are the leading cause of blindness [7]. According to this study, in a sample of 5000 urban Greek people over 59 years of age, 57.1% of glaucoma cases were found to be undiagnosed. A study of 3654 predominantly white Australians (90.2% over 60 years of age and 24% over 80 years of age) found that 51% of glaucoma cases were again due to lack of early detection.

This motivated us in crafting ANSAN, which bridges the gap between potential patients and cures and helps them detect glaucoma early. In this paper, we present ANSAN, its design, and how it is going to help us detect glaucoma early and prevent people from becoming visually impaired. Glaucoma is the leading cause of blindness according to the World Health Organization [1] and the fact that there is no cure, early detection, and treatment with our app can significantly decrease the blindness count. The mobile app can also significantly improve the scale at which hospitals work by attending to more patients efficiently through the app.

There are 5 main types of glaucoma [8]:

- Open Angle Glaucoma
- Normal-tension Glaucoma
- Angle Closure Glaucoma
- Congenital Glaucoma
- Secondary Glaucoma

As an initial step, our app will detect the presence or absence of glaucoma. In future versions, we plan to even detect the type of glaucoma with annotations on the image.

In recent years, the emergence of AI-powered apps and their potential in providing innovative solutions for disease management has raised the bar on the quality of apps that patients and healthcare professionals expect. There have been several solutions to glaucoma pre-screening like the Yanbao App [2] and e-Paarvai, an app for cataract pre-screening [9], but what's unique in our solution is the end-to-end relationship that we build between doctors and patients.

Connecting people from rural areas with comparatively limited or no access to technology through the camps conducted by hospitals where our app will be used to quickly identify potential patients with glaucoma and notify the hospital is the novelty of our app. Through this paper, we propose ANSAN, a robust app that we believe will definitely meet the expectations of healthcare professionals and users and make this world free from blindness due to an individual being unaware of the presence of glaucoma. ANSAN is built with **Flutter** for front-end. Back end for the first approach to store data in local hospital server, is built using **Node.js, Express.js and MySQL** for data storage whereas for the cloud based approach it is integrated into Flutter with Google's **Firebase**, Fireabase Auth for authentication, Firestore for data storage, Firebase ML for deploying the **ViT** based machine learning model and Firebase storage for image storage. We'll talk about this in detail in this paper.

We'd like to mention that ANSAN was shortlisted in the Google Solutions Challenge, 2024 [3] for the India Regional Bootcamp that was conducted on January 30th, 2024, in Chennai. In this paper, we present the design and architecture of ANSAN.

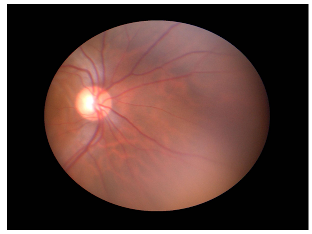

Fig. 1: Healthy Eye

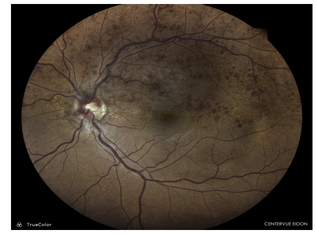

Fig. 2: Glaucoma Eye

II. Literature Survey

[11] states that a six-layer CNN model incorporating dropout mechanisms effectively differentiates between glaucomatous and non-glaucomatous patterns in eye images. This model was tested on the ORIGA and SCES datasets, achieving Area Under the Curve (AUC) values of 0.822 and 0.882, respectively, outperforming several state-of-the-art methods.

[12] evaluates traditional image processing techniques alongside machine learning and deep learning methods. The study found that while traditional methods like K-Nearest Neighbor (KNN) achieved an accuracy of 98%, the VGG-16 deep learning model outperformed these methods with a remarkable accuracy of 99.6% on the Bin Rushed dataset, underscoring the potential of deep learning in medical image classification.

Additionally, [17] presents DenseNet201, a model with 201 layers pre-trained on ImageNet, which was applied to the ACRIMA dataset. DenseNet201 achieved a high accuracy of 97% and an F1-score of 0.969, further solidifying the efficacy of deep learning models in glaucoma detection.

Hybrid models that combine various deep learning techniques have been explored to address data scarcity and improve detection accuracy. [13] introduces CAPSGAN, a hybrid model combining a Generative Adversarial Network (GAN) for synthetic image generation with a modified Capsule Network (CAPSNET) for classification. This approach preserved crucial spatial and orientational details in images, outperforming traditional CNN models.

Similarly, [20] proposes CDED-Net, a model focused on the joint segmentation of the optic disc (OD) and optic cup (OC) in fundus images. CDED-Net eliminates pre- and post-processing steps, achieving state-of-the-art results on the DRISHTI-GS, RIM-ONE, and REFUGE datasets. The model's efficiency in OD and OC segmentation highlights its potential for glaucoma screening.

Transfer learning has proven instrumental in enhancing glaucoma detection models by leveraging pre-trained networks. [14] explores this approach by employing ResNet-152, which achieved the highest accuracy of 86.9% on the Large-scale Attention-based Glaucoma (LAG) dataset, demonstrating the effectiveness of transfer learning in improving model performance.

[21] introduces the Xception model, which achieved superior performance with a training accuracy of 97.63% and a validation accuracy of 98.11%. The study concluded that transfer learning models like Xception offer significant advantages in medical image classification and suggested further exploration of additional datasets and model improvements.

Although deep learning models have demonstrated superior performance, traditional image processing techniques remain relevant, especially in resource-constrained environments. [15]implements digital image processing techniques in MATLAB, achieving a high accuracy of 94.61% on the ACRIMA database, thereby demonstrating the feasibility of using MATLAB for clinical applications in glaucoma detection.

Moreover, [16] introduces a 2-D Compact Variational Mode Decomposition (2-D-C-VMD) algorithm for classifying glaucoma stages. The method achieved a classification accuracy of 98.11% with tenfold cross-validation, highlighting its effectiveness in early-stage glaucoma detection.

Emerging techniques such as Self-Organized Operational Neural Networks (Self-ONNs) and novel segmentation and fusion methods have shown promise in glaucoma detection. [18] compares Self-ONNs with conventional deep CNNs across multiple datasets, demonstrating superior detection performance and reduced computational complexity, making Self-ONNs a promising option in scenarios with limited data and high computational demands.

In addition, [19] introduces a segmentation and fusion technique utilizing green channel fundus images. The proposed algorithm effectively detected retinal veins and exudates, making it suitable for rapid processing and clinical use.

This literature survey highlights the rapid advancements in glaucoma detection through machine learning and deep learning techniques. Deep learning models, particularly CNNs and hybrid models, have shown significant potential in improving diagnostic accuracy. Transfer learning approaches have further enhanced model performance, while traditional image processing techniques remain valuable in specific contexts. Emerging techniques like Self-ONNs and advanced segmentation methods offer promising directions for future research, pushing the boundaries of glaucoma detection.

III. Methodology

The methodology we propose is, consider a scenario where a medical camp for glaucoma awareness is conducted by a hospital in the underprivileged areas by the hospital that uses ANSAN. The frontline workers will use our ANSAN app to register the probable patient to the app, get the medical history of the patient as well as possible. After that, two images of the left eye and two images of the right eye are taken using the handheld fundus camera connected to the mobile device and the images are uploaded to the cloud. The deep learning model processes the image of the eyes and predicts the chance of the presence of glaucoma, prediction is recorded and then

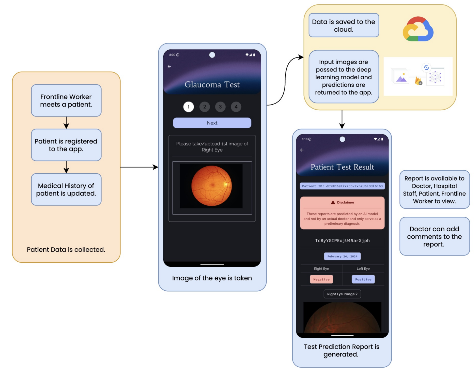

Fig. 3: High-Level Overview of the Methodology proposed

a report is generated which will be immediately accessible to doctors, patients (if they have an email ID or a phone number) and everyone who has access to it. Based on the prediction, the frontline worker can redirect the patient to the doctor and inform them of the patient ID which they will have to tell the doctor so that the doctor can also view the report. This way we increase the pace at which we can spread awareness of glaucoma and prevent it by pre-screening which will predict the case accurately. Another point is that if the patient is technically literate, they can use the app to pre-screen themselves.

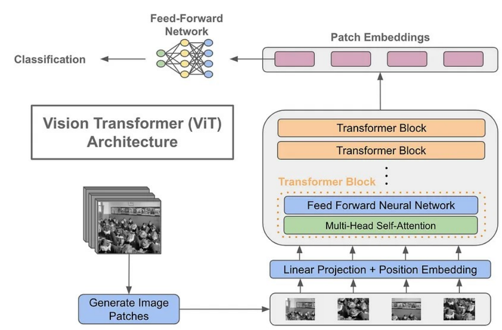

Fig. 4: Workflow of Vision Transformer Deep Learning Model [22]

The Vision Transformer (ViT) represents a paradigm shift in image analysis by adapting the transformer architecture, which was originally designed for sequential data in natural language processing, to the domain of computer vision. Instead of relying on convolutional operations to extract local features, the ViT divides an image into fixed-size patches and treats each patch as an individual token. These tokens are then embedded linearly, enriched with positional encodings, and processed through multiple transformer layers. This architecture allows the model to capture global dependencies across the entire image from the very first layer, enabling it to learn complex relationships between distant parts of the image.

One of the distinguishing characteristics of the ViT is its reliance on large-scale datasets for effective training. Unlike convolutional neural networks (CNNs), which inherently incorporate inductive biases like locality and translation invariance, the ViT operates without these biases, making it more flexible but also more dependent on extensive data to achieve comparable performance. When trained on sufficient data, the ViT can surpass traditional CNNs in tasks such as image classification, demonstrating the effectiveness of the self-attention mechanism in capturing rich contextual information across the entire image.

A. Design

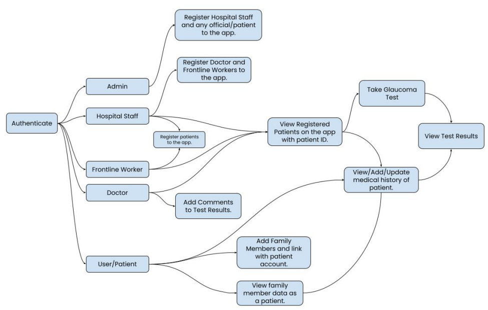

Fig. 5: Workflow of Application

We will start by explaining the roles and use cases that we have identified for ANSAN and then move on in explaining the implementation architecture.

B. Use Cases

Users of the Application(Actors):

1) Administrator
2) Hospital Staff
3) Doctor
4) Front-line Worker
5) Patient with access to technology
6) Patient without access to technology

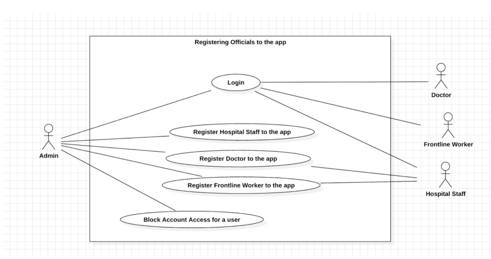

Fig. 6: Use Case 1 (Registering officials to the app)

Admin logs in to the app using his credentials. Admin can register hospital staff, doctors, and frontline workers to the app. Admin can block account access of a user. This feature is there to restrict access to users with bad intentions or prevent fraudulent activities.

Hospital Staff logs in and can register new doctors and frontline workers to the app. When a user is registered to the app, login credentials are mailed to the user's registered mail ID. On first time login, the user will have to undergo a two-step email-OTP based login process.

Patients with access to technology can self-register to the app. Those who do not have access to technology, can take part in camps or visit hospitals directly, and then either frontline workers, hospital staff, or doctors can register the patient to the app.

Frontline workers, hospital staff and doctors can view a patient's bio data by entering the unique Patient ID assigned to the patient which will be available in the patient's profile. Doctors can make use of the family members data of the patient to deduce any genetic reasons for the presence of Glaucoma in the patient.

Before the fundus images can be captured by a [10] handheld fundus camera or uploaded, medical history questionnaires can be completed by the officials or can also be self-completed by the patient. Officials can capture the fundus images of the left and right eyes via a [10] handheld fundus camera or can upload respective images and then a glaucoma diagnosis report is generated by the app that has the result of the pre-screening.

Officials and patients can view the glaucoma diagnosis report of the patient. Doctors can add comments to the diagnosis report. Patients can be contacted by the hospital based on the final doctor's diagnosis opinion. Officials and patients can update their respective medical history data.

One point to note here is admin has access to every single use case mentioned above.

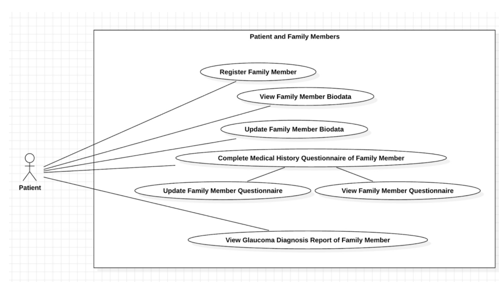

Fig. 7: Use Case 2(Patients and their family members)

Patients can register their family members and link to the patient's account. They can view a family member's biodata, update a family member's biodata, complete the medical history questionnaires of a family member, view the medical history questionnaires of a family member, and update the medical history questionnaires of a family member. They can view glaucoma diagnosis reports of a family member.

C. Architecture

We propose two architectures. One is with the data located inside hospital servers and the other is data stored and operated through Google Cloud. Certain hospitals might not want their data to be stored in the Cloud and might want a solution that can store this data in their servers. This is the main reason for the two designs being proposed.

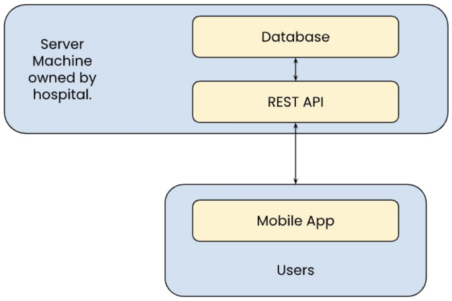

Fig. 8: Case 1: Data in local Hospital Servers.

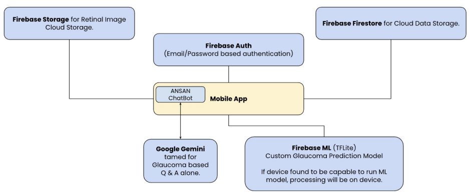

Fig. 9: Case 2: Data Stored in Google Cloud with Firebase.

TABLE I: Comparison of Data Storage Options

Feature	**Data in Local Hospital Owned Servers**	**Data Stored in Google Cloud with Firebase**
Data Storage	Relational database model (RDBMS)	NoSQL model
Software Security	Ensured via token-based authentication with libraries such as paseto package in Node.js.	Firebase Auth library provides guaranteed security.
Hardware Security	Requires hospitals to implement security measures for hardware where the data is stored.	Relies on Google Cloud's security measures which are guaranteed to be secure and easy to maintain.
Cost	Purchasing hardware for servers will require significant	Pay-as-you-go pricing leading to less dependency on hardware.

The mobile app is developed using Flutter, a cross-platform framework, ensuring compatibility and seamless functionality across both Android and iOS devices and can be extended to Web, Desktop, Windows, Linux and Mac OS users too easily with a single code-base. This enables broader accessibility for users, regardless of their preferred mobile operating system.

IV. Parameter Settings

V. Dataset Description

In our study, we used images merging two different datasets:

1) **AKSHI** - The dataset is collected from the Sathyan Eye Care and Glaucoma Foundation, Coimbatore.
2) **CHAKSU** - A benchmark dataset consisting of over 1300 images and is the largest Indian-ethnicity-specific fundus image database.

TABLE II: Hyperparameters and Values

Hyperparameter	**Value**
Number of Classes	2
Batch Size	32
Type of Normalization	Batch
Type of Activation	ReLU
Optimizer	Adam
Learning Rate	0.001
Patience	10
Epochs	100
Image Input Size	(150, 150, 3)
Loss Function	Binary Crossentropy
Dropout Rate	0.5

The fundus camera used to capture the images records various parts of the eye such as the retina, optic nerve head, macula, retinal blood vessels, choroid, and vitreous. The dataset was split into two different ratios (70-15-15 and 80-10-10) for training, testing, and validation to examine the effect of training ratios on the results produced by the deep learning model.

TABLE III: Class Distribution

Class Name	**Number of Images**
Healthy	1098
Glaucoma	536

VI. Results

Eight deep learning models—CNN, RNN, CNN-LSTM, ResNet, ViT, DNN+VAE, LeNet, and YOLOv7—were implemented for comparison to identify the best model for glaucoma detection.

These models were built using tools like NumPy for computations, Matplotlib for visualization, and scikit-learn for preprocessing and evaluation. TensorFlow and its high-level API Keras were used for building and training deep learning models while PyTorch (with Torchvision) offered flexibility in handling neural networks, particularly for computer vision applications. OpenCV-Python was used for image and video processing, crucial for models dealing with visual data. Additionally, labelImg was employed for annotating images, enabling the creation of labeled datasets. Development and experimentation were conducted in VS Code, while Google Colab was used for its GPU support, ensuring efficient training and fast model iteration.

We evaluated all the eight deep-learning models using two different dataset split ratio: 70-15-15 and 80-10-10.The results are summarized below.

TABLE IV: Model Performance (70-15-15 Split)

Metric	**CNN**	**RNN**	**CNN-LSTM**	**ResNet**	**ViT**	**DNN+VAE**	**LeNet**	**YOLOv7**
Accuracy	0.88	0.90	0.88	0.89	**0.90**	0.82	0.86	0.85
Precision	0.89	0.89	0.87	0.95	**0.97**	0.89	0.87	0.87
Recall	0.98	0.96	0.96	0.96	**0.98**	0.91	0.93	0.73
F1-Score	0.90	0.93	0.91	0.92	**0.93**	0.90	0.90	0.78
AUC-ROC	0.88	0.90	0.83	0.93	**0.95**	0.87	0.89	0.82
Cohen's Kappa	0.66	**0.77**	0.71	0.73	0.75	0.70	0.69	0.62
F2-Score	0.94	0.95	0.94	0.94	**0.96**	0.91	0.92	0.84
Diagnostic odds ratio	113.1	112.8	66.82	75.71	**273.33**	37.75	42.45	31.62

The results of the model evaluation clearly indicate that the Vision Transformer (ViT) outperforms the other deep-learning

TABLE V: Model Performance (80-10-10 Split)

Metric	**CNN**	**RNN**	**CNN-LSTM**	**ResNet**	**ViT**	**DNN+VAE**	**LeNet**	**YOLOv7**
Accuracy	0.87	0.89	0.85	0.85	**0.93**	0.90	0.87	0.85
Precision	0.93	0.90	0.84	0.89	**0.95**	0.90	0.89	0.87
Recall	0.97	0.93	0.96	0.98	**0.98**	0.95	0.92	0.90
F1-Score	0.90	0.91	0.89	0.90	**0.94**	0.92	0.91	0.88
AUC-ROC	0.91	0.88	0.79	0.86	**0.95**	0.88	0.88	0.87
Cohen's Kappa	0.63	0.74	0.64	0.65	**0.81**	0.77	0.71	0.62
F2-Score	0.94	0.92	0.93	0.94	**0.97**	0.94	0.92	0.92
Diagnostic odds ratio	56.04	57.5	45.05	84.85	**343.76**	82.09	44.62	40.06

models across nearly all metrics, making it the most effective model for glaucoma detection. In both dataset splits—70-15-15 and 80-10-10—ViT consistently achieves the highest scores in critical performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC, which are essential for evaluating the reliability and robustness of a medical diagnostic system.

In terms of accuracy, ViT maintains a performance of 90

The F1-scores for ViT (0.93 and 0.94) show that it strikes a strong balance between precision and recall, making it well-suited for situations where both false positives and false negatives need to be minimized. The AUC-ROC scores (0.95 in both splits) further demonstrate ViT's ability to differentiate between glaucoma and non-glaucoma cases effectively, even across varying classification thresholds. Additionally, ViT's Cohen's Kappa score (0.75 and 0.81) suggests a substantial agreement between the model's predictions and the true labels, confirming the reliability of its results.

Another critical metric, the Diagnostic Odds Ratio (DOR), reveals ViT's significant superiority over the other models. In both splits, ViT's DOR (273.33 in the 70-15-15 split and 343.76 in the 80-10-10 split) is much higher than the other models, indicating that it is substantially better at distinguishing between true positive and true negative cases.

Although other models like RNN and DNN+VAE show reasonably strong performance, they do not match ViT's consistent excellence across all metrics. For instance, RNN performs well in the 70-15-15 split with high accuracy and recall, but it falls short of ViT in precision and diagnostic effectiveness. Similarly, while DNN+VAE offers a competitive performance, it lacks the diagnostic robustness that ViT demonstrates, especially in the larger dataset split of 80-10-10.

In summary, ViT stands out as the most reliable and effective model for glaucoma detection, demonstrating superior accuracy, precision, and diagnostic power in both dataset configurations. Its ability to balance between identifying true positives while minimizing false positives and negatives makes it the optimal choice for real-world deployment in medical diagnostics. This strong performance, particularly in critical medical metrics, highlights its potential for broader applications in disease detection and healthcare, where accuracy and reliability are of utmost importance.

The superior performance of the Vision Transformer (ViT) compared to other models can be attributed to its unique architecture, which is well-suited for capturing complex patterns in image data. Unlike traditional convolutional neural networks (CNNs) that rely on localized feature extraction, ViT leverages the self-attention mechanism, allowing it to analyze the entire image globally from the outset. This enables the model to capture long-range dependencies and intricate relationships between different parts of the image, which are crucial for identifying subtle indicators of glaucoma in medical imaging. Additionally, the transformer architecture's flexibility in processing high-dimensional data without the inherent biases of convolutional operations makes it more adaptable, especially when trained on large datasets. ViT's reliance on positional encodings also helps it retain spatial information, enhancing its ability to differentiate between healthy and glaucomatous eyes more effectively than CNNs, RNNs, or other hybrid models. This global attention mechanism, combined with ViT's ability to scale efficiently with data, explains its outperformance in both precision and recall, as well as its overall diagnostic effectiveness compared to more conventional deep learning models.

The findings from this study suggest that the Vision Transformer (ViT) model, due to its superior performance in detecting glaucoma, could have significant real-world applications, particularly in early diagnosis and screening in healthcare settings. In practical terms, this model can be integrated into telemedicine platforms or mobile healthcare applications, like ANSAN, to facilitate mass screening, especially in underprivileged areas where access to ophthalmologists is limited. The high accuracy, precision, and recall demonstrated by ViT mean that it can reliably detect glaucoma at an early stage, reducing the likelihood of both false positives and false negatives. This can streamline the referral process, ensuring that only patients who truly need further diagnostic procedures are directed to specialists, thus optimizing healthcare resources.

For future research, the success of ViT in this study opens new avenues for the application of transformer-based models in other medical imaging tasks beyond glaucoma, such as detecting diabetic retinopathy, macular degeneration, or other retinal diseases. Furthermore, the adaptability of ViT to various datasets suggests that with larger and more diverse training data, its performance could improve even further. One possible direction for future research could involve exploring hybrid models that combine ViT's global attention mechanism with CNN's localized feature extraction to leverage the strengths of both architectures. Additionally, deploying this system in real-time clinical environments would allow researchers to test its efficacy under varied conditions, such as different image qualities, lighting conditions, and patient demographics. Future studies could also investigate how transfer learning with pre-trained ViT models might further enhance performance across different medical imaging tasks, offering broader applications in healthcare.

VII. Analysis and Conclusion

The Vision Transformer (ViT) consistently outperformed other models across various performance metrics in both data splits. Therefore, we implemented the ViT model in the backend of our mobile application to detect the presence or absence of glaucoma.

VIII. Conclusion

The development of ANSAN is a step forward in leveraging technology to prevent blindness due to glaucoma. By enabling early detection and facilitating a strong connection between patients and healthcare providers, ANSAN has the potential to make a significant impact on public health. The scalability of the app ensures it can be implemented widely, especially in regions with limited access to advanced healthcare facilities.

IX. Acknowledgement

The authors wish to express their gratitude to Prof. Senthil Kumar Thangavel, in the Department of CSE at Amrita School of Computing, Coimbatore for his invaluable mentorship, and Dr. Sathyan Parthasarathi, Director of Sathyan Eye Hospital and Coimbatore Glaucoma Foundation, Ramanathapuram, Coimbatore, as well as to Dr. P. Ram Gopal from Medivista Eye Centre at Secunderabad, for their support.

TABLE VI: Author Contributions

Name of the Author	Contribution
Ashwin Narayanan S	Mobile Application & Server Architecture
Viswa A V	Dataset Collection & Deep Learning Model
Ashwin R Sharma	Deep Learning Model
Harinandan N	Deep Learning Model
Murugesh JLP	Deep Learning Model
K Vaisakhkrishnan	Server Architecture
Selvanayaki Kolandapalayam Shanmugam	Expert Opinion in DL Architectures
Somasundaram K	Opinion on Mathematical Aspects of DL Models

References

[1] World Health Organization (WHO). “Blindness and vision impairment”, *https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment*, accessed May 5, 2024.

[2] F. Guo *et al.*, ”Yanbao: A Mobile App Using the Measurement of Clinical Parameters for Glaucoma Screening,” *IEEE Access*, vol. 6, pp. 77414-77428, 2018, doi: 10.1109/ACCESS.2018.2882946.

[3] Google Solutions Challenge, 2024 *https://developers.google.com/community/gdsc-solution-challenge*, accessed May 5, 2024.

[4] Github Repository of Relational Model of app, *https://github.com/Ashrockzzz2003/ansan_v2*, accessed May 5, 2024.

[5] Github Repository of NoSQL Model of app, *https://github.com/Ashrockzzz2003/ansan_v3*, accessed May 5, 2024.

[6] D. L. Demmin and S. M. Silverstein, ”Visual Impairment and Mental Health: Unmet Needs and Treatment Options,” *Clin Ophthalmol.*, vol. 14, pp. 4229-4251, Dec. 2020, doi: 10.2147/OPTH.S258783.

[7] C. W. McMonnies, ”Glaucoma history and risk factors,” *Journal of Optometry*, vol. 10, no. 2, pp. 71-78, 2017, doi: 10.1016/j.optom.2016.02.003.

[8] National Eye Institute Article, *https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/glaucoma/types-glaucoma*, accessed May 5, 2024.

[9] Vasan, C. S., Gupta, S., Shekhar, M., Nagu, K., Balakrishnan, L., Ravindran, R. D., ... Subburaman, G. B. B. (2023). Accuracy of an artificial intelligence-based mobile application for detecting cataracts: Results from a field study. Indian Journal of Ophthalmology, 71(8), 2984-2989.

[10] L. Lu *et al.*, ”Diagnostic accuracy of handheld fundus photography: A comparative study of three commercially available cameras,” *PLOS Digit Health.*, vol. 1, no. 11, p. e0000131, Nov. 2022, doi: 10.1371/journal.pdig.0000131.

[11] A. Saxena, A. Vyas, L. Parashar, and U. Singh, ”A glaucoma detection using convolutional neural network,” in *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2020, pp. 815-820, doi: 10.1109/ICESC48915.2020.9155913.

[12] A. Pandey, P. Patre, and J. Minj, ”Detection of glaucoma disease using image processing, soft computing and deep learning approaches,” in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2020, pp. 1-7, doi: 10.1109/I-SMAC49090.2020.9243362.

[13] J. Manikandan and R. Yogeshkumar, ”Glaucoma Disease Detection Using Hybrid Deep Learning Model,” in *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, 2023, pp. 1-7, doi: 10.1109/ICONSTEM56992.2023.10179261.

[14] A. Sallam *et al.*, ”Early detection of glaucoma using transfer learning from pre-trained cnn models,” in *2021 International Conference of Technology, Science and Administration (ICTSA)*, 2021, pp. 1-5, doi: 10.1109/ICTSA52272.2021.9478513.

[15] D. Almeida-Galárraga *et al.*, ”Glaucoma detection through digital processing from fundus images using MATLAB,” in *2021 Second International Conference on Information Systems and Software Technologies (ICI2ST)*, 2021, pp. 39-45, doi: 10.1109/ICI2ST53863.2021.9538831.

[16] D. Parashar and D. Agrawal, ”2-D compact variational mode decomposition-based automatic classification of glaucoma stages from fundus images,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-10, 2021, doi: 10.1109/TIM.2021.3086883.

[17] S. Ovreiu, E. A. Paraschiv, and E. Ovreiu, ”Deep learning digital fundus images: Glaucoma detection using DenseNet,” in *2021 13th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, 2021, pp. 1-4, doi: 10.1109/ECAI52376.2021.9515031.

[18] O. C. Devecioglu *et al.*, ”Real-time glaucoma detection from digital fundus images using self-onns,” *IEEE Access*, vol. 9, pp. 140031-140041, 2021, doi: 10.1109/ACCESS.2021.3119309.

[19] A. R. Shaik, K. R. Chandra, B. E. Raju, and P. R. Budumuru, ”Glaucoma identification based on segmentation and fusion techniques,” in *2021 International Conference on Advances in Computing, Communication, and Control (ICAC3)*, 2021, pp. 1-4, doi: 10.1109/ICAC353642.2021.9740859.

[20] M. Tabassum *et al.*, ”CDED-Net: Joint segmentation of optic disc and optic cup for glaucoma screening,” *IEEE Access*, vol. 8, pp. 102733-102747, 2020, doi: 10.1109/ACCESS.2020.2996988.

[21] N. S. Shadin *et al.*, ”Performance analysis of glaucoma detection using deep learning models,” in *2022 International Conference on Innovations in Science, Engineering and Technology (ICISET)*, 2022, pp. 190-195, doi: 10.1109/ICISET54393.2022.9844237.

[22] https://towardsdatascience.com/using-transformers-for-computer-vision-6f764c5a078b

[23] Viswa, A. V., Pravardhana, V. D., Thangavel, S. K., & Jeyakumar, G. (2024, May). Fall Detection for Elderly People using LiDAR Sensor. In 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT) (pp. 1-6). IEEE.

[24] A. Baskar and T. Gireesh Kumar, “Facial Expression Classification Using Machine Learning Approach: A Review,” Advances in intelligent systems and computing, pp. 337–345, Jan. 2018, doi:https://doi.org/10.1007/978-981-10-3223-3_32.

[25] S. Sruthy, L. Parameswaran and A. P. Sasi, ”Image fusion technique using DT-CWT,” 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), Kottayam, India, 2013, pp. 160-164, doi: 10.1109/iMac4s.2013.6526400.

[26] R. Karthika and L. Parameswaran, “Study of Gabor Wavelet for Face Recognition Invariant to Pose and Orientation,” Advances in intelligent systems and computing, pp. 501–509, Dec. 2015, doi:https://doi.org/10.1007/978-81-322-2671-0_48.