Segmentation of Optic Disc and Cup for Glaucoma Detection using U-Net and Bayesian Network U-Net Models

Mahankali Prathyusha Lahari

November,2024

Contents

Abstract

Glaucoma is one of the most prevalent causes of irreversible blindness, necessitating early detection and treatment. This project applies deep learning models, specifically U-Net and Bayesian Network U-Net (BNN U-Net), to segment the optic disc and cup in retinal images from the REFUGE2 dataset, which can aid in glaucoma diagnosis. U-Net is a widely used architecture in medical image segmentation, known for its effectiveness and simplicity. BNN U-Net enhances U-Net by integrating Bayesian inference to capture predictive uncertainty, providing a measure of model confidence. Both models were evaluated based on accuracy, loss, and DICE coefficient, and the results highlight the BNN U-Net's advantage in interpretability and confidence. This study underscores the value of uncertainty-aware segmentation models for clinical use in glaucoma detection.

Chapter 1

Introduction

1.1 Motivation

Glaucoma is the leading cause of irreversible blindness worldwide.The increasing prevalence of glaucoma necessitates reliable diagnostic tools for early detection. Often asymptomatic for years, this disease can progress significantly before patients become aware of the loss of visual function. Critical examination of the optic nerve through ophthalmoscopy or using fundus images is a crucial component of glaucoma detection before the onset of vision loss. The vertical cup-to-disc ratio (VCDR) is a key structural indicator for glaucoma, as thinning of the superior and inferior neuro-retinal rim is a hallmark of the disease. However, manual assessment of fundus images is both time-consuming and subject to variability based on clinician expertise and interpretation. In this study, we develop a robust and accurate automated system employing Image processing techniques for segmentation and classification and infer the result through Bayesian Network model.

Primary open-angle glaucoma (POAG) is a progressive condition where tissue loss at the neuro-retinal edge of the optic disc leads to an increase in the optic cup's size. This results in a larger cup-to-disc ratio (CDR), which is a key indicator in diagnosing glaucoma. Variations in CDR can be influenced by individual differences and other eye conditions. There's the importance of optic disc (OD) and cup-to-disc ratio (CDR) in diagnosing glaucoma. Glaucoma causes the optic cup to enlarge due to tissue loss in the neuro-retinal rim, altering the CDR.

Traditional methods for optic disc and cup segmentation are often manual and time-consuming, leading to potential inaccuracies. Deep learning models, particularly U-Net, have shown promise in automating this process. By comparing U-Net with BNN U-Net, which leverages uncertainty quantification, this study aims to enhance segmentation accuracy and reliability.

1.2 Objectives

- **Develop a U-Net model** for segmenting the optic disc and cup in retinal images.
- **Implement a Bayesian Network U-Net model** to evaluate the impact of uncertainty on segmentation performance.

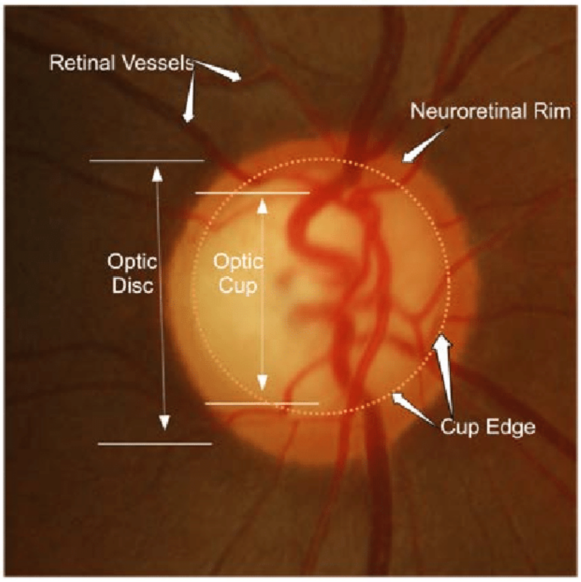

Figure 1.1: Anatomical Structure of Optic Disc.

- **Compare the results** of both models to identify strengths and weaknesses in segmentation accuracy.

1.3 Problem Definition

The primary challenge addressed in this study is the accurate segmentation of the optic disc and cup from retinal images(neuro retinal-rim width) to assist in glaucoma diagnosis. This involves designing effective model architectures and evaluating their performance based on established metrics.

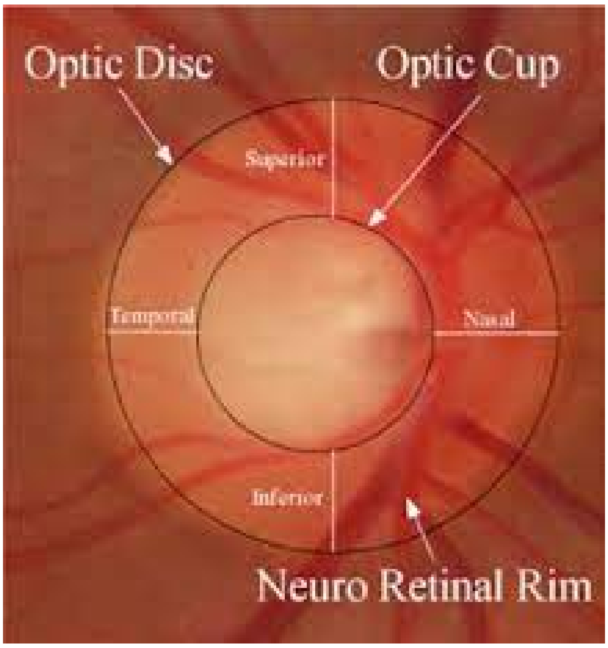

Figure 1.2: Neuro Retinal Rim Width.

Chapter 2

Literature Review

Recent years have seen significant advances in deep learning for medical image analysis, particularly with convolutional neural networks (CNNs) and their derivatives. U-Net, introduced by Ronneberger et al., has become the gold standard for image segmentation in the medical field due to its encoder-decoder structure with skip connections. This structure allows U-Net to retain high-resolution features, essential for precise segmentation of small structures in biomedical images. Bayesian deep learning, advanced by Blundell et al., introduces probabilistic layers that can model prediction uncertainty. In clinical diagnostics, models with uncertainty quantification have shown to be beneficial, as they allow clinicians to evaluate the confidence of automated predictions. Studies combining U-Net with Bayesian approaches have shown promise in improving robustness and reliability in various medical imaging applications, motivating the use of BNN U-Net in this project.

There have been many models implemented for glaucoma diagnosis using convolutional neural networks (CNNs), transfer learning, segmentation-based methods, feature extraction-based methods, and hybrid approaches. Results showed high accuracy in glaucoma diagnosis, ranging from 88-99%. Specifically, CNNs achieved 90-98% accuracy, transfer learning reached 92-97%, and segmentation-based methods attained 95-98%. Hybrid approaches demonstrated superior performance. However, challenges persist, including class imbalance and variability in imaging protocols. Future research should address these limitations, explore new architectures, and integrate clinical data to enhance diagnostic accuracy. Notable studies include Fu et al. (2020) achieving 96.4% accuracy with CNNs, Li et al. (2021) reporting 94.5% accuracy with transfer learning, and Chen et al. (2020) demonstrating 95.2% accuracy with U-Net.

Sr. No.	Authors	Year	Model	Datasets	Results
1	Yu et al.	2019	Pre-trained U-Net, ResNet	RIGA, DRISHTI-GS, RIM-ONE	Dice 97.38% (Disc), Dice 88.77% (Cup)
4	Liao et al.	2019	ResNet	ORIGA	Accuracy 0.88
5	Serte et al.	2019	Ensemble ResNet-50	HRF, DRISHTI-GS1, RI-MONE	Accuracy 53%, AUC 83%
6	Juneja et al.	2019	U-Net	DRISHTI-GS	Accuracy 95.8% (OD), 93.0% (OC)
8	Thakoor et al.	2019	Pre-trained CNN	Local dataset of 737 images	Accuracy 96.27%
9	Maheshwari et al.	2020	AlexNet	RIM-ONE	Accuracy 98.90%, Specificity 97.50%
10	Lima et al.	2020	CNN	RIM-ONE r3	Accuracy 91%
14	Elangovan and Nath	2020	CNN	Multiple datasets	Accuracy 96.64%
18	Chaudhary and Pachori	2021	Ensemble ResNet Models	RIM-ONE, ORIGA	Accuracy 91.1%, specificity 94.3%
21	Veena et al.	2022	CNN	DRISHTI-GS	Accuracy 98% (OD)

Table 2.1: Literature Review Summary

Chapter 3

Methodology

3.1 Dataset

The REFUGE2 dataset was used, containing retinal images labeled with masks for the optic disc and cup regions. The dataset was preprocessed by resizing each image to 256x256 pixels, which is a common resolution for CNNs that balances computational efficiency with sufficient detail for segmentation tasks. The dataset was split into training and validation sets, with 80% for training and 20% for validation to ensure robust model assessment.

3.2 Model Architectures

3.2.1 U-Net

The U-Net architecture is widely used for image segmentation tasks, especially in medical imaging, due to its ability to precisely localize and segment complex structures.,The U-Net model follows an encoder-decoder structure.Here's an in-depth look at each component of U-Net's architecture and how it contributes to the model's effectiveness in segmentation tasks:

- **Encoder:** Uses a series of convolutional layers and max-pooling to downsample the image while extracting features.

- **Decoder:** Uses upsampling layers to reconstruct the original resolution, with skip connections that pass high-resolution features from the encoder to corresponding layers in the decoder. This helps retain spatial information, enhancing segmentation accuracy.

- **Dropout Layers:** To reduce overfitting, dropout layers are included in the encoder, particularly effective given the dataset's limited size.

3.2.2 Bayesian Network U-Net (BNN U-Net)

Bayesian Network can help in classification by modeling the probabilistic relationships between features (e.g., cup-to-disc ratio, optic disc size, neuro-retinal rim width) and the target condition (e.g., glaucoma or non-glaucoma). It uses the conditional probabilities of each feature contributing to the final classification. After training, the network can classify new cases by calculating the probability that a given set of features corresponds to a certain class (glaucoma or normal), making it useful for diagnosis in medical image analysis.we can apply a Bayesian Network in the fundus image processing pipeline by incorporating it into the diagnosis and decision-making phase after the image segmentation.

The Bayesian Network helps in making probabilistic inferences based on the segmented image features, improving decision-making by incorporating multiple risk factors. The BNN U-Net extends U-Net by adding Monte Carlo Dropout during inference. This means dropout layers remain active during predictions, enabling multiple forward passes on the same input. By averaging these outputs, the model can generate an uncertainty estimate, indicating confidence in its predictions. This architecture is advantageous for clinical applications as it identifies areas where the model is less certain, potentially guiding further review by medical professionals

3.3 Training Process

Both models were trained using the Adam optimizer and binary cross-entropy loss function. A batch size of 16 and a total of 10 epochs were used for training, with a validation split of 20%. Data augmentation techniques, such as rotation and flipping, were applied to enhance model generalization.Data augmentation techniques, such as random rotations and horizontal flipping, were applied to increase the diversity of training samples, thus enhancing model robustness

3.4 Evaluation Metrics

The models were assessed using the following metrics:

- **Accuracy:** Measures the proportion of correct predictions, indicating overall performance.
- **Loss:** Evaluates the model's optimization; lower loss signifies better fit to the data.
- **DICE Coefficient:** Measures overlap between the predicted and ground truth masks, with a higher score indicating better segmentation quality. This metric is particularly suited to segmentation tasks in medical imaging.

3.5 Mathematical Representation for Segmentation by Bayesian U-Net Framework

For retinal fundus image segmentation for glaucoma detection, employing a Bayesian U-Net model allows for both precise segmentation and quantification of uncertainty.

Objective :

- **Input:** Retinal fundus image I.
- **Output:** Segmentation map S indicating optic disc and cup regions.

3.5.1 Segmentation Mask Representation

The segmentation mask S for a retinal fundus image I is defined as the union of the optic disc and cup regions:

$$S = S_{\text{disc}} \cup S_{\text{cup}}$$

Where:

- S_{disc}: Binary segmentation mask for the optic disc.
- S_{cup}: Binary segmentation mask for the optic cup.

For a pixel i in the image:

$$S_{\text{disc}}(i) = \begin{cases} 1 & \text{if pixel } i \text{ belongs to the optic disc} \\ 0 & \text{otherwise} \end{cases}$$

$$S_{\text{cup}}(i) = \begin{cases} 1 & \text{if pixel } i \text{ belongs to the optic cup} \\ 0 & \text{otherwise} \end{cases}$$

3.5.2 Cup-to-Disc Ratio (CDR)

The Cup-to-Disc Ratio (CDR) is a key metric for glaucoma detection. It is defined as the ratio of the optic cup area to the optic disc area. Let A_{cup} and A_{disc} denote the areas of the optic cup and optic disc, respectively, calculated from the segmentation mask S:

$$A_{\text{cup}} = \sum_i S_{\text{cup}}(i), \quad A_{\text{disc}} = \sum_i S_{\text{disc}}(i)$$

Then, the CDR is given by:

$$\text{CDR} = \frac{A_{\text{cup}}}{A_{\text{disc}}}$$

Where:

- A_{cup}: Number of pixels classified as the optic cup.
- A_{disc}: Number of pixels classified as the optic disc.

3.5.3 CDR Embedded in Segmentation Representation

The segmentation S can be expressed with the CDR as an additional feature:

$$S = \{S_{\text{disc}}, S_{\text{cup}}, \text{CDR}\}$$

This includes:

1. S_{disc}: Binary mask for the optic disc region.
2. S_{cup}: Binary mask for the optic cup region.
3. **CDR**: Ratio of the cup area to the disc area, derived from S_{cup} and S_{disc}.

3.5.4 Bayesian Segmentation Formulation

The Bayesian formulation for segmentation includes uncertainty in the prediction of S. The posterior probability of S, given the retinal fundus image I and training data $\mathcal{D}$, is:

$$P(S|I, \mathcal{D}) = \int P(S|I, \theta)P(\theta|\mathcal{D})\, d\theta$$

Where:

- $P(S|I, \theta)$: Likelihood of the segmentation S, given the image I and model parameters θ.
- $P(\theta|\mathcal{D})$: Posterior distribution of the model parameters θ, given the training data $\mathcal{D}$.

3.5.5 Monte Carlo Approximation

The integral above can be approximated using Monte Carlo Dropout:

$$P(S|I, \mathcal{D}) \approx \frac{1}{T}\sum_{t=1}^{T} P(S|I, \theta_t)$$

Where:

- T: Number of stochastic forward passes with dropout enabled.
- θ_t: Model parameters sampled during the t-th pass.

3.5.6 Loss Function

A common choice for the loss function in segmentation tasks is the Dice loss, which measures the overlap between the predicted segmentation and the ground truth:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2\sum_i p_i g_i}{\sum_i p_i + \sum_i g_i}$$

Where:

- p_i: Predicted probability for pixel i.
- g_i: Ground truth label for pixel i.

3.5.7 Uncertainty Quantification

The variance across the T predictions provides an estimate of the model's uncertainty at each pixel:

$$\text{Uncertainty}(i) = \frac{1}{T} \sum_{t=1}^{T} (p_i^{(t)} - \bar{p}_i)^2$$

Where:

- $p_i^{(t)}$: Predicted probability for pixel i at the t-th pass.
- $\bar{p}_i$: Mean predicted probability for pixel i over all passes.

The Bayesian U-Net framework not only provides accurate segmentations of the optic disc and cup but also highlights regions of uncertainty, enhancing clinical reliability in glaucoma diagnosis.

Chapter 4

Results

4.1 U-Net Results

The U-Net model achieved an accuracy of 98.26% on the validation set, with a corresponding loss of 0.0456.

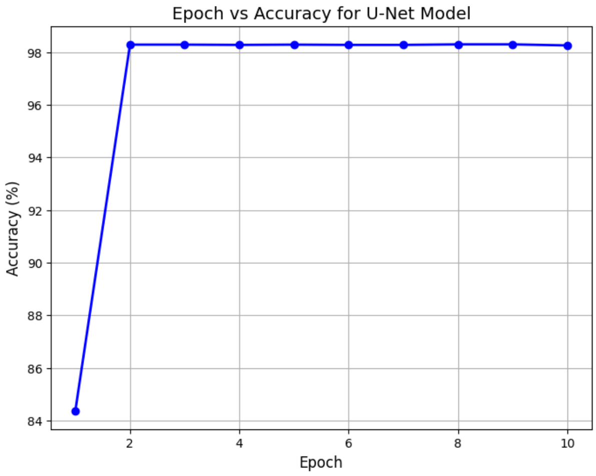

Figure 4.1: Accuracy v/s Epoch in Deep Learning U-NET Model

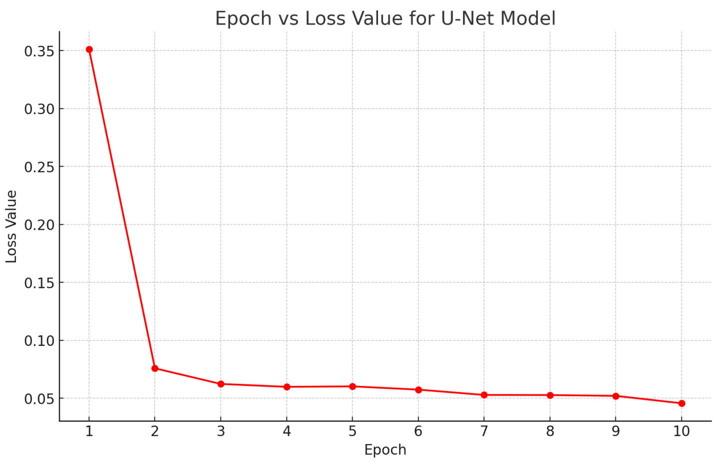

Figure 4.2: Loss v/s Epoch in Deep Learning U-NET Model

4.2 BNN U-Net Results

The BNN U-Net demonstrated an accuracy of 98.32% and a loss of 0.0422, along with an average DICE score of 0.9947, indicating excellent segmentation performance.

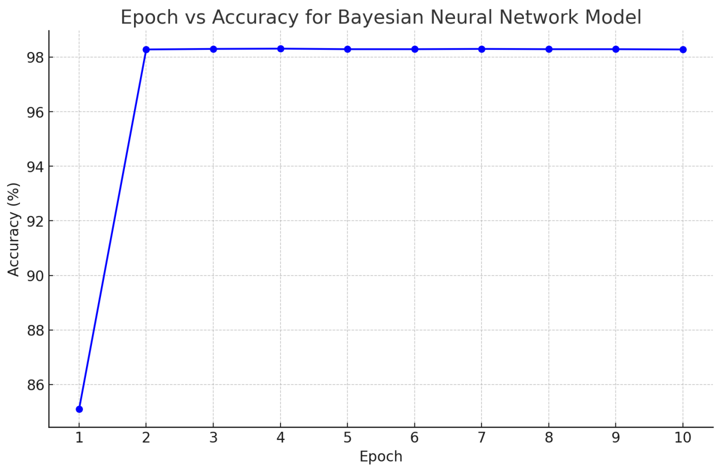

Figure 4.3: Accuracy v/s Epoch in Bayesian Neural Network Model

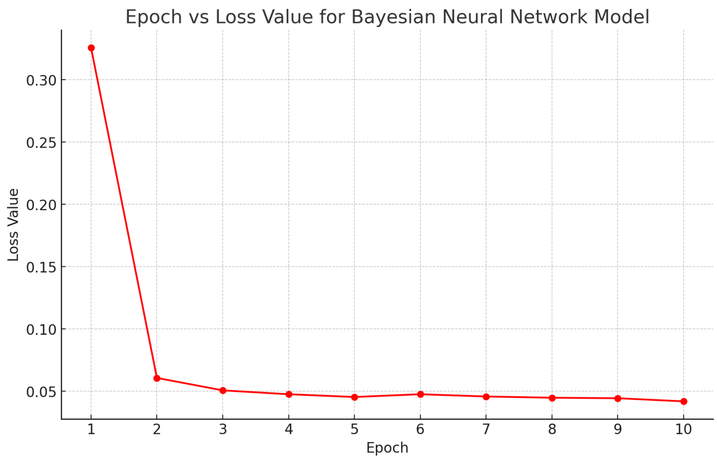

Figure 4.4: Loss v/s Epoch in Bayesian Neural Network Model

4.3 Comparative Analysis

Both models exhibited high accuracy and DICE scores; however, the BNN U-Net's ability to quantify uncertainty adds a layer of reliability to the predictions. This can be crucial in clinical settings where understanding model confidence is essential.Average DICE Score obtained is 0.9947.

Chapter 5

Discussion

The results demonstrate that both the U-Net and BNN U-Net models are effective for segmenting the optic disc and cup. While U-Net provides high accuracy, BNN U-Net's uncertainty estimates can enhance the interpretability of model predictions. Future work could explore integrating these models into clinical workflows to assist ophthalmologists in diagnosing glaucoma.

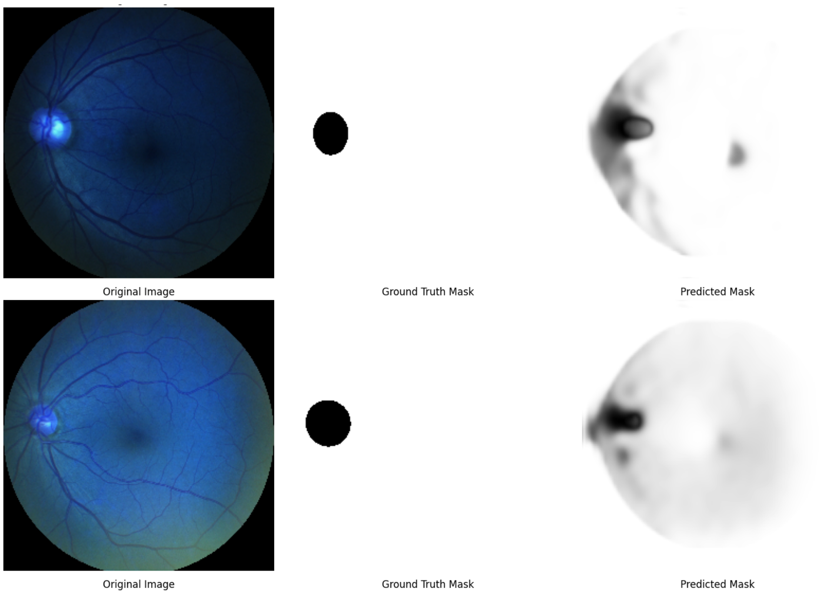

Figure 5.1: Predicted Image U-Net Model

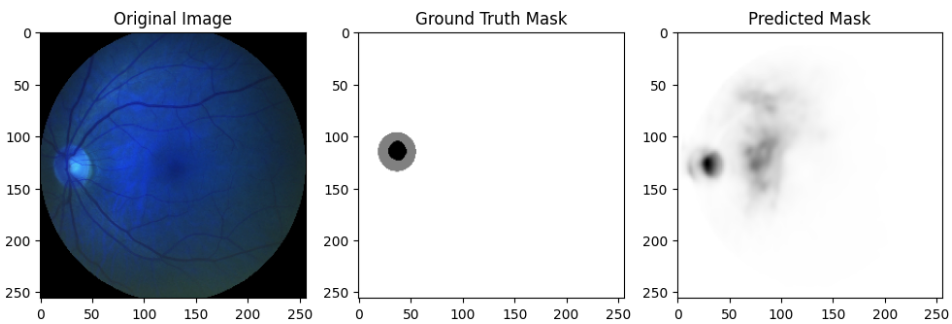

Figure 5.2: Predicted Image BNN Model

Chapter 6

Conclusion

This study successfully developed and compared U-Net and BNN U-Net models for segmenting the optic disc and cup in retinal images. Both models achieved high accuracy, with BNN U-Net providing additional interpretability through uncertainty quantification. These findings emphasize the potential of BNN U-Net for improving diagnostic reliability in glaucoma detection and demonstrate that deep learning models can significantly aid ophthalmologists in clinical practice

6.1 Future Work

Future research could expand upon this work by:

- **Exploring Larger Datasets:** Adding more diverse images to further improve model generalization.
- **Ensemble Methods**: Combining the strengths of U-Net and BNN U-Net could enhance both accuracy and reliability.
- **Real-time Clinical Testing:** Implementing these models in clinical settings to observe their impact on diagnostic workflows and patient outcomes.
- **Uncertainty Calibration**: Developing techniques to calibrate the uncertainty estimates, making them even more interpretable for clinical use.

Bibliography

[1] Yu S., Xiao D., Frost S., Kanagasingam Y. Robust optic disc and cup segmentation with deep learning for glaucoma detection. Comput. Med. Imaging Graph. 2019.

[2] Shoukat, A., Akbar, S., Hassan, S. A., Iqbal, S., Mehmood, A., Ilyas, Q. M. (2023). Automatic diagnosis of glaucoma from retinal images using deep learning approach. Diagnostics.

[3] Kashyap, R., Nair, R., Gangadharan, S. M. P., Botto-Tobar, M., Farooq, S., Rizwan, A. (2022). Glaucoma Detection and Classification Using Improved U-Net Deep Learning Model. Healthcare.

[4] Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)* (Vol. 48, pp. 1050–1059). PMLR.

[5] Xiong H, Liu S, Sharan RV, Coiera E, Berkovsky S. Weak label based Bayesian U-Net for optic disc segmentation in fundus images. Artif Intell Med. 2022 Apr; Epub 2022 Feb 26.

[6] Zhou, M. et al. Diabetes mellitus as a risk factor for open-angle glaucoma: A systematic review and meta-analysis.(2014).