

## **Abstract**

**Glaucoma is an eye condition brought on by optic nerve damage. If glaucoma is not promptly and effectively treated, visual loss may result. An ophthalmologist's observation is required in this situation to verify it. An expert doctor's observation takes a long time and is inconsistent because it is subjective. To detect glaucoma early on, automate fundus image processing, save time with optical disc and cup segmentation, and calculate Cup to Disc Ratio (CDR), a computer-aided diagnostic (CAD) system was created. The use of models based on Vision Transformers (ViT), Convolutional Neural Networks (CNN), and a mix of the two has been suggested in a number of earlier research. The CNN-based encoderdecoder model has a huge size and performs calculations slowly, while the ViT model has issues with the number of model computations increasing as the image size also rises. The Dense Prediction Transformer (DPT), which can handle data in parallel, was therefore chosen as the segmentation technique; it uses transformers as encoders and convolution as decoders. This work uses a case study of the optical disc and cup segmentation on fundus images using the ORIGA dataset to discuss the DPT's implementation. As a result, the DPT model slightly outperforms Segmenter with a difference in mIoU = 4.9%.**

**Keywords: glaucoma, segmentation, disc, cup, Dense Prediction Transformer**