Abstract
Glaucoma is an eye disorder that occurs due to damage to the optic nerve which can cause blindness, if not treated early. Because glaucoma rarely shows early symptoms in the sufferer, it requires regular observation of the fundus retina by an ophthalmologist, to find out if this eye disorder appears. Doctors' observations are subjective, so they are inconsistent and take a long time. Thus, a computer-aided diagnostic (CAD) system was built to automate the analysis of retinal fundus images, to detect glaucoma at an early stage, consistently and time efficiently, through optical disc and cup segmentation and cup-to-disc ratio (vCDR) calculation. CAD systems can also be used as decision support by doctors. Several previous studies have been proposed using models based on Convolution Neural Network (CNN), Vision Transformer (ViT) and a combination of both. However, the CNN-based encoder-decoder model has a large size and is slow to compute, and the ViT model has the problem that the computational amount of the model increases as the image size increases. Therefore, the segmentation method used is the Swin-Transformer-based encoder-decoder model, Swin-Unet, which has the advantage of linear computation and is more efficient. This research will discuss the implementation of Swin-Unet in a case study of optical disk and cup segmentation in an eye fundus image with the REFUGE dataset. CDR calculation with a threshold of 0,63 resulted in 94% accuracy. The Iou score results using the REFUGE dataset managed to get a score of 84% for the disc part and 80% for the cup part.

Keywords: Glaucoma, Segmentation, CDR, Disc, Cup, Swin-Unet.