

1. Pendahuluan

1.1 Latar Belakang

Visual memainkan peran penting dalam bidang visi komputer untuk tujuan pemecahan masalah. Segmentasi citra adalah komponen penting dalam berbagai sistem pemahaman visual, termasuk yang digunakan untuk analisis citra medis. Analisis citra medis menawarkan wawasan yang berharga tentang deteksi dan pengobatan penyakit dalam sistem perawatan kesehatan. Salah satu masalah khusus yang terkait dengan segmentasi citra medis adalah segmentasi optik disk dan cup secara akurat, yang terbukti bermanfaat dalam mendiagnosis glaukoma. Glaukoma mempengaruhi sebagian besar populasi global, dengan lebih dari 64 kasus yang dilaporkan pada tahun 2013 saja. Proyeksi menunjukkan bahwa jumlah ini diperkirakan akan terus meningkat, mencapai sekitar 80 juta dan 11,8 juta kasus pada tahun 2020 dan 2040. Glaukoma menempati urutan kedua di antara penyebab utama kehilangan penglihatan di seluruh dunia, dan secara signifikan berdampak pada sekitar 4,5 juta orang yang mengalami kebutaan akibat kondisi ini dalam skala global [1].

Sejumlah penelitian telah berfokus pada subjek segmentasi citra, termasuk pemanfaatan *U-Net* untuk segmentasi citra medis [2]–[4]. Jaringan *U-Net*, yang didasarkan pada jaringan saraf konvolusi (CNN), mencakup struktur *encoder-decoder* dengan *skip connection*. Kemajuan terbaru dalam visi komputer telah memperkenalkan *Vision Transformer*, khususnya untuk aplikasi seperti segmentasi citra [5]–[7]. Ada juga pendekatan yang menggabungkan metodologi *CNN* dan *Transformer* untuk segmentasi citra medis [8]–[10]. Namun, perlu dicatat bahwa model *encoder-decoder* berbasis *CNN* menunjukkan ukuran yang besar dan kecepatan komputasi yang lebih lambat, sedangkan model *ViT* menghadapi tantangan di mana kebutuhan komputasi meningkat seiring dengan citra berukuran lebih besar [11].

Algoritma *Swin-Transformer* adalah pengembangan terbaru berdasarkan *Vision Transformer* yang diusulkan oleh Liu dkk [12]. Tidak seperti *ViT*, yang tidak memiliki *feature map* hirarkis seperti yang digambarkan pada Gambar 3, pendekatan *Swin-Transformer* menggunakan representasi hirarkis *feature map*. Hal ini melibatkan pembuatan patch yang lebih kecil dan secara progresif menggabungkan patch yang berdekatan dengan setiap *layer*, sehingga cocok untuk tugas segmentasi. Keuntungan dari model ini terletak pada kemampuannya untuk mencapai segmentasi tingkat piksel yang lebih tepat karena pengambilan informasi yang lebih baik ketika memulai dengan patch yang lebih kecil [12].

Dalam penelitian ini, sistem segmentasi optik disk dan cup baru berdasarkan citra fundus retina diusulkan. Sistem yang diusulkan didasarkan pada *swin-transformer* yang memiliki keunggulan komputasi yang lebih efisien, karena mekanisme *self-attention* dilakukan di *local windows*. Kontribusi utama dari penelitian ini adalah sebagai berikut: (i) sistem segmentasi optik disk dan cup dikembangkan dengan menggunakan pendekatan *Swin-Transformer*, yang berbeda dengan penelitian-penelitian sebelumnya yang umumnya menggunakan *CNN* dan *Transformer*; (ii) sistem segmentasi optik disk dan cup berbasis *swin-transformer* yang dikembangkan memberikan performa segmentasi yang cukup sebanding dengan metode yang sudah ada untuk dataset yang sama.

1.2 Tujuan

Penelitian ini mencakup poin-poin utama berikut:

- Mengimplementasikan sistem segmentasi optik disk dan cup pada citra fundus retina menggunakan *Swin-Unet*.
- Mengetahui performa *Swin-Unet* dalam segmentasi optik disk dan cup.

1.3 Organisasi Tulisan

Penelitian ini memiliki struktur yang terdiri dari beberapa bagian. Bagian pertama merupakan yang memaparkan pentingnya penelitian ini. Bagian kedua penjelasan tentang studi sebelumnya tentang segmentasi citra dan aplikasi praktisnya, bersama dengan tinjauan literatur yang relevan dan penelitian yang sedang berlangsung. Bagian ketiga membahas secara ekstensif pendekatan *Swin-Unet*. Bagian keempat memaparkan hasil pengujian dan analisis yang dilakukan. Terakhir, terdapat bagian kesimpulan yang merupakan ringkasan dari temuan dan penilaian terhadap hasil penelitian ini.