

1. Pendahuluan

1.1 Latar Belakang

Kanker merupakan salah satu penyakit dengan tingkat kematian tertinggi di dunia, dengan peringkat ke-2 di Amerika Serikat [1]. Di Indonesia, sebesar 2,2 miliar rupiah dikeluarkan untuk menangani sekitar 1,3 juta kasus kanker [2]. Namun, seiring berkembangnya teknologi, risiko kematian kanker cenderung menurun sejak 1991 sebesar 32%, dan sekitar 3,5 juta kematian telah dihindari pada tahun 2019. Penurunan ini disebabkan perkembangan pada teknik operasi, terapi tertarget, dan deteksi dini. Salah satu metode deteksi dini adalah dengan melakukan pemindaian tulang seluruh tubuh untuk memonitor persebaran kanker [1].

Bone Scan Index (BSI) adalah skala yang digunakan untuk menghitung persebaran kanker pada tulang. Indeks ini merepresentasikan berapa banyak daerah pada tulang yang telah terinfeksi tumor dan persebaran dari metastasis [3]. Tetapi sebelum dapat menghitung BSI, proses segmentasi tulang perlu dilakukan. EXINI telah membuat sebuah aplikasi *software* untuk melakukan segmentasi dan perhitungan BSI sekaligus [4]. Dalam kasus ini juga telah dilakukan beberapa eksperimen pada segmentasi citra sidik tulang dan menunjukkan hasil yang baik [5]–[7]. Namun model seperti *Constrained Local Model* dan *Active Shape Model* ini masih menggunakan *landmark points* sebagai anotasi pada fase pelatihan modelnya, belum menggunakan *mask* segmentasi.

Salah satu model segmentasi semantik adalah DeepLabv3+ dengan *backbone ResNet* yang mengandalkan mekanisme konvolusi untuk mengekstraksi fitur lokal pada proses segmentasi [8]. Pendekatan lain, *Vision Transformer*, berfokus pada mekanisme *self-attention* untuk mengekstraksi konteks global dari sebuah citra [9], [10]. Salah satu model dengan basis ini adalah *Segmenter*, yang menggunakan pendekatan transformer dan menunjukkan performa yang menyaingi *DeepLab* [11]. Model lain, *DANet*, menggabungkan kedua pendekatan dengan menggunakan *backbone ResNet* sebagai bagian konvolusinya dan dua modul atensi sebagai bagian atensi, model ini mencoba untuk mengekstraksi konteks lokal dan global dari suatu citra [12].

Pada penelitian ini, diusulkan sistem segmentasi citra sidik tulang menggunakan *Dual Attention Model*. *DANet* dipilih menjadi metode basis karena menggabungkan kedua pendekatan yaitu konvolusi dan atensi. Perbandingan juga dilakukan pada *Segmenter* [11] dan *DeepLabv3+* [13] untuk menganalisis kekuatan dan kelemahan dari model. Beberapa varian dari setiap model juga diimplementasikan. Model-model ini dipilih sebagai pembanding karena masing-masing merepresentasikan pendekatan *transformer* dan konvolusi.

1.2 Tujuan

Penelitian ini mencakup poin-poin utama berikut:

- Mengimplementasikan sistem segmentasi citra sidik tulang berbasis model *Dual Attention Model*.
- Mengimplementasikan sistem pembanding *Segmenter* dan *DeepLabv3+*.
- Mengetahui hasil performa *DANet* dan membandingkannya dengan model pembanding.

1.3 Organisasi Tulisan

Penelitian ini diorganisir menjadi beberapa bagian. Bagian pertama menjelaskan pendahuluan yang memaparkan pentingnya penelitian ini, lalu diikuti oleh bagian kedua yang menjelaskan studi terkait yang relevan dengan penelitian ini. Bagian ketiga menjelaskan metode yang akan diimplementasikan sebagai basis model pada penelitian ini. Bagian keempat memaparkan hasil eksperimen dan analisis dari perbandingan model-model tersebut. Terakhir, pada bagian kelima terdapat bagian kesimpulan yang merupakan ringkasan dari seluruh kegiatan penelitian ini.