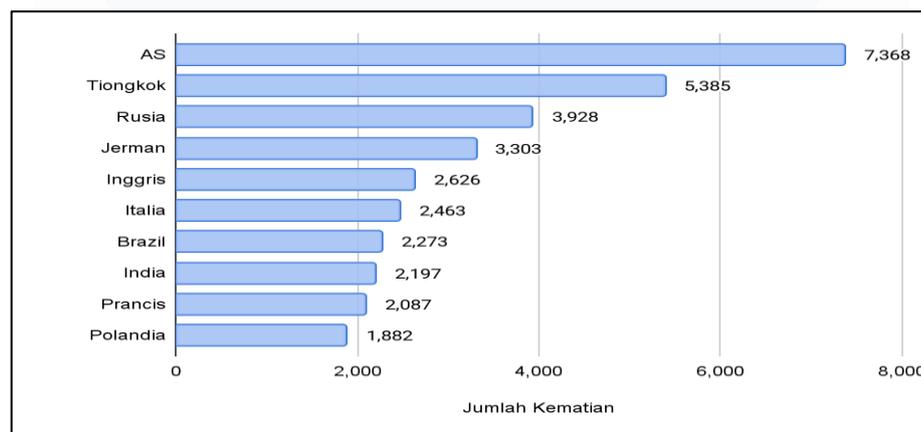


BAB I

PENDAHULUAN

1.1. Latar Belakang

Kulit merupakan lapisan terluar tubuh yang berfungsi sebagai pelindung utama dari pengaruh lingkungan [1]. Namun, karena kulit harus merespons berbagai faktor eksternal, seperti sentuhan, tekanan, dan perubahan suhu. Hal ini dapat meningkatkan risiko terjadinya penyakit serius, seperti kanker kulit. Statistik kanker global (*World Cancer Research Fund, 2022*) menunjukkan beban kanker yang signifikan di seluruh dunia, dengan total 331.722 kasus baru kanker kulit [2]. Angka kejadian kanker kulit sendiri berkisar antara 5,9% hingga 7,8% dari keseluruhan kasus kanker yang tercatat setiap tahunnya [3]. Grafik jumlah kematian akibat kanker kulit tahun 2022 ditampilkan pada Gambar 1.1.



Gambar 1. 1 Grafik Jumlah Kematian Akibat Kanker Kulit Tahun 2022 [2]

Gambar 1.1 menunjukkan grafik yang menampilkan jumlah kematian akibat kanker kulit di 10 negara dengan angka tertinggi pada tahun 2022 [2]. Amerika Serikat mencatat jumlah kematian tertinggi, yaitu 7.368 kasus, diikuti oleh Tiongkok dengan 5.385 kasus dan Rusia dengan 3.928 kasus. Grafik ini menunjukkan perbedaan jumlah kematian antarnegara, dengan Amerika Serikat memiliki angka tertinggi dibandingkan negara lainnya. Sementara itu, di Indonesia, *GoodStats* menunjukkan bahwa pada tahun 2020 terdapat 18.000 kasus kanker kulit dengan angka kematian mencapai 3.000 kasus [4].

Paparan sinar matahari dalam durasi lama, terutama tanpa perlindungan yang memadai, dapat merusak struktur kulit dan memperbesar kemungkinan terjadinya kanker kulit [5]. Oleh karena itu, kanker kulit menjadi salah satu jenis kanker yang paling umum terjadi di dunia, menempati urutan kelima dalam prevalensinya, dan merupakan salah satu penyakit yang cukup mematikan [6]. Organisasi Kesehatan Dunia (WHO) dan Organisasi Buruh Internasional (ILO) juga menyoroti bahwa peningkatan jumlah kasus kanker kulit sejalan dengan kenaikan suhu bumi dan intensitas pancaran sinar matahari [7]. Oleh karena itu, sangat disarankan untuk melakukan tindakan pencegahan, seperti penggunaan tabir surya, terutama di daerah-daerah dengan paparan sinar matahari yang tinggi, guna mengurangi risiko kanker kulit. Secara umum, kanker kulit terbagi menjadi dua jenis utama, yaitu melanoma dan non-melanoma [8]. Meskipun kasus melanoma lebih jarang dibandingkan non-melanoma, jenis ini jauh lebih berbahaya karena dapat menyebar dengan cepat ke organ lain jika tidak segera dideteksi dan ditangani pada tahap awal.

Dalam mendeteksi kanker kulit, deteksi dini sangat penting untuk membedakan kanker kulit dari kelainan kulit lainnya. Untuk itu, dibutuhkan sistem yang dapat mengidentifikasi jenis kanker kulit lebih awal, guna mengurangi risiko keterlambatan dalam penanganan [1]. Saat ini, deteksi kanker kulit masih banyak bergantung pada pemeriksaan visual oleh dokter atau melalui biopsi, dengan memakan banyak waktu dan biaya [9]. Untuk mengatasi keterbatasan ini, teknologi mulai digunakan dalam dunia medis, terutama dalam analisis gambar kulit [10]. Dalam konteks ini, KDD (*Knowledge Discovery in Databases*) memainkan peran penting dengan memanfaatkan teknik *data mining* untuk menemukan pola tersembunyi dalam data besar yang digunakan untuk mendeteksi kanker kulit secara lebih akurat dan efisien. KDD terdiri dari beberapa tahapan, seperti pemilihan data, pra-pengolahan, transformasi, *data mining*, dan evaluasi hasil. Secara umum, KDD juga dikenal sebagai proses penemuan pengetahuan dalam basis data [11]. Dengan memanfaatkan KDD, diharapkan deteksi kanker kulit dapat dilakukan dengan lebih cepat dan tepat, meningkatkan kemungkinan diagnosis yang lebih baik.

Deep Learning merupakan salah satu cabang dari *Machine Learning* yang algoritmanya dikembangkan berdasarkan inspirasi dari cara kerja otak manusia [12]. Seiring perkembangannya, berbagai arsitektur *deep learning* mulai dikembangkan untuk menangani jenis data dan permasalahan yang berbeda. Salah satu algoritma *deep learning* yang banyak digunakan dalam pemrosesan data berurutan adalah Transformer [13]. Algoritma ini menggunakan mekanisme *self-attention* dalam arsitekturnya dan dirancang untuk memproses data secara efisien dalam bentuk urutan. Dengan struktur *encoder-decoder*, *Transformer* mampu melakukan prediksi kata secara berurutan serta mentransformasikan satu urutan menjadi urutan lainnya [14]. Seiring perkembangan teknologi, arsitektur ini kemudian diadaptasi ke dalam bidang visi komputer melalui *Vision Transformer (ViT)*. ViT merupakan pengembangan terbaru yang menerapkan prinsip *Transformer* untuk data visual, dengan memanfaatkan teknik *patch embedding*, *multi-head attention*, dan *multi-layer perceptron* guna menghasilkan analisis citra yang lebih efektif dan mendalam [15].

Pengembangan lebih lanjut dari ViT adalah *Swin Transformer*, yang dirancang dengan struktur hierarkis dan mekanisme *shifted window* untuk meningkatkan efisiensi serta akurasi dalam pengolahan citra [16]. Dengan membatasi *self-attention* pada jendela lokal yang bergeser di setiap *layer*, *Swin Transformer* mampu menangkap informasi global dengan kompleksitas komputasi yang lebih rendah, dan telah terbukti unggul dalam berbagai tugas visi komputer seperti klasifikasi gambar, deteksi objek, dan segmentasi semantik. Selain itu, terdapat juga pengembangan model lain seperti *Data-efficient Image Transformers (DeiT)*, yang menonjol dalam efisiensi penggunaan data dan komputasi [17]. DeiT dirancang agar lebih ringan dan cepat tanpa mengorbankan akurasi, serta tersedia dalam beberapa varian seperti *DeiT Base* dan *DeiT Tiny*, sehingga dapat disesuaikan dengan kebutuhan komputasi yang berbeda.

Beberapa penelitian telah mengembangkan berbagai model deteksi kanker kulit menggunakan teknologi *deep learning* berbasis *Transformer*. Salah satunya adalah *Vision Transformer (ViT)* yang digunakan dalam klasifikasi kanker kulit

dengan pendekatan *multi-scale attention* dan *ensemble learning*, yang berhasil mencapai akurasi 95.05% pada dataset ISIC2018 [17]. Selain itu, model *ConvNeXtV2* dan *focal self-attention* juga telah digunakan untuk menangani dataset tidak seimbang [18], yang menghasilkan akurasi 93.60%, serta *precision* 91.69% dan *recall* 90.05% pada dataset ISIC 2019. Kombinasi model ViT, CNN, *Xception*, dan *Explainable AI (Grad-CAM)* juga menunjukkan hasil yang baik [19], dengan akurasi 96.74% pada dataset HAM10000 dan ISIC. Model hybrid yang menggabungkan CNN dan ViT dengan *focal loss* terbukti efektif dalam mengatasi ketidakseimbangan kelas pada dataset ISIC 2018 [20]. Selain itu, *transfer learning* yang digabungkan dengan *channel attention* dan ViT menunjukkan akurasi 99% pada multimodal dataset [21], memperlihatkan peningkatan signifikan dibandingkan metode sebelumnya.

Model *Separable Vision Transformer (SVT)* dengan data *augmentation* dan *preprocessing* juga mencapai akurasi 95.6%, serta sensitivitas 96.7% dan spesifisitas 95% pada dataset HAM10000, ISBI-2017, dan ISIC [22]. Selanjutnya, model gabungan dengan swin menggunakan dataset HAM-10000, berhasil meningkatkan akurasi klasifikasi menjadi 99.10%, dengan sensitivitas 99.27% dan spesifisitas 99.80% [23]. Terakhir, model *SkinSwinViT* [24], berbasis *Swin Transformer*, menunjukkan akurasi 97.88% dan *recall* 97.75% pada dataset ISIC2018, memperlihatkan keunggulannya dibandingkan dengan model lainnya dalam menangani data yang tidak seimbang. Hasil-hasil ini menggambarkan kemajuan pesat dalam pengembangan model deteksi kanker kulit berbasis *Transformer*, dengan ViT dan Swin memiliki potensi besar untuk meningkatkan akurasi, sensitivitas, dan kemampuan generalisasi sistem deteksi kanker kulit otomatis. Selain,

Penelitian ini akan mengembangkan model *Transformer* dengan melakukan komparasi model ViT, DeiT, dan Swin untuk mendeteksi kanker kulit dengan menggunakan dataset *International Skin Imaging Collaboration (ISIC) Tahun 2017* [22]. Selain itu, penelitian ini juga mengadopsi metodologi KDD (*Knowledge Discovery in Databases*) untuk memastikan proses analisis data yang sistematis,

mulai dari pemilihan data, pra-pengolahan, transformasi, *data mining*, hingga evaluasi hasil yang efektif dalam mendukung diagnosis kanker kulit [11]. Pada penelitian terdahulu yang dijadikan acuan untuk penelitian ini, yaitu peneliti telah mengembangkan model ViT dengan pendekatan *layer normalization* serta penerapan *focal loss*, dengan dataset ISIC 2017 dan berhasil memperoleh akurasi sebesar 86% serta nilai AUC sebesar 0.83 dalam klasifikasi melanoma [25]. Serta, penelitian yang mengusulkan arsitektur *Swin-SimAM*, yakni kombinasi antara *Swin Transformer* dan modul perhatian *SimAM*, yang menunjukkan performa dengan nilai AUC mencapai 0.90 dan juga menggunakan dataset ISIC 2017 [26]. Sebagai bentuk pengembangan dari studi sebelumnya yang berfokus pada evaluasi kinerja arsitektur *Vision Transformer* seperti ViT dan *Swin Transformer*, penelitian ini turut mengimplementasikan arsitektur *DeiT* untuk memperkaya analisis performa model *Vision Transformer* dalam konteks klasifikasi citra medis. Sehingga, penelitian ini diharapkan dapat mengoptimalkan hasil akurasi dari penelitian terdahulu dan dapat memberikan kontribusi terhadap pengembangan sistem diagnosis otomatis berbasis *deep learning* yang lebih presisi dan efisien.

Tantangan utama dalam penelitian ini terletak pada masih terbatasnya pemanfaatan teknologi yang optimal dalam proses deteksi melanoma, yang umumnya masih dilakukan secara manual atau menggunakan model yang belum mencapai akurasi tinggi. Hal ini dapat menyebabkan keterlambatan diagnosis dan penanganan yang kurang tepat. Oleh karena itu, penelitian ini bertujuan untuk mengatasi permasalahan tersebut dengan mengembangkan dan membandingkan tiga algoritma *deep learning*, yaitu *Vision Transformer* (ViT), *Swin Transformer*, dan *Data-efficient Image Transformer* (DeiT). Komparasi ini dilakukan untuk mengidentifikasi model yang paling efektif dalam mendeteksi kanker kulit melanoma secara otomatis, sehingga diharapkan dapat meningkatkan akurasi diagnosis dan efisiensi dalam proses deteksi dini penyakit tersebut.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, rumusan masalah dalam penelitian ini antara lain:

1. Bagaimana hasil rancangan model deteksi kanker kulit dengan memanfaatkan model ViT, Swin, dan DeiT pada citra medis?
2. Bagaimana hasil peningkatan akurasi *train*, *validation* dan *testing* model Transformer dalam meningkatkan performa deteksi kanker kulit?
3. Bagaimana hasil kinerja model ViT, Swin, dan DeiT dalam mendeteksi kanker kulit menggunakan metrik evaluasi AUC, spesifisitas sensitivitas, dan akurasi.

1.3. Tujuan

Tujuan penelitian ini antara lain:

1. Menganalisis hasil rancangan model deteksi kanker kulit berbasis citra medis dengan memanfaatkan arsitektur ViT, Swin, dan DeiT untuk menghasilkan sistem yang efektif dalam mengidentifikasi kanker kulit.
2. Menganalisis hasil peningkatan akurasi *train*, *validation* dan *testing* model *Transformer* untuk meningkatkan performa deteksi kanker kulit.
3. Mengevaluasi kinerja model ViT, Swin, dan DeiT dalam mendeteksi kanker kulit menggunakan metrik AUC, spesifisitas, sensitivitas, dan akurasi guna menentukan model paling optimal untuk diagnosis medis berbasis citra.

1.4. Manfaat

Penelitian ini diharapkan dapat memberikan kontribusi baik dari sisi teoritis maupun praktis. Manfaat yang diharapkan dari penelitian ini sebagai berikut:

1. Manfaat Teoritis

Penelitian ini memberikan kontribusi dalam pengembangan ilmu di bidang *deep learning*, terutama terkait penggunaan model *Transformer* seperti ViT, Swin Transformer, dan DeiT dalam pengolahan citra medis. Hasil penelitian ini dapat

menjadi referensi baru dalam memahami kelebihan model Transformer dalam tugas klasifikasi gambar kulit.

2. Manfaat Praktis

Manfaat praktis dari penelitian ini, yaitu dapat membantu dalam pengembangan sistem pendukung diagnosis kanker kulit yang lebih akurat dan efisien. Dengan menggunakan model *deep learning* berbasis *Transformer*, proses deteksi dapat dilakukan secara otomatis dan cepat, sehingga dapat mendukung tenaga medis dalam pengambilan keputusan dan meningkatkan kualitas layanan kesehatan, khususnya dalam bidang medis.