

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Kanker kulit merupakan salah satu jenis kanker yang paling umum dengan tingkat insidensi yang terus meningkat, sehingga memberikan beban berat bagi sistem kesehatan (Smak Gregoor dkk., 2023). Hal tersebut didukung oleh data statistik yang dikeluarkan oleh GLOBOCAN pada tahun 2018 yang disusun oleh International Agency for Research on Cancer (IARC) menunjukkan bahwa kanker kulit melanoma dan non melanoma masing – masing menempati posisi 19 dan 5 teratas sebagai kanker paling umum yang terjadi (Bray dkk., 2018). Pada tahun yang sama World Health Organization (WHO) telah mencatat terdapat 14 juta pasien baru dan 9.6 juta orang meninggal diakibatkan oleh kanker, dengan kanker kulit menjadi jenis kanker yang paling signifikan dalam meningkatkan angka kematian (Kassem dkk., 2020). Di Indonesia, kanker kulit menempati urutan ketiga setelah kanker rahim dan kanker payudara. Insiden kanker kulit dijumpai 5,9 – 7,8 % dari semua jenis kanker per tahun. Kanker kulit yang paling banyak di Indonesia adalah karsinoma sel basal (65,5%), diikuti karsinoma sel skuamosa (23%), melanoma malignant (7,9%) dan kanker kulit lainnya (Wilvestra dkk., 2018). International Agency for Research on Cancer (IARC) menyimpulkan pada tahun 2009 sinar ultraviolet (UV) adalah penyebab kanker kulit pada manusia, radiasi UV membantu perkembangan kanker melanoma dan kanker non melanoma (karsinoma sel basal dan karsinoma sel skuamosa), hal ini lebih umum terjadi pada individu dengan kulit sensitif terhadap sinar matahari, dan lokasi yang lebih dekat dengan garis khatulistiwa (Fauziyyah dkk., 2023).

Untuk dapat mengetahui apakah suatu tipe jenis kulit kanker kulit adalah dengan cara diagnosis langsung oleh dokter dengan proses biopsi dan mikroskopis. Proses biopsi dilakukan dengan cara mengambil potongan kecil dari bagian sel kanker untuk kemudian dicek dan diuji mendetail oleh dokter atau ahli dermatologis. Dengan menggunakan teknik pengujian ini membutuhkan waktu

yang cukup lama untuk seorang ahli dermatologis dan memiliki resiko kecelakaan pada saat proses biopsi (Nurlitasari dkk., 2022). Dengan berkembangnya teknologi terdapat cara alternatif lain untuk melakukan klasifikasi kanker kulit, penerapan *machine learning* dan *deep learning* dipercaya dapat membantu masyarakat dalam klasifikasi kanker kulit.

Secara umum, *machine learning* dan *deep learning* dapat melakukan *self-training* tanpa adanya pengkodean yang berulang oleh manusia. *Deep Learning* merupakan *machine learning* Tingkat lanjut yang membutuhkan sekumpulan data inisial yang disebut dengan dataset, untuk memprediksi hasil dari data. *Deep Learning* akan menghasilkan *output* berdasarkan data latihan dan tes. Setelah melalui proses evaluasi, *deep learning* dapat melakukan prediksi terhadap data (Purwono dkk., 2022). Salah satu algoritma *deep learning* yang paling sering digunakan adalah *convolutional neural network* (Purwono dkk., 2022; Sulistya dkk., 2023). *Convolutional Neural Network* telah dimanfaatkan untuk menyelesaikan komputasi tingkat tinggi mengenai tugas visual yang sulit dan biasanya berurusan dengan *image classification*, *segmentation*, *object detection*, *video processing*, *natural language processing*, dan *speech recognition* (Purwono dkk., 2022).

Convolutional Neural Network sendiri memiliki kumpulan lapisan arsitektur seperti *convolutional layer*, *pooling layer*, dan *fully connected layer* yang biasanya disebut sebagai model *convolutional neural network*. Terdapat beberapa model *convolutional neural network* yang telah dilatih sebelumnya diantaranya *AlexNet*, *ResNet50*, *GoogLeNet*, *VGG16*, *ResNet101*, *VGG19*, *InceptionV3*, *InceptionResNetV2*, *DenseNet*, *CGG19*, dan *MobileNet* (Sulistya dkk., 2023). Salah satu teknik yang biasa dilakukan dalam mengembangkan sebuah model *convolutional neural network* adalah *transfer learning*. Menurut Sulistya dkk (2023) menyatakan bahwa *transfer learning* adalah sebuah proses untuk menggunakan model yang telah dilatih sebelumnya dalam permasalahan lain untuk menyelesaikan permasalahan baru. *transfer learning* dapat sangat membantu dalam

mengembangkan sebuah model karena tidak perlu membangun sebuah model dari awal. Oleh karena itu, algoritma *convolutional neural network* dan teknik *transfer learning* dapat dimanfaatkan untuk membuat sebuah model dalam melakukan klasifikasi kanker kulit.

Sebelumnya telah terdapat beberapa penelitian serupa mengenai pengembangan model untuk klasifikasi kanker kulit seperti menggunakan SVM dan KNN berhasil mendapatkan metrik evaluasi *accuracy* 70,61% dengan KNN (Faruk dan Nafi'iyah, 2020). Lalu penelitian menggunakan *VGG16* dengan jumlah sampel pelatihan sebanyak 6.594 gambar dari dataset Kaggle yang berjudul “*Skin Cancer Malignant vs Benign*” berhasil mendapatkan metrik evaluasi *accuracy* sebesar 93,18% (Hasan dkk., 2021). Penelitian menggunakan arsitektur *convolutional neural network* berhasil mendapatkan *accuracy* sebesar 96% dengan jumlah dataset sebanyak 2.000 gambar (Khan dkk., 2021).

Penelitian selanjutnya menggunakan arsitektur *ResNet152* dengan dataset ISIC 2017 memiliki total data sebesar 2.742 gambar berhasil mendapatkan metrik evaluasi *accuracy* sebesar 90,4%, *sensitivity* 82%, *specificity* 92,5%, dan *balance accuracy* 87,2% (Jojoa Acosta dkk., 2021). Terdapat pula penelitian menggunakan arsitektur *Alexnet* untuk melakukan klasifikasi *benign* dan *malignant*, dengan menggunakan dataset ISIC 2019 dengan total data sebesar 25.331 gambar berhasil mendapatkan metrik evaluasi *accuracy* sebesar 81,26%, *precision* 84,66%, *recall* 86,71 dan *f1-score* 82,14% (Saputra & Ezar Al Rivan, 2023).

Penelitian yang dilakukan oleh Poornima dkk (2021) menggunakan dataset ISIC dengan arsitektur *VGG16* berhasil mendapatkan *accuracy* sebesar 97%. Kemudian penelitian untuk klasifikasi kanker kulit dengan menggunakan dataset ISIC 2020 dengan total data sebesar 33.126 gambar berhasil mendapatkan metrik evaluasi *accuracy* sebesar 98,39% dengan cara menggabungkan ketiga model yang telah dilatih sebelumnya yaitu *ResNet*, *VGGNet*, dan *MobileNet* (Singha & Roy, 2022). Dan terakhir penelitian yang dilakukan oleh Rashid dkk (2022) menggunakan model arsitektur *MobileNetV2* dengan memakai dataset ISIC 2020

berhasil mendapatkan *average accuracy* sebesar 98,20% dengan memanfaatkan metrik evaluasi *accuracy, precision, recall, fl-score*, dan AUC-ROC.

Untuk penelitian kali ini, dataset yang akan digunakan untuk dijadikan bahan penelitian adalah dataset ISIC 2020. Dalam dataset tersebut terdapat sepuluh jenis penyakit kulit diantaranya *unknown, nevus, melanoma, seborrheic keratosis, lentigo NOS, lichenoid keratosis, solar lentigo, cafe-au-lait macule*, dan *atypical melanocytic proliferation*. Dari sepuluh jenis penyakit tersebut terbagi lagi menjadi dua tipe kelas yaitu *benign* dan *malignant*, dengan tujuan dari dataset ini adalah untuk melakukan klasifikasi diantara keduanya. Dataset tersebut merupakan dataset keluaran terbaru yang dibuat oleh International Skin Imaging Collaboration (ISIC) yang menyediakan koleksi dataset terbesar, terdiri dari gambar digital yang dibutuhkan oleh para peneliti dalam melakukan diagnosis penyakit kulit dengan menggunakan teknik *artificial intelligence* dengan bersifat *open source* (Duman & Tolan, 2021). Dataset ISIC 2020 diambil dari berbagai sumber seperti Hospital Clinic de Barcelona, Medical University of Vienna, Memorial Sloan Kettering Cancer Center, Melanoma Institute Australia, University of Queensland, dan University of Athens Medical School. Dataset tersebut juga lebih bervariasi dibandingkan dengan dataset yang telah dikeluarkan oleh ISIC sebelumnya, dengan total 33.126 gambar. Gambar latihan tersebut terbagi menjadi dua kategori yaitu 32.542 gambar termasuk *benign* dan 584 gambar adalah *malignant*. Selain itu, beberapa dataset pada website Kaggle yang relevan dengan penelitian ini bersumber pada dataset ISIC.

Melihat ketidakseimbangan data antara kelas *benign* dan *malignant* yang terdapat pada dataset ISIC 2020, diperlukannya analisis lebih jauh untuk memilih model dengan arsitektur *convolutional neural network* yang telah dilatih sebelumnya dan metrik evaluasi yang akan dipakai. Bila mengacu dari penelitian yang telah dikaji sebelumnya, arsitektur *MobileNet* memiliki performa yang lebih baik bila dibandingkan dengan arsitektur lain. Hal ini juga didukung oleh penelitian yang dilakukan oleh Duman dan Tolan (2021) bahwa arsitektur *MobileNet*

meskipun memiliki ukuran yang lebih kecil, tetapi dapat memiliki performa yang lebih baik bila dibandingkan dengan beberapa arsitektur lain seperti *ResNet* dan *NasNet*. Berdasarkan pemaparan tersebut maka dalam penelitian kali ini, arsitektur yang akan digunakan adalah *MobileNet* dengan variasi *MobileNetV3Large* yang dapat melakukan pelatihan lebih kompleks. Serta metrik evaluasi yang akan digunakan dalam penelitian kali ini adalah *accuracy*, *precision*, *recall*, *f1-score*, dan AUC-ROC.

Untuk membantu pelatihan model terdapat beberapa metode tambahan yang akan dilakukan seperti transformasi morfologi, teknik mengatasi ketidakseimbangan data, pemilihan optimisasi, dan penerapan *mixed precision*. Transformasi morfologi dengan menggunakan operasi *black-hat* dilakukan untuk menemukan dan menghapus rambut atau *noise* yang menghalangi gambar dengan tujuan untuk membantu pelatihan model dalam mengenali penyakit, hal ini berpacu pada penelitian yang dilakukan oleh Khan dkk (2021) dengan membandingkan model yang dilatih pada dataset yang menerapkan *black-hat* dan pada dataset yang tidak menerapkan *black-hat*, dimana model yang dilatih pada dataset yang diterapkan *black-hat* memiliki performa yang lebih baik.

Untuk mengatasi ketidakseimbangan data antara kelas *benign* dan *malignant*, penelitian ini akan menggunakan teknik *oversampling*. *Oversampling* adalah suatu cara untuk menyeimbangkan data dengan melakukan penambahan data pada kelas minoritas, alasan menggunakan teknik ini mengacu pada penelitian yang dilakukan oleh Werner dkk (2023) mengenai perbandingan teknik untuk *preprocessing* data, dimana *oversampling* menjadi teknik yang paling sering digunakan dan baik dalam menjaga keutuhan data. Lalu optimisasi yang akan digunakan adalah optimisasi adam, hal ini mengacu pada penelitian yang dilakukan oleh Valova dkk (2020) untuk mencari optimisasi terbaik dalam melakukan *image classification* pada dataset yang tidak seimbang, hasilnya optimisasi adam dapat memberikan performa yang baik dalam pelatihan. Dan terakhir adalah penerapan *mixed precision*, cara kerja dari *mixed precision* adalah dengan menggabungkan

kombinasi *float32* dan *float16*, dimana *float16* digunakan untuk meningkatkan performa dan mempercepat waktu pelatihan saat pelatihan dan *float32* digunakan untuk menyimpan variabel untuk menjaga kestabilan numerik (Nandakumar dkk., 2020).

Berdasarkan latar belakang tersebut, maka melalui penelitian yang berjudul **“Klasifikasi Kanker Kulit Menggunakan *Convolutional Neural Network* Dengan Implementasi *Mixed Precision*”** berharap dapat memberikan evaluasi performa dari model yang menerapkan *mixed precision* dan model yang tidak menerapkan *mixed precision*.

1.2 Rumusan Masalah Penelitian

Berdasarkan latar belakang dan identifikasi masalah maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana performa dari model setelah menggunakan *mixed precision* dalam pelatihannya?
2. Bagaimana perbedaan waktu pelatihan model setelah menggunakan *mixed precision* bila dibandingkan dengan model yang belum diterapkan *mixed precision*?

1.3 Tujuan Penelitian

Penelitian ini memiliki tujuan penelitian yang merupakan hasil jawaban dari rumusan masalah. Adapun tujuan dari penelitian ini adalah:

1. Untuk mengetahui performa model setelah diterapkannya metode *mixed precision*.
2. Untuk mengetahui perbedaan waktu pelatihan model antara model tanpa *mixed precision* dan model yang telah diimplementasikan *mixed precision*.

1.4 Batasan Masalah

Untuk membatasi permasalahan dan menentukan fokus tujuan dari penelitian, maka dibuatkanlah Batasan permasalahan dengan diantaranya adalah sebagai berikut:

1. Model hanya akan melakukan klasifikasi dari dua tipe kelas saja yaitu *benign* dan *malignant*. Model tidak akan melakukan klasifikasi jenis penyakit dari masing-masing kelas.

1.5 Manfaat/Signifikansi Penelitian

Adapun manfaat yang dapat dihasilkan dari penelitian ini diantaranya adalah sebagai berikut:

1. Penelitian ini dapat membantu diagnosis medis. Dengan menggunakan model *convolutional neural network* yang telah dikembangkan, dapat memberikan bantuan untuk klasifikasi kanker kulit pada tahap awal, membantu dokter dalam proses pengambilan keputusan dan meningkatkan peluang kesembuhan pasien.
2. Membantu untuk menciptakan dasar perumusan kebijakan kesehatan yang lebih efektif. Dengan menerapkan teknologi berbasis kecerdasan buatan untuk klasifikasi penyakit khususnya penyakit kulit.
3. Model dapat membantu untuk melakukan diagnosis tahap awal terhadap penyakit kulit pengguna, sehingga dapat membantu pengguna untuk mengambil keputusan sebelum berkonsultasi dengan pihak medis.
4. Penelitian ini dapat meningkatkan kesadaran masyarakat mengenai kesehatan kulit, dengan melakukan diagnosis penyakit kulit yang mudah yaitu memakai telepon cerdas, hal tersebut dapat membantu masyarakat untuk lebih sadar akan kesehatan kulit mereka.

1.6 Sistematika Penulisan Skripsi

Sistematika penulisan skripsi bertujuan untuk menjadi pedoman dalam menulis agar dapat lebih terarah, maka skripsi ini dibagi menjadi beberapa bab. Adapun struktur organisasi skripsi ini adalah sebagai berikut:

BAB I Pendahuluan

Pada Pendahuluan berisi tentang: latar belakang penelitian, rumusan masalah penelitian, tujuan penelitian, batasan penelitian, manfaat/signifikansi penelitian dan struktur organisasi skripsi.

BAB II Kajian Pustaka

Pada Kajian Pustaka berisi tentang: membahas mengenai teori dan konsep yang relevan dengan penelitian. Membahas mengenai hasil penelitian terdahulu mengenai penelitian yang relevan, posisi teoritis peneliti yang berkenaan dengan masalah.

BAB III Metode Penelitian

Pada Metode Penelitian berisi tentang hal-hal prosedural dalam melakukan penelitian, pada penelitian ini pendekatan yang akan digunakan adalah dengan menggunakan **kuantitatif**. Pendekatan ini memiliki beberapa bagian seperti desain penelitian, partisipan, populasi dan sampel, instrumen penelitian, prosedur penelitian, dan analisis data.

BAB IV Temuan dan Pembahasan

Pada bagian ini akan menjelaskan mengenai hasil temuan untuk menyelesaikan rumusan masalah yang telah ditetapkan sebelumnya dengan menggunakan pendekatan **kuantitatif**,

BAB V Kesimpulan, Implikasi dan Rekomendasi

Kesimpulan, implikasi dan rekomendasi memuat tentang kesimpulan dari hasil analisis temuan penelitian, serta implikasi dan rekomendasi bagi para pembaca dan pengguna hasil penelitian.