

BAB 1 PENDAHULUAN

1.1. Latar Belakang

Penyakit kulit adalah salah satu masalah kesehatan yang umum terjadi di seluruh dunia, dengan berbagai penyebab seperti infeksi bakteri, jamur, reaksi alergi, atau gangguan autoimun. Sebagai organ terbesar tubuh manusia, kulit rentan terhadap berbagai jenis gangguan yang dapat memengaruhi kualitas hidup seseorang dan membutuhkan intervensi medis yang cepat. Penyakit kulit dapat berkembang dari gejala ringan hingga kondisi kronis yang membutuhkan diagnosis dan penanganan segera untuk mencegah komplikasi lebih lanjut. Berdasarkan data dari World Health Organization (WHO), penyakit kulit, khususnya *melanoma* yang merupakan salah satu jenis kanker kulit, menjadi salah satu penyebab utama kematian akibat penyakit kulit, terutama di negara-negara dengan paparan sinar *ultraviolet* yang tinggi [1]. Namun, dalam praktiknya, proses diagnosis manual masih menghadapi tantangan berupa keterbatasan jumlah tenaga medis spesialis, waktu yang diperlukan untuk analisis visual, serta potensi ketidakkonsistenan antarpraktisi. Hal ini menjadi hambatan tersendiri terutama di daerah dengan akses terbatas terhadap layanan dermatologi. Pengenalan dan klasifikasi penyakit kulit berbasis citra merupakan langkah penting dalam mendukung proses diagnosis yang lebih cepat dan akurat. Oleh karena itu, diperlukan pendekatan berbasis teknologi untuk membantu proses diagnosis yang lebih cepat, akurat, dan efisien.

Dalam beberapa tahun terakhir, teknologi kecerdasan buatan (*Artificial Intelligence/AI*) telah menunjukkan potensi besar dalam berbagai bidang, termasuk bidang medis. AI memungkinkan pengolahan data dalam jumlah besar secara efisien dan memberikan hasil analisis yang mendalam, sehingga dapat mendukung pengambilan keputusan klinis. Salah satu penerapan AI yang menonjol di bidang medis adalah analisis citra medis, yang bertujuan untuk membantu proses diagnosis dan klasifikasi penyakit secara lebih akurat. Dalam konteks ini, metode berbasis deep learning, seperti

Convolutional Neural Network (CNN), menjadi salah satu pendekatan yang banyak digunakan karena kemampuannya dalam mengenali pola kompleks pada citra medis. CNN telah terbukti sangat efektif dalam klasifikasi citra, termasuk citra dermatoskopi yang digunakan untuk mendeteksi penyakit dan kanker kulit [2]. CNN mampu mengekstraksi fitur-fitur penting dari citra secara otomatis, sehingga dapat mengurangi ketergantungan pada fitur manual yang sering kali memerlukan keahlian khusus.

Salah satu arsitektur CNN yang telah menunjukkan performa unggul dalam berbagai tugas klasifikasi citra adalah EfficientNet. EfficientNet memperkenalkan metode '*compound scaling*' yang secara serentak mengatur kedalaman, lebar, dan resolusi jaringan menggunakan satu parameter sederhana, memungkinkan peningkatan efisiensi dan akurasi model tanpa memerlukan sumber daya komputasi yang berlebihan.

Sebagai pengembangan dari arsitektur ini, EfficientNetV2 diperkenalkan untuk mengatasi kelemahan pada proses pelatihan model skala besar dan untuk mempercepat konvergensi. EfficientNetV2 menggabungkan teknik fused-MBConv (yang menggabungkan depthwise convolution dengan standard convolution pada lapisan awal) dengan strategi progressive learning, yakni memulai pelatihan pada citra resolusi rendah dan secara bertahap meningkatkan resolusi citra selama proses pelatihan. Pendekatan ini memungkinkan model untuk mencapai akurasi tinggi dengan waktu pelatihan yang lebih singkat dan efisiensi penggunaan memori yang lebih baik dibandingkan pendahulunya [3]. Hal ini menjadikannya pilihan yang tepat untuk tugas klasifikasi penyakit kulit menggunakan citra dermatoskopi, di mana detail-detail kecil pada citra sangat penting untuk membedakan antara lesi jinak dan ganas.

Namun, meskipun teknologi CNN telah banyak digunakan, terdapat gap antara kondisi saat ini dengan kondisi yang diharapkan. Sebagian besar penelitian sebelumnya hanya berfokus pada dataset tertentu dengan jumlah data yang terbatas, sehingga model yang dihasilkan sering kali kurang generalisasi ketika diterapkan pada dataset baru [4]. Di samping itu, beberapa model mengalami kesulitan dalam mengenali lesi kulit dari kelas minoritas,

seperti melanoma atau akiec, yang padahal justru penting secara klinis. Oleh karena itu, diperlukan pengembangan arsitektur dan strategi pelatihan yang lebih adaptif terhadap variasi data dan distribusi kelas yang tidak seimbang. Dalam penelitian ini, *dataset* yang digunakan adalah HAM10000 (*Human Against Machine with 10000 training images*), yang merupakan salah satu dataset terbesar dan paling beragam untuk deteksi melanoma dan lesi kulit lainnya. *Dataset* HAM10000 terdiri dari 10.015 citra dermatoskopi yang dikumpulkan dari berbagai sumber untuk merepresentasikan berbagai jenis lesi kulit, termasuk *melanoma*, *keratosis seboroik*, dan *nevus* [5]. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model klasifikasi kanker kulit berbasis CNN dengan menggunakan arsitektur *EfficientNetV2*, yang diharapkan dapat memberikan akurasi tinggi dan mampu mengatasi keterbatasan generalisasi pada dataset yang berbeda.

Motivasi utama dari penelitian ini adalah untuk menjawab kebutuhan akan sistem pendukung diagnosis yang tidak hanya akurat, tetapi juga efisien dan adaptif terhadap keragaman data citra medis. Penelitian terkait penggunaan CNN dalam klasifikasi penyakit kulit telah menunjukkan hasil yang menjanjikan, namun masih terdapat tantangan dalam meningkatkan akurasi dan kemampuan generalisasi model, terutama saat diterapkan pada beragam jenis data citra medis. Mengingat pentingnya diagnosis yang cepat dan akurat dalam mendukung penanganan penyakit kulit, termasuk kanker kulit, pendekatan berbasis pengolahan citra dengan deep learning, seperti CNN, menjadi sangat relevan. Selain itu, penggunaan citra dermatoskopi yang dapat memperjelas detail permukaan kulit memberikan peluang besar untuk meningkatkan presisi dalam pengenalan dan klasifikasi penyakit kulit. Dengan latar belakang tersebut, penelitian ini diusulkan untuk mengembangkan model klasifikasi yang lebih akurat dan andal, sehingga dapat memberikan kontribusi signifikan dalam mendukung teknologi berbasis AI di bidang medis, khususnya dalam sistem pendukung diagnosis penyakit kulit yang efektif dan praktis.

1.2. Perumusan Masalah

Berdasarkan latar belakang penelitian yang sudah ada, dapat dipahami bahwa model yang akan dibangun merupakan model yang mengidentifikasi dan mengklasifikasi beberapa penyakit-penyakit kulit yang ada pada citra dermastokopi. Citra dermastokopi merupakan gambar yang dihasilkan melalui proses pemeriksaan kulit menggunakan dermastokop. Dermatoskop adalah perangkat optik yang memungkinkan dokter atau ahli kulit untuk melihat struktur kulit yang tidak terlihat dengan mata telanjang, seperti pola pigmen, pembuluh darah, atau kelainan lainnya pada lapisan kulit.

Pada penelitian ini diasumsikan bahwa, model arsitektur EfficientNet dapat diimplementasikan secara optimal. Selain itu, *Dataset* HAM10000 diasumsikan merepresentasikan variasi jenis lesi kulit yang relevan untuk tugas klasifikasi. Lingkup penelitian ini berfokus pada pengembangan model klasifikasi penyakit kulit ke dalam 7 kategori lesi kulit yaitu *melanoma* (mel), *melanocytic nevi* (nv), *basal cell carcinoma* (bcc), *actinic keratosis* (akiec), *benign keratosis* (bkl), *dermatofibroma* (df), dan *vascular lesion* (vasc). Evaluasi model dilakukan menggunakan metrik evaluasi standar, seperti akurasi, *precision*, *recall*, dan *F1-score*.

Tantangan utama dalam klasifikasi penyakit kulit berbasis citra dermatoskopi adalah kompleksitas pola visual pada lesi kulit, variasi kualitas citra, serta distribusi data yang tidak merata. Convolutional Neural Network (CNN) dipilih sebagai pendekatan utama karena kemampuannya dalam mengenali pola visual yang kompleks dan menangani data citra secara efisien. EfficientNet telah terbukti unggul dalam berbagai tugas klasifikasi citra medis, termasuk analisis citra dermatoskopi, sehingga diharapkan dapat menjadi solusi yang tepat untuk mengatasi tantangan dalam penelitian ini.

Berdasarkan permasalahan di atas, penelitian ini dirumuskan ke dalam beberapa poin berikut:

1. Bagaimana pengaruh kualitas dataset citra dermatoskopi, seperti variasi resolusi, noise, dan jumlah data, terhadap performa model klasifikasi EfficientNet?
2. Bagaimana performa arsitektur EfficientNet dalam hal akurasi, *presicion*, *recall*, dan *F1-score* terhadap dataset HAM10000?

1.3. Tujuan

Berdasarkan rumusan masalah yang ada dalam penelitian ini, tujuan dari penelitian tugas akhir ini adalah:

1. Menganalisis pengaruh kualitas dataset citra dermatoskopi, seperti variasi resolusi, noise, dan jumlah data, terhadap performa model klasifikasi.
2. Mengevaluasi performa model EfficientNet dalam hal akurasi, *presicion*, *recall*, dan *F1-score* yang dapat dicapai menggunakan dataset HAM10000.

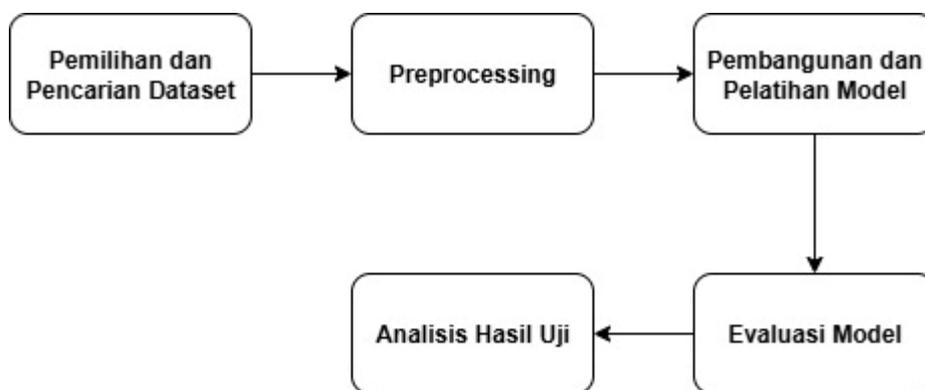
1.4. Batasan Masalah

Dalam suatu penelitian, penting untuk menetapkan batasan masalah guna memperjelas ruang lingkup dan fokus dari studi yang dilakukan. Batasan ini diperlukan agar proses analisis, eksperimen, dan evaluasi dapat dilakukan secara terarah serta sesuai dengan tujuan penelitian. Oleh karena itu, penelitian ini menetapkan beberapa batasan sebagai berikut:

1. Dataset yang digunakan adalah HAM10000 (Human Against Machine with 10000 training images), yang terdiri dari 7 kelas jenis penyakit kulit, yaitu: akiec, bcc, bkl, df, mel, nv, dan vasc [13].
2. Model arsitektur yang digunakan adalah EfficientNetV2B3, dengan perbandingan performa juga dilakukan terhadap model EfficientNetB0 dan EfficientNetV2B0.
3. Evaluasi performa model dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score, serta analisis melalui confusion matrix untuk mengetahui pola kesalahan klasifikasi antar kelas.

1.5. Metode Penelitian

Metode penelitian yang digunakan dalam tugas akhir ini terdiri dari beberapa tahapan utama. Pertama, pengumpulan dan eksplorasi dataset HAM10000, termasuk proses pengelompokan citra berdasarkan label penyakit kulit. Kedua, dilakukan preprocessing citra pada dataset berupa resizing, normalisasi, augmentasi, dan pemisahan subset untuk menyeimbangkan distribusi kelas. Tahapan berikutnya adalah pelatihan model menggunakan tiga arsitektur CNN berbasis EfficientNet (B0, V2B0, V2B3) yang dilakukan dalam dua tahap: feature extraction dan fine-tuning. Model dilatih menggunakan Focal Loss, optimizer Adam, dan callback seperti EarlyStopping dan ModelCheckpoint.



Gambar 1.1 Diagram Alur Penelitian

Evaluasi dilakukan menggunakan data uji berdasarkan metrik akurasi, precision, recall, F1-score, serta analisis confusion matrix. Hasil dari setiap skenario dibandingkan untuk melihat pengaruh arsitektur dan strategi pelatihan terhadap performa klasifikasi.