

# BAB 1 PENDAHULUAN

## 1.1. Latar Belakang

Kanker adalah penyakit proliferasi yang tidak terkendali yang disebabkan oleh sel-sel yang bermutasi yang berkembang di bawah seleksi alam [1]. Sel-sel ini dapat berasal dari hampir semua bagian tubuh manusia, yang terdiri dari triliunan sel. Kanker paru-paru muncul ketika pertumbuhan sel abnormal tersebut terjadi di paru-paru [2]. Mengingat prevalensi dan tingkat kematiannya yang tinggi, deteksi dini dan akurat dari kanker paru-paru sangat penting untuk meningkatkan hasil pasien dan mengurangi beban keseluruhan pada sistem perawatan kesehatan.

Deep learning, khususnya Convolutional Neural Networks (CNN), telah muncul sebagai teknologi transformatif dalam analisis citra medis, mengotomatisasi tugas-tugas diagnostik yang kompleks dan mengurangi beban kerja penyedia layanan kesehatan. CNN telah menunjukkan keberhasilan yang luar biasa dalam berbagai aplikasi pendeteksian kanker dengan menganalisis gambar histopatologi, yang sangat penting untuk memahami struktur seluler jaringan kanker. Di antara arsitektur CNN, EfficientNet, yang diperkenalkan oleh Tan dan Le (2019) [3], dikenal karena keseimbangannya antara akurasi tinggi dan efisiensi komputasi, yang dicapai melalui *compound scaling* yang secara optimal menyesuaikan kedalaman, lebar, dan resolusi model. Efisiensi ini membuat EfficientNet menjadi pilihan yang cocok untuk tugas pencitraan medis di mana sumber daya komputasi mungkin terbatas.

Para peneliti telah menggunakan teknik *deep learning* dan pembelajaran mesin dalam penelitian sebelumnya untuk menganalisis data medis, terutama gambar histopatologi, untuk deteksi kanker. Hatuwal dan kolega [4] menggunakan arsitektur CNN dengan tiga lapisan konvolusi dan satu lapisan

yang terhubung penuh untuk deteksi kanker paru-paru berdasarkan gambar histopatologi. Meskipun pendekatan ini mencapai akurasi yang tinggi, kesederhanaan model ini dapat membatasi kemampuannya untuk menangkap pola yang kompleks pada jaringan kanker. Anjum dan kolega [5] mengembangkan penelitian ini dengan membandingkan berbagai varian EfficientNet untuk klasifikasi kanker paru-paru menggunakan gambar histopatologi, yang pada akhirnya mengidentifikasi EfficientNetB2 sebagai model yang paling efektif.

Dalam penelitian lain yang tidak terbatas pada kanker paru-paru, Kallipolitis dan kolega [6] mengeksplorasi model ensemble berdasarkan EfficientNet untuk klasifikasi citra kanker, dengan menggabungkan Grad-CAM untuk meningkatkan kemampuan interpretasi keputusan klasifikasi ensemble. Namun, penelitian ini menerapkan augmentasi gambar yang terbatas, yang mungkin telah membatasi kemampuan generalisasi model pada gambar histopatologi yang beragam. Abioye dan kolega. [7] mengevaluasi model EfficientNetV2 pada citra kanker payudara jinak, dengan EfficientNetV2L menunjukkan akurasi tertinggi. Namun, penelitian mereka terbatas pada kasus jinak pada pembesaran 40x, yang dapat membatasi generalisasi untuk kasus ganas atau pembesaran lainnya. Selain itu, augmentasi yang dilakukan juga terbatas, sehingga berpotensi membatasi kemampuan adaptasi model terhadap gambar histopatologi yang beragam.

Penelitian ini menjawab kesenjangan yang diidentifikasi dalam penelitian sebelumnya dengan memanfaatkan ensemble varian EfficientNetV2 untuk klasifikasi kanker paru-paru menggunakan gambar histopatologi. Dengan menggabungkan varian EfficientNetV2 dan menerapkan augmentasi gambar yang komprehensif, penelitian ini berupaya untuk meningkatkan ketahanan model, akurasi, dan generalisasi secara khusus untuk deteksi kanker paru-paru. Selain itu, Grad-CAM akan meningkatkan kemampuan interpretasi, memberikan wawasan ke dalam proses pengambilan keputusan model. Penelitian ini bertujuan untuk

membangun kerangka kerja diagnosis kanker paru-paru yang kuat melalui teknik-teknik ini, meningkatkan akurasi dan keandalan analisis citra histopatologi dan pada akhirnya berkontribusi pada alur kerja diagnostik yang lebih baik dan strategi perawatan pasien.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang di atas, rumusan masalah dapat dijabarkan sebagai berikut:

- a) Bagaimana *augmentation* gambar histopatologi dapat meningkatkan kinerja model klasifikasi kanker paru-paru?
- b) Bagaimana pengaruh *Ensemble Learning EfficientNetV2* terbaik dalam meningkatkan kinerja model?
- c) Bagaimana Teknik interpretasi *Grad-CAM* dapat digunakan untuk meningkatkan *explainability* atas keputusan klasifikasi model ensemble?

## **1.3. Tujuan dan Manfaat**

Penelitian ini bertujuan untuk meningkatkan akurasi, ketahanan model, dan generalisasi dalam klasifikasi citra histopatologi kanker paru-paru dengan memanfaatkan ensembel varian EfficientNetV2 serta teknik augmentasi data yang lebih komprehensif. Dengan menerapkan metode Grad-CAM, penelitian ini juga berupaya meningkatkan interpretabilitas model, sehingga keputusan klasifikasi dapat lebih transparan dan dapat dipercaya dalam lingkungan medis.

Manfaat dari penelitian ini meliputi pengembangan model klasifikasi yang lebih andal dalam mendeteksi berbagai jenis kanker paru-paru, yang pada akhirnya dapat membantu tenaga medis dalam mendiagnosis penyakit dengan lebih akurat dan efisien. Selain itu, hasil penelitian ini diharapkan dapat berkontribusi terhadap penelitian lebih lanjut dalam bidang kecerdasan buatan untuk aplikasi medis, khususnya dalam analisis citra histopatologi.

## **1.4. Batasan Masalah**

Dalam penelitian ini, beberapa batasan diterapkan sebagai berikut:

- a) Penelitian ini hanya menggunakan dataset LC25000, dengan fokus pada tiga kelas kanker paru: lung adenocarcinoma, lung benign tissue, dan lung squamous cell carcinoma, berjumlah 15.000 citra.
- b) Model hanya diuji menggunakan dataset tanpa validasi eksternal pada dataset lain atau citra histopatologi dari rumah sakit/institusi medis. Hal ini disebabkan keterbatasan akses terhadap data medis yang lebih luas.
- c) Augmentasi data hanya menggunakan teknik dasar seperti rotasi, pergeseran lebar, shear, zoom, flipping vertikal, dan penyesuaian kecerahan, tanpa eksplorasi teknik lanjutan, karena keterbatasan sumber daya komputasi dan waktu pelatihan yang diperlukan.

### **1.5. Metode Penelitian**

Penelitian ini dilakukan dengan pendekatan studi literatur, simulasi, analisis statistik, dan implementasi model deep learning. Studi literatur digunakan untuk memahami penerapan EfficientNetV2 dan teknik ensemble learning dalam klasifikasi citra histopatologi kanker paru serta metode interpretabilitas model dengan Grad-CAM.

Dataset yang digunakan adalah LC25000, dengan fokus pada tiga kelas kanker paru: lung adenocarcinoma, lung benign tissue, dan lung squamous cell carcinoma. Setelah dilakukan preprocessing untuk menghapus duplikasi data, dataset dibagi menjadi 60% data latih, 20% data validasi, dan 20% data uji. Augmentasi data diterapkan menggunakan rotasi, shear, zoom, flipping vertikal, dan penyesuaian kecerahan untuk meningkatkan generalisasi model.

Implementasi model menggunakan EfficientNetV2-B0 hingga B3, yang dilatih dengan transfer learning dari ImageNet, Adam optimizer (learning rate 0.001, batch size 128, 10 epoch), serta teknik ensemble learning (hard voting, soft voting, weighted voting) untuk meningkatkan akurasi.

Evaluasi dilakukan dengan metrik akurasi, presisi, recall, F1-score, serta confusion matrix untuk menganalisis kesalahan klasifikasi. Selain itu, Grad-CAM digunakan untuk mengevaluasi area citra yang menjadi fokus model dalam pengambilan keputusan. Hasil penelitian diharapkan dapat menghasilkan model klasifikasi kanker paru yang akurat dan transparan untuk diterapkan dalam sistem pendukung keputusan medis..

### 1.6. Jadwal Pelaksanaan

Jadwal pelaksanaan akan menjadi acuan dalam mengevaluasi tahap-tahap pekerjaan seperti yang tertuang dalam milestone yang sudah ditetapkan. Jadwal pada Table 1.1 berfungsi sebagai panduan dalam kegiatan untuk mengetahui apa yang harus dilakukan dan kapan harus dilakukan.

**Tabel 1.1. Contoh Jadwal Pelaksanaan Tugas Akhir.**

| No. | Deskripsi Tahapan                      | Bulan<br>1 | Bulan<br>2 | Bulan<br>3 | Bulan<br>4 | Bulan<br>5 | Bulan<br>6 |
|-----|--|------------|------------|------------|------------|------------|------------|
| 1   | Studi Literatur                        |            |            |            |            |            |            |
| 2   | Pengumpulan Data                       |            |            |            |            |            |            |
| 3   | Perancangan Sistem                     |            |            |            |            |            |            |
| 4   | Pelatihan model                        |            |            |            |            |            |            |
| 5   | Pengujian dan Analisa hasil penelitian |            |            |            |            |            |            |
| 6   | Penyusunan Laporan/Buku TA             |            |            |            |            |            |            |