

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pisang merupakan jenis tanaman yang paling banyak diproduksi pada negara Indonesia, berdasarkan data dari Badan Pusat Statistik (BPS), produksi pisang menyumbang sekitar 64% dari seluruh produksi buah nasional di Indonesia pada tahun 2024 dengan total produksi pisang sebesar 96.865.994 kuintal atau 9.686.599 ton dalam tahun 2024 (Badan Pusat Statistik, 2025). Tanaman pisang menjadi sumber pendapatan bagi petani lokal karena pisang dapat tumbuh subur dalam berbagai kondisi tanah dan tahan terhadap kekeringan, selain itu hampir semua bagian dari tanaman pisang dapat digunakan seperti daun dan batang untuk pakan ternak dan pembungkus tanaman, jantung pisang untuk sayur, dan buahnya untuk konsumsi. Tanaman pisang dapat menjadi sumber pangan alternatif karena pisang dapat panen hampir sepanjang tahun sehingga menjadi sumber pangan yang stabil (Henuk dkk, 2020).

Dalam menanam tumbuhan pisang, terdapat ancaman seperti penyakit akibat jamur dan bakteri seperti penyakit yang terdapat pada daun pisang seperti *cordana*, *pestalotipsis*, dan *sigatoka* yang dimana dapat mengancam hasil dari produksi panen pisang sehingga kualitas dari pisang yang dapat menurun dan juga penurunan jumlah hasil produksi panen dari pisang sehingga dapat merugikan perekonomian bagi petani pisang dan gangguan pada ketahanan pangan bagi masyarakat jika penyakit pada tanaman pisang tidak ditangani dengan tepat (Henuk dkk, 2020).

Terdapat tantangan dalam mengidentifikasi penyakit pada tanaman pisang secara manual yaitu dibutuhkan seseorang dengan keahlian khusus dalam mengidentifikasi penyakit pada tanaman pisang, lalu proses yang dilakukan membutuhkan waktu yang lama, dan seringkali terjadi kesalahan identifikasi penyakit sehingga dibutuhkan suatu pendekatan dalam mengidentifikasi penyakit pada tanaman pisang agar proses identifikasi penyakit yang dilakukan menjadi lebih efektif.

Pada perkembangan teknologi yang pesat ini, pendekatan *deep learning* telah membuat peluang baru pada beberapa bidang aplikasi, seperti analisis citra medis, prediksi penyakit, sistem pendukung keputusan medis, dan robot bedah (Yan, 2025). Salah satu langkah yang dilakukan untuk mencegah terjadinya serangan hama dan penyakit pada tanaman pisang pada penelitian ini yaitu dengan mengimplementasikan suatu model dengan pendekatan *deep learning* yang dapat mengidentifikasi penyakit pada daun pisang.

Beberapa peneliti yang menggunakan pendekatan *deep learning* untuk mengidentifikasi penyakit pada tanaman dengan metode yang berbeda-beda seperti pada penelitian oleh (Srivastav dkk, 2024), menggunakan *Convolutional Neural Network* (CNN) untuk mengidentifikasi penyakit pada daun pisang dengan akurasi tertinggi sebesar 96.52%. Penelitian oleh (Tiwari, 2024) yaitu menggunakan model *Vision Transformer* (ViT) untuk identifikasi penyakit pada daun pisang dengan akurasi sebesar 96.88%, dengan ViT-B16 menjadi model akurat yang mengalahkan *Graph Convolutional Networks* (GCN), *EfficientNet-B3*, *MobileNet-V2*, dan *ResNet-cut*. Pada penelitian oleh (Das dkk, 2020) menggunakan beberapa metode

untuk mendeteksi penyakit pada daun tomat dengan akurasi tertinggi didapatkan menggunakan metode *Support Vector Machine* (SVM) dengan akurasi sebesar 87.6%, lalu metode *Random Forest* sebesar 70.05%, dan pada metode *Logistic Regression* sebesar 67.3%. Selain menggunakan metode tunggal, adapun penelitian oleh (Thakur dkk, 2023) yang menggunakan metode *hybrid* CNN-ViT untuk mendeteksi penyakit pada daun tanaman dengan akurasi sebesar 98.86% yang dimana mendapatkan akurasi tertinggi dibandingkan penelitian dengan metode lainnya dan dijadikannya metode *hybrid* CNN-ViT menjadi metode yang digunakan dalam penelitian ini.

Model CNN memiliki kemampuan secara otomatis mengekstrak fitur dari gambar, menangani jumlah data yang banyak atau besar, mendapatkan akurasi yang tinggi pada tugas identifikasi gambar yang kompleks, dan dapat dengan mudah untuk beradaptasi pada gambar baru dan tugas yang spesifik, sehingga membuat model CNN lebih unggul dibandingkan metode pemrosesan gambar tradisional (Rayen dkk, 2024). Model CNN memiliki kekurangan seperti kesulitan untuk memprediksi secara akurat pada lingkungan yang kompleks, fokus pada fitur yang tidak relevan, dan kurang efektif dalam menangani *noise* yang tinggi dan kondisi dunia nyata yang bervariasi (Smit dkk, 2024), selain itu model CNN memiliki kekurangan seperti rentan terhadap *overfitting*, kesulitan menangani duplikasi data, dan dapat menghasilkan hasil yang bias jika data kelas tidak seimbang (Sunarjo dkk, 2023).

Model ViT memiliki keunggulan dalam menangkap depensi jarak jauh, hanya memerlukan bias induktif yang sedikit, dan mendukung pemrosesan paralel,

sehingga fitur-fitur ini memungkinkan ViT mengungguli model tradisional dalam tugas identifikasi gambar (Wang dkk, 2025). Selain itu, model ViT juga memiliki kekurangan pada modelnya yang rumit sehingga dapat melemahkan proses pengambilan keputusan yang menyebabkan peta atensi menjadi ambigu (P. dan B., 2024).

Metode *hybrid* CNN-ViT memiliki kekurangan yaitu adanya ketergantungan pada *pre-training* dengan *dataset* besar sehingga jika dilatih dari awal tanpa adanya *pre-training* maka kinerja model akan menurun, lalu terdapat modul *Multi-head Self Attention* (MSA) pada Transformer yang membutuhkan banyak parameter dan komputasi yang tinggi sehingga dapat menyebabkan *overfitting* pada *dataset* yang kecil (Chen dkk, 2024). Sementara keunggulan dari metode *hybrid* CNN-ViT yaitu dapat menggunakan pengenalan pola yang dimiliki CNN yang dimana unggul dalam ekstraksi fitur lokal seperti tepi daun dan pola bercak pada daun, lalu pengenalan pola yang dimiliki ViT yang dimana unggul dalam menangkap hubungan global antar bagian gambar seperti daun yang berwarna kuning secara menyeluruh dan bercak yang tersebar pada seluruh bagian daun, sehingga dapat meningkatkan ketahanan terhadap variasi pada citra dan mendapatkan akurasi yang tinggi (P. dan B., 2024). Penggunaan metode *hybrid* CNN-ViT dapat meningkatkan sensitivitas terhadap perubahan halus dan meningkatkan akurasi identifikasi gambar pada model (Zhao dkk, 2024). Pada penelitian (Syihad dkk, 2023) menggunakan model CNN untuk mengidentifikasi penyakit pada daun pisang dengan akurasi sebesar 94% dan penelitian (Saini, 2024) menggunakan model ViT untuk mengidentifikasi penyakit pada daun pisang dengan akurasi sebesar 96.88%

yang dimana hanya menggunakan model tunggal seperti CNN atau ViT saja sehingga akurasi yang didapatkan masih belum maksimal, sehingga dengan digunakannya metode *hybrid* CNN-ViT yang menggabungkan keunggulan ekstraksi fitur lokal dari model CNN dan ekstraksi fitur global dari model ViT maka diharapkan dapat meningkatkan akurasi model dalam mengidentifikasi penyakit pada daun pisang, selain itu penggunaan arsitektur *hybrid* CNN-ViT masih sangat terbatas khususnya dalam konteks identifikasi penyakit pada daun pisang. Pada penelitian (Thakur dkk, 2023) telah diusulkan metode ConViTX yaitu model *hybrid* yang menggabungkan arsitektur CNN dan ViT dengan jumlah parameter sebesar 0,85 juta dan mendapatkan akurasi sebesar 98.86% dalam mengidentifikasi penyakit pada daun tanaman, tetapi akurasi yang diperoleh masih belum maksimal, sehingga dengan digunakannya varian ViT-B16 dalam metode *hybrid* CNN-ViT yang menurut (Dosovitskiy dkk, 2021) merupakan varian model yang optimal dengan kompleksitas yang tidak terlalu tinggi, digunakannya varian ViT-B16 karena *dataset* yang digunakan tidak terlalu besar sehingga model dengan kompleksitas yang tidak terlalu tinggi seperti ViT-B16 atau ViT-Base sehingga model tersebut cocok dengan *dataset* yang tidak terlalu besar, tidak seperti varian ViT-Large ataupun ViT-Huge yang dimana memiliki kompleksitas yang sangat tinggi dibandingkan varian ViT-Base, sehingga diharapkan dapat meningkatkan akurasi pada identifikasi penyakit pada daun pisang.

Berdasarkan latar belakang yang telah disebutkan, maka penelitian ini akan mengimplementasikan sebuah model *hybrid* CNN-ViT dengan penggunaan varian ViT yaitu ViT-B16 yang diharapkan dapat mengidentifikasi penyakit pada daun

pisang secara akurat karena digabungkannya kedua keunggulan ekstraksi fitur lokal dari model CNN dan keunggulan ekstraksi fitur global dari model ViT dan penggunaan varian ViT-B16 yang dimana merupakan varian ViT yang optimal dan memiliki kompleksitas yang tidak terlalu tinggi, sehingga petani pisang nantinya dapat mengidentifikasi penyakit pada daun pisang menggunakan pendekatan *deep learning* dan dapat memproduksi hasil panen pisang yang berkualitas dengan jumlah panen yang banyak.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diurai, penelitian ini memiliki rumusan masalah yang dapat dispesifikasikan dalam bentuk pernyataan sebagai berikut:

1. Bagaimana mengimplementasikan metode *hybrid* pada model CNN-ViT dengan menggunakan varian dari ViT yaitu ViT-B16 untuk mengidentifikasi penyakit pada citra daun pisang dengan akurat?
2. Bagaimana pengaruh implementasi metode *hybrid* pada model CNN-ViT terhadap peningkatan akurasi, presisi, *recall*, dan *F1-score* dibandingkan pada model tunggal CNN atau ViT dalam identifikasi penyakit pada citra daun pisang?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, penelitian ini memiliki tujuan sebagai berikut:

1. Mengimplementasikan metode *hybrid* pada model CNN-ViT dengan menggunakan varian dari ViT yaitu ViT-B16 untuk mengidentifikasi penyakit pada citra daun pisang dengan akurat.

2. Mengukur metode *hybrid* pada model CNN-ViT terhadap peningkatan akurasi, presisi, *recall*, dan *F1-score* dibandingkan pada model tunggal CNN atau ViT dalam identifikasi penyakit pada citra daun pisang.

1.4 Manfaat Penelitian

Terdapat manfaat penelitian yang dapat dispesifikasikan kedalam pertanyaan sebagai berikut:

1. Pemodelan *hybrid* CNN-ViT dapat mengoptimalkan ekstraksi fitur spasial dan kontekstual, sehingga dapat menghasilkan model yang lebih akurat dan optimal.
2. Hasil identifikasi pada *hybrid* CNN-ViT memberikan informasi penyakit yang dialami oleh pohon pisang dengan mendeteksi penyakit melalui daun pisang.

1.5 Batasan Masalah

Terdapat batasan masalah dalam penelitian ini guna membatasi penelitian ke arah yang lebih spesifik, adapun batasan masalah yang berupa pertanyaan sebagai berikut:

1. Penelitian ini berfokuskan pada pengimplementasian model *hybrid* CNN-ViT dalam mengidentifikasi penyakit pada daun pisang dengan empat kelas yaitu *cordana*, *healthy*, *pestalotiopsis*, dan *sigatoka*.
2. *Dataset* yang digunakan yaitu berasal dari situs Kaggle dengan judul “Banana Leaf Spot Diseases Dataset” yang memiliki empat kelas yaitu *cordana*, *healthy*, *pestalotiopsis*, dan *sigatoka*.