BAB I
PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi komputasi dan ketersediaan data berskala besar telah mendorong perkembangan pesat dalam bidang kecerdasan buatan, khususnya pada ranah deep learning. Salah satu pendekatan yang paling berpengaruh dalam analisis data visual adalah *Convolutional Neural Network* (CNN). Model ini dirancang untuk meniru cara kerja sistem visual manusia dalam mengenali pola, tepi, tekstur, serta struktur objek pada citra. Berbeda dengan metode *machine learning* konvensional yang memerlukan proses ekstraksi fitur secara manual, CNN mampu melakukan pembelajaran fitur secara otomatis dan bertingkat melalui lapisan konvolusi, aktivasi, dan pooling. Menurut Darmawan et al. (2025), kemampuan CNN dalam membangun representasi fitur hierarkis memungkinkan model mengenali karakteristik sederhana pada lapisan awal dan fitur yang lebih kompleks pada lapisan yang lebih dalam. Hal ini menjadikan CNN sebagai metode yang sangat efektif dalam berbagai tugas seperti klasifikasi citra, deteksi objek, segmentasi gambar, pengenalan wajah, analisis video, serta integrasi dalam sistem berbasis perangkat bergerak dan *Internet of Things* (Kusnadi et al., 2024).

Perkembangan CNN tidak terlepas dari inovasi arsitektur yang terus diperbarui untuk meningkatkan akurasi dan efisiensi komputasi. Beberapa arsitektur yang banyak digunakan dalam penelitian dan aplikasi industri antara lain R*esNet, DenseNet, Inception, MobileNet*, dan *VGG* (Prayogo et al., 2025). Di antara berbagai arsitektur tersebut, VGG menempati posisi penting karena strukturnya yang relatif sederhana namun memiliki kedalaman jaringan yang cukup untuk mengekstraksi fitur kompleks secara optimal. VGG menggunakan kernel konvolusi berukuran kecil secara berulang, sehingga mampu memperdalam jaringan tanpa meningkatkan kompleksitas parameter secara drastis. Salah satu varian yang paling banyak digunakan adalah VGG-19, yang terdiri dari 19 lapisan pembelajaran dan dikenal memiliki kemampuan representasi fitur visual yang sangat detail.

Kedalaman VGG-19 memungkinkan model menangkap pola tekstur, gradasi warna, hingga bentuk objek secara lebih komprehensif dibandingkan model yang lebih dangkal. Selain itu, struktur arsitekturnya yang konsisten memudahkan proses modifikasi serta integrasi dengan teknik tambahan seperti transfer learning. *Transfer learning* memungkinkan model yang telah dilatih pada dataset berskala besar untuk diaplikasikan pada dataset baru dengan proses pelatihan ulang yang lebih efisien. Penelitian oleh Raharja et al. (2025) menunjukkan bahwa VGG-19 mampu mencapai akurasi 96,9 persen dalam klasifikasi citra medis berbasis radiografi. Sementara itu, Handayani et al. (2024) melaporkan akurasi 94,7 persen dalam klasifikasi penyakit tanaman berbasis citra daun. Dalam bidang industri manufaktur, Wijanarko et al. (2025) juga menunjukkan bahwa VGG-19 mampu mendeteksi cacat produk dengan tingkat

akurasi di atas 95 persen. Hasil-hasil tersebut menunjukkan bahwa VGG-19 memiliki stabilitas performa yang tinggi pada berbagai domain permasalahan.

Meskipun memiliki performa yang kuat, arsitektur VGG-19 tetap memiliki keterbatasan, terutama dalam hal efisiensi dan kemampuan memfokuskan perhatian pada fitur yang benar-benar relevan. Dalam banyak kasus, model CNN dapat memproses seluruh informasi dalam citra tanpa membedakan bagian mana yang lebih penting untuk proses klasifikasi. Untuk mengatasi permasalahan tersebut, berbagai penelitian mengembangkan mekanisme perhatian atau attention mechanism. Mekanisme ini dirancang untuk meningkatkan sensitivitas model terhadap fitur signifikan sekaligus mengurangi pengaruh informasi yang kurang relevan. Beberapa modul perhatian yang telah dikembangkan meliputi *Squeeze-and-Excitation*, *Self-Attention*, *Residual Attention Network*, *Spatial Transformer Network*, serta *Convolutional Block Attention Module* atau *CBAM* (Firmansyah et al., 2025).

CBAM menjadi salah satu modul perhatian yang populer karena pendekatannya yang sederhana namun efektif. Modul ini bekerja melalui dua tahapan utama, yaitu *channel attention* dan *spatial attention*. *Channel attention* bertujuan untuk menentukan saluran fitur mana yang paling berkontribusi terhadap proses klasifikasi, sedangkan *spatial attention* berfokus pada lokasi atau area penting dalam citra. Dengan kombinasi kedua mekanisme tersebut, model mampu meningkatkan akurasi tanpa perlu menambahkan parameter yang terlalu besar. Menurut Maulana et al. (2024), integrasi CBAM pada arsitektur CNN dapat meningkatkan kemampuan diskriminatif model secara signifikan. Penelitian oleh Oktaviani et al. (2025) menunjukkan bahwa penerapan CBAM pada klasifikasi tumor otak menghasilkan akurasi 98,8 persen. Sementara itu, Prakoso et al. (2024) melaporkan peningkatan akurasi hingga 97,6 persen dalam deteksi kerusakan mesin industri setelah menambahkan modul CBAM.

Dalam sektor pertanian, penerapan teknologi *deep learning* memiliki urgensi yang tinggi, terutama dalam deteksi penyakit tanaman. Penyakit daun yang tidak teridentifikasi sejak dini dapat menyebar dengan cepat dan menyebabkan kerugian ekonomi yang besar. Proses identifikasi manual sering kali memerlukan keahlian khusus serta waktu yang tidak singkat, sehingga kurang efisien pada skala produksi besar. Implementasi CNN dalam sistem berbasis perangkat seluler memungkinkan proses identifikasi dilakukan secara otomatis dan *real-time*, sehingga petani dapat segera mengambil tindakan pencegahan atau penanganan (Setyawan et al., 2023). Dengan dukungan arsitektur yang lebih dalam seperti VGG-19 serta penambahan modul perhatian seperti CBAM, sistem deteksi penyakit daun diharapkan mampu menghasilkan klasifikasi yang lebih akurat dan tahan terhadap variasi pencahayaan, latar belakang, serta kondisi lingkungan.

Selain peran pentingnya dalam bidang pertanian, daun juga memiliki nilai strategis dalam mendukung keberlanjutan pangan dan lingkungan. Anindita et al. (2024) menyatakan bahwa daun memiliki potensi sebagai sumber protein nabati alternatif yang dapat dimanfaatkan untuk meningkatkan ketahanan pangan global. Di sisi lain, Haryanto et al. (2023) menunjukkan bahwa beberapa jenis daun memiliki kemampuan

menyerap fosfat dari air limbah, sehingga berkontribusi dalam upaya pengendalian pencemaran lingkungan secara ramah dan berkelanjutan.

Berdasarkan berbagai kajian tersebut, integrasi modul CBAM ke dalam arsitektur VGG-19 menjadi pendekatan yang relevan dan inovatif untuk meningkatkan performa klasifikasi penyakit daun tanaman. Kombinasi kedalaman jaringan VGG-19 dengan kemampuan selektif CBAM dalam memfokuskan perhatian pada fitur penting diharapkan menghasilkan model yang lebih adaptif, presisi, dan robust terhadap variasi data. Pengembangan model ini tidak hanya memberikan kontribusi teoretis dalam pengayaan arsitektur deep learning, tetapi juga memiliki implikasi praktis dalam mendukung pertanian berbasis teknologi, peningkatan produktivitas, serta keberlanjutan sistem pangan di era digital.

Berdasarkan pada uraian yang ada di atas membuat peneliti tertarik untuk melakukan penelitian dengan judul: **Deteksi Penyakit Daun Tanaman dengan Menggunakan VGG19 yang Dilengkapi dengan Mekanisme Perhatian CBAM**

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian yang telah diuraikan sebelumnya, perkembangan teknologi *deep learning* khususnya pada arsitektur *Convolutional Neural Network* menunjukkan potensi besar dalam meningkatkan akurasi klasifikasi citra. Namun, model CNN konvensional masih memiliki keterbatasan dalam memfokuskan perhatian pada fitur yang paling relevan dari suatu citra, sehingga berpotensi menurunkan performa klasifikasi pada kondisi data yang kompleks. Oleh karena itu, penelitian ini merumuskan permasalahan utama: Bagaimana pengaruh penerapan modul perhatian *Convolutional Block Attention Module* (CBAM) terhadap peningkatan akurasi klasifikasi penyakit daun tanaman pada arsitektur VGG19?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah dan tidak menyimpang dari tujuan utama, maka ruang lingkup penelitian dibatasi pada beberapa aspek berikut:

1. Objek penelitian difokuskan pada deteksi dan klasifikasi penyakit daun pada tanaman pisang, jeruk, dan kentang berdasarkan citra digital daun.
2. Model yang digunakan dalam penelitian ini adalah arsitektur VGG19 yang diintegrasikan dengan modul perhatian CBAM.
3. Data yang digunakan merupakan data sekunder yang diperoleh dari platform Kaggle.
4. Proses penelitian difokuskan pada tahap pelatihan, pengujian, dan evaluasi performa model berdasarkan metrik akurasi.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menganalisis pengaruh integrasi modul perhatian CBAM terhadap peningkatan kinerja model VGG19 dalam mengklasifikasikan penyakit daun tanaman.

2. Mengembangkan model klasifikasi penyakit daun berbasis *deep learning* yang memiliki tingkat akurasi lebih tinggi melalui optimalisasi mekanisme perhatian.
3. Mengevaluasi efektivitas kombinasi VGG19 dan CBAM dalam meningkatkan kemampuan model dalam mengekstraksi fitur yang relevan pada citra daun.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Teoretis
   Memberikan kontribusi ilmiah dalam pengembangan arsitektur CNN berbasis modul perhatian, khususnya dalam penerapan CBAM pada model VGG19 untuk klasifikasi citra penyakit tanaman.
2. Manfaat Praktis
   Menghasilkan model deteksi penyakit daun yang lebih akurat sehingga dapat membantu proses identifikasi penyakit secara lebih cepat dan efisien.
3. Manfaat Akademis
   Menjadi referensi bagi peneliti selanjutnya dalam mengimplementasikan modul perhatian pada model CNN, khususnya dalam penelitian yang berkaitan dengan sistem deteksi dan pengendalian penyakit tanaman berbasis citra digital.

1.6 Tinjauan Pustaka

1.6.1 *Deep Learning*

Deep Learning merupakan cabang dari *machine learning* yang menggunakan jaringan saraf tiruan berlapis banyak untuk mempelajari pola kompleks dari data dalam jumlah besar. Pendekatan ini memungkinkan sistem melakukan pembelajaran fitur secara otomatis tanpa perlu proses ekstraksi fitur manual. Menurut Santosa et al. (2023), *deep learning* bekerja dengan membangun representasi data secara bertingkat, di mana setiap lapisan jaringan mengekstraksi karakteristik yang semakin abstrak dan kompleks.

Selanjutnya, Wibowo et al. (2024) menjelaskan bahwa keunggulan utama *deep learning* terletak pada kemampuannya menangani data tidak terstruktur seperti citra, audio, dan teks dengan tingkat akurasi yang tinggi. Dukungan perkembangan perangkat keras seperti *Graphics Processing Unit* serta ketersediaan *big data* turut mempercepat proses pelatihan model. Dengan karakteristik tersebut, *deep learning* telah banyak diterapkan dalam berbagai bidang seperti analisis citra medis, sistem rekomendasi, pengenalan suara, serta sistem klasifikasi berbasis citra digital.

1.6.2 *Convolutional Neural Network* (CNN)

Convolutional Neural Network atau CNN merupakan salah satu arsitektur dalam *deep learning* yang secara khusus dirancang untuk mengolah data berbentuk grid seperti gambar. CNN memanfaatkan operasi konvolusi untuk mengekstraksi fitur lokal seperti tepi, tekstur, dan pola tertentu pada citra. Menurut Prasetyo et al. (2024), CNN memiliki kemampuan untuk melakukan ekstraksi fitur secara otomatis melalui beberapa lapisan seperti *convolution layer*, *activation function*, *pooling layer*, dan *fully connected layer*.

Lebih lanjut, Hidayat et al. (2023) menyatakan bahwa keunggulan CNN dibandingkan metode klasifikasi tradisional terletak pada kemampuannya mempertahankan informasi spasial pada citra. Dengan struktur tersebut, CNN mampu menghasilkan performa yang tinggi dalam berbagai tugas seperti klasifikasi gambar, deteksi objek, serta segmentasi citra.

1.6.3 VGG19

VGG19 merupakan salah satu arsitektur CNN yang dikembangkan untuk meningkatkan kedalaman jaringan guna memperoleh representasi fitur yang lebih detail. Arsitektur ini menggunakan kernel berukuran kecil secara berulang untuk memperdalam jaringan tanpa meningkatkan kompleksitas parameter secara signifikan. Menurut Ramadhan et al. (2025), VGG19 memiliki 19 lapisan pembelajaran yang memungkinkan model menangkap pola visual secara lebih komprehensif dibandingkan arsitektur yang lebih dangkal.

Selain itu, Putri et al. (2024) menjelaskan bahwa struktur VGG19 yang konsisten memudahkan penerapan teknik *transfer learning*, sehingga model dapat digunakan kembali pada dataset yang berbeda dengan proses pelatihan ulang yang lebih efisien. VGG19 telah banyak digunakan dalam penelitian klasifikasi citra medis, deteksi penyakit tanaman, serta identifikasi objek industri karena stabilitas dan akurasinya yang tinggi.

1.6.4 *Convolutional Block Attention Module* (CBAM)

Convolutional Block Attention Module atau CBAM merupakan modul perhatian yang dirancang untuk meningkatkan kemampuan CNN dalam memfokuskan perhatian pada fitur yang relevan. Modul ini bekerja melalui dua mekanisme utama, yaitu *channel attention* dan *spatial attention*. Menurut Nugroho et al. (2024), *channel attention* berfungsi untuk menentukan saluran fitur yang paling berpengaruh terhadap proses klasifikasi, sedangkan *spatial attention* menyoroti area penting dalam citra.

Selanjutnya, Saputra et al. (2025) menyatakan bahwa integrasi CBAM ke dalam arsitektur CNN mampu meningkatkan akurasi model tanpa menambah kompleksitas parameter secara signifikan. Dengan kombinasi mekanisme perhatian tersebut, model menjadi lebih selektif dalam memproses informasi sehingga menghasilkan performa klasifikasi yang lebih optimal dan stabil pada berbagai kondisi data.