BAB I
PENDAHULUAN

1.1. Latar Belakang

Indonesia merupakan negara yang terletak di kedua sisi garis khatulistiwa, yang merupakan lokasi ideal untuk menanam kopi karena cuaca dan kondisi tanaman yang stabil di wilayah *Bean Belt*. Kopi merupakan tanaman penting di Indonesia, berperan penting dalam perekonomian negara dan merupakan barang ekspor utama yang menghasilkan devisa. Pada tahun 2021, total produksi kopi di Indonesia mencapai 762.000 ton. Di antaranya perkebunan BUMN menyumbang 0.49% atau 3.700 ton, sedangkan perkebunan swasta menyumbang 0.18% atau 1.400 ton [1].

Ekspor kopi Indonesia hanya mampu memenuhi 4% permintaan Uni Eropa karena banyak faktor. Menurut studi yang dilakukan oleh *Center for Indonesia Policy Studies* (CIPS) di bidang pangan, setidaknya ada dua penyebab rendahnya produksi kopi di Indonesia. Pasalnya, pohon kopi yang sudah tua lebih rentan terserang penyakit sehingga dapat mengganggu pertumbuhan dan produktivitas [2].

Perubahan fisik pada daun tanaman kopi dapat terjadi karena berbagai faktor, seperti serangan hama dan kondisi cuaca yang tidak dapat diprediksi. Penyakit seperti karat dan tungau laba-laba merah juga dapat menyerang daun tanaman kopi, terutama karena kelembapan, hujan, dan jamur *Hemileia Vastatrix*. Untuk mengidentifikasi atau memvisualisasikan penyakit-penyakit ini, ada banyak kemajuan teknologi inovatif, termasuk tes konvolusi langsung dan tes laboratorium, telah dikembangkan. Namun, penting untuk mengenai kondisi patologis agar dapat mengenalinya melalui pengamatan langsung [3].

Proses klasifikasi melibatkan pemilahan setiap data gambar ke dalam kategori tertentu berdasarkan karakteristiknya. Ini adalah teknik penting yang digunakan untuk pengenalan gambar. Klasifikasi adalah metode *data mining* yang memungkinkan identifikasi item dalam dataset dan mengklasifikasikannya menurut kelas atau levelnya [4]. Tujuan utama klasifikasi adalah untuk memprediksi kelas target untuk setiap *instance* dalam kumpulan data. Kelas target ini biasanya dikenal sebagai dataset pelatihan, dimana algoritma klasifikasi menganalisis data yang prosesnya dikenal sebagai pelatihan. Selama fase pelatihan, algoritma menghasilkan keluaran model yang dapat digunakan untuk mengklasifikasikan data baru [5].

Saat ini, *deep learning* terus menjadi pendekatan *machine learning* yang memfasilitasi pengenalan dan deteksi otomatis objek dalam gambar digital [6]. Penulis dalam penelitian ini

menggunakan metode *deep learning*, khususnya *Convolutional Neural Network* (CNN) yang banyak digunakan untuk mengidentifikasi dan mengklasifikasikan objek yang berbeda dalam input gambar yang berbeda secara bersamaan.

CNN mencakup banyak model arsitektur, termasuk VGG-19, yang akan digunakan dalam penelitian ini. Terlepas dari keragaman dalam arsitektur CNN, semuanya memiliki tujuan yang sama. Beberapa desain yang paling umum digunakan untuk tantangan klasifikasi gambar antara lain VGG, ResNet, AlexNet, GoogleNet, MobileNet, dan lainnya [7].

Penelitian sebelumnya telah menggunakan metode CNN untuk mengklasifikasikan penyakit pada daun. Dalam salah satu penelitian ini [5], gambar penyakit daun anggur diklasifikasikan menggunakan algoritma segmentasi *K-Means Clustering*. Ekstraksi fitur dilakukan dengan metode *transfer learning* VGG-16, dan klasifikasi dilakukan dengan menggunakan algoritma CNN. Dataset yang digunakan terdiri dari 4.000 gambar daun anggur yang diperoleh dari *website* Kaggle, dengan tambahan 100 data gambar yang bersumber dari Google yang digunakan sebagai data uji di luar dataset tersebut. Model CNN yang digunakan mencapai akurasi 97.25% dengan menggunakan data uji, sedangkan akurasi 95% dicapai dengan menggunakan data uji di luar dataset.

Dalam studi [8], peneliti menggunakan model VGG dan ResNet50 untuk memprediksi umur daun teh, menggunakan algoritma CNN untuk memproses gambar digital. Para peneliti juga mengumpulkan 600 gambar dan menggunakan pengoptimal RMSprop, *learning rate* 0.01, *batch size* 32, dan 100 *epochs* untuk menyegmentasi 1800 gambar. Keakuratan sistem dievaluasi melalui data pengujian dan keakuratannya mencapai 97.5% pada data uji.

Pada penelitian [9], teknik CNN digunakan untuk mengklasifikasikan citra daun mangga. Pengambilan citra diambil dari tiga kategori mangga manalagi, dan 90% data digunakan untuk *training*, sedangkan 10% digunakan untuk validasi. Data *training* diulang 60 kali untuk menghasilkan akurasi terbaik. Data validasi memiliki akurasi 89.20%, sedangkan data pelatihan memiliki akurasi sebesar 97.725%.

Menurut penelitian [10], penyakit daun padi dibagi menjadi tiga jenis, yaitu *Bacterial Leaf Blight*, *Brown Spot*, dan *Leaf Fungus*. Metode *transfer learning* dengan model terlatih ResNet101 diterapkan dalam penelitian ini, bersamaan dengan penambahan lapisan arsitektur seperti lapisan *dense*, *dropout*, dan lapisan *batch normalization*. Data validasi menunjukkan tingkat akurasi sebesar 100%, dengan nilai *loss* 5.61%.

Pada artikel sebelumnya [2], *platform* pemrograman MATLAB dan metode CNN digunakan untuk mendeteksi penyakit tanaman kopi pada data citra. Penelitian ini menggunakan 300 data citra yang dikategorikan menjadi tiga kelas, yaitu *health*, *rust*, dan *red spider mite*. Penelitian ini memperoleh tingkat akurasi 80.56% pada data latih yang menggunakan 240 data citra, sedangkan akurasi 81.60% berhasil dicapai pada data uji dengan menggunakan 60 data citra.

Penelitian ini bertujuan untuk mendeteksi citra penyakit pada daun tanaman kopi menggunakan *platform* pemrograman *Python* dengan menggunakan salah satu model CNN yaitu VGG-19 yang dikenal dengan kemampuan klasifikasi citra yang sangat baik dan pernah menjadi juara pada *ImageNet Challenge* 2014. VGG-19 terdiri dari 16 lapisan konvolusi, dan 3 lapisan *fully-connected*, yang menjadikannya jaringan yang kuat. Sedangkan CNN terdiri dari tiga lapisan, yaitu *convolutional*, *pooling*, dan *fully-connected layer* [11].

Perbedaan antara penelitian yang diusulkan dan penelitian sebelumnya adalah penelitian sebelumnya menggunakan MATLAB sebagai *platform* mereka untuk mengimplementasikan program, dan juga artikel tersebut tidak melakukan *contouring* untuk mengklasifikasikan data citra, dan juga pada penelitian sebelumnya hanya menggunakan satu model, yaitu AlexNet. Alasan penulis memasukkan model ResNet50 hanya pada bab hasil dan pembahasan adalah karena selain untuk mengatasi masalah gradien dengan menambah *output* dari *layer* sebelumnya ke *layer* berikutnya, ResNet50 juga dapat menghapus struktur jaringan yang redundan untuk mempercepat pengolahan data, yang mana dapat menyebabkan masalah dengan gradien. Hasilnya, penulis melakukan perbandingan antara dua model, VGG-19 dan ResNet50. VGG-19 memiliki beberapa keunggulan dibandingkan dengan ResNet50 dan model yang digunakan dalam jurnal penelitian utama, AlexNet, diantaranya model VGG-19 menggunakan banyak lapisan konvolusional dengan filter kecil dan penyatuan lapisan untuk meningkatkan efektivitas klasifikasi citra. VGG-19 juga menggunakan fungsi ReLU untuk meminimalkan kesalahan *backpropagation* dan meningkatkan kinerja jaringan, dan menggunakan lapisan terakhir, yaitu *softmax* untuk mencegah *overfitting*.

1.2. Rumusan Masalah

Dari uraian latar belakang diatas, maka dapat dibuat rumusan masalah penelitian sebagai berikut:

1. Bagaimana hasil evaluasi dari data kontur dan model VGG-19 pada penelitian yang diajukan untuk meningkatkan nilai keakuratannya dalam klasifikasi penyakit daun kopi?

2. Bagaimana perbandingan model VGG-19 pada penelitian yang diajukan dengan model AlexNet pada penelitian sebelumnya?

1.3. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Untuk mendapatkan hasil evaluasi terbaik dari data kontur dan model VGG-19 agar meningkatkan nilai keakuratannya dalam mengklasifikasi penyakit daun kopi.
2. Untuk mendapatkan perbandingan dari model VGG-19 pada penelitian yang diajukan dengan model AlexNet pada penelitian sebelumnya.

1.4. Batasan Masalah

Dalam pengerjaan Tugas Akhir, permasalahan yang dibahas terbatas pada beberapa hal berikut ini:

1. Menggunakan data kontur dan model VGG-19 pada proses evaluasi data dalam klasifikasi penyakit daun kopi untuk mendapatkan akurasi yang lebih baik.
2. Dataset yang digunakan diambil dari situs Kaggle yang berjudul "*RoCoLe : A Robusta Leaf Images Dataset*" yang dapat diakses pada link berikut (https://www.kaggle.com/datasets/diegopgonzlez/rocoleoriginal)
3. Total dataset yang diambil dari situs Kaggle tersebut berjumlah 1560 tetapi berdasarkan penelitian sebelumnya, hanya diambil 100 citra per kelas dengan kriteria tertentu yaitu citra dengan fokus yang baik pada objek utama tanpa ada objek lain di sekitarnya.
4. Pengolahan data akan dilakukan menggunakan bahasa pemrograman *Python*.
5. Menggunakan *Google Colaboratory*.