

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan salah satu negara dengan jumlah penduduk tinggi di dunia, Indonesia menempati urutan keempat sebagai negara dengan jumlah penduduk terbanyak. Menurut Badan Pusat Statistik (BPS), pada Februari 2022, terdapat 135,61 juta penduduk Indonesia yang aktif bekerja. Dari jumlah tersebut mayoritas nya, yaitu 38,7 juta jiwa, terserap di sektor pertanian, yang menjadikannya sektor pekerjaan paling besar di Indonesia (Kusnandar, 2023). Salah satu komoditas pertanian yang paling banyak diproduksi di Indonesia yaitu padi. Indonesia dikenal sebagai lumbung padi Asia Tenggara bahkan dunia. Sebutan itu muncul karena produksinya yang tinggi. Data FAO mencatat produksi padi di Indonesia sebesar 54,6 juta ton, sementara produksi beras sebesar 36,45 juta ton (Dihni, 2022). Tetapi berdasarkan data BPS ,proyeksi luas panen dan hasil produksi padi untuk tahun 2023 menunjukkan penurunan. Luas panen diperkirakan sekitar 10,20 juta hektar, turun sebesar 2,45% dari tahun 2022 dan produksi padi diperkirakan mencapai 53,63 juta ton GKG, mengalami penurunan sekitar 2,05%. Penurunan tersebut bisa terjadi karena beberapa faktor, seperti serangan hama dan penyakit pada tanaman padi, kekeringan, bencana alam banjir, dan lain sebagainya. Food and Agriculture Organization (FAO) mencatat bahwa sekitar 20-40% kegagalan produksi pangan dunia disebabkan oleh hama dan penyakit, sementara International Rice Research Institute (IRRI) menyatakan bahwa petani mengalami kegagalan panen sekitar 37% dari tanaman padi setiap tahunnya akibat serangan tersebut (Khoiruddin, Junaidi, & Saputra, 2022).

Tanda-tanda adanya hama, penyakit bakteri pada tanaman padi dapat berupa pertumbuhan yang terhambat, perubahan warna pada daun atau batang tanaman, adanya bercak atau luka pada tanaman, dan terjadinya kerontokan pada daun. Hama dan penyakit bakteri padi dapat mengakibatkan gangguan dalam proses fotosintesis. Serangan penyakit pada fase generatif tanaman dapat memiliki dampak serius, termasuk mengganggu pengisian biji padi sehingga berdampak pada berat dan kualitas

hasil panen yang rendah. Kondisi ini bisa menjadi masalah yang dihadapi oleh para petani di Indonesia, karena menurut data BPS 2023, sebaran petani Indonesia masih didominasi oleh Gen X atau yang sudah berusia sekitar 40 hingga 55 tahun, dimana seiring bertambahnya usia dapat terjadi penurunan kemampuan penglihatan, ini dapat menjadi masalah jika para petani harus memantau kondisi padi secara langsung melalui pengamatan mata saja, yang dapat terjadi kesalahan. Oleh karena itu diperlukan tindakan untuk dapat membantu para petani agar bisa mendeteksi penyakit dan hama yang menyerang daun padi secara akurat, agar bisa mencegahnya sedini mungkin dan tidak sampai parah, sehingga dapat meningkatkan panen produksi padi lebih baik lagi. Dalam menghadapi tantangan ini, teknologi pemrosesan citra dan pembelajaran mesin menjadi alat yang efektif dalam proses klasifikasi citra jenis penyakit tanaman daun padi. Padi yang sehat dan sakit dapat diidentifikasi sejak awal melalui kondisi daunnya. Proses klasifikasi ini bertujuan untuk menghasilkan model yang dapat secara akurat mengenali dan mengklasifikasikan jenis citra yang sesuai dengan ciri-ciri yang sama atau berbeda.

Salah satu pendekatan yang efektif dan umum digunakan dalam klasifikasi jenis penyakit daun pada tanaman padi adalah metode *Support Vector Machine* (SVM) dengan menggunakan hasil ekstraksi fitur yang diperoleh melalui *deep learning*. Pendekatan ini masuk pada konsep *deep learning*, yakni subbidang dari *machine learning* yang memanfaatkan algoritma berdasarkan prinsip matematika dan beroperasi mirip dengan proses pemikiran manusia. Dalam ranah pengolahan citra, terdapat metode *Convolutional Neural Network* (CNN) yang merupakan metode *deep learning*, yang digunakan untuk mendapatkan fitur mendalam (*deep features*) dalam penelitian ini (Akram, Fayakun, & Ramza, 2023). CNN merupakan jenis jaringan saraf tiruan yang dikembangkan dari metode *multilayer perceptron* (MLP) yang didesain untuk mengolah data dua dimensi atau data citra dengan piksel yang besar. CNN memiliki kemampuan dalam meniru cara kerja saraf manusia dalam pengenalan citra sehingga dapat mengekstraksi pola-pola penting dari citra (Riti & Tandjung, 2022). CNN mengambil citra sebagai input, mengkonversinya ke dalam bentuk vektor, dan

melibatkan beberapa operasi dasar seperti konvolusi dan pooling. (Jaafari, Douzi, Douzi, & Hssina, 2021). *Deep feature* merujuk pada representasi fitur yang dihasilkan oleh lapisan-lapisan dalam arsitektur jaringan saraf. Jaringan saraf konvolusional memiliki kemampuan untuk secara otomatis mempelajari dan mengekstraksi fitur-fitur yang semakin kompleks seiring dengan peningkatan kedalaman lapisan. Proses ekstraksi *deep feature* ini melibatkan transformasi gambar mentah menjadi representasi yang lebih abstrak, seperti tepi, tekstur, dan pola yang spesifik. Hasil dari ekstraksi ini adalah vektor berdimensi tinggi. Vektor ini biasanya memiliki dimensi lebih dari satu, namun dapat diratakan menjadi vektor satu dimensi, yaitu melalui lapisan flatten, yang kemudian digunakan sebagai data masukan dalam SVM untuk melakukan klasifikasi. Arsitektur CNN cukup beragam, seperti AlexNet, GoogleNet, ResNet, VGG, dan DenseNet, dan lain sebagainya.

Sementara itu, SVM atau *Support Vector Machine*, dikenal sebagai metode dalam *supervised learning classification* dan *Support Vector Regression*. SVM mampu menangani permasalahan klasifikasi dan regresi dengan pendekatan linear maupun non-linear. (Yohannes & Rivian, 2022). Terdapat beberapa jenis proses pelatihan yang digunakan dalam metode SVM, seperti *Sequential Minimum Optimization (SMO)*, chunking (*Quadratic programming*), dan sebagainya. Dibandingkan dengan metode training SVM lainnya, SMO (*Sequential Minimal Optimization*) memiliki kelebihan dalam efisiensi dan kecepatan pelatihan SVM, terutama untuk dataset besar. SMO beroperasi dengan memecah masalah *quadratic programming (QP)* utama menjadi serangkaian submasalah QP yang lebih kecil. Pendekatan ini dilakukan untuk mengatasi kekurangan dari algoritma SVM yang mungkin mengalami masalah pada *quadratic programming (QP)* (Nurhidayat, Asmunin, & Suyatno, 2021). SMO memilih untuk menyelesaikan submasalah optimisasi yang paling kecil pada setiap iterasinya. Hal ini dilakukan dengan memanfaatkan dua *lagrange multipliers* tanpa memerlukan matriks besar, yang kemudian dioptimalkan secara bersama-sama untuk mencapai nilai optimal yang dibutuhkan guna memperbarui nilai SVM (Kurniadi & Haviana, 2016).

Berkat pendekatan ini, SMO mampu menghasilkan proses pelatihan SVM yang cepat dan efisien.

Beberapa penelitian terdahulu telah menerapkan metode SVM sebagai alat klasifikasi dan CNN sebagai alat ekstraksi fitur, seperti dalam penelitian” Plant Disease Identification and Classification Using Convolutional Neural Network and SVM” (Kibriya, Abdullah, & Nasrullah, 2022), “Classification of Plant Diseases Using Machine and Deep Learning” (Lamba, Gigras, & Dhull, 2021), dan “Deep Feature Base Rice Life Disease Identification Using Support Vector Machine” (Sethy, Barpanda, Rath, & Bahera, 2020). Penelitian pertama berfokus pada identifikasi penyakit daun dari berbagai tanaman dengan menggunakan CNN untuk ekstraksi fitur dan klasifikasi menggunakan SVM berbasis ECOC (Error-Correcting Output Codes) untuk mengubah masalah klasifikasi multi-kelas menjadi serangkaian masalah klasifikasi biner. Metode ini berhasil mencapai tingkat akurasi tinggi, 98,9% untuk daun anggur dengan VGG19 dan 98,8% untuk daun apel dengan AlexNet.

Penelitian kedua juga mengadopsi pendekatan yang serupa dengan menguji performa berbagai algoritma machine learning dan deep learning untuk klasifikasi penyakit tanaman menggunakan empat dataset berbeda: padi, cabai, tomat, dan kentang. Ekstraksi fitur dilakukan dengan metode Auto Color Correlogram (ACC), dan lima belas algoritma machine learning diuji. Pada dataset cabai, algoritma random forest memiliki akurasi terbaik mencapai 0.997, disusul dengan Sequential Minimal Optimization (SMO) dengan tingkat akurasi yang dekat, mencapai 0.996. Di sisi lain, dataset padi menampilkan perbandingan yang berbeda, dengan SMO dan random forest memiliki akurasi tinggi yang sama dengan akurasi masing-masing mencapai 0.867. Untuk dataset tomat, algoritma LibSVM mendapatkan akurasi terbaik 0.982, diikuti oleh SMO dengan akurasi 0.98. Sedangkan pada dataset kentang, algoritma LibSVM mendapatkan hasil terbaik dengan akurasi mencapai 0.993, diikuti oleh SMO dengan akurasi 0.99.

Penelitian ketiga memfokuskan pada masalah identifikasi penyakit daun padi di wilayah barat Odisha, India, menggunakan dataset berisi 5932 sampel. Metode yang

diterapkan serupa dengan penelitian sebelumnya, yaitu menggunakan arsitektur deep learning CNN untuk ekstraksi fitur dari gambar daun padi, dengan menggunakan berbagai arsitektur seperti AlexNet, VGG16, VGG19, GoogleNet, ResNet50, dan sebagainya. Setelah ekstraksi fitur dilakukan, langkah selanjutnya adalah klasifikasi menggunakan algoritma Support Vector Machine (SVM). Hasil penelitian menunjukkan bahwa kombinasi fitur mendalam dari DenseNet201 dengan SVM memberikan kinerja klasifikasi terbaik dengan nilai F1 score sebesar 0.9852, diikuti oleh arsitektur ResNet101 dengan SVM (F1 score = 0.9839), dilanjutkan dengan hasil arsitektur ResNet50 dengan SVM (F1 score=0.9838), kemudian hasil dari arsitektur AlexNet dengan SVM (F1 score=0.9763). Terakhir, hasil dari arsitektur VGG19 dengan SVM (F1 score=0,9761).

Dengan demikian, berdasarkan latar belakang tersebut, penelitian ini menggunakan metode *Convolutional Neural Networks* (CNN) untuk mengekstraksi fitur mendalam dari gambar daun padi dan menggunakan metode klasifikasi *Support Vector Machine* (SVM) dengan optimasi *Sequential Minimal Optimization* (SMO). Pendekatan ini tidak hanya memungkinkan proses yang terintegrasi dan efisien, tetapi juga berpotensi meningkatkan akurasi klasifikasi citra. Pertama, CNN digunakan untuk mengekstraksi fitur-fitur penting dari gambar daun padi secara otomatis seperti tepi, tekstur, warna dan pola lainnya dari gambar. Kemudian, fitur-fitur yang diekstraksi oleh CNN tersebut diolah lebih lanjut dan digunakan sebagai input untuk pelatihan model SVM menggunakan algoritma optimasi SMO. SVM kemudian memanfaatkan fitur-fitur ini untuk membangun model klasifikasi yang optimal. Dengan pendekatan ini, CNN bertanggung jawab untuk mengekstraksi fitur-fitur dari gambar yang menghasilkan ekstraksi *deep feature*, sementara SVM bertanggung jawab untuk mempelajari pola dari fitur-fitur tersebut dan membuat keputusan klasifikasi. Oleh karena itu, penelitian ini akan menggali lebih dalam penerapan metode tersebut pada klasifikasi penyakit daun tanaman padi, yang berjudul "KLASIFIKASI PENYAKIT DAUN PADI MENGGUNAKAN SUPPORT VECTOR MACHINE BERDASARKAN FITUR MENDALAM (DEEP FEATURE)". Penelitian ini

difokuskan pada penggunaan metode SVM dengan optimasi SMO sebagai klasifikasi utama. Lalu, Arsitektur CNN yang dipilih untuk mengekstrak citranya adalah AlexNet dan VGG19. Pemilihan ini, karena AlexNet dan VGG19 telah terbukti memiliki kemampuan yang baik dalam mengekstraksi fitur-fitur kompleks dari citra berdasarkan hasil penelitian sebelumnya yang mencapai akurasi tinggi. Selanjutnya, dataset diambil secara langsung di sawah padi daerah Kota Surabaya. Dataset ini terdiri dari 3 jenis kelompok, yaitu bacterial leaf blight, hama ulat, dan normal (tidak berpenyakit). Hasil penelitian ini diharapkan dapat memberikan model akurasi dan performa yang baik dalam klasifikasi citra penyakit daun padi.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka penulis dapat merumuskan beberapa permasalahan yang akan dibahas dalam penelitian ini sebagai berikut:

1. Seberapa efektif metode deep learning CNN dengan arsitektur AlexNet jika dibandingkan dengan arsitektur VGG19 dalam mengekstraksi fitur mendalam (deep features) dari gambar daun padi?
2. Bagaimana akurasi dan performa algoritma *Support Vector Machine* (SVM) yang dilatih menggunakan metode optimasi *Sequential Minimal Optimization* (SMO) dalam mengklasifikasikan penyakit daun padi berdasarkan fitur yang diekstraksi oleh CNN (AlexNet dan VGG19)?

1.3 Tujuan

Adapun penelitian ini bertujuan untuk:

1. Mengukur efektivitas metode *deep learning* pada arsitektur AlexNet yang dibandingkan dengan arsitektur VGG19 dalam mengekstraksi fitur mendalam (*deep features*) dari gambar daun padi.
2. Menganalisis akurasi dan performa algoritma *Support Vector Machine* (SVM) yang dilatih menggunakan metode optimasi *Sequential Minimal Optimization* (SMO)

dalam mengklasifikasikan penyakit daun padi berdasarkan fitur yang diekstraksi oleh CNN (AlexNet dan VGG19).

1.4 Manfaat

Penelitian ini memiliki beberapa manfaat sebagai berikut :

1. Bagi penulis :
 - Untuk membantu dalam pemahaman dan memperluas pengetahuannya dalam penerapan metode Support Vector Machine (SVM) dalam konteks klasifikasi citra, serta algoritma deep feature melalui Convolutional Neural Network (CNN).
 - Penelitian ini diharapkan dapat menjadi referensi yang bermanfaat bagi peneliti lain yang tertarik untuk melakukan penelitian mengenai klasifikasi citra jenis penyakit daun padi atau yang sejenisnya.
2. Bagi Pengguna (petani dan pihak terkait)
 - Membantu dalam mengidentifikasi jenis penyakit pada tanaman daun padi, sehingga dapat melakukan penanggulangan penyakit tanaman padi secara dini.

1.5 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan masalah yang perlu diperhatikan agar penelitian tetap terfokus pada masalah yang dikaji. Batasan-batasan ini mencakup:

1. Penelitian ini fokus pada klasifikasi 3 jenis kelas yaitu dua jenis kondisi daun padi berpenyakit (*Bacterial leaf blight* dan hama ulat) dan kondisi daun padi tidak berpenyakit atau normal. Menggunakan dataset berupa citra yang diambil langsung dari sawah di Kota Surabaya.
2. Dataset citra daun padi diambil pada kondisi cuaca cerah, sekitar pukul 10.00 hingga 14.00 WIB, supaya mendapatkan pencahayaan yang baik.