

ANALISIS PERBANDINGAN TINGKAT AKURASI ALGORITMA CNN DAN SVM DALAM KLASIFIKASI PADA DAUN GEDI, DAUN PEPAYA DAN DAUN UBI

Sri Mulyana, Inna Muthmainnah

Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan
Jl. William Iskandar Pasar V Medan, Indonesia
srimulyana@mhs.unimed.ac.id

ABSTRAK

Daun merupakan salah satu aplikasi penting dalam pengenalan pola dan pengolahan citra. Identifikasi dan klasifikasi daun dapat digunakan dalam berbagai bidang, seperti pertanian, penginderaan jauh, dan penelitian lingkungan. Penelitian ini mengeksplorasi perbandingan tingkat akurasi antara dua algoritma klasifikasi, yaitu *Convolutional Neural Network* (CNN) dengan arsitektur VGG16 dan *Support Vector Machine* (SVM), dalam konteks klasifikasi daun Gedi, daun Pepaya, dan daun Ubi. Data citra daun dikumpulkan dan dibagi menjadi kelompok pelatihan, validasi, dan pengujian. CNN diimplementasikan dengan arsitektur VGG16 untuk mengeksplorasi potensi penggunaan arsitektur yang lebih kompleks untuk memahami fitur daun. Hasilnya menunjukkan tingkat akurasi yang sangat baik untuk CNN, mencapai 100%, sedangkan SVM memiliki tingkat akurasi 44,07%.

Kata kunci: CNN, SVM, VGG16, Daun, Algoritma

1. PENDAHULUAN

Klasifikasi daun adalah salah satu aplikasi penting dalam pengenalan pola dan pengolahan citra. Identifikasi dan klasifikasi daun dapat digunakan dalam berbagai bidang, seperti pertanian, penginderaan jauh, dan penelitian lingkungan. Untuk memungkinkan masyarakat untuk lebih mudah mengidentifikasi daun, metode *Support Vector Machine* (SVM) dan *Convolutional Neural Network* (CNN) akan digunakan untuk membedakan daun gedi, pepaya, dan ubi[1]. Analisis perbandingan antara algoritma *Convolutional Neural Network* (CNN) dan *Support Vector Machine* (SVM) dalam klasifikasi daun gedi, daun pepaya, dan daun ubi dapat memberikan wawasan tentang keunggulan relatif dan tingkat akurasi keduanya.

Dengan menggunakan CNN dan SVM, dataset gambar daun dari tiga jenis tumbuhan (seratus daun gedi, ratusan daun pepaya, dan ratusan daun ubi) dengan seratus gambar daun beresolusi tinggi untuk masing-masing genus menjadi lebih mudah untuk diidentifikasi. Kemudahan ini disebabkan oleh fakta bahwa perbandingan cnn dan svm dapat mengidentifikasi objek dalam citra dalam berbagai posisi yang mungkin (translation invariance)[2].

Pengolahan citra digital ialah teknologi yang tepat dalam melakukan pengolahan suatu citra. metode yang cukup banyak digunakan adalah metode *Convolutional Neural Network* (CNN). CNN adalah salah satu metode *deep learning* yang dapat mengklasifikasi dengan memanfaatkan sebuah layer konvolusi untuk mengenali pola dan fitur didalam citra, termasuk klasifikasi tingkat kematangan buah naga. Metode *Convolutional Neural Network* (CNN) memiliki hasil yang sangat akurat dalam pengenalan citra digital.[3] Dalam klasifikasi daun, *Support Vector Machine* (SVM) adalah algoritma pembelajaran mesin yang efektif untuk memisahkan data dengan

menggunakan hyperplane (batas Keputusan) yang optimal di antara dua kelas yang berbeda.

Penelitian[4] menggunakan metode SVM (*Support Vector Machine*) dan CNN digunakan untuk menentukan penyakit yang memengaruhi daun tomat. Pengujian dilakukan dengan menggunakan 200 sampel citra daun tomat, di mana 160 citra digunakan sebagai data latih dan 40 citra sebagai data uji. Hasil pengujian menunjukkan bahwa metode CNN memiliki rata-rata persentase akurasi sebesar 97,5%, presisi sebesar 95,45%, dan recall sebesar 95%, dan tingkat kesalahan sebesar 5%. Sementara itu, SVM menghasilkan akurasi rata-rata sebesar 95%, presisi sebesar 90,83%, recall sebesar 90%, dan tingkat kesalahan sebesar 10%. Berdasarkan hasil yang telah diuji, bahwa dalam penelitian ini, CNN adalah klasifikasi yang lebih baik dibandingkan SVM. Terdapat juga penelitian terdahulu menggunakan metode CNN untuk klasifikasi daun herbal, yang dilakukan oleh [5], dengan hasil akurasi yang didapat mencapai 98%.

Tujuan penelitian ini membahas tentang perbandingan Algoritma CNN dan Algoritma SVM untuk membandingkan kinerja dari waktu eksekusi dan hasil akurasi dari masing-masing algoritma. Uji kedua algoritma untuk pengklasifikasian jenis daun akan dilakukan untuk membandingkannya. Uji ini akan mencakup pengklasifikasian daun Gedi, daun pepaya, dan daun ubi. Penelitian ini bertujuan untuk mengevaluasi kinerja relatif kedua algoritma tersebut dalam tugas klasifikasi ini dan menentukan algoritma yang lebih efektif dan akurat dalam mengklasifikasikan jenis-jenis daun tersebut

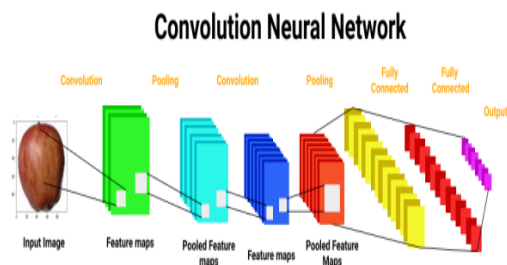
2. TINJAUAN PUSTAKA

2.1. *Convolutional Neural Network* (CNN)

Convolutional Neural Network (CNN) dirancang secara khusus untuk mengolah informasi yang

biasanya berbentuk dua dimensi, seperti suara atau gambar[6]

Convolutional Neural Network (CNN) atau yang sering disebut sebagai CNN, juga salah satu algoritma dalam deep learning yang berkembang dari konsep *MultiLayer Perceptron* (MLP). CNN didesain khusus untuk memproses data yang umumnya berbentuk dua dimensi, seperti gambar atau suara.[7]. CNN umumnya digunakan untuk melakukan pengenalan objek atau pemandangan. Proses operasi linear dalam CNN melibatkan konvolusi dan bobot, di mana bobot ini diwakili dalam bentuk matriks empat dimensi yang terdiri dari kumpulan kernel. Dalam konteks ini, konvolusi merujuk pada operasi pergeseran kernel melalui data gambar untuk mengekstrak fitur dan pola yang relevan. Dengan menggunakan metode ini, CNN dapat efektif mengidentifikasi informasi penting dalam gambar untuk tujuan klasifikasi atau pengenalan objek[8].



Gambar 1. Arsitektur metode CNN.

2.1.1. Convolution Layer

Proses manipulasi citra dengan menggunakan masker eksternal melibatkan operasi konvolusional untuk mereduksi dimensi citra dan mengekstrak fitur penting melalui encoding. Convolutional Layer merupakan hasil perkalian elemen dengan kesesuaian di antara dua vektor, berperan kritis dalam transformasi citra untuk menghasilkan representasi yang fokus pada fitur utama.

2.1.2. Pooling Layer

Pooling layer merupakan komponen kunci dalam *Convolutional Neural Network* (CNN). Fungsinya adalah melakukan pengurangan ukuran citra dengan tujuan meningkatkan invariansi terhadap posisi fitur. Dalam konteks algoritma CNN, metode pooling yang paling umum digunakan adalah max pooling. Max pooling mengambil nilai maks dari grid atau area kecil pada convolutional layer untuk merangkum citra, sehingga ukurannya menjadi lebih kecil.

2.1.3. Fully Connected Layer

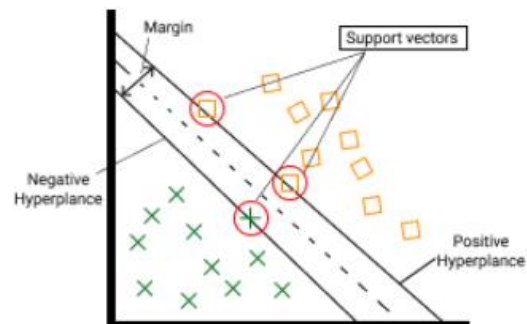
Fully Connected Layer merupakan lapisan yang digunakan dalam *MultiLayer Perceptron* (MLP) dan bertujuan untuk mentransformasi dimensi data agar dapat dijalankan melalui proses klasifikasi secara linear. Neuron-neuron dari convolution layer diubah

menjadi data satu dimensi dan kemudian dimasukkan ke dalam Fully Connected Layer.

2.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) ialah cara untuk mengelompokkan atau mengkategorikan data. SVM bekerja dengan mencoba membuat jarak antar kelompok semaksimal mungkin, seperti membuat jalan antara dua kawanan yang sejauh mungkin. Tujuannya adalah mencari garis atau batas terbaik untuk memisahkan kelompok-kelompok tersebut, dan garis ini sejauh mungkin dari setiap kelompok. SVM mencoba mencari cara terbaik untuk membuat batasan keputusan yang optimal dan efisien[9]

SVM adalah metode dalam supervised learning yang sering digunakan untuk klasifikasi, baik *Support Vector Regression* maupun *Support Vector Classification*. Algoritma SVM bekerja dengan menemukan hyperplane terbaik yang memaksimalkan jarak antara kelas-kelas yang berbeda. Hyperplane dalam SVM adalah sebuah fungsi yang digunakan untuk memisahkan satu kelas dari kelas lainnya. Pada klasifikasi antar kelas di ruang 2 dimensi, fungsi ini disebut garis, sedangkan di ruang 3 dimensi disebut bidang. Untuk klasifikasi pada ruang dimensi yang lebih tinggi dari 3 dimensi, fungsi tersebut disebut hyperplane.



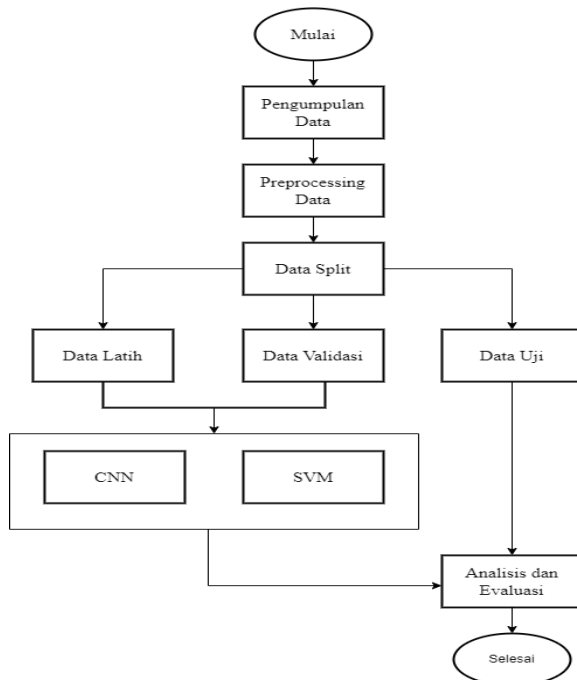
Gambar. 2. Ilustrasi metode SVM

Hyperplane pada SVM dapat dilihat pada Gambar 2, di mana hyperplane terletak di tengah-tengah antara kelas-kelas yang berbeda. Dalam algoritma SVM, objek data yang posisinya paling luar dari kelompoknya dan paling dekat dengan hyperplane disebut support vector. Karena SVM bertujuan untuk menemukan hyperplane terbaik dengan cara memaksimalkan jarak antara kelas-kelas yang berbeda, support vector ini digunakan untuk menentukan hyperplane yang paling optimal.

3. METODE PENELITIAN

Pada penelitian terdapat beberapa tahap yang akan dilakukan dimulai dari pengumpulan dataset, preprocessing data, dan pembagian dataset menjadi data latih, data validasi, dan data uji. Selanjutnya, dilakukan proses klasifikasi mencari akurasi dan menganalisis perbedaan akurasi. Pada penelitian yang dilakukan terdapat beberapa langkah untuk

mendapatkan hasil yang diinginkan seperti yang ditunjukkan pada Gambar 3.



Gambar 3. Langkah Metode Penelitian

3.1. Pengumpulan Data

Pengumpulan dataset dilakukan dengan mengambil sampel secara langsung dari tiga jenis daun utama, yaitu daun gedi, daun pepaya, dan daun ubi. Setiap sampel difoto dengan detail untuk memastikan representasi yang baik dalam analisis selanjutnya. Langkah ini menjadi dasar penting dalam memahami dan mengklasifikasikan perbedaan antara ketiga jenis daun dalam konteks penelitian.



Gambar 4. Daun Gedi



Gambar 5. Daun Pepaya



Gambar 6. Daun Ubi

Dapat dilihat pada Gambar 4, 5 dan 6 merupakan gambar sampel dataset tiga jenis daun yaitu daun gedi, daun pepaya dan daun ubi yang akan digunakan pada penelitian ini.

3.2. Preprocessing Data

Preprocessing Data adalah sekumpulan teknik yang digunakan untuk membersihkan dan menormalkan data yang tidak teratur dan tidak konsisten. Hal ini untuk membersihkan *noise*, *missing value*, dan data yang tidak *balance*[7].

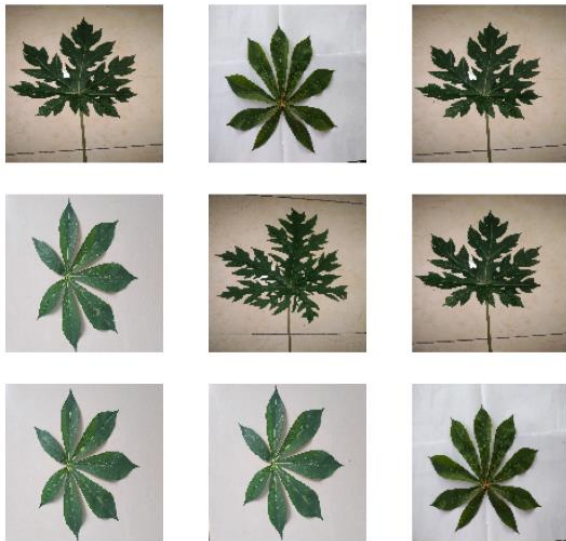
Dalam penelitian ini dilakukan preprocessing data dengan melakukan fungsi augmentasi pada setiap data latih dan data validasi yang dapat dilihat pada Gambar 7 dan Gambar 8. Augmentasi data adalah teknik yang umumnya digunakan untuk memperluas dataset pelatihan dengan membuat variasi dari data yang sudah ada. Augmentasi data merupakan langkah dalam pemrosesan data gambar di mana gambar-gambar dimodifikasi atau diubah sedemikian rupa sehingga mesin dapat mengidentifikasi perbedaan antar gambar yang diubah, tetapi manusia masih dapat mengenali bahwa esensi gambar tetap sama[10].

Augmentasi Data Training



Gambar 7. Hasil Augmentasi Data Training

Augmentasi Data Validasi



Gambar 8. Hasil Augmentasi Data Validasi

Berdasarkan Gambar 7 dan 8 menunjukkan hasil dari augmentasi data training dan data validasi, hal ini dapat bermanfaat baik untuk evaluasi model pada *Convolutional Neural Networks* (CNN) maupun *Support Vector Machines* (SVM) dalam konteks pengenalan gambar.

3.3. Implementasi Model

VGG16, singkatan dari "Visual Geometry Group 16", adalah arsitektur Convolutional Neural Network (CNN) yang terkenal dalam bidang penglihatan komputer. Dikenal karena desainnya yang sederhana dan konsekuen, VGG16 memanfaatkan lapisan konvolusi dan max pooling dengan filter 3x3 piksel untuk mengekstrak fitur-fitur hierarkis dari gambar. Kegunaan utamanya terletak pada kemampuannya dalam klasifikasi gambar yang kompleks. Dengan 16 lapisan, termasuk lapisan fully connected, VGG16 dapat mengidentifikasi dan memahami pola dan struktur gambar dengan baik, menjadikannya efektif untuk tugas-tugas seperti klasifikasi objek dalam gambar. Meskipun membutuhkan sumber daya komputasi yang signifikan untuk pelatihan, keunggulan VGG16 terletak pada kemampuannya untuk menangani tugas-tugas penglihatan yang memerlukan pemahaman fitur-fitur yang lebih kompleks pada tingkat hierarkis.

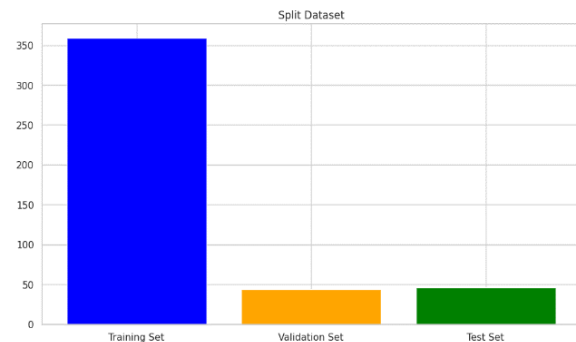
3.4. Data Split

Setelah dilakukan augmentasi data, proses selanjutnya adalah proses split data, pada penelitian ini data dibagi menjadi 3 yaitu data latih, data validasi dan data uji dengan perbandingan data latih 80%, data validasi 10%, dan data uji 10%.

4. HASIL DAN PEMBAHASAN

Dalam tahap preproses data, gambar-gambar akan diubah ukurannya menjadi 128x128 piksel untuk

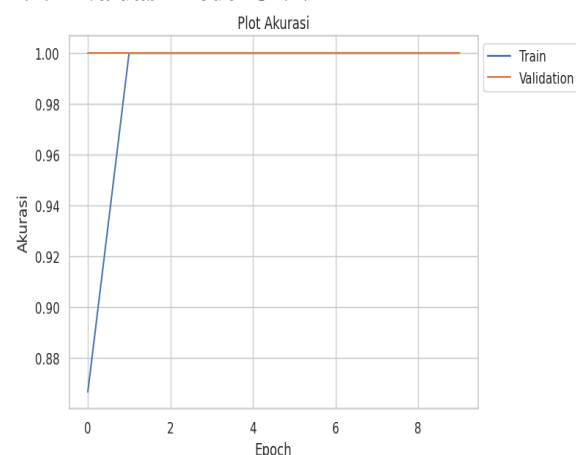
memastikan konsistensi dimensi dan mempersiapkan data input yang sesuai dengan arsitektur model, seperti VGG16 pada Convolutional Neural Network (CNN). Dataset berisi kumpulan citra daun ubi, daun gedi dan daun pepaya dibagi menjadi 3 yaitu data latih, data validasi dan data uji dengan perbandingan data latih 80%, data validasi 10%, dan data uji 10%.



Gambar 9. Hasil Pembagian Data

Dari total 499 data yang tersedia, data dibagi menjadi tiga subset berdasarkan pembagian yang telah ditentukan. Data training terdiri dari 359 data, digunakan untuk melatih model machine learning agar dapat belajar dari sebagian besar dataset. Data validasi terdiri dari 44 data, yang digunakan untuk menguji performa model secara berkala selama proses pelatihan untuk memonitor dan menyesuaikan parameter model. Sementara itu, data uji terdiri dari 46 data, yang digunakan untuk evaluasi akhir terhadap model yang telah dilatih, memastikan bahwa model mampu melakukan prediksi dengan baik pada data baru yang belum pernah dilihat sebelumnya. Dengan pembagian ini, keseluruhan proses pengembangan model dapat dilakukan secara sistematis dan terstruktur.

4.1. Evaluasi Model CNN

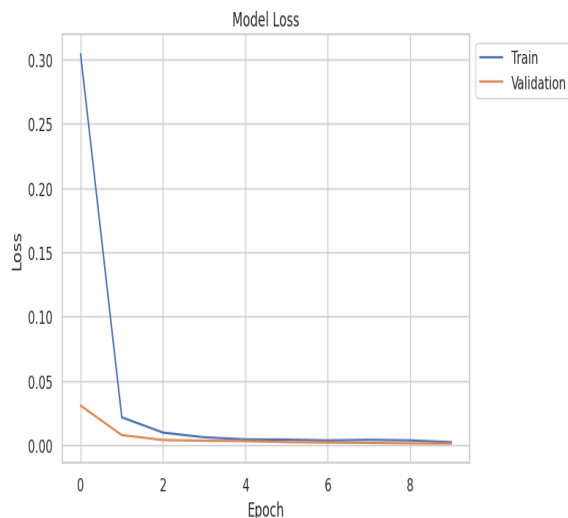


Gambar 10. Hasil Model Akurasi dan Validation

Dalam penelitian ini, penggunaan algoritma Convolutional Neural Network (CNN) dilakukan dengan melakukan 10 epoch. Hasilnya, diperoleh nilai akurasi tertinggi dan nilai validasi. Pada Gambar 4,

model menunjukkan nilai akurasi mencapai 100%, sedangkan pada data uji, nilai akurasi juga mencapai 100%. Hal ini menunjukkan bahwa model telah mencapai titik optimal karena grafiknya konvergen.

Pada Gambar 10, model loss dari hasil pelatihan data dengan menggunakan 10 epoch menunjukkan nilai loss sebesar 0%.



Gambar 11. Hasil Model Loss Validation

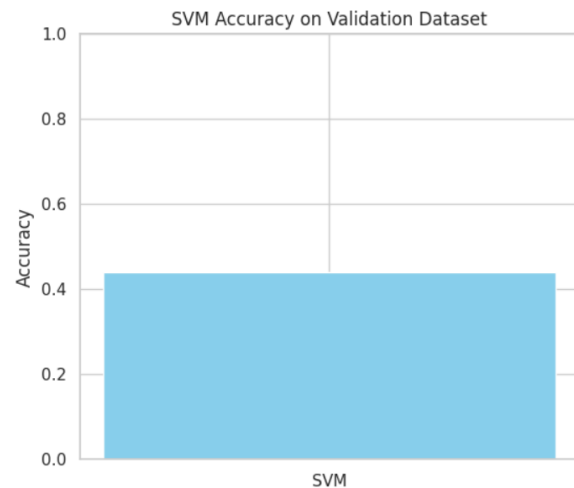
Dari hasil tersebut, terdapat beberapa faktor yang dapat mempengaruhi nilai akurasi ketika menggunakan metode convolutional neural network. Faktor-faktor tersebut meliputi nilai parameter setiap layer yang digunakan, bentuk arsitektur pada layer konvolusi citra, nilai learning rate, dan jumlah epoch yang digunakan selama pelatihan data.

4.2. Evaluasi Model SVM

Setelah mengimplementasikan model CNN, Peneliti juga mengimplementasikan algoritma CNN guna untuk mengkomparasi kedua algoritma tersebut. Dalam permodelan SVM dilakukan ekstraksi fitur dari layer CNN model yang telah dilatih pada saat mengimplementasikan CNN yakni block5_pool. Setelah itu Dengan menggunakan model ekstraktor fitur, kita menghasilkan representasi fitur dari dataset pelatihan (train_dataset) dan validasi (valid_dataset). Ini adalah representasi fitur yang diekstrak dari lapisan 'block5_pool' dari model VGG16. Berikutnya mereshape untuk mendapatkan vektor satu dimensi untuk setiap sampel. Ini diperlukan karena model SVM memerlukan input berupa vektor satu dimensi. Setelah direshape, membuat model SVM dengan menggunakan make_pipeline. Dalam pipeline, pertama-tama, fitur-fitur dinormalisasi menggunakan StandardScaler() untuk memastikan bahwa setiap fitur memiliki skala yang serupa.

Kemudian, model SVM (svm.SVC) dengan kernel linear (linier) diaplikasikan. Parameter lainnya seperti C sebesar 1,0 dan gamma otomatis. Model SVM yang telah dilatih sebelumnya diuji pada dataset

validasi dengan menggunakan metode predict pada fitur yang telah diekstrak dan diratakan. Setelah mendapatkan prediksi dari model SVM, akurasi dihitung dengan membandingkan prediksi dengan label sebenarnya dari dataset validasi. Akurasi diukur menggunakan fungsi accuracy_score dari scikit-learn, dengan hasil 44.07%.



Gambar 12. Hasil akurasi SVM

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, dapat disimpulkan metode CNN lebih efektif dalam klasifikasi daun Gedi, daun Pepaya, dan daun Ubi dibandingkan dengan SVM. Tingkat akurasi yang mencapai 100% pada CNN menunjukkan kemampuannya yang luar biasa dalam memahami pola dan fitur pada data gambar. Oleh karena itu, untuk tugas klasifikasi jenis daun, penggunaan CNN lebih disarankan. Sedangkan hasil akurasi SVM hanya sebesar 44.07%. memberikan hasil yang kurang baik, penggunaannya mungkin lebih sesuai untuk tugas klasifikasi yang lebih sederhana atau dengan data yang lebih terstruktur. Dengan demikian, pemilihan algoritma klasifikasi harus disesuaikan dengan kompleksitas dan karakteristik data yang dihadapi.

DAFTAR PUSTAKA

- [1] B. W. Kurniadi, H. Prasetyo, G. L. Ahmad, B. Aditya Wibisono, and D. Sandya Prasvita, *Analisis Perbandingan Algoritma SVM dan CNN untuk Klasifikasi Buah*. 2021.
- [2] S. Ilahiyah and A. Nilogiri, "Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network".
- [3] A. Daun Herbal Menggunakan, K. Anam, and A. Saleh, "Autentikasi Daun Herbal Menggunakan Convolutional Neural Network dan Raspberry Pi (Authentication of Herbal Leaves Using Convolutional Neural Network and Raspberry Pi)," 2020.
- [4] S. Faisal, T. F. M Butarbutar, P. Sirait, and J. SIFO Mikroskil, "Implementasi CNN dan SVM

- untuk Identifikasi Penyakit Tomat via Daun,” *OKTOBER 2019 IJCCS*, vol. 20, pp. 1–5.
- [5] J. Khatib Sulaiman, A. Regita Azzahra, H. Darwis, and D. Widyawati, “Klasifikasi Daun Herbal Menggunakan Metode CNN dan Naïve Bayes dengan Fitur GLCM,” *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 4, p. 2183.
- [6] R. Pujiati and N. Rochmawati, “Identifikasi Citra Daun Tanaman Herbal Menggunakan Metode Convolutional Neural Network (CNN),” *Journal of Informatics and Computer Science*, vol. 03, 2022.
- [7] A. A. Jaelani, F. Y. Supratman, and N. Ibrahim, “PERANCANGAN APLIKASI UNTUK KLASIFIKASI KLON DAUN TEH SERI GAMBUNG (GMB) MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK,” 2020. [Online]. Available: <https://api.semanticscholar.org/CorpusID:229069985>
- [8] E. Ratnasari, “Pengenalan Jenis Buah pada Citra Menggunakan Pendekatan Klasifikasi Berdasarkan Fitur Warna Lab dan Tekstur Co-Occurrence,” *Inform: Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 1, Dec. 2016, doi: 10.25139/inform.v1i2.846.
- [9] Nithya Devaraj, *Feature Selection using Genetic Algorithm to improve SVM Classifier*. LAP LAMBERT Academic Publishing, 2019.
- [10] S. Muhammad and A. T. Wibowo, “KLASIFIKASI TANAMAN AGLAONEMA BERDASARKAN CITRA DAUN MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN).”