



Analisis Performansi Pendekatan Machine Learning Pada Deteksi Penyakit Daun Tanaman Kopi

**Yodhi Yuniarthe¹, Rosyana Fitria Purnomo², Hilda Dwi Yunita³,
Fatimah Fahurian⁴, Ahmad Ikhwan⁵**

^{1,5} Program Studi Informatika, Fakultas Komputer, Universitas Mitra Indonesia

^{2,3,4} Program Studi Sistem Informasi, Fakultas Komputer, Universitas Mitra Indonesia

¹yodhi@umitra.ac.id, ²rosyanapurnomo@umitra.ac.id, ³hildadwiunita@umitra.ac.id,

⁴fatimah_fahurian@umitra.ac.id, ⁵ikhwan69@umitra.ac.id

Article Info

Article history:

Received September 3, 2025

Revised Desember 3, 2025

Accepted Januari 13, 2026

Keywords:

Coffee Classification

Image Processing

Machine Learning

Plant Disease Detection

ABSTRACT

Detection and identification of plant diseases is critical to the success and efficiency of agricultural production. Plant disease outbreaks are becoming more frequent throughout the world, and the presence of these diseases in cultivated plants has a significant impact on productivity. Therefore, researchers are focusing on developing effective and reliable plant disease detection methods. Thus, farmers can take advantage of early detection of this disease to minimize future losses. This article discusses machine learning approaches as well as decision trees, K-nearest neighbors, naive Bayes, support vector machines (SVM), and random forests for detecting coffee leaf diseases using leaf images. The above-mentioned classifications were researched and compared to determine the most suitable plant disease prediction model with the highest accuracy. Compared with other classification algorithms, the SVM algorithm achieves the highest accuracy of 99.75%. All the models trained above will be used by farmers to quickly identify and classify new diseases in images as a preventive strategy. As a preventive measure, farmers can detect and classify new diseases in images early.

Corresponding Author:

Yodhi Yuniarthe

Universitas Mitra Indonesia

Jl. ZA. Pagar Alam No.7, Gedong Meneng, Kec. Rajabasa, Kota Bandar Lampung

Email: yodhi@umitra.ac.id



1. PENDAHULUAN

Deteksi dan identifikasi penyakit tanaman sangat penting untuk penerapan langkah-langkah perlindungan tanaman yang efektif dan terarah dalam produksi pertanian [1]. Wabah penyakit tanaman semakin sering terjadi di seluruh dunia, dan prevalensi penyakit pada tanaman budidaya memiliki dampak signifikan terhadap produktivitas. Akibatnya, para peneliti fokus pada pengembangan metode yang efektif dan andal untuk mendeteksi penyakit tanaman [2]. Penyakit tanaman dapat dideteksi menggunakan prosedur laboratorium atau pendekatan pemrosesan gambar [3]. Metode laboratorium seperti uji molekuler dan serologis diperlukan untuk mendiagnosis penyakit tanaman [4]. Meskipun

pendekatan molekuler berbasis DNA dan metode serologis telah meningkatkan diagnosis penyakit tanaman, teknik pemrosesan gambar yang dikombinasikan dengan pembelajaran mesin kini dapat mendiagnosis penyakit tersebut dengan cepat dan efektif pada tingkat lanjut dengan presisi tinggi pada gambar berwarna [5]. Berkat metodologi berbasis pembelajaran mesin, pendekatan berbasis gambar dalam pertanian presisi telah berkembang pesat. Dalam beberapa tahun terakhir, hal ini dimungkinkan oleh ketersediaan pengukuran berkualitas tinggi yang dikombinasikan dengan algoritma canggih dan kemampuan yang meningkat untuk menggabungkan berbagai sumber gambar. Foto-foto ini dapat berasal dari berbagai sumber, seperti citra satelit, sensor, atau bahkan kamera. Efektivitas berbagai algoritma pembelajaran mesin dalam mengklasifikasikan gambar daun berdasarkan ciri-ciri tanaman umum seperti bentuk, urat, tekstur, dan kombinasi banyak faktor diklasifikasikan [6].

Pendekatan machine learning yang paling umum digunakan adalah supervised learning dan unsupervised learning. Supervised learning melatih sistem menggunakan dataset gambar yang dilabeli. Data tersebut telah dilabeli dengan klasifikasi penyakit yang sesuai atau ketiadaan klasifikasi tersebut [7]. Secara umum, dataset yang lebih besar menghasilkan hasil yang lebih akurat. Namun, memperoleh dataset biasanya memerlukan biaya (waktu komputasi atau upaya eksperimental). Oleh karena itu, terdapat tradeoff antara akurasi dan upaya untuk mendapatkan dataset pelatihan yang lebih besar. [8]. Sebagian besar metode pembelajaran mesin saat ini tidak cukup tangguh untuk menjembatani kesenjangan antara metode deteksi penyakit tanaman di dunia nyata. Salah satu tantangan terberat adalah bahwa beberapa penyakit sangat mirip di antara spesies tanaman dan dapat muncul secara bersamaan pada satu tanaman. Hal ini memengaruhi pemrosesan gambar, teknik pembelajaran mesin, dan spesialis manusia. Oleh karena itu, pendekatan yang dipilih harus relevan untuk mendeteksi penyakit target [9]. Banyak proyek penelitian berusaha mengembangkan model yang mampu mendeteksi dan mengklasifikasikan penyakit daun pada tahap awal menggunakan teknik pembelajaran mesin serta mengambil langkah-langkah pencegahan yang diperlukan [10]–[13]. Penelitian ini menerapkan algoritma klasifikasi pembelajaran mesin terawasi seperti NB [14], [15], KNN [14]–[16], DT [14][15], SVM [7], [14]–[16], dan RF [14], [15] untuk deteksi dan klasifikasi penyakit, serta perbandingan berbagai algoritma. Penelitian ini juga menyediakan metode-metode yang memberikan hasil paling akurat dibandingkan dengan algoritma lain. Tanaman kopi digunakan dalam penelitian ini untuk mengkaji beberapa penyakit daun. Di Indonesia, keberlanjutan rantai pasok kopi sangat penting; negara ini merupakan salah satu produsen kopi terbesar di dunia. Keberlanjutan rantai pasok kopi tidak hanya penting bagi sektor kopi, tetapi juga bagi mata pencaharian ribuan orang [17]. Penyakit daun kopi terutama disebabkan oleh faktor abiotik (akibat kekurangan nutrisi, kelembapan, dan suhu) dan biotik (bakteri, jamur, nematoda, dan virus). Patogen pada daun kopi merusak daun. Gambar 1 menunjukkan beberapa gangguan yang paling umum dan kondisi sehat; *Cercospora leaf spot*, yang juga disebut *grey leaf spot*, dimulai sebagai bintik nekrotik kecil pada daun dan berkembang menjadi lesi berbentuk persegi panjang. Bintik-bintik tersebut menggelap dan akhirnya berubah menjadi abu-abu. Penggunaan benih hibrida tahan penyakit dan fungisida daun adalah dua strategi pengelolaan. Kemunculan bintik-bintik coklat mendefinisikan karat. Bintik-bintik ini terdapat di kedua permukaan daun saat bintik-bintik pecah dan melepaskan spora merah berdebu. Bintik-bintik terbentuk pada malai, tongkol, dan daun setelah penyakit parah, membuatnya kuning. Miner mirip dengan karat, tetapi daunnya robek. Awalnya, lesi elips abu-abu kehijauan pada daun dapat dideteksi. Sementara daun sehat terlihat segar, bersih, dan tidak bercak.



Gambar 1. Menunjukkan gambar daun kopi yang sehat dan yang terkena gangguan yang umum.

Tujuan utama penelitian ini adalah untuk menentukan akurasi dan kesesuaian metode klasifikasi penyakit yang ada pada daun kopi. Meskipun metode klasifikasi telah diverifikasi menggunakan dataset gambar, banyak peneliti tanaman lebih memilih untuk mengumpulkan dan mencatat data secara manual. Analisis yang disajikan dalam makalah ini bertujuan untuk memberikan rekomendasi kepada peneliti pertanian, dan secara lebih luas kepada peneliti penyakit, mengenai algoritma klasifikasi yang berkinerja baik dalam hal prediksi penyakit dan akurasi klasifikasi saat diterapkan pada dataset pengukuran mentah. Selain itu, sebagai langkah pencegahan, petani akan menggunakan semua model yang telah dilatih untuk mengidentifikasi dan mengklasifikasikan penyakit gambar baru dengan cepat. Petani dapat menggunakan alat ini untuk membantu dalam deteksi dini dan klasifikasi penyakit gambar yang muncul sebagai langkah pencegahan. Penelitian ini disusun sebagai berikut: Bagian 2, metode penelitian. Bagian 3, hasil dan pembahasan. Bagian 4 memberikan kesimpulan dan rekomendasi untuk studi lebih lanjut.

2. METODE PENELITIAN

Penelitian ini bertujuan untuk membandingkan dan menganalisis kemampuan klasifikasi dan prediksi dari lima metode klasifikasi yang umum digunakan (pohon keputusan, mesin vektor pendukung, naive bayes, k-nearest neighbor, dan hutan acak) menggunakan dataset kopi. Bagian ini akan menjelaskan proses kami, dimulai dari persiapan data dan melalui teknik klasifikasi serta kriteria evaluasi. Bagian berikutnya akan menjelaskan temuan eksperimen secara rinci.

2.1. Persiapan Eksperimen Dan Data

Tahap pertama dalam metode pembelajaran mesin untuk mendeteksi penyakit tanaman adalah persiapan data, yang meliputi pengumpulan dan prapemrosesan gambar daun untuk membentuk kumpulan data gambar. Kumpulan data gambar, khususnya untuk gambar penyakit kopi, dibuat menggunakan kamera smartphone. Subkumpulan data tanaman kopi berisi 2000 gambar dengan empat label kelas penyakit seperti cercospora, karat, hama, dan sehat, masing-masing dengan 500 gambar dan ukuran piksel 2350 x 4100. Gambar kemudian dipotong, dan ukuran piksel disamakan menjadi 128 x 128. Gambar kemudian disegmentasi dengan memutar dan memperbesar, dengan probabilitas memutar 70% dan memperbesar 50% untuk setiap gambar digital. Proses augmentasi menghasilkan sebanyak 2.000 gambar digital, dengan 500 gambar digital per kelas. Gambar 2 menampilkan setiap kelas gambar. Seluruh dataset dibagi secara acak menjadi skenario pelatihan-validasi-pengujian untuk setiap kelas (Cercospora, karat, miner, dan sehat) guna menilai akurasi klasifikasi pada data yang belum dilihat. Selain itu, sampel yang ditingkatkan, yang dipisahkan secara terpisah, juga dipertimbangkan. Seperti yang disebutkan dalam Tabel 1, pembagian ini adalah 50-20-30 (50% pelatihan, 20% validasi, dan 30% pengujian) dan 60-20-20 (60% pelatihan, 20% validasi, dan 20% pengujian). Hal ini memungkinkan eksperimen tentang pengaruh data pelatihan terhadap akurasi model yang dihasilkan.

Tabel 1. Data set pelatihan-validasi-pengujian yang menampilkan jumlah sampel untuk setiap pembagian data.

Class	Split Data					
	Train	Val	Test	Train	Val	Test
	50	20	30	60	20	20
Cercospora	500	200	300	600	200	200
Rust	500	200	300	600	200	200
Miner	500	200	300	600	200	200
Healthy	500	200	300	600	200	200
Total	2000	800	1200	2400	800	800



Gambar 2. Contoh sampel gambar daun kopi dari proses yang ditingkatkan.

2.2. Metode Klasifikasi

Lima metode klasifikasi yang berbeda dievaluasi dalam penelitian ini untuk mengembangkan model klasifikasi penyakit pada daun kopi. Lima metode klasifikasi yang digunakan dalam studi ini adalah:

2.2.1 Decision Tree (DT)

Pohon keputusan adalah struktur hierarkis yang terdiri dari simpul inti yang mewakili atribut, cabang yang menunjukkan hasil uji observasi, dan simpul daun yang mewakili label kelas. Dalam struktur pohon ini, simpul akar adalah simpul teratas. Pohon keputusan dieksplorasi dari simpul akar hingga simpul daun untuk mengidentifikasi kelas observasi [15].

2.2.2 Support Vector Machine (SVM)

SVM adalah teknik pembelajaran terawasi yang sering digunakan untuk menyelesaikan masalah klasifikasi atau regresi. Teknik ini melakukan klasifikasi linier pada dua kelas dalam bentuk aslinya. Klasifikasi dilakukan dengan membatasi ruang fitur menggunakan hiperplane pemisah. SVM juga mampu melakukan klasifikasi non-linier dengan menggunakan kernel. Kernel digunakan untuk mentransformasi ruang fitur asli secara efisien menjadi ruang fitur berdimensi tinggi atau tak terhingga, yang memungkinkan pembentukan hiperplane non-linier yang sangat kompleks. SVM mampu menangani dataset yang sangat kompleks sambil menunjukkan kemampuan generalisasi yang sangat baik [16].

2.2.3 Naiv Bayes (NB)

Naive Bayes adalah variasi dari klasifikasi Bayes yang didasarkan pada prinsip Bayes. Metode ini didasarkan pada probabilitas a priori dari pola-pola data dan penugasan probabilitas a posteriori ke label kelas. Probabilitas posterior menentukan nilai kemungkinan maksimum data yang termasuk dalam label kelas tertentu menggunakan hipotesis ini. Probabilitas ini dihitung menggunakan Teorema Bayes dengan mengalikan probabilitas kondisional setiap fitur. Meskipun hipotesis ini jarang benar dalam situasi dunia nyata, ia telah terbukti sangat sukses dalam banyak kasus klasifikasi. [14].

2.2.4 K-Nearest Neighbor (KNN)

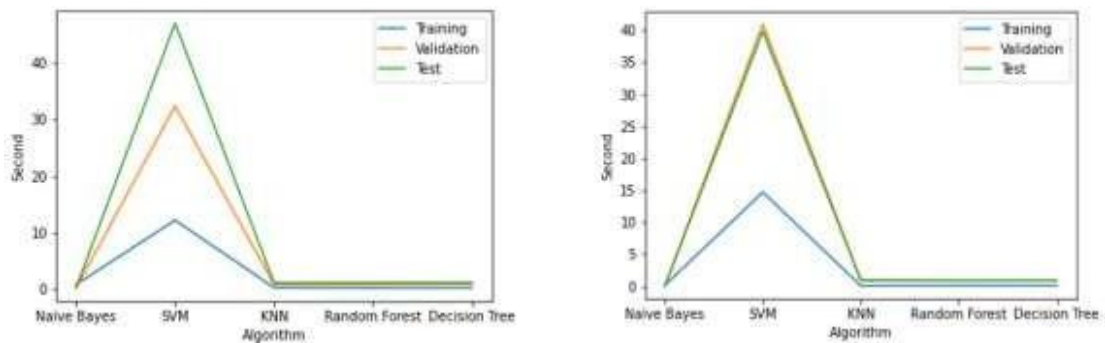
Metode k-nearest neighbor (k-NN) adalah strategi sederhana yang secara konsisten menghasilkan hasil yang superior. Ini adalah pendekatan nonparametrik berbasis contoh untuk pembelajaran yang relatif lambat. Metode ini diterapkan pada masalah klasifikasi dan regresi. k-NN digunakan untuk menentukan kelas klasifikasi di mana suatu item baru yang belum diberi label termasuk. Hal ini dilakukan dengan memilih nilai ganjil untuk 'k' (di mana k adalah jumlah tetangga yang dipertimbangkan) dan menghitung jarak antara titik data terdekat ke objek menggunakan jarak Euclidean, jarak Hamming, jarak Manhattan, atau jarak Minkowski. Setelah perhitungan jarak, 'k' tetangga terdekat dipilih, dan suara mereka menentukan kelas objek baru [7].

2.2.4 Random Forest (RF)

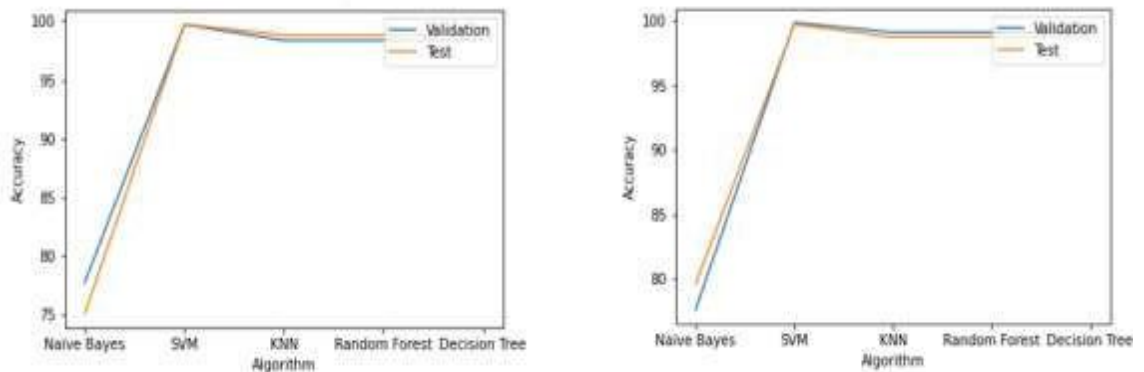
Random forests adalah metode pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. Selama proses pelatihan, metode ini menghasilkan hutan pohon keputusan. Hutan acak tidak rentan terhadap overfitting, berbeda dengan pohon keputusan, dan dapat menangani baik data numerik maupun kategorikal [15].

3. HASIL DAN PEMBAHASAN

Bagian ini mengevaluasi kinerja berbagai pendekatan klasifikasi, termasuk DT, SVM, NB, KNN, dan RF, menggunakan dataset untuk identifikasi penyakit pada daun kopi. Klasifikasi ini diimplementasikan dalam Python 3.8 pada sistem operasi Windows 10. Prosesor x64 berkecepatan 3.7 GHz dengan RAM 8.00 GB. Perangkat lunak Python digunakan bersama dengan perpustakaan machine learning Python scikit-learn. Lima metode klasifikasi dilatih dan dioptimalkan menggunakan pembagian 60-20-20 untuk distribusi pelatihan, validasi, dan pengujian. Prosedur yang sama digunakan untuk pembagian 50-20-30. Waktu total yang diperlukan untuk pelatihan hingga pengujian CPU sekitar 100 detik dalam kedua kasus. Durasi proses ini ditampilkan pada Gambar 3. Pelatihan merupakan aspek yang paling memakan waktu dalam seluruh proses implementasi model. Akurasi model validasi dan pengujian ditampilkan pada Gambar 4. Secara keseluruhan, skenario 60-20-20 mencapai akurasi 98,75% untuk DT, 99,75% untuk SVM, 75% untuk NB, 98,83% untuk KNN, dan 98,83% untuk RF. Sementara skenario 60-20-20 mencapai 98,83% untuk DT, 99,75% untuk SVM, 79,6% untuk NB, 98,75% untuk KNN, dan 98,75% untuk RF. Tabel x menunjukkan hasil akurasi model validasi dan pengujian



Gambar 3. Waktu Eksekusi setiap model 50-20-30 (kiri) dan 60-20-20 (kanan)



Gambar 4. Akurasi Validasi-Uji 50-20-30 (kiri) dan 60-20-20 (kanan)

Tabel 2. Akurasi Pelatihan-Validasi-Uji

Classification Method	50-20-30		60-20-20	
	Validation	Test	validation	test
Naive Bayes	77.75	75.0929	77.625	79.625

SVM	99.75	99.75	99.875	99.75
KNN	98.375	98.8333	99.125	98.75
Random Forest	98.375	98.8333	99.125	98.75
Decision Tree	98.375	98.8333	99.125	98.75

Dalam membandingkan hasil, SVM menunjukkan kinerja terbaik pada dataset daun kopi dengan akurasi rata-rata 99%, namun memiliki waktu eksekusi terlama. Sementara itu, semua metode kecuali Naive Bayes mencapai akurasi rata-rata 98% dengan waktu eksekusi yang lebih singkat pada dataset daun kopi. Secara alami, sebagian besar metode klasifikasi akan bekerja optimal pada dataset daun kopi. Skema pembagian data pelatihan-validasi-pengujian 50-20-30 dan 60-20-20 tidak berpengaruh pada akurasi dan waktu eksekusi.

4. KESIMPULAN

Deteksi penyakit pada tanaman sangat penting untuk menjaga hasil panen dan kelangsungan usaha pertanian, terutama pada tanaman bernilai tinggi seperti kopi. Karena semakin sering terjadi wabah penyakit di berbagai daerah, diperlukan metode pengenalan penyakit yang cepat, tepat, dan dapat dipercaya.

Dari penelitian ini, beberapa algoritma pembelajaran mesin seperti decision tree, K-nearest neighbor, naive Bayes, support vector machine (SVM), dan random forest berhasil digunakan untuk mendeteksi penyakit pada daun kopi dengan memanfaatkan gambar daun sebagai data utama. Hasilnya menunjukkan bahwa model SVM memiliki performa terbaik dengan akurasi mencapai 99,75%, lebih baik dibandingkan algoritma lainnya. Temuan ini menunjukkan bahwa SVM sangat efektif dalam mengenali pola visual pada daun kopi. Semua model yang dikembangkan dalam penelitian ini bisa menjadi alat bantu yang berguna bagi para petani, membantu mereka mendeteksi penyakit secara dini dan akurat.

Dengan sistem deteksi penyakit berbasis machine learning, para petani bisa bertindak lebih cepat dalam mencegah penyakit, mengurangi kerugian, dan meningkatkan cara merawat tanaman mereka. Dengan demikian, penelitian ini mendukung pengembangan pertanian cerdas yang relevan untuk menjaga produktivitas perkebunan kopi di masa depan.

REFERENCES

- [1] S. Thomas *et al.*, "Benefits Of Hyperspectral Imaging For Plant Disease Detection And Plant Protection: A Technical Perspective," *J. Plant Dis. Prot.*, vol. 125, no. 1, pp. 5–20, 2018, doi: 10.1007/s41348-017-0124-6.
- [2] V. Singh, N. Sharma, and S. Singh, "A Review Of Imaging Techniques For Plant Disease Detection," *Artif. Intell. Agric.*, vol. 4, pp. 229–242, 2020, doi: 10.1016/j.aiaa.2020.10.002.
- [3] R. Manavalan, "Automatic Identification Of Diseases In Grains Crops Through Computational Approaches: A Review," *Comput. Electron. Agric.*, vol. 178, p. 105802, 2020, doi: 10.1016/j.compag.2020.105802.
- [4] K. Golhani, S. K. Balasundram, G. Vadamalai, and B. Pradhan, *A Review Of Neural Networks In Plant Disease Detection Using Hyperspectral Data*, vol. 5, no. 3. China Agricultural University, 2018.
- [5] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A Comparative Study Of Fine-tuning Deep Learning Models For Plant Disease Identification," *Comput. Electron. Agric.*, vol. 161, pp. 272–279, 2019, doi: 10.1016/j.compag.2018.03.032.
- [6] M. M. Singh, "A Survey On Different Methods For Medicinal Plants Identification and Classification System," *Rev. Gestão Inovação e Tecnol.*, vol. 11, no. 4, pp. 3191–3202, 2021, doi: 10.47059/revistageintec.v11i4.2362.
- [7] H. Kaur and V. Kumari, "Predictive Modelling And Analytics For Diabetes Using A Machine

- Learning Approach,” *Appl. Comput. Informatics*, vol. 18, no. 1–2, pp. 90–100, 2022, doi: 10.1016/j.aci.2018.12.004.
- [8] H. Wei, S. Zhao, Q. Rong, and H. Bao, “Predicting The Effective Thermal Conductivities Of Composite Materials And Porous Media By Machine Learning Methods,” *Int. J. Heat Mass Transf.*, vol. 127, pp. 908–916, 2018, doi: 10.1016/j.ijheatmasstransfer.2018.08.082.
- [9] A. M. Abdu, M. M. Mokji, and U. U. Sheikh, “Machine learning for plant disease detection: An investigative comparison between support vector machine and deep learning,” *IAES Int. J. Artif. Intell.*, vol. 9, no. 4, pp. 670–683, 2020, doi: 10.11591/ijai.v9.i4.pp670-683.
- [10] S. Poornima, S. Kavitha, S. Mohanavalli, and N. Sripriya, “Detection And Classification Of Diseases In Plants Using Image Processing And Machine Learning Techniques,” *AIP Conf. Proc.*, vol. 2095, 2019, doi: 10.1063/1.5097529.
- [11] S. V. Militante, B. D. Gerardo, and N. V. Dionisio, “Plant Leaf Detection And Disease Recognition Using Deep Learning,” *2019 IEEE Eurasia Conf. IOT, Commun. Eng. ECICE 2019*, pp. 579–582, 2019, doi: 10.1109/ECICE47484.2019.8942686.
- [12] M. Sardogan, A. Tuncer, and Y. Ozen, “Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm,” *UBMK 2018 - 3rd Int. Conf. Comput. Sci. Eng.*, pp. 382–385, 2018, doi: 10.1109/UBMK.2018.8566635.
- [13] A. Picon, A. Alvarez-Gila, M. Seitz, A. Ortiz-Barredo, J. Echazarra, and A. Johannes, “Deep Convolutional Neural Networks For Mobile Capture Device-based Crop Disease Classification In The Wild,” *Comput. Electron. Agric.*, vol. 161, pp. 280–290, 2019, doi: 10.1016/j.compag.2018.04.002.
- [14] Z. Huang, A. Qin, J. Lu, A. Menon, and J. Gao, “Maize Leaf Disease Detection and Classification Using Machine Learning,” *Proc. - IEEE Congr. Cybermatics 2020 IEEE Int. Conf. Internet Things, iThings 2020, IEEE Green Comput. Commun. GreenCom 2020, IEEE Cyber, Phys. Soc. Comput. CPSCom 2020 IEEE Smart Data, SmartD*, pp. 870–877, 2020, doi: 10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics50389.2020.00150.
- [15] M. Morgan, C. Blank, and R. Seetan, “Plant Disease Prediction Using Classification Algorithms,” *IAES Int. J. Artif. Intell.*, vol. 10, no. 1, pp. 257–264, 2021, doi: 10.11591/ijai.v10.i1.pp257-264.
- [16] D. Radovanovic and S. Dukanovic, “Image-Based Plant Disease Detection: A Comparison of Deep Learning and Classical Machine Learning Algorithms,” *2020 24th Int. Conf. Inf. Technol. IT 2020*, no. February, pp. 1–4, 2020, doi: 10.1109/IT48810.2020.9070664.
- [17] M. Bashiri, B. Tjahjono, J. Lazell, J. Ferreira, and T. Perdana, “The Dynamics Of Sustainability Risks In The Global Coffee Supply Chain: A Case Of Indonesia–UK,” *Sustain.*, vol. 13, no. 2, pp. 1–20, 2021, doi: 10.3390/su13020589.