

Literature Review : Pendekatan Random Forest untuk Deteksi Penyakit pada Tanaman Buah

Bagas Alamsyah Putra^{1*}, Chairil Syahrain², Rahmah Yunita³, Reyon Lau Jiemin⁴

¹Fakultas Teknik, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspiptek No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan, Banten 15310, Indonesia

Email: ^{1*}bagasalamsyahp@gmail.com, ²chairilsyahrain24@gmail.com,
³avocadopopo@gmail.com, ⁴bursin6969@gmail.com

(* : coresponding author)

Abstrak– Penyakit pohon buah-buahan dapat menyebabkan kerugian yang signifikan baik dalam kualitas maupun kuantitas produksi pertanian. Deteksi dini penyakit tanaman sangat penting untuk mengurangi dampak negatif dan memastikan produksi berkelanjutan. Penelitian ini menyelidiki efektivitas algoritma random forest dalam mendeteksi penyakit pohon buah-buahan berdasarkan berbagai karakteristik seperti warna daun, bentuk daun, tekstur, dan pola kerusakan. Algoritma Random Forest dipilih karena kemampuannya yang unggul dalam menangani kumpulan data berdimensi besar dan mengurangi risiko overfitting. Data berupa gambar daun dan buah yang sudah terinfeksi dan tidak terinfeksi dikumpulkan dan dianalisis untuk melatih dan menguji model. Hasil penelitian menunjukkan bahwa algoritma random forest dapat mencapai tingkat akurasi yang tinggi dalam mendeteksi penyakit pada pohon buah-buahan dan memiliki kinerja yang stabil pada berbagai jenis penyakit dan kondisi tanaman. Studi ini menunjukkan bahwa pendekatan acak berbasis hutan dapat menjadi solusi efektif untuk deteksi penyakit tanaman secara cepat dan akurat serta dapat diterapkan lebih lanjut untuk mendukung sistem pertanian berbasis teknologi.

Kata Kunci: Deteksi Penyakit, Tanaman Buah-Buahan, Random Forest, Pertanian Cerdas, Pengolahan Citra

Abstract– Fruit tree diseases can cause significant losses in both the quality and quantity of agricultural production. Early detection of plant diseases is crucial to mitigate these negative impacts and ensure sustainable production. This study investigates the effectiveness of the Random Forest algorithm in detecting diseases in fruit trees based on various features, such as leaf color, leaf shape, texture, and damage patterns. The Random Forest algorithm was chosen for its superior ability to handle high-dimensional datasets and its capacity to reduce overfitting risks. Images of infected and non-infected leaves and fruits were collected and analyzed to train and test the model. The findings indicate that the Random Forest algorithm can achieve high accuracy levels in detecting diseases in fruit trees, with stable performance across various disease types and plant conditions. This study demonstrates that the Random Forest-based approach can be an effective solution for rapid and accurate plant disease detection and can be further implemented to support technology-based agricultural systems.

Keywords: Disease Detection, Fruit Trees, Random Forest, Smart Agriculture, Image Processing

1. PENDAHULUAN

Sektor pertanian memegang peranan penting dalam memenuhi kebutuhan pangan, serta sebagai penopang ekonomi di banyak negara, khususnya di daerah dengan iklim tropis yang mendukung budidaya tanaman buah. Namun, produktivitas tanaman buah sering kali terganggu oleh berbagai jenis penyakit yang dapat menyerang bagian tanaman, seperti daun, batang, dan buah. Penyakit-penyakit ini tidak hanya menurunkan kualitas dan kuantitas hasil panen tetapi juga mengancam keberlanjutan produksi apabila tidak segera ditangani. Oleh karena itu, deteksi dini terhadap penyakit tanaman sangat diperlukan untuk meminimalkan dampak negatif dan mempertahankan hasil produksi secara optimal.

Selama ini, identifikasi penyakit pada tanaman umumnya dilakukan secara manual oleh petani atau ahli pertanian berdasarkan pengamatan fisik. Metode ini memiliki beberapa keterbatasan, di antaranya keterbatasan waktu, biaya, serta kemungkinan kesalahan dalam diagnosis karena gejala yang tampak sering kali mirip antar berbagai jenis penyakit. Untuk mengatasi masalah ini, pengembangan metode otomatis berbasis teknologi untuk deteksi penyakit pada tanaman menjadi kebutuhan mendesak. Salah satu pendekatan yang banyak dikembangkan adalah teknik pembelajaran mesin (machine learning), yang memungkinkan identifikasi penyakit secara lebih cepat, tepat, dan konsisten.

Algoritma Random Forest merupakan salah satu teknik pembelajaran mesin yang populer karena kemampuannya dalam menangani dataset berdimensi tinggi serta efektifitasnya dalam klasifikasi yang kompleks. Algoritma ini bekerja dengan menggabungkan sejumlah pohon keputusan (decision trees) yang bekerja secara paralel, sehingga dapat mengurangi risiko overfitting dan menghasilkan prediksi yang lebih akurat. Beberapa penelitian telah menunjukkan potensi algoritma Random Forest dalam klasifikasi gambar, termasuk untuk deteksi penyakit tanaman. Karakteristik yang dimiliki tanaman, seperti warna daun, bentuk, tekstur, dan pola kerusakan akibat penyakit, dapat diekstraksi dari citra dan digunakan sebagai fitur dalam model Random Forest.

Penelitian ini bertujuan untuk mengembangkan model deteksi penyakit pada tanaman buah dengan memanfaatkan algoritma Random Forest. Data citra daun dan buah yang terinfeksi dan tidak terinfeksi akan dianalisis dan diolah sebagai input dalam pelatihan model. Diharapkan, hasil penelitian ini dapat berkontribusi pada peningkatan efektivitas sistem deteksi penyakit tanaman secara otomatis serta mendukung penerapan teknologi cerdas dalam pertanian.

2. METODE PENELITIAN

2.1 CNN (Convolutional Neural Network)

Jaringan saraf konvolusional (Convolutional Neural Network) adalah jenis jaringan yang sering digunakan untuk mengenali pola dalam gambar. CNN adalah teknik yang terinspirasi oleh cara mamalia – manusia menghasilkan persepsi visual. Jaringan saraf konvolusional bukanlah jaringan saraf secara umum. Lapisan konvolusional berisi neuron-neuron yang tersusun sehingga membentuk filter panjang dan tinggi.

2.2 Random Forest

Random forest adalah algoritma pembelajaran mesin yang umum digunakan, yang merupakan merek dagang Leo Breiman dan Adele Cutler, yang menggabungkan output dari beberapa pohon keputusan untuk mencapai satu hasil. Kemudahan penggunaan dan fleksibilitasnya telah mendorong adopsinya, karena menangani masalah klasifikasi dan regresi.

2.3 Metode Clustering Kmeans

Metode K-Means adalah sebuah algoritma yang dipakai untuk pengelompokan suatu data. Algoritma ini termasuk dalam unsupervised learning dan non-hierarchy. K-Means Clustering merupakan algoritma yang sering digunakan pada Machine Learning dan Data Mining.

2.4 Deep Learning

Deep learning adalah metode kecerdasan buatan (AI) yang menggunakan algoritma untuk mengajarkan komputer memproses data seperti cara kerja otak manusia. Deep learning merupakan bagian dari machine learning yang menggunakan artificial neural networks (ANN) atau jaringan saraf tiruan.

2.5 Pengumpulan Data

Pengumpulan data menggunakan kumpulan data yang diekstrak dari platform Kaggle, untuk mengetahui semua tentang penyakit tanaman di Bangladesh. Dataset ini berisi 31.100 gambar berbagai penyakit dengan ukuran total 7 GB. Dataset yang akan digunakan pada penelitian ini akan memperhitungkan hanya 2 kelas dari 15 kelas, khususnya fokus pada penyakit daun jagung.

2.6 Preprocessing Data

Kumpulan data yang diekstraksi dari Kaggle kemudian menjalani prapemrosesan data. Preprocessing merupakan bagian dari deep learning dimana kumpulan data berupa gambar akan diproses dengan model deep learning yaitu model Convolutional Neural Network.

2.7 Analisis Hasil

Klasifikasi gambar dengan studi kasus pada daun jagung menggunakan model jaringan saraf konvolusional dalam menunjukkan akurasi dengan nilai sebesar 0,9990 dengan 50 epoch. Kumpulan data yang digunakan untuk pengujian ini mungkin cukup untuk membuat data pengujian dan pelatihan.

3. ANALISA DAN PEMBAHASAN

Bagian ini berisi kesimpulan, hasil dan pembahasan topik penelitian.

Tabel 1. Hasil Temuan Artikel Relevan

N o.	Author/ Tahun	Metode Penelitian	Kelebihan	Kekurangan	Faktor Pengaruh Keberhasilan	Manfaat Sistem
1	Luthfiana Ratnawati dan Dwi Ratna Sulistyaningrum 2019	Random Forest	Random Forest memiliki akurasi yang baik , relatif kuat terhadap outlier dan noise, lebih cepat daripada bagging dan boosting, sederhana dan mudah untuk diparalelkan.	Memiliki resiko overfitting pada data tertentu, Konsumsi memori tinggi untuk dataset besar, Kurang akurat pada data yang berukuran besar.	Tingkat akurasi yang beragam pada hasil uji coba klasifikasi yang dipengaruhi oleh kombinasi proporsi data latih dan data uji, nilai akurasi yang tinggi pada penggunaan pohon keputusan yang semakin banyak dapat meningkatkan performa klasifikasi Random Forest.	Memper memudahkan identifikasi penyakit, membantu pengambilan keputusan terkait penanganan yang sesuai sehingga dapat mencegah penyebaran penyakit dan menurunkan resiko penurunan pada produksi.
2	Imam Yafi, Ahmad Fitry Boy,S.Kom., M.Kom., Dedi Setiawan, S.Kom., M.Kom. 2020	Metode Dempster Shafer	Memudahkan diagnosa penyakit tanaman secara cepat dan akurat tanpa perlu tenaga ahli yang harus selalu di lapangan	Memiliki keterbatasan untuk mendeteksi gejala baru yang belum terdata dalam pengetahuan sistem, sehingga harus dilakukan update secara berkala.	Ketepatan dalam penentuan nilai densitas dan evidence untuk setiap gejala agar perhitungan Dempster Shafer akurat.	Meningkatkan efisiensi proses deteksi penyakit tanaman sehingga memakan waktu yang lebih cepat dan mendukung peningkatan kualitas dan produksi.
3	Riksa Adenia, Agus Eko Minarno, Yufis Azhar 2022	Random Forest	Random Forest memiliki akurasi yang baik, hasil gabungan dari beberapa pohon keputusan hasil lebih	Kinerja model bergantung pada kualitas dan jumlah dataset yang digunakan, proses training model membutuhkan waktu komputasi yang lama dan	Data gambar yang jelas dan relevan dengan augmentasi dan preprocessing yang tepat sangat mempengaruhi hasil dari model dalam mendeteksi	Mengurangi waktu yang diperlukan untuk mendeteksi penyakit tanaman dan juga memberikan hasil yang cukup akurat

			tahan pada overfitting.	membutuhkan sumber daya yang tinggi.	penyakit tanaman.	dan bisa di andalkan.
4	Muh. Iqbal Rizky Nurfirdaus, Muhamad Rafi, Bayu Seno Adjie 2023	Convolutional Neural Network	CNN bisa mengolah data gambar secara efektif, dapat beradaptasi pada berbagai jenis penyakit tanaman dan klasifikasi citra lainnya.	Terdapat overfit pada model decision tree dan random forest, dan juga underfit pada model algoritma svm.	Dataset yang variatif dan representatif yang mempengaruhi kemampuan model dalam mengenali model yang berbeda.	Penggunaan random forest dapat mendeteksi dini penyakit pada tanaman dan memungkinkan intervensi yang tepat waktu.
5	Andhika Bagas Prakosa, Hendry, Radius Tanone 2023	Convolutional Neural Network	Akurasi tinggi pada klasifikasi gambar, dapat mengidentifikasi dan membedakan fitur penting pada gambar melalui berbagai lapisan pemrosesan.	Membutuhkan dataset yang besar untuk dapat mencapai hasil yang optimal, Waktu komputasi yang lama, terutama pada model dengan parameter besar dan jumlah epoch yang tinggi.	Confusion Matrix atau matriks kebingungan dalam deep learning digunakan untuk menunjukkan data hasil prediksi yang terbagi ke dalam beberapa kelas.	Meningkatkan produktivitas dan hasil panen yang baik dengan adanya klasifikasi penyakit pada tanaman.

4. KESIMPULAN

Pendekatan Random Forest terbukti efektif dalam mendeteksi penyakit pada tanaman buah dengan tingkat akurasi yang tinggi, terutama dalam mengklasifikasikan berbagai jenis penyakit berdasarkan ciri-ciri visual seperti warna, bentuk, tekstur, dan pola kerusakan pada daun dan buah. Algoritma ini mampu menangani dataset berdimensi tinggi serta mengurangi risiko overfitting, sehingga hasil prediksi menjadi lebih akurat dan konsisten.

Meskipun demikian, hasil deteksi masih dapat ditingkatkan dengan memperhatikan proses pra-pengolahan citra, seperti segmentasi dan ekstraksi fitur, untuk memastikan bahwa data yang digunakan benar-benar mewakili gejala penyakit yang relevan. Pendekatan ini menunjukkan potensi besar dalam pengembangan sistem deteksi penyakit tanaman secara otomatis, yang dapat diterapkan untuk mendukung pertanian berbasis teknologi cerdas dalam mengoptimalkan produktivitas dan menjaga kualitas hasil pertanian.

REFERENCES

- Badan Perencanaan Pembangunan Nasional, "Apel dalam Tentang Budidaya Pertanian," 2000.
- D. Putra, Pengolahan Citra Digital. Yogyakarta: Andi, 2010.
- S. Archana and A. Sahayadhas, "Automatic Rice Leaf Disease Segmentation Using Image Processing Techniques," Int. J. Eng. Technol., vol. 7, pp. 182–185, 2018.
- I. Nedjar, M. E. H. Daho, and N. Settout, "Random Forest Based Classification of Medical X-Ray Images Using a Medical Genetic Algorithm for Feature Selection," J. Mech. Medicin Biol., vol. 15, no. 2, pp. 1–8, 2015.

- R. Nunir, Pengolahan Citra Digital dengan Pendekatan Algoritmik. Bandung: Informatika, 2004
- H. Triwidodo, S. Wiyono, and P. B. Ayuwati, "Teknik Pembibitan dan Organisme Pengganggu Bibit Durian Menoreh Kuning di Kecamatan Kalibawang, Kulon Progo," *Agrovigor: Jurnal Agroekoteknologi*, vol. 13, no. 1, pp. 43–50, Mar. 2020, doi: 10.21107/agrovigor.v13i1.6061.
- H. Triwidodo and M. H. Tanjung, "Hama Penyakit Utama Tanaman Bawang Merah (*Allium Ascalonicum*) dan Tindakan Pengendalian di Brebes, Jawa Tengah," *Agrovigor: Jurnal Agroekoteknologi*, vol. 13, no. 2, pp. 149–154, Sep. 2020, doi: 10.21107/agrovigor.v13i2.7131.
- D. Dahang et al., "PENGENDALIAN PENYAKIT GANODERMA PADA KELAPA SAWIT DENGAN MENGGUNAKAN JAMUR ENDOFITIK HENDERSONIA," vol. 5, no. 2, pp. 548–559, 2021, doi: 10.31764/jmm.v5i2.4090.
- P. Purwadi and A. H. Nasyuha, "Implementasi Teorema Bayes Untuk Diagnosa Penyakit Hawar Daun Bakteri (Kresak) Dan Penyakit Blas Tanaman Padi," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 4, p. 777, Aug. 2022, doi: 10.30865/jurikom.v9i4.4350.
- Rosyani Perani, dkk., Menggunakan YOLO dalam deteksi objek: Tinjauan Literatur Sistematis, *Jurnal AI dan SPK: Jurnal Kecerdasan Buatan dan Sistem Pendukung Keputusan*, Volume 1, No. 1 Juni 2023.
- Rosyani Perani, dkk., Tinjauan literatur sistematis: sistem diagnostik pakar Pasien gigi menggunakan metode string biner langsung: *Jurnal Ilmu Komputer, Teknik dan Multimedia*, Volume 1, No. 1 April 2023.
- Rosyani Perani, dkk., MENDETEKSI PLAT NOMOR KENDARAAN DENGAN METODE YOLO (You Only Look Once) DAN SINGLE SHOT DETECTOR (SSD).
- Rosyani Perani, dkk., *Jurnal Manajemen, Ekonomi, Hukum, Kewirausahaan, Kesehatan, Pendidikan dan Informatika (MANEKIN)* Volume 1, No. 04, Juni Tahun 2023 ISSN 2985-4202 (media online) Hal 152-157 Erni | <https://journal.mediapublikasi.id/index.php/manekin> | Page 152 Sistem Pakar Diagnosa Penyakit Kulit Dengan Menggunakan Metode Forward Chaining
- Nugroho Agung Fajar, dkk., DAMPAK VIDEO PENDIDIKAN KESEHATAN TENTANG ASUHAN KEPERAWATAN HIPERTENSI TERHADAP TINGKAT PENGETAHUAN HIPERTENSI PADA TAHAP PENDIDIKAN, *Journal of Nursing Sciences (NSJ)*, volume 4, edisi Desember 2022.
- Syafii Abdullah, Wasis Haryono, APLIKASI PEMROGRAMAN EKSTRIM DALAM PEMBANGUNAN GAME EDUKASI ASMAUL HUSNA, SIFAT ALLAH DAN NAMA NABI MENGGUNAKAN APLIKASI CONSTRUCT 2, *Journal of Artificial Intelligence*. Penerbangan 3, tidak. 1 Februari 2022.
- Zakaria Hadi, dkk., menggunakan APLIKASI MICROSOFT EXCEL GUI SEBAGAI ALAT BELAJAR PEMBUATAN DATABASE, *AMMA: Jurnal Pengabdian Masyarakat*, Volume 1, No. 2 Maret (2022).