

**KLASIFIKASI DAUN HERBAL MENGGUNAKAN
WATERSHED DAN RANDOM FOREST**

SKRIPSI



Oleh:

LINTANG SARI PUTRI WARDHANI

NPM. 21081010247

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"

JAWA TIMUR

2024

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan salah satu negara yang memiliki tingkat keanekaragaman hayati yang tinggi atau biasa disebut megabiodiversitas di dunia. Hal ini disebabkan oleh letak geografis Indonesia yang membentuk bioregion yang memisahkan antara biogeografi flora-fauna Asia dan Australia. Selain itu, Indonesia memiliki ekosistem yang didominasi oleh hutan hujan tropis serta perairannya yang menyebabkan keanekaragaman hayati di Indonesia melimpah. Keanekaragaman hayati di Indonesia dibedakan menjadi beberapa tingkatan yaitu tingkat gen, spesies, dan ekosistem. Seperti halnya pada tanaman obat keluarga atau biasa disebut Daun Herbal, Daun Herbal termasuk pada tingkatan spesies dikarenakan setiap tanaman obat memiliki spesies yang berbeda dan setiap spesies memiliki khasiat yang berbeda pula.

Sejak dahulu, masyarakat Indonesia telah memanfaatkan tanaman obat keluarga sebagai pengobatan. Daun Herbal terbukti ampuh untuk mengobati penyakit manusia selama ribuan tahun yang lalu. Di Indonesia, terdapat kurang lebih 1.200 tanaman obat yang telah dimanfaatkan oleh masyarakat. Tanaman obat diolah menjadi jamu tradisional yang diracik alami tanpa menggunakan bahan kimia, khususnya masyarakat di Jawa atau diolah menjadi obat-obatan herbal. Daun Herbal sering disebut sebagai obat-obat ringan karena dapat mengobati beberapa penyakit seperti demam dan batuk. Secara umum, Daun Herbal biasa dimanfaatkan sebagai ramuan untuk gangguan kesehatan ringan, minuman kebugaran, memelihara kesehatan, dan juga dapat meningkatkan gizi. Ada beberapa bagian pada tanaman obat yang dapat dimanfaatkan seperti akar, rimpang, umbi, bunga, buah, biji, batang, dan daun. Sebagai bahan pembuatan obat, daun memiliki banyak manfaat selain karena mudah ditemui, penggunaannya juga relatif mudah.

Pada era modern seperti saat ini, masyarakat masih banyak yang memilih pengobatan tradisional. Pengobatan modern seringkali memiliki harga yang mahal dan tidak terjangkau bagi sebagian masyarakat terutama yang tinggal di pedesaan atau daerah-daerah terpencil, dimana mereka kesulitan untuk mengakses layanan kesehatan seperti puskesmas dan rumah sakit. Penggunaan obat-obatan kimia juga memiliki efek samping, sehingga masyarakat mencari pengobatan alternatif yang lebih aman dan alami. Selain itu, Daun Herbal mudah ditemukan baik di pasar tradisional atau pada

pekarangan rumah. Penelitian yang telah dilakukan (susilo, 2016) mengatakan bahwa pemahaman masyarakat tentang kesehatan akan mempengaruhi pemakaian obat herbal yang berasal dari tanaman dengan metode tradisional yang telah dimanfaatkan sejak zaman nenek moyang bangsa indonesia. Who (world health organization) telah menganjurkan gaya hidup sehat dengan kembali ke alam. Dengan adanya anjuran tersebut, pemerintah mulai kembali mengembangkan Daun Herbal. Daun Herbal harus dijadikan sebagai sarana pembelajaran sejak dini untuk mengenal berbagai macam tanaman obat terkhusus pada macam-macam daun tanaman obat. Edukasi terkait Daun Herbal dapat membentuk karakter kemandirian dalam upaya pengobatan keluarga.

Dalam pengenalan tanaman obat menggunakan metode tradisional seringkali tidak efisien karena banyak tanaman obat memiliki ciri morfologi yang mirip, terutama pada daun. Hal ini membuat sulit membedakan jika hanya berdasarkan pengamatan visual. Tumbuhan yang sama seringkali memiliki nama yang berbeda-beda disetiap daerah. Selain itu, deskripsi fisik yang bergantung pada pengamatan visual seringkali kurang presisi seperti “daun lebar”, “batang berkayu”, atau “bunga kecil” akan memiliki interpretasi yang berbeda setiap orang. Penggunaan metode tradisional dapat mengakibatkan kesalahan identifikasi daun tanaman obat dan berpotensi menyebabkan bahaya jika penggunaannya tidak tepat. Selain itu, identifikasi yang berulang-ulang dapat membuang waktu. Tantangan ini membuka peluang untuk mengembangkan teknologi berbasis komputer untuk klasifikasi daun, khususnya menggunakan metode pengolahan citra digital. Di era digital, teknologi pengolahan citra semakin banyak digunakan untuk mengidentifikasi berbagai jenis objek, termasuk daun tanaman. Pengolahan citra memungkinkan identifikasi yang lebih cepat dan akurat, sehingga menjadi alternatif untuk mendukung pengenalan dan klasifikasi daun tanaman obat keluarga.

Beberapa penelitian pernah dilakukan untuk dengan teknologi ini, seperti klasifikasi daun herbal berdasarkan fitur invariant moment dan fitur geometri yang di klasifikasi menggunakan knn dan naive bayes. Kemudian penelitian menggunakan random forest berdasarkan fitur histogram of gradients yang mendapatkan akurasi keseluruhan sebesar 85,33%. Tantangan utama meliputi kesulitan dalam segmentasi batas daun yang presisi dan ekstraksi fitur yang representatif. Oleh karena itu, penelitian ini bertujuan untuk mengklasifikasi daun Daun Herbal menggunakan metode watershed untuk segmentasi dan random forest untuk klasifikasi. Penelitian ini diharapkan dapat

memberikan hasil yang akurat dan andal dalam mengenali daun Daun Herbal berdasarkan karakteristik morfologinya.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, maka rumusan masalah yang teridentifikasi meliputi:

1. Bagaimana kinerja metode Random Forest dalam mengklasifikasi jenis daun tanaman obat keluarga setelah proses segmentasi Watershed?
2. Bagaimana pengaruh segmentasi Watershed terhadap akurasi klasifikasi jenis daun tanaman obat keluarga?

1.3 Tujuan

Berdasarkan rumusan masalah tersebut, adapun tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mengetahui kinerja metode Random Forest dalam mengklasifikasi jenis daun tanaman obat keluarga setelah proses segmentasi Watershed
2. Mengetahui pengaruh segmentasi Watershed terhadap akurasi klasifikasi jenis daun tanaman obat keluarga?

1.4 Manfaat

Berdasarkan pada tujuan penelitian yang telah disebutkan sebelumnya, manfaat yang dapat dihasilkan dari penelitian ini adalah sebagai berikut:

1. Membantu masyarakat dan ahli herbal dalam mengenali jenis daun herbal dengan cepat dan tepat berdasarkan karakteristik morfologinya.
2. Mengurangi waktu dan kesalahan dalam proses identifikasi manual dengan memanfaatkan teknologi pengolahan citra digital.
3. Menyediakan informasi yang lebih terstruktur dan mudah diakses terkait berbagai jenis daun herbal untuk kebutuhan pengobatan tradisional.

1.5 Batasan Masalah

Untuk mencegah pembahasan yang terlalu jauh dari pembahasan utama, berikut adalah batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Menggunakan 10 jenis daun DAUN HERBAL yaitu daun belimbing wuluh, daun jambu biji, daun jeruk nipis, daun kemangi, daun lidah buaya, daun nangka, daun pandan, daun pepaya, daun seledri, dan daun sirih.
2. Dilakukan satu kali pengujian dengan menggunakan segmentasi Watershed dengan metode klasifikasi yaitu Random Forest.
3. Menggunakan metode K-Fold Cross Validation untuk evaluasi kinerja model.

4. Menggunakan bahasa python dalam proses pembuatan dan pengembangannya.

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Pada penelitian dengan yang dilakukan oleh Nurul Hidayah, dkk “Klasifikasi Daun Herbal Berdasarkan Fitur Bentuk dan Tekstur Menggunakan KNN” digunakan ekstraksi fitur HOG dan LBP untuk mengklasifikasikan 15 jenis daun herbal. Akurasi HOG mencapai 92,67% dan LBP 88,67%.

Penelitian yang dilakukan oleh Slamet Fifin Alamsyah “Implementasi Deep Learning Untuk Klasifikasi Tanaman Daun Herbal Berdasarkan Ciri Daun Berbasis Android” mengatakan bahwa Masih sedikit yang mampu mengenali tumbuhan obat, seperti anak SD, SMP dan SMA, masih banyak yang tidak mengetahui nama dari tumbuhan obat, oleh karena itu penulis melakukan penelitian untuk membuat aplikasi klasifikasi tanaman Daun Herbal yang mampu mengenali jenis tanaman Daun Herbal berdasarkan daun hanya dengan menggunakan perangkat mobile yang mana bisa digunakan dengan mudah untuk mengetahui jenis tanaman Daun Herbal, hanya dengan mengambil foto daun dari tanaman Daun Herbal dapat diketahui jenis tanaman Daun Herbal, sehingga dibutuhkan pendekatan untuk penyelesaian masalah ini. Pendekatan dalam penyelesain masalah ini menggunakan machine learning (ML), salah satu cabang artificial intelligence (AI) yang populer, dimana mesin mampu belajar seperti layaknya pikiran manusia. ML sendiri mempunyai bidang keilmuan baru yaitu deep learning, dimana mesin mampu melakukan pembelajaran lebih dalam, pada metode deep learning ada metode yang cocok digunakan untuk mengklasifikasikan sebuah citra yaitu metode Convolutional Neural Network (CNN), kelebihan dari CNN adalah mampu melakukan proses pembelajaran fitur-fitur dari citra secara mandiri yang disebut dengan feature learning, berbeda dengan feature extraction yang harus mendapatkan fitur-fitur dari citra terlebih dahulu sebelum melakukan klasifikasi. CNN digunakan untuk membedakan jenis tanaman dengan memberikan label dari daun tanaman Daun Herbal. Pada penelitian ini menggunakan 10 kelas jenis tanaman Daun Herbal yaitu teh hijau, tapak dewa, sirsak, semanggi, mengkudu, mahoni, kumis kucing, jambu biji, blimbing wuluh, bayam merah, Pengujian terhadap data pelatihan menghasilkan akurasi 75% dan data pengujian menghasilkan akurasi 80%.

Penelitian yang dilakukan oleh Rizki Ramadhani dkk, “Klasifikasi Tanaman Obat Berdasarkan Citra Daun Menggunakan Backpropagation Neural Networks”.

Penelitian ini menggunakan algoritma Backpropagation Neural Networks (BNN) untuk mengklasifikasikan citra daun herbal dengan akurasi mencapai 93,75%. Proses melibatkan segmentasi citra dan ekstraksi ciri bentuk.

Kemudian juga penelitian yang dilakukan oleh Raishudin Jafar, dkk yang mengatakan bahwa Daun adalah salah satu bagian tumbuhan yang mempunyai manfaat bagi manusia khususnya untuk kesehatan tubuh. Daun dapat digunakan sebagai obat herbal yang dapat menjadi alternatif yang dapat membantu dalam meningkatkan imun serta daya tahan tubuh. Akan tetapi tidak semua daun memiliki khasiat sebagai obat, maka dari itu pengetahuan mengenai jenis-jenis daun berkhasiat obat menjadi suatu hal yang penting. Tujuan dari penelitian yaitu untuk mengembangkan sistem klasifikasi citra daun berkhasiat obat dengan menerapkan model jaringan syaraf tiruan Extreme Learning Machine (ELM). Untuk mendukung algoritma ELM digunakan ekstraksi ciri morfologi yang dapat memberikan informasi mengenai karakteristik bentuk dari objek yang ada. Extreme Learning Machine (ELM) disebut juga sebagai pendekatan jaringan syaraf tiruan yang menggunakan satu hidden layer. Pada tahap klasifikasi algoritma Extreme Learning Machine (ELM) dapat menentukan nilai bobot antara input neuron dan hidden layer secara random sehingga pola pembelajaran menjadi lebih cepat. Berdasarkan hasil uji precision, recall dan accuracy menghasilkan nilai precision sebesar 90.67%, nilai recall sebesar 89.47% dan accuracy sebesar 90%. Sehingga, berdasarkan hasil tersebut dapat dikatakan bahwa model ELM yang dibangun dapat mengklasifikasikan citra jenis daun berkhasiat obat dengan baik.

2.2 Daun Herbal

Daun herbal adalah bagian tanaman yang sering digunakan dalam pengobatan tradisional karena kandungan senyawa aktif seperti flavonoid, alkaloid, dan tanin. Di Indonesia, tanaman obat keluarga (Daun Herbal) memainkan peran penting dalam menjaga kesehatan masyarakat. Daun herbal memiliki beragam bentuk, tekstur, dan pola yang menjadi ciri khas tiap spesies. Penelitian pada daun herbal penting untuk mendukung identifikasi, klasifikasi, dan optimalisasi manfaatnya, terutama di tengah perkembangan teknologi modern.

Dalam klasifikasi daun herbal, pengenalan karakteristik morfologi daun menjadi langkah awal yang sangat penting. Morfologi ini meliputi bentuk tepi daun, pola urat, warna, dan tekstur. Namun, keragaman bentuk daun sering kali menjadi tantangan dalam proses identifikasi manual. Oleh karena itu, diperlukan teknologi berbasis komputer untuk meningkatkan efisiensi dan akurasi klasifikasi.

Daun herbal merupakan bagian tanaman yang secara tradisional dimanfaatkan dalam pengobatan untuk mengatasi berbagai penyakit. Kandungan bioaktif seperti flavonoid, alkaloid, tanin, dan saponin menjadi alasan utama penggunaannya. Di Indonesia, keanekaragaman hayati yang melimpah memberikan potensi besar untuk pemanfaatan daun herbal, terutama dari tanaman obat keluarga (Daun Herbal). Selain digunakan sebagai bahan jamu dan ramuan tradisional, daun herbal juga banyak dikembangkan untuk produk farmasi modern.

Keanekaragaman bentuk, tekstur, pola, dan warna daun menjadi ciri khas yang membedakan satu jenis tanaman dengan lainnya. Namun, identifikasi manual daun herbal sering kali sulit dilakukan karena adanya kemiripan morfologi antara beberapa jenis tanaman. Penelitian terhadap daun herbal penting untuk memberikan solusi identifikasi yang lebih akurat, terutama dengan bantuan teknologi modern seperti pengolahan citra.

Dalam klasifikasi daun herbal, pengenalan karakteristik morfologi menjadi langkah awal yang krusial. Bentuk tepi daun, pola urat, dan tekstur permukaan menjadi indikator penting dalam proses ini. Namun, pendekatan manual membutuhkan waktu lama dan rentan terhadap kesalahan, sehingga teknologi berbasis komputer menjadi solusi potensial. Sistem otomatis memungkinkan klasifikasi lebih cepat, efisien, dan presisi, serta mendukung pelestarian ilmu pengobatan tradisional di era modern.

2.3 Machine Learning

Machine learning adalah salah satu bidang dari kecerdasan buatan yang bertujuan untuk memungkinkan komputer belajar dari data tanpa perlu diinstruksikan secara eksplisit. Dalam konteks pengolahan citra, machine learning memungkinkan identifikasi objek secara otomatis berdasarkan pola yang telah dipelajari dari dataset. Machine learning sangat relevan dalam penelitian klasifikasi daun herbal, karena kompleksitas data citra memerlukan algoritma yang dapat menangkap pola non-linear dan variabel multivariat.

Salah satu aplikasi utama machine learning dalam pengolahan citra adalah klasifikasi. Dengan memanfaatkan algoritma seperti Random Forest, model dapat dilatih untuk mengenali karakteristik unik daun herbal, seperti bentuk, tekstur, dan pola urat. Algoritma machine learning bekerja dengan mengekstraksi fitur dari citra, mengolahnya melalui proses pelatihan, dan membuat prediksi berdasarkan data baru. Keunggulan machine learning adalah kemampuannya untuk memproses data dalam jumlah besar dengan kecepatan tinggi dan akurasi yang baik. Dalam penelitian ini,

machine learning memainkan peran penting sebagai inti dari proses klasifikasi. Penggabungan metode segmentasi dan ekstraksi fitur yang kuat dengan algoritma klasifikasi canggih memungkinkan pengembangan sistem klasifikasi daun herbal yang efektif dan andal.

Dalam pengolahan citra, machine learning digunakan untuk mengenali pola, memproses data, dan melakukan klasifikasi objek secara otomatis. Algoritma seperti Random Forest telah terbukti efektif dalam menangani dataset kompleks dan beragam. Penggunaan machine learning pada klasifikasi daun herbal memungkinkan pemrosesan dalam jumlah besar dengan kecepatan yang jauh lebih tinggi dibandingkan metode manual. Melalui ekstraksi fitur seperti tekstur, bentuk, dan pola urat, model machine learning dapat melatih algoritma untuk mengenali karakteristik unik dari masing-masing spesies daun.

2.4 Segmentasi Watershed

Segmentasi Watershed adalah metode pengolahan citra yang digunakan untuk memisahkan objek dalam gambar berdasarkan intensitas piksel. Metode ini bekerja seperti air yang mengalir ke lembah, di mana setiap wilayah rendah membentuk segmen yang terpisah. Dalam penelitian ini, segmentasi Watershed digunakan untuk memisahkan daun herbal dari latar belakang atau objek lainnya dalam citra.

Salah satu keunggulan utama dari metode Watershed adalah kemampuannya menangkap detail batas objek, termasuk bentuk tepi daun yang kompleks. Hal ini membuat metode ini sangat cocok untuk aplikasi pengolahan citra daun herbal, terutama dalam mengidentifikasi tepi atau area daun yang tumpang tindih. Namun, Watershed memiliki kelemahan seperti over-segmentation, yang dapat terjadi karena noise atau variasi kecil pada intensitas piksel.

Untuk mengatasi kelemahan tersebut, preprocessing seperti smoothing atau thresholding sering dilakukan sebelum segmentasi Watershed diterapkan. Teknik ini membantu mengurangi noise sehingga hasil segmentasi menjadi lebih bersih dan sesuai dengan kebutuhan analisis. Selain itu, kombinasi dengan teknik morfologi dapat meningkatkan akurasi segmentasi.

Dalam penelitian ini, segmentasi Watershed menjadi langkah awal yang penting sebelum ekstraksi fitur dilakukan. Kualitas segmentasi sangat menentukan akurasi keseluruhan sistem, karena fitur-fitur yang dihasilkan sangat bergantung pada seberapa baik objek dapat dipisahkan.

Dengan hasil segmentasi yang baik, daun herbal dapat diidentifikasi dengan lebih akurat menggunakan metode machine learning pada tahapan berikutnya. Hal ini menunjukkan pentingnya peran segmentasi Watershed dalam mendukung keberhasilan sistem klasifikasi secara keseluruhan.

2.5 Random Forest

Random Forest adalah algoritma machine learning berbasis ensemble yang menggabungkan banyak decision tree untuk membuat prediksi yang lebih akurat. Dalam penelitian ini, Random Forest digunakan untuk mengklasifikasikan daun herbal berdasarkan fitur-fitur yang telah diekstraksi dari citra. Algoritma ini terkenal karena stabilitas dan akurasinya, terutama dalam menangani dataset yang kompleks.

Cara kerja Random Forest dimulai dengan membangun beberapa decision tree dari subset data yang dipilih secara acak. Setiap tree memberikan prediksi, dan hasil akhirnya ditentukan melalui voting mayoritas untuk klasifikasi. Pendekatan ini membuat Random Forest tahan terhadap overfitting, yang sering menjadi masalah dalam algoritma berbasis tree.

Keunggulan lain dari Random Forest adalah kemampuannya untuk menentukan fitur mana yang paling relevan untuk klasifikasi. Dalam penelitian ini, fitur seperti bentuk daun, pola urat, dan tekstur permukaan menjadi indikator penting yang dipertimbangkan oleh model untuk membedakan jenis daun herbal.

Random Forest juga memiliki kecepatan pemrosesan yang cukup baik, bahkan untuk dataset berukuran besar. Hal ini menjadikannya algoritma yang efisien untuk digunakan dalam pengolahan citra daun herbal. Selain itu, kemampuan algoritma ini dalam menangani variabel non-linear dan multivariat membuatnya sangat cocok untuk tugas klasifikasi dengan data kompleks.

Dengan menggunakan Random Forest, penelitian ini bertujuan untuk menghasilkan sistem klasifikasi yang tidak hanya akurat tetapi juga dapat diandalkan untuk berbagai jenis daun herbal. Algoritma ini memainkan peran kunci dalam mengintegrasikan hasil segmentasi dengan proses klasifikasi akhir.

BAB III

METODOLOGI PENELITIAN

3.1 Prosedur Penelitian dan Objek Penelitian

Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi daun herbal dengan memanfaatkan metode Watershed untuk segmentasi gambar dan algoritma Random Forest untuk klasifikasi. Proses penelitian diawali dengan identifikasi kebutuhan sistem, pengumpulan data, hingga implementasi dan evaluasi. Pendekatan ini memungkinkan sistem yang dihasilkan untuk dapat berfungsi secara efektif dalam mengenali jenis daun herbal dengan akurasi tinggi.

Objek penelitian dalam studi ini adalah daun-daun herbal yang biasa digunakan dalam pengobatan tradisional, seperti daun sirih, daun mint, dan daun binahong. Daun-daun ini memiliki karakteristik visual yang berbeda sehingga menjadi tantangan dalam proses pengolahan citra dan klasifikasi. Variasi bentuk, warna, dan tekstur pada daun herbal memerlukan metode yang tepat agar dapat dikenali secara akurat.

Prosedur penelitian dirancang secara sistematis dengan tahapan yang saling berkesinambungan. Peneliti memulai dengan akuisisi data berupa gambar daun, dilanjutkan dengan pra-proses data untuk meningkatkan kualitas gambar. Tahapan ini diikuti oleh segmentasi gambar menggunakan metode Watershed untuk memisahkan objek daun dari latar belakangnya. Fitur-fitur yang relevan dari daun diekstraksi dan digunakan dalam proses klasifikasi menggunakan Random Forest.

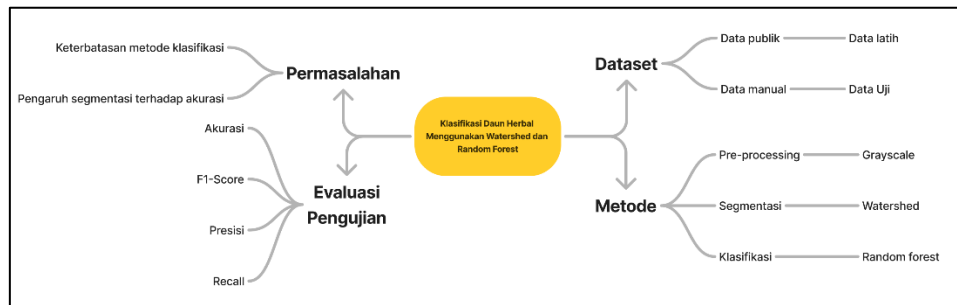
Metode yang digunakan dalam penelitian ini dipilih berdasarkan keunggulannya. Watershed dikenal efektif untuk segmentasi gambar dengan kontur yang kompleks, sementara Random Forest memberikan hasil klasifikasi yang stabil dan akurat meskipun data memiliki dimensi tinggi. Kombinasi kedua metode ini diharapkan mampu menghasilkan sistem klasifikasi daun herbal yang andal.

Prosedur penelitian diakhiri dengan pengujian dan evaluasi sistem. Sistem diuji menggunakan dataset terpisah untuk mengukur akurasi, kecepatan proses, dan kemampuan generalisasi. Evaluasi ini memberikan gambaran performa sistem serta potensi pengembangannya untuk aplikasi nyata.

3.2 Desain Sistem

Desain sistem pada penelitian ini melibatkan berbagai komponen yang saling terintegrasi untuk memastikan alur proses berjalan dengan baik. Sistem ini dirancang

untuk menerima input berupa gambar daun herbal dan memberikan output berupa jenis daun yang terklasifikasi dengan akurat. Proses ini terdiri dari beberapa tahap utama: akuisisi data, pra-proses data, segmentasi, ekstraksi fitur, dan klasifikasi.



Komponen pertama adalah modul akuisisi data yang berfungsi untuk mengumpulkan gambar daun herbal. Gambar-gambar ini kemudian diproses dalam modul pra-proses data untuk meningkatkan kualitasnya, seperti penghapusan noise dan peningkatan kontras. Modul ini memastikan bahwa gambar yang digunakan pada tahap segmentasi memiliki kualitas yang optimal.

Tahap segmentasi menggunakan metode Watershed untuk memisahkan objek daun dari latar belakang. Hasil segmentasi ini menjadi input untuk modul ekstraksi fitur. Fitur-fitur seperti bentuk, warna, dan tekstur daun diekstraksi untuk merepresentasikan karakteristik unik masing-masing jenis daun herbal.

Setelah fitur diekstraksi, modul klasifikasi menggunakan algoritma Random Forest untuk menentukan jenis daun. Random Forest dipilih karena kemampuannya menangani dataset dengan dimensi yang tinggi dan memberikan hasil yang stabil. Algoritma ini bekerja dengan membangun banyak pohon keputusan selama pelatihan dan memberikan prediksi berdasarkan hasil voting dari seluruh pohon.

Desain sistem ini juga mencakup pengujian dan evaluasi untuk memastikan performanya. Parameter seperti akurasi klasifikasi, kecepatan proses, dan kemampuan generalisasi menjadi indikator utama yang digunakan untuk mengevaluasi keberhasilan sistem.

3.3 Akuisisi Data

Tahap akuisisi data dilakukan dengan mengumpulkan gambar daun herbal dari berbagai sumber yang relevan dan mendukung kebutuhan penelitian. Data ini diperoleh melalui dua cara utama: pengambilan gambar secara langsung dan pemanfaatan dataset publik. Pengambilan gambar secara langsung dilakukan dengan menggunakan kamera beresolusi tinggi untuk memastikan setiap detail visual dari daun dapat tertangkap dengan jelas. Gambar-gambar ini diambil di bawah pengaturan yang terkendali, seperti

pencahayaannya yang stabil dan latar belakang seragam, untuk meminimalkan noise yang tidak diinginkan.

Dataset publik juga menjadi sumber data penting yang melengkapi variasi data yang diperoleh secara langsung. Dataset ini sering kali mencakup ribuan gambar dengan berbagai jenis daun dan kondisi pencahayaan yang berbeda, sehingga memberikan keragaman dalam data. Penggunaan dataset publik memungkinkan sistem untuk belajar dari data yang mencakup lebih banyak variasi bentuk, tekstur, dan warna daun, yang akan meningkatkan kemampuan generalisasi sistem dalam mengenali daun dari kondisi dunia nyata yang berbeda.

Dalam proses akuisisi, perhatian khusus diberikan pada kondisi pengambilan gambar seperti sudut kamera, tingkat pencahayaan, dan latar belakang. Gambar yang buram, tidak terfokus, atau memiliki latar belakang yang kompleks dapat memengaruhi akurasi segmentasi dan klasifikasi secara signifikan. Oleh karena itu, langkah penyaringan dilakukan untuk memastikan hanya gambar dengan kualitas yang memenuhi kriteria tertentu yang akan digunakan pada proses lebih lanjut. Selain itu, pengambilan gambar dilakukan dari berbagai sudut untuk menangkap karakteristik daun secara keseluruhan, termasuk variasi bentuk pinggir daun dan pola urat.

Setiap gambar yang diakuisisi kemudian diberi label berdasarkan jenis daun herbalnya. Pemberian label ini menjadi langkah penting dalam proses pelatihan model, karena model klasifikasi membutuhkan data yang terstruktur dan terlabel untuk belajar mengenali pola-pola yang membedakan satu jenis daun dengan jenis lainnya. Labeling dilakukan secara manual oleh ahli botani untuk memastikan akurasi, terutama dalam identifikasi daun yang memiliki kemiripan bentuk dengan jenis lainnya.

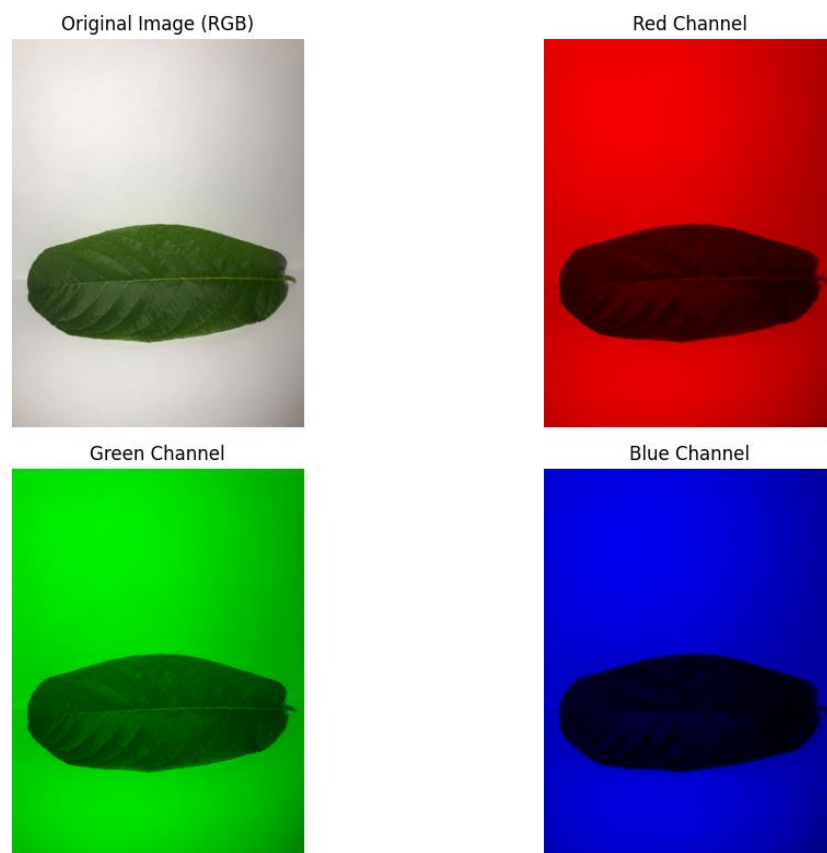
Proses akuisisi diakhiri dengan penyimpanan data dalam format standar, seperti JPG atau PNG, yang dipilih karena kompatibilitasnya dengan perangkat lunak pengolahan citra yang digunakan dalam penelitian ini. Selain itu, metadata seperti lokasi pengambilan gambar, waktu, dan kondisi pencahayaan juga dicatat untuk analisis lebih lanjut. Dengan langkah ini, dataset yang dihasilkan tidak hanya berkualitas tinggi, tetapi juga siap digunakan dalam proses pra-proses data, segmentasi, dan klasifikasi.

3.4 Pra Proses Data

Pra-proses data merupakan salah satu tahap paling krusial dalam pengolahan citra digital untuk memastikan gambar yang digunakan memiliki kualitas optimal. Tahap ini bertujuan untuk mempersiapkan data mentah menjadi bentuk yang lebih terstruktur, bersih, dan siap untuk diproses pada langkah-langkah selanjutnya. Sebagai

langkah awal, dilakukan percobaan menggunakan citra daun jambu biji. Berbagai teknik diterapkan selama pra-proses untuk mengatasi masalah yang mungkin timbul akibat kualitas data yang tidak konsisten, seperti ukuran yang berbeda-beda, noise, atau kontras yang rendah.

Langkah pertama adalah resizing, yaitu menyesuaikan ukuran gambar agar seragam di seluruh dataset. Penyesuaian ukuran ini penting untuk memudahkan pemrosesan di tahap segmentasi dan klasifikasi, terutama ketika model yang digunakan membutuhkan input dengan dimensi tetap. Ukuran gambar dipilih sedemikian rupa agar tidak terlalu besar sehingga menghambat efisiensi pemrosesan, namun tetap cukup detail untuk menangkap informasi penting dari tekstur dan pola daun.



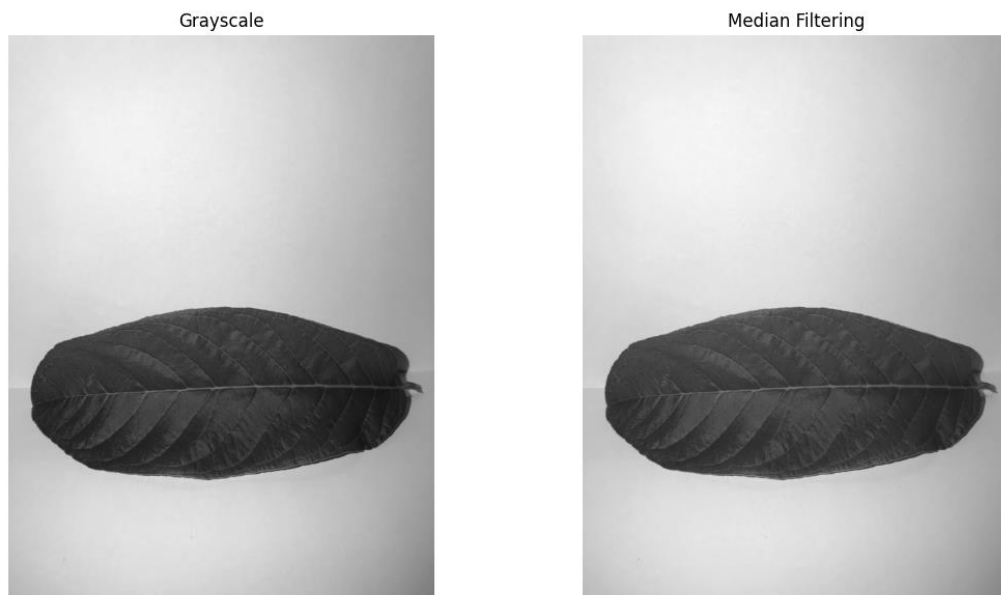
Gambar. Citra RGB

Langkah berikutnya adalah konversi gambar ke skala abu-abu (grayscale). Proses ini bertujuan untuk mengurangi dimensi data tanpa menghilangkan informasi penting yang diperlukan untuk analisis lebih lanjut. Dengan menghilangkan informasi warna, sistem dapat fokus pada elemen tekstur, pola urat, dan bentuk tepi daun. Selain itu, konversi ke grayscale mengurangi kompleksitas perhitungan, sehingga proses segmentasi dan klasifikasi menjadi lebih cepat.



Gambar. Mengubah RGB ke Grayscale

Untuk mengatasi noise, digunakan filter median, yang merupakan salah satu metode paling efektif dalam penghapusan noise tanpa merusak detail penting pada gambar. Filter ini bekerja dengan mengganti nilai piksel dengan median dari piksel di sekitarnya, sehingga dapat menghilangkan bercak-bercak kecil atau gangguan visual lainnya yang sering muncul pada gambar hasil akuisisi. Noise seperti ini, jika tidak ditangani, dapat memengaruhi kemampuan metode segmentasi untuk mendeteksi batas objek secara akurat.



Gambar. Median Filtering

Selain itu, dilakukan peningkatan kontras menggunakan metode histogram equalization. Metode ini bertujuan untuk mempertegas perbedaan antara objek daun

dan latar belakang, sehingga batas-batas daun menjadi lebih jelas. Peningkatan kontras sangat penting untuk segmentasi berbasis Watershed, yang mengandalkan informasi gradien pada gambar untuk menentukan batas objek. Dengan meningkatkan kontras, metode ini dapat bekerja lebih optimal dalam membedakan daun dari latar belakang atau objek lain yang tidak relevan.



Gambar. Citra Menggunakan Histogram Equalization

Hasil dari tahap pra-proses adalah gambar yang telah bersih, tajam, dan memiliki kualitas visual yang lebih baik. Setiap gambar yang telah diproses berada dalam kondisi optimal untuk digunakan pada tahap segmentasi dan klasifikasi. Dengan data yang dipersiapkan secara hati-hati melalui pra-proses ini, performa keseluruhan sistem, mulai dari segmentasi hingga klasifikasi akhir, dapat ditingkatkan secara signifikan. Langkah ini memastikan bahwa sistem bekerja dengan data yang representatif dan bebas dari gangguan teknis yang dapat memengaruhi hasil akhir.

3.5 Segmentasi Menggunakan Watershed

Segmentasi menggunakan metode Watershed merupakan salah satu pendekatan berbasis pengolahan citra yang efektif dalam memisahkan objek utama dari latar belakang. Metode ini sering digunakan dalam pengolahan citra medis dan botani, termasuk dalam penelitian ini untuk mendeteksi dan mengekstrak daun herbal dari latar belakangnya. Prinsip dasar metode Watershed didasarkan pada konsep topografi gambar, di mana gambar dianggap sebagai medan ketinggian dengan daerah terang sebagai puncak dan daerah gelap sebagai lembah.

Tahap awal dalam segmentasi Watershed adalah menghitung gradien intensitas pada gambar hasil pra-proses. Gradien ini memberikan informasi tentang perubahan

intensitas piksel, sehingga membantu mendeteksi tepi objek pada gambar. Dengan menghitung gradien, perbedaan antara daun dan latar belakang menjadi lebih jelas. Namun, gradien saja sering kali menghasilkan noise atau tepi yang terputus, sehingga diperlukan teknik tambahan untuk memperbaiki hasilnya.

Langkah berikutnya adalah pembuatan marker internal dan eksternal, yang merupakan inti dari algoritma Watershed. Marker internal digunakan untuk menandai area dalam objek (daun), sementara marker eksternal digunakan untuk menandai area di luar objek (latar belakang). Marker ini membantu algoritma untuk memulai segmentasi dari titik-titik tertentu yang telah ditentukan, sehingga proses segmentasi menjadi lebih terarah dan akurat. Marker internal biasanya dibuat dengan cara thresholding, sementara marker eksternal dibuat dengan cara invers dari marker internal. Setelah marker dibuat, algoritma Watershed diterapkan pada gambar gradien dengan bantuan marker tersebut. Proses ini menyerupai pengisian air pada lembah-lembah dalam medan topografi. Air mengalir dari marker internal dan eksternal hingga bertemu pada batas antara objek dan latar belakang. Hasil akhirnya adalah batas-batas yang jelas antara daun herbal dan elemen lain pada gambar. Untuk menghindari over-segmentasi, yaitu ketika satu objek dibagi menjadi beberapa segmen kecil, dilakukan pemrosesan tambahan seperti penggabungan segmen kecil atau smoothing tepi hasil segmentasi. Hasil segmentasi berupa gambar biner, di mana piksel yang termasuk dalam daun herbal diberi nilai 1 (putih), sedangkan piksel lainnya diberi nilai 0 (hitam). Gambar biner ini kemudian digunakan untuk mengekstrak fitur dari daun, seperti bentuk, tekstur, dan pola urat, yang akan menjadi masukan untuk tahap klasifikasi. Dengan segmentasi Watershed yang diterapkan secara efektif, daun dapat dipisahkan dari latar belakang dengan presisi tinggi, sehingga meningkatkan akurasi klasifikasi pada tahap berikutnya.

3.6 Klasifikasi Menggunakan Random Forest

Tahap klasifikasi merupakan inti dari penelitian ini, di mana daun herbal yang telah disegmentasi akan diidentifikasi berdasarkan jenisnya. Dalam penelitian ini, algoritma Random Forest digunakan sebagai metode klasifikasi utama karena keandalannya dalam menangani data dengan dimensi tinggi dan kemampuannya untuk mengatasi overfitting. Random Forest adalah algoritma berbasis ensemble learning yang terdiri dari banyak pohon keputusan (decision trees) yang bekerja secara paralel untuk menghasilkan prediksi akhir.

Proses klasifikasi diawali dengan ekstraksi fitur dari gambar daun yang telah melalui tahap segmentasi. Fitur-fitur yang diekstrak meliputi karakteristik bentuk,

tekstur, warna, dan pola urat daun. Fitur bentuk mencakup metrik seperti luas, keliling, dan rasio panjang terhadap lebar. Fitur tekstur diukur menggunakan matriks co-occurrence (GLCM), yang memberikan informasi tentang pola kekasaran atau keseragaman pada permukaan daun. Fitur warna, meskipun terbatas pada gambar grayscale, tetap dapat mencakup variasi intensitas yang relevan.

Setelah fitur diekstraksi, dataset disiapkan dengan pembagian data menjadi set pelatihan dan pengujian. Set pelatihan digunakan untuk melatih model Random Forest, sementara set pengujian digunakan untuk mengevaluasi performa model. Selama pelatihan, algoritma Random Forest membangun beberapa pohon keputusan secara acak dengan menggunakan subset data dan subset fitur. Tiap pohon membuat prediksi independen, dan hasil akhir ditentukan melalui voting mayoritas dari semua pohon. Kelebihan utama Random Forest adalah kemampuannya untuk menangani ketidakseimbangan data dan mengurangi overfitting. Karena setiap pohon hanya menggunakan sebagian data untuk pelatihan, model tetap memiliki generalisasi yang baik pada data yang tidak terlihat. Selain itu, algoritma ini juga menyediakan metrik penting seperti feature importance, yang menunjukkan kontribusi masing-masing fitur terhadap prediksi akhir. Informasi ini dapat digunakan untuk mengidentifikasi fitur mana yang paling relevan dalam membedakan jenis daun.

Hasil klasifikasi disajikan dalam bentuk metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Jika akurasi model masih dianggap kurang memadai, proses tuning parameter dilakukan untuk mengoptimalkan performa. Beberapa parameter yang disesuaikan mencakup jumlah pohon dalam hutan, kedalaman maksimum pohon, dan kriteria pemisahan. Dengan pendekatan ini, Random Forest dapat memberikan hasil klasifikasi yang andal untuk berbagai jenis daun herbal yang menjadi fokus penelitian.

3.7 Skenario Uji Coba Sistem

Untuk memastikan sistem bekerja dengan baik, diperlukan serangkaian uji coba yang terstruktur dan mencakup berbagai skenario. Uji coba dilakukan untuk mengevaluasi performa setiap komponen sistem, mulai dari akurasi segmentasi hingga kemampuan klasifikasi dalam mengidentifikasi jenis daun herbal. Tujuan utama dari skenario uji coba ini adalah untuk mengukur keandalan sistem dalam kondisi nyata serta mengidentifikasi kelemahan yang perlu diperbaiki.

Tahap pertama adalah pengujian segmentasi menggunakan metode Watershed. Pengujian ini dilakukan dengan membandingkan hasil segmentasi dengan ground truth, yaitu gambar yang telah diberi label manual oleh ahli. Metrik yang digunakan meliputi

Intersection over Union (IoU) untuk mengukur sejauh mana hasil segmentasi sesuai dengan ground truth, serta waktu komputasi untuk menilai efisiensi algoritma. Pengujian ini dilakukan pada berbagai kondisi gambar, termasuk variasi pencahayaan dan latar belakang, untuk memastikan metode segmentasi tetap robust.

Selanjutnya, dilakukan pengujian klasifikasi menggunakan dataset yang telah dibagi menjadi set pelatihan dan pengujian. Performansi model Random Forest dievaluasi menggunakan metrik seperti akurasi, presisi, recall, dan F1-score. Selain itu, confusion matrix digunakan untuk mengidentifikasi jenis daun yang sering salah diklasifikasi. Analisis ini membantu menemukan pola kesalahan, misalnya jika model kesulitan membedakan daun dengan bentuk atau tekstur yang mirip.

Selain pengujian teknis, sistem juga diuji dalam skenario yang menyerupai kondisi nyata. Contohnya adalah uji coba dengan gambar daun yang diambil dari lingkungan berbeda, seperti daun yang ditempatkan di luar ruangan dengan pencahayaan alami atau gambar daun yang sebagian tertutup oleh objek lain. Skenario ini bertujuan untuk mengevaluasi kemampuan sistem dalam menangani tantangan dunia nyata yang mungkin tidak muncul pada dataset terstruktur.

Pengujian terakhir melibatkan pengguna akhir, seperti ahli botani atau praktisi herbal, yang diminta untuk menggunakan sistem dan memberikan umpan balik. Feedback ini mencakup aspek usability, seperti kemudahan antarmuka, serta keandalan sistem dalam memberikan hasil klasifikasi. Informasi ini digunakan untuk melakukan iterasi pada desain dan fungsionalitas sistem, sehingga dapat lebih sesuai dengan kebutuhan pengguna.

Hasil uji coba disajikan dalam bentuk laporan yang mencakup temuan, metrik performa, serta rekomendasi perbaikan. Laporan ini menjadi dasar untuk pengembangan lebih lanjut, baik dari segi teknis maupun penerapan praktis. Dengan skenario uji coba yang komprehensif, penelitian ini dapat menghasilkan sistem klasifikasi daun herbal yang andal dan aplikatif.