**IDENTIFIKASI Jenis Daun Tanaman Obat Berdasarkan Fitur Tekstur dan Bentuk Menggunakan Klasifikasi K-Nearest Neighbor**

Rian Setyawan1)

1, 2)Fakultas Informatika dan Universitas Gunadarma

Jl. Margonda Raya 100 Depok 16424

e-mail: [riansetyawan72@gmail.com](mailto:riansetyawan72@gmail.com)1)

**ABSTRAK**

*Tanaman obat bisa dijadikan sebagai alternatif pengobatan yang alami, selain obat-obatan kimia. Namun karena terlalu banyak jumlah dan jenis tumbuhan serta kurangnya pengetahuan, sehingga identifikasi juga akan sulit. Maka dibutuhkan sistem identifikasi secara otomatis menggunakan teknologi image processing supaya proses identifikasi menjadi lebih mudah, tepat, akurat dan efisien. Penelitian ini mengusulkan identifikasi tanaman obat berdasarkan fitur tekstur Grey Level Co-occurrence Matrix (GLCM)dan fitur bentuk metric dan eccentricity. Untuk proses identifikasi jenis daun dilakukan dengan K- Nearest Neighbors (KNN). Ada 3 jenis tanaman obat yang digunakan dalam penelitian ini. Dari hasil pengujian yang dilakukan pada aplikasi ini, hasil tertinggi diperoleh nilai k=5 dengan jumlah data pelatihan sebanyak 160 citra dan data uji coba sebanyak 30 citra yang mempunyai nilai rata-rata recall, precision, dan f-measure yaitu sebesar 97%.*

**Kata Kunci**: Tanaman Obat, Ekstraksi Fitur, Grey Level Co-occurrence Matrix, K- Nearest Neighbors.

**ABSTRACT**

Medicinal plants can be used as an alternative natural treatment, in addition to chemical drugs. But because there are too many numbers and types of plants and a lack of knowledge, identification will also be difficult. Then an automatic identification system is needed using image processing technology so that the identification process becomes easier, precise, accurate and efficient. This study proposes the identification of medicinal plants based on the features of the Gray Level Co-occurrence Matrix (GLCM) texture and metric and eccentricity features. For the process of identifying leaf types carried out by K-Nearest Neighbors (KNN). There are 3 types of medicinal plants used in this study. From the results of the tests conducted on this application, the highest results obtained by the value k = 5 with the number of training data as many as 160 images and trial data as many as 30 images that have an average value of recall, precision, and f-measure which is 97%..

**Keywords**: Medical Plants, Feature Extraction, Grey Level Co-occurrence Matrix, K- Nearest Neighbors.

# Pendahuluan

I

NDONESIA adalah negara megabiodiversity dimana Indonesia memiliki kekayaan tanaman obat yang sangat potensial untuk dikembangkan. Indonesia memiliki 75% kekayaan tanaman meliputi 30.000 jenis tanaman dari 40.000 jenis tanaman di dunia. Dari jumlah tersebut, 9.606 jenis tanaman diantaranya berkhasiat sebagai obat. Tanaman obat adalah salah satu jenis tanaman yang memiliki peran sangat penting dalam hidup manusia. Banyaknya jumlah dan jenis tanaman obat, membuat masyarakat menjadi kesulitan dalam hal membedakan jenis tanaman obat karena kurangnya pengetahuan, sehingga banyak masyarakat yang lebih memilih untuk menggunakan obat-obatan kimia. Untuk memberikan informasi kepada masyarakat, dibutuhkan sistem pengenalan tanaman obat yang mampu melakukan identifikasi dan pengenalan tanaman obat. Informasi yang didapat dapat berupa citra digital yang kemudian dianalisis dan diproses oleh sistem. Sistem dapat mengidentifikasi citra dari daun tanaman obat dan melakukan pengenalan suatu pola atau karakteristik dari objek tersebut [1].

Peneltian yang membahas tentang identifikasi pada suatu citra daun juga telah dilakukan. Reni dkk [2] pada tahun 2018 menggunakan metode *Artificial* *Neural* *Network* untuk mengidentifikasi jenis daun dari sejumlah jenis tanaman obat. Penelitian ini diterapkan algoritam *Back Propagation* dan menunjukan hasil yang baik yaitu persentase sebesar 93.6%. Pada penelitian selanjutnya yaitu Puji dkk [3] pada tahun 2018 menggunakan metode *Support* *Vector* *Machine* untuk mengklasifikasi penyakit daun kentang. Pada peneltian ini diterapkan *preprocessing* sebelum melakukan segmentasi dan klasifikasi, persentase kebenaran dari hasil peneltian yaitu sebesar 87%.

Pada penelitian kali ini diusulkan penggunakan metode *Gray Level Cooccurrences Matrix* (GLCM) pada tahap ekstraksi fitur. Metode GLCM pertama kali diusulkan oleh Haralick pada tahun 1973 yang ditujukan untuk ektraksi fitur tekstur pada suatu citra [4].

Pada penelitian ini terdapat lima bagian, yaitu pendahuluan pada bagian pertama, kemudian pada bagian kedua merupakan landasan teori, pada bagian ketiga merupakan penjelasan tentang metode yang diusulkan, yang keempat adalah pembahasan tentang hasil penelitian, dan bagian akhir merupakan kesimpulan dari penelitian yang dilakukan.

# Landasan Teori

## Tanaman obat

Tanaman obat adalah jenis-jenis tanaman yang memiliki fungsi dan berkhasiat sebagai obat dan dipergunakan untuk penyembuhan maupun mencegah berbagai penyakit. Tanaman obat mengandung zat aktif yang bisa mengobati penyakit tertentu atau jika tidak memiliki kandungan zat aktif tertentu tapi memiliki kandungan efek resultan atau sinergi dari berbagai zat yang mempunyai efek mengobati. Penggunaan tanaman obat sebagai obat bisa dengan cara diminum, ditempel, dan dihirup sehingga kegunaannya dapat memenuhi konsep kerja reseptor sel dalam menerima senyawa kimia atau rangsangan.

Salah satu jenis tanaman obat adalah.jambu biji. Jambu bji memiliki kandungan nutrisi pada daun sehingga dapat dimanfaatkan untuk mengobati diare, mengobati maag, menghilangkan kembung, mengatasi jerawat, mengobati sariawan, menurunkan kolesterol, dan meningkatkan metabolisme tubuh. Gambar 1 merupakan daun jambu biji. Terdapat juga daun kersen dan daun sirih yang merupakan jenis tanaman obat. Gambar 2 merupakan daun kersen dan gambar 3 merupakan gambar daun sirih.

Pada peneltian ini digunakan *dataset* yang terdiri dari data latih dan data uji yang berbentuk citra. Data latih didapat dari *google search image* sedangkan data uji didapat dari pengambilan gambar dengan menggunakan kamera *smartphone*.



Gambar 1. Daun Jambu Biji



Gambar 2. Daun Kersen

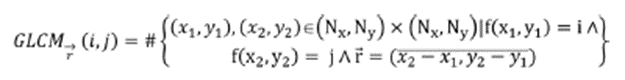


Gambar 3. Daun Sirih

## GLCM (Gray Level Co-occurrence Matrix)

*Gray Level Co-Occurence Matrix* (GLCM) merupakan metode untuk melakukan ektraksi fitur tekstur. Metode ini membedakan tekstur objek satu dengan objek lainnya dapat menggunakan ciri statistik orde dua. Ciri orde dua didasarkan pada probabilitas hubungan ketetanggaan antara dua piksel pada jarak dan orientasi sudut tertentu. Metode GLCM pertama kali diusulkan oleh Haralick pada tahun 1973.

GLCM→(i,j) didefinisikan sebagai 𝑟 jumlah piksel dengan j1, ..., L yang terjadi pada ofset vektor r terhadap piksel dengan nilai i1, ..., L, yang dapat dinyatakan dalam persamaan (1).

 (1)

Sebagai ilustrasi, ketetanggaan piksel dapat dipilih ke arah timur(kanan). Salah satu cara untuk merepresentasikan hubungan ini yaitu berupa (1,0), yang menyatakan hubungan dua piksel yang sejajar secara horizontal dengan piksel bernilai 1 diikuti dengan piksel bernilai 0, sehingga jumlah kelompok piksel yang memenuhi hubungan tersebut dihitung.



Gambar 4. Contoh penentuan awal matriks GLCM

Matriks pada Gambar 2.7 dinamakan matrix framework. Matriks ini kemudian diolah menjadi matriks simetris dengan cara menambahkan dengan hasil tranposnya, seperti yang terlihat pada Gambar 2.8.



Gambar 5. Matriks framework menjadi matriks simetris

Untuk menghilangkan ketergantungan pada ukuran citra, nilai-nilai elemen GLCM perlu dinormalisasi sehingga jumlahnya bernilai 1. Dengan demikian, hasil normalisasi dari matriks GLCM pada Gambar 2.7.



Gambar 6. Normalisasi matriks GLCM

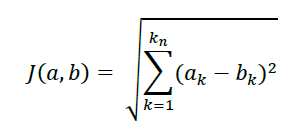
## KNN (K-Nearest Neighbor)

KNN merupakan salah satu algoritma pembelajaran mesin sederhana. Hal ini hanya didasarkan pada gagasan bahwa suatu objek yang dekat satu sama lain juga akan memiliki karakteristik yang mirip. Ini berarti jika kita mengetahui ciri-ciri dari salah satu objek, maka kita juga dapat memprediksi objek lain berdasarkan tetangga terdekatnya. KNN adalah improvisasi lanjutan dari teknik klasifikasi Nearest Neighbor. Hal ini didasarkan pada gagasan bahwa setiap contoh baru dapat diklasifikasikan oleh suara mayoritas dari k tetangga, di mana k adalah bilangan bulat positif, dan biasanya dengan jumlah kecil. Algoritma klasifikasi KNN memprediksi kategori tes sampel sesuai dengan sampel pelatihan k yang merupakan tetangga terdekat dengan sampel uji dan memasukkan ke dalam kategori yang memiliki kategori probabilitas terbesar.

Dalam pengenalan pola, algoritma KNN adalah metode yang digunakan untuk mengklasifikasikan objek berdasarkan contoh pelatihan terdekat di ruang fitur. KNN adalah jenis *insctance*-*based* *learning*, atau *lazy* *learning* dimana fungsi ini hanya didekati secara lokal dan semua perhitungan ditangguhkan sampai klasifikasi.

Metode klasifikasi KNN memiliki beberapa tahap, yang pertama nilai k yang merupakan jumlah tetangga terdekat yang akan menentukan query baru masuk ke kelas mana ditentukan. Tahap kedua, k tetangga terdekat dicari dengan cara menghitung jarak titik *query* dengan titik *training*. Tahap ketiga, setelah mengetahui jarak masing-masing titik *training* dengan titik *query*, kemudian lihat nilai yang paling kecil. Tahap keempat ambil k nilai terkecil selanjutnya lihat kelasnya. Kelas yang paling banyak merupakan kelas dari queri baru.

Dekat atau jauhnya jarak titik dengan tetangganya bisa dihitung dengan menggunakan *Euclidean* *Distance*. *Euclidean* *distance* direpresentasikan dengan menggunakan persamaan (2).

 (2)

dimana J(a,b) adalah jarak Euclidean antara vector a dan vektor b, ak dan bk adalah komponen ke k dari vektor k dan kn adalah jumlah komponen pada vektor a dan vektor b.

# Metode Penelitian

Data yang digunakan pada penelitian ini adalah citra daun tanaman obat. Citra yang dikumpulkan terdiri dari 3 jenis daun yaitu daun jambu biji, daun kersen, dan daun sirih. Citra yang telah dikumpulkan dibagi menjadi dua bagian yaitu data latih dan data uji. Data latih yang digunakan sebanyak 160 citra sedangkan data uji terdiri dari 30 citra dan terbagi menjadi 10 citra tiap kategori.

Pada penelitian ini diusulkan penggunaan metode *Gray Level Co-Occurence Matrix* (GLCM) untuk ekstraksi fitur tekstur dan penggunaan parameter *metric* dan *eccentricity* untuk ekstraksi fitur bentuk pada citra tanaman obat. Tahap awal dari penelitian ini adalah akuisisi citra untuk digunakan sebagai citra pengujian dan pelatihan. Selanjutnya memasuki tahapan *preprocessing* yaitu mengubah resolusi citra (*resizing*), tranformasi warna citra, segmentasi citra untuk memperoleh nilai bentuk citra, mengubah citra kedalam citra keabuan dan melakukan proses ektraksi ciri (*feature* *extraction*) terhadap citra menggunakan metode *Gray Level Co-Occurrence Matrix* (GLCM) untuk memperoleh nilai tekstur citra, dan tahap identifikasi menggunakan metode *K-Nearest Neighbor* (KNN). Setelah seluruh tahap dilakukan maka aplikasi menghasilkan keluaran berupa informasi jenis tanaman obat. Adapun tahapan tersebut dapat dilihat dalam bentuk arsitektur umum pada gambar 7.



Gambar 7. Arsitektur Umum

Pada tahap akuisisi citra dilakukan pemilihan gambar untuk data latih dan data uji. Gambar untuk data latih yang telah diperoleh selanjutnya dirapihkan terlebih dahulu dengan menggunakan *Photoshop*. Input pada tahap ini merupakan data uji yang diambil dengan mengambil gambar pada daun yang diletakan pada kertas kosong. Karena ukuran gambar berbeda-beda, maka dilakukan proses *precosessing* dengan merubah citra ukuran menjadi ukuran 200x250 pixel.

Setelah didapatkan hasil *preprocessing*, dilakukan transformasi warna pada citra dari RGB menjadi CIELab. Proses transformasi warna dilakukan dengan tujuan membuat hasil citra segmentasi sempurna. Selanjutnya dilakukan proses segmentasi untuk mendapatkan citra biner (hitam dan putih). Proses ini dilakukan untuk mendapatkan fitur bentuk yaitu nilai *metric* dan *eccentricity* dari citra daun yang akan diklasifikasi. Tahap selanjutnya adalah ekstraksi fitur tekstur dengan menggunakan metode GLCM. Pada tahap ini citra dilakukan dengan terlebih dahulu mengubah citra dari RGB menjadi citra *grayscale*. Kemudian dari citra *grayscale* dilakukan proses ekstraksi fitur yaitu dengan mengambil nilai *contrast*, *correlation*, *energy* dan *homogeneity* pada citra daun.

Bagian terakhir pada peneltian ini yaitu klasifikasi. Pada penelitian ini digunakan klasifikasi KNN (*K-Nearest Neighbor*). Fitur hasil proses ekstraksi citra digunakan sebagai proses klasifikasi. KNN bekerja dengan cara mengklasifikasikan suatu objek yang memiliki kemiripan paling dekat dengan objek lainnya. KNN memiliki atribut yang diinisalisasikan sebagai k, yaitu jumlah nilai tetangga yang dijadikan acuan pada klasifikasi KNN. Pada penelitian ini nilai k yang digunakan yaitu 1,3 dan 5.

# Hasil dan Analisa

Pada pengujian ini ditujukan untuk mengetahui performa dari metode ekstraksi fitur tektur dan bentuk untuk pengenalan jenis daun tanaman obat dengan menggunakan metode klasifikasi KNN. Data training yang digunakan memiliki jumlah yang berbeda-beda. Terdapat kelas yang memiliki data training sebanyak 60 citra, seperti daun jambu biji dan daun sirih. Namun, terdapat pula yang hanya memiliki data training sebanyak 40 citra yaitu daun kersen. Sedangkan untuk data uji, terdiri dari 30 citra dan terbagi menjadi 10 citra tiap kelas.

Pada penelitian ini pengujian dilakukan dengan nilai k yang berbeda yaitu 1, 3 dan 5 pada masing-masing citra. Perhitungan yang digunakan untuk mengetahui hasil dari performa yaitu *recall*, *precision* dan *f-measure*.

1. Percobaan dilakukan dengan nilai k=1, hasil uji identifikasi mempunyai nilai yang berbeda-beda dapat dilihat pada Tabel 1.

TABEL I

HASIL RECALL, PRECISSION DAN F-MEASURE DENGAN NILAI K = 1

|  |  |  |  |
| --- | --- | --- | --- |
| Jenis Daun | Jambu Biji | Kersen | Sirih |
| Jambu Biji | **9** | 1 | 0 |
| Kersen | 2 | **8** | 0 |
| Sirih | 1 | 0 | **9** |
| *Recall* | 90% | 80% | 90% |
| *Precision* | 75% | 89% | 100% |
| *F-Measure* | 82% | 84% | 95% |

Dari hasil uji coba pada tabel 1, nilai *recall* yang tertinggi adalah daun jambu biji dan sirih yaitu sebesar 90%, sedangkan nilai *recall* terendah adalah daun kersen yaitu sebesar 80%. Untuk nilai *precision* yang tertinggi adalah daun sirih yaitu sebesar 100%, sedangkan nilai *precision* yang terendah adalah daun jambu biji yaitu sebesar 75%. Untuk nilai *f-measure* tertinggi adalah daun sirih yaitu sebesar 95%, sedangkan nilai *f-measure* terendah adalah daun jambu biji yaitu sebesar 82%. Hasil akhir yang diperoleh dari performa klasifikasi jenis daun jambu biji, kersen, dan sirih mempunyai nilai rata-rata *recall* 87%, *precision* 88% dan *f-measure* 87%.

1. Percobaan dilakukan dengan nilai k=3, dimana nilai *recall*, *precision* dan *f-measure* mengalami penurunan pada jenis daun kersen, seperti yang terlihat pada Tabel 2.

TABEL II

HASIL RECALL, PRECISSION DAN F-MEASURE DENGAN NILAI K = 3

|  |  |  |  |
| --- | --- | --- | --- |
| Jenis Daun | Jambu Biji | Kersen | Sirih |
| Jambu Biji | **9** | 1 | 0 |
| Kersen | 3 | **7** | 0 |
| Sirih | 0 | 1 | **9** |
| *Recall* | 90% | 70% | 90% |
| *Precision* | 75% | 78% | 100% |
| *F-Measure* | 82% | 74% | 95% |

Pada hasil penelitian yang terlihat pada tabel 2, nilai *recall* yang tertinggi adalah daun jambu biji dan sirih yaitu sebesar 90%, sedangkan nilai *recall* terendah adalah daun kersen yaitu sebesar 70%. Untuk nilai *precision* yang tertinggi adalah daun sirih yaitu sebesar 100%, sedangkan nilai *precision* yang terendah adalah daun jambu biji yaitu sebesar 75%. Untuk nilai *f-measure* tertinggi adalah daun sirih yaitu sebesar 95%, sedangkan nilai *f-measure* terendah adalah daun kersen yaitu sebesar 74%. Hasil akhir yang diperoleh dari performa klasifikasi jenis daun jambu biji, kersen, dan sirih mempunyai nilai rata-rata *recall* 83%, *precision* 84%, dan *f-measure* 84%. .

1. Percobaan dilakukan dengan nilai k=5, dimana nilai *recall*, *precision* dan *f-measure* mengalami kenaikan pada semua kategori, seperti yang terlihat pada Tabel 3.

TABEL III

HASIL RECALL, PRECISSION, DAN F-MEASURE DENGAN NILAI K = 5

|  |  |  |  |
| --- | --- | --- | --- |
| Jenis Daun | Jambu Biji | Kersen | Sirih |
| Jambu Biji | **10** | 0 | 0 |
| Kersen | 1 | **9** | 0 |
| Sirih | 0 | 0 | **10** |
| *Recall* | 100% | 90% | 100% |
| *Precision* | 91% | 100% | 100% |
| *F-Measure* | 95% | 95% | 100% |

Pada hasil penelitian yang terlihat pada tabel 3, nilai *recall* yang tertinggi adalah daun jambu biji dan sirih yaitu sebesar 100%, sedangkan nilai *recall* terendah adalah daun kersen yaitu sebesar 90%. Untuk nilai *precision* yang tertinggi adalah daun kersen dan sirih yaitu sebesar 100%, sedangkan nilai *precision* yang terendah adalah daun jambu biji yaitu sebesar 91%. Untuk nilai *f-measure* tertinggi adalah daun sirih yaitu sebesar 100%, sedangkan nilai *f-measure* terendah adalah daun jambu biji dan kersen yaitu sebesar 95%. Hasil akhir yang diperoleh dari performa klasifikasi jenis daun jambu biji, kersen, dan sirih mempunyai nilai rata-rata *recall*, *precision*, dan *f-measure* yaitu sebesar 97%. Sehingga dapat dikatakan bahwa proses identifikasi jenis daun yang dilakukan dengan nilai k=5 sangat efektif. .

# Kesimpulan

Berdasarkan penelitian dapat disimpulkan bahwa hasil evaluasi performa identifikasi pada jenis daun tanaman obat mendapatkan hasil yang berbeda pada setiap pengujian yang telah dilakukan. Pada penelitian ini hasil evaluasi tertinggi dicapai pada saat pengujian dengan nilai k=5, yaitu mempunyai nilai *recall*, *precision*, dan *f-measure* sebesar 97%. Sedengkan performa hasil evaluasi terendah dicapai pada saat pengujian dengan nilai k=1, yaitu mempunyai nilai *recall* 87%, *precision* 88%, dan *f-measure* 87%.

Untuk penelitian selanjutnya, dapat dikembangkan dengan menggunakan metode jenis klasifikasi lain serta mencoba mengembangkan ekstraksi ciri fitur tekstur atau bentuk yang lain, seperti menambahkan fitur warna dan juga menambahkan macam-macam jenis daun tanaman yang lain.

Daftar Pustaka

1. F. S. Ni'mah, T. Sutojo, and D. R. I. M. Setiadi, "Identifikasi Tumbuhan Obat Herbal Berdasarkan Citra Daun Menggunakan Algoritma Gray Level Co-occurence Matrix dan K-Nearest Neighbor", *Jurnal Teknologi dan Sistem Komputer*, vol. 6, no. 2, Apr. 2018.
2. R. Rahmadewi, V. Efelina, and E. Purwanti, “Identifikasi Jenis Tumbuhan Menggunakan Citra Daun Berbasis Jaringan Saraf Tiruan (Artificial Neural Networks), *Jurnal Media Elektro*, Karawang, 2018.
3. P.U. Rakhmawati, Y.M. Pranoto, and E. Setyati, “Klasifikasi Penyakit Daun Kentang Berdasarkan Fitur Tekstur Dan Fitur Warna Menggunakan Support Vector Machine”, *Seminar Nasonal Teknologi dan Rekayasa*, Surabaya, 2018.
4. Z.E.Fitri, “ Klasifikasi Trombosit Pada Citra Hapusan Darah Tepi Berdasarkan Gray Level Co-occurance Matrix Menggunakan Backpropagation”, tesis magister, Dept. Teknik Elektro, Institut Teknologi Sepuluh November, Surabaya, Indonesia, 2017.
5. H.Z. Zahro, "Analisis Tekstur Untuk Identifikasi Tumbuhan Obat Menggunakan Klasifikasi Support Vector Machine”, *Industri Inovatif*, vol. 6, no. 2, Sept. 2016.