

Gaussian Mixture Model Dalam Proses Pengenalan Daun Untuk Mengidentifikasi Tanaman Herbal

Gaussian Mixture Model In Leaf Recognition Process To Identify Herbal Plants

1st Dewi Intan Savrylia
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

dewiintansav@student.telkomuniversit
y.ac.id

2nd Anggunmeka Luhur Prasasti
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

anggunmeka@telkomuniversity.ac.id

3rd Marisa W. Paryasto
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia

marisaparyasto@telkomuniversity.ac.id

Abstrak—Tanaman herbal merupakan salah satu obat alternatif yang digunakan dalam bidang Kesehatan untuk mengobati penyakit. Tanaman herbal dapat ditemukan di sekitar rumah atau di lingkungan terdekat. Dengan keterbatasan pengetahuan manusia mengenai tanaman herbal, seiring kemajuan teknologi maka penelitian ini dibuat untuk mendeteksi daun tanaman herbal menggunakan Image Processing yang memanfaatkan teknologi yang semakin maju. Untuk pengenalan daun tanaman herbal digunakan segmentasi *Gaussian Mixture Model* sebagai clustering dan memanfaatkan image processing, dibutuhkan dataset daun tanaman herbal agar dapat mengenali daun tanaman herbal yang ingin di deteksi sebagai data latih. Dalam penelitian ini menghasilkan rata rata tingkat kecocokan sebesar 79.2% dengan waktu pemrosesan 4.44 detik. Selanjutnya hasilnya akan ditampilkan pada website sebagai hasil outputnya.

Kata kunci— *website, gaussian mixture model, clustering, tanaman herbal.*

Abstract—Herbal plants are one of the alternative medicines used in the health sector to treat diseases. Herbal plants can be found around the house or in the immediate environment. With limited human knowledge about herbal plants, as technology advances, this research was made to detect herbal plant leaves using Image Processing which utilizes increasingly advanced technology. For the introduction of herbal plant leaves, Gaussian Mixture Model segmentation is used as

clustering and utilizes image processing, it takes a dataset of herbal plant leaves to recognize the leaves of herbal plants that want to detect as training data. In this study, the average match rate was 79.2% with a processing time of 4.44 seconds. Furthermore, the results will be displayed on the website as the output.

Keywords— *website, gaussian mixture model, clustering, herbal plants.*

I. PENDAHULUAN

Tanaman adalah salah satu makhluk hidup yang sangat penting untuk kelangsungan kehidupan selain manusia dan hewan. Tanaman sangat bermanfaat bagi kehidupan makhluk hidup, begitupun dengan manusia dan hewan yang ketiganya saling bergantung satu sama lain [1]. Dengan adanya pandemi *Corona Virus Disease* atau *Covid-19* ini tanaman herbal dapat menjadi salah satu obat alternatif yang digunakan untuk meningkatkan imunitas tubuh. Banyaknya tanaman herbal dan minimnya pengetahuan mengenai tanaman herbal yang sangat berguna untuk pengobatan. Dengan kemajuan teknologi, gambar dapat di deteksi sesuai dengan kebutuhannya, pada penelitian ini memanfaatkan *machine learning*, selain itu algoritma yang digunakan adalah *Gaussian Mixture Model* dalam proses *clustering*, *Gray Level Co-Occurrence* sebagai ekstraksi fitur yang sangat penting dalam klasifikasi tanaman herbal karena proses ini bertujuan untuk mengubah warna

agar warna citra hanya pada tingkat keabuan saja. Selain itu proses ini juga bertujuan untuk menyederhanakan model citra dan mempermudah proses perhitungan [2], dan *website* sebagai hasil *output* beserta tingkat kecocokannya.

II. KAJIAN TEORI

A. Tanaman Herbal

Tanaman Herbal adalah tanaman yang mempunyai khasiat untuk pengobatan penyakit tertentu. Tanaman herbal sudah menjadi pengobatan tradisional sejak zaman dahulu sebelum mengenal bahan kimia yang dapat

gunakan sebagai obat [3]. Tanaman herbal juga dapat dimanfaatkan sebagai obat alternatif yang dapat digunakan untuk pengobatan makhluk hidup. Bagian tanaman herbal yang dapat dimanfaatkan untuk pengobatan penyakit salah satunya adalah bagian daun. Dengan mengenali tanaman herbal salah satunya bisa dilihat dari pola tulang daun yang bisa dilakukan selain menggunakan *Gaussian Mixture Model* dapat menggunakan segmentasi dengan *curvature points* [4]. Beberapa contoh tanaman herbal yang dimanfaatkan pada daunnya yaitu:



Gambar 1 Daun Teh
Gambar 3 Daun Sirih



Gambar 2 Daun Mengkudu



Gambar 4 Daun Kersen



Gambar 5 Betadine

B. Gaussian Mixture Model

Gaussian Mixture Model (GMM) adalah salah satu metode *clustering* untuk melakukan segmentasi. *Gaussian Mixture Model* merupakan salah satu model yang baik dan banyak digunakan untuk *clustering* [4]. GMM terdiri dari beberapa komponen fungsi fungsi *Gaussian*, komponen fungsi ini terdiri dari *Threshold* yang berbeda dikarenakan menghasilkan *multi-model density*. *Gaussian Mixture Model* termasuk *Unsupervised Learning*.

Penggunaan *Gaussian Mixture Model* dalam *Object Segmentation on Gaussian Mixture Model and Conditional Random Fields* menunjukkan bahwa GMM kompetitif untuk segmentasi objek[5]. *Gaussian Mixture Model* juga merupakan model statistik dari distribusi probabilitas yang didapatkan dari nilai bobot setiap distribusi Gaussian. Sehingga, *Gaussian Mixture Model* merupakan metode yang sangat tepat untuk perhitungan baik dengan parameter ataupun tidak.

$$f(x) = \frac{\exp\left(\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu)\right)}{2\pi\sqrt{|\Sigma|}} \quad (1)$$

Keterangan:

μ = nilai rata rata

Σ = Co variance

x = data point

C. Image Processing

Image Processing atau Pengolahan Citra adalah sistem dimana proses dilakukan dengan masukan (*input*) berupa gambar yang menghasilkan (*output*) berupa gambar yang lebih baik [6]. *Image Processing* ini dilakukan untuk memperbaiki kualitas gambar, dengan berkembangnya teknologi maka *Image Processing* dapat digunakan untuk memperoleh informasi yang dapat digunakan untuk suatu keperluan tertentu yang menggunakan gambar atau citra. Dengan adanya *image processing* dapat digunakan dalam hal penelitian. Pada *Image Processing* ada yang dinamakan dengan *Image Enhancement*, dimana *Image Enhancement* merupakan salah satu proses awal dalam *image processing* yang berfungsi untuk meningkatkan kualitas citra yang ada. beberapa proses yang termasuk dalam bagian *Image Enhancement* meliputi perubahan kecerahan citra, peningkatan kontras, peregangkan kontras, konversi histogram citra, pelunakan citra, penajaman, deteksi tepi, pemerataan histogram dan perubahan geometri [7]. *Image Enhancement* ini dapat diterapkan di banyak bidang, seperti implementasi di bidang militer, bidang medis, bidang hukum, bidang industri, bidang hiburan dan masih banyak lagi. Penggunaan utama *Image Enhancement* di setiap

bidang adalah untuk mendapatkan informasi yang jelas[8].

D. *Gray Level Co-Occurrence*

Gray Level Co-occurrence Matrix (GLCM) adalah metode analisis tekstur atau ekstraksi ciri. *Gray Level Co-occurrence Matrix* salah satu teknik tertua untuk analisis tekstur [9]. *Gray Level Co-occurrence Matrix* memiliki dua parameter yang sangat penting yaitu jarak dan arah. Jarak direpresentasikan sebagai piksel dan orientasi suatu sudut direpresentasikan dalam derajat. Ada empat arah sudut interval yang umum digunakan dalam ekstraksi fitur GLCM yaitu, 0°;45°;90° dan 135°, jarak antar piksel ditentukan sebesar 1 piksel [10]. Pada ekstraksi fitur GLCM terdapat enam ciri, yaitu *Contrast*, *Correlation*, *Dissimilarity*, *Homogeneity*, *ASM* dan *Energy*. Pada penelitian ini digunakan empat ekstraksi ciri yaitu:

1. *Contrast*: Fitur yang digunakan untuk menghitung tingkat perbedaan abu-abu dalam suatu gambar, semakin besar perbedaannya semakin tinggi kontrasnya dan sebaliknya. *Contrast* didefinisikan sebagai berikut:

$$\sum_i \sum_j (i - j)^2 p(i - j)$$

(2)

Keterangan:

i=baris pada matriks

j= kolom pada matriks

(i, j) =matriks GLCM

p= probabilitas kemunculan

2. *Correlation*: Fitur ini digunakan untuk menampilkan bagaimana korelasi referensi piksel dengan

tetangganya. *Correlation* didefinisikan sebagai berikut

(3)

$$\sum_i \sum_j \frac{(i - \mu_i)(i - \mu_j)p(i - j)}{\sigma_i \sigma_j}$$

3. *Homogeneity*: Fitur homogenitas digunakan untuk menghitung homogenitas abu abu suatu gambar.

Homogenitas didefinisikan sebagai berikut:

(4)

$$\sum_i \sum_j \frac{p(i,j)}{1 + (1 - j)^2}$$

4. *Dissimilarity*: Fitur ini digunakan ketidakmiripan pada suatu tekstur pada suatu gambar atau kelompok.

Dissimilarity didefinisikan sebagai berikut:

(5)

$$\sum_i \sum_j |i - j|p(i,j)$$

E. Support Vector Machine

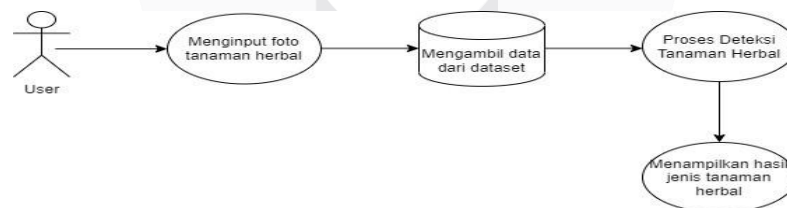
Support Vector Machine (SVM) merupakan salah satu metode algoritma *Machine Learning* yang digunakan untuk melakukan prediksi, dalam hal klasifikasi ataupun regresi. *Support Vector Machine* memiliki prinsip dasar yaitu *classifier*. *Classifier* sendiri adalah klasifikasi yang secara linier dapat dipisahkan. Selain menerapkan prinsip *classifier*. Salah satu keunggulan SVM adalah mempunyai kemampuan generalisasi yang tinggi, sehingga proses komputasinya memiliki tingkat akurasi yang tinggi [11]. *Support Vector Machine* banyak digunakan dalam hal klasifikasi. *Support Vector Machine* termasuk kategori *supervised learning*. Pada penelitian ini algoritma *Support Vector Machine* digunakan sebagai algoritma pendukung untuk mengoptimalkan hasil akurasi

oleh pemakai atau user yang terhubung ke internet. Website dapat berisikan teks, gambar, suara, animasi, video yang didalamnya terdapat *Hypertext Transfer Protocol* (HTTP) untuk mengakses website dibutuhkan perangkat lunak dan internet yang terhubung, agar dapat memuat isi dari website atau web. Pada penelitian ini *website* digunakan sebagai tampilan dalam menampilkan hasil pengenalan tanaman herbal, untuk dapat memperoleh hasil deteksi digunakan sebuah *website* yang dapat diakses pada lokal di laptop. Dengan menggunakan website, pengguna dapat melihat nama daun, nama latin, daerah persebaran dan tingkat kecocokan.

F. Website

Website atau Web adalah layanan yang dapat digunakan atau didapatkan

A. Desain Sistem



GAMBAR 6
GAMBARAN UMUM SISTEM

Gambaran umum tersebut menjelaskan alur proses untuk mendapatkan hasil deteksi tanaman herbal, dengan dataset yang ada dan algoritma yang digunakan adalah *Gaussian Mixture Model* sebagai *clustering* dan menggunakan *Gray Level Co-occurrence* untuk mengekstraksi fiturnya selanjutnya gambar akan diubah menjadi warna abu abu untuk di proses

dan masuk ke dalam algoritma *Gaussian Mixture Model*.

B. Diagram Alir *Training Data*

Berdasarkan perancangan sistem secara umum, sistem digambarkan dalam bentuk diagram alir pada gambar berikut.

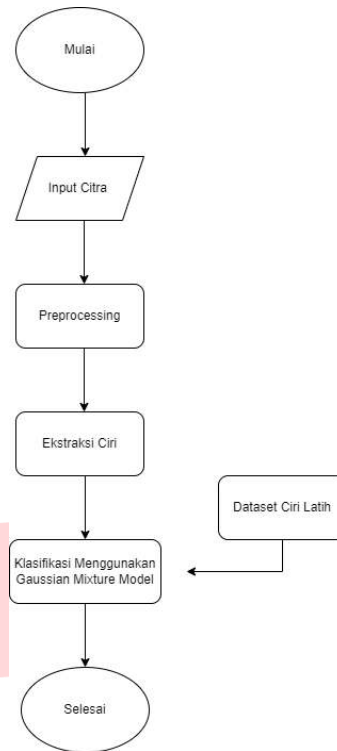


GAMBAR 7
DIAGRAM ALIR SISTEM *TRAINING*

Pada Gambar 6 menunjukkan bahwa merupakan proses dalam training dalam proses deteksi daun herbal. Proses ini dilakukan dengan menginput citra yang diambil selanjutnya akan dilakukan *processing* dimana citra akan dilakukan perubahan warna citra dari citra RGB ke *grayscale*. Setelah itu dilakukan proses ekstraksi ciri menggunakan *Gray Level Co-Occurrence*, setelah melalui proses ekstraksi maka hasil ekstraksi tersebut menghasilkan angka – angka yang dimana akan di simpan ke dataset berupa file csv dimana data latih yang digunakan sebanyak 254. Kemudian dataset tersebut akan digunakan untuk proses *testing* dengan menggunakan data uji sebanyak 51.

C. Diagram Alir *Website*

Selain proses training, dilakukan proses *testing* pada *website*. Berikut merupakan diagram alir sistem pada proses *testing website*.



GAMBAR 8
DIAGRAM *TESTING WEBSITE*

Proses ini dilakukan dengan menginput citra yang diambil melalui camera selanjutnya akan dilakukan *processing* dimana citra akan dilakukan perubahan warna citra dari citra RGB ke *grayscale*. Setelah itu dilakukan proses ekstraksi ciri menggunakan *Gray Level Co-Occurrence* Selanjutnya citra akan di proses dengan dataset yang ada untuk mendapatkan hasil yang sesuai dengan dataset yang ada, pengujian ini menggunakan 51 sample data selanjutnya citra akan di proses menggunakan algoritma *Gaussian Mixture Model* dimana *Gaussian Mixture Model (GMM)* merupakan salah satu metode *clustering* untuk melakukan segmentasi.

IV. HASIL DAN PEMBAHASAN

A. Pengujian Fitur Ekstraksi Ciri

Pada pengujian ini dilakukan percobaan pada empat fitur yang digunakan dalam proses deteksi daun, yaitu *Homogeneity*, *Contrast*, *Correlation*, *Dissimilarity*. Berikut merupakan pengujian menggunakan empat fitur pada *Gray Level Co-Occurrence*.

TABEL 1
HASIL PENGUJIAN EKSTRAKSI CIRI

Kombinasi 4 parameter	Akurasi
Dissimilarity, Correlation, Homogeneity, Contrast	88%

Dengan melakukan pengujian empat fitur tekstur pada *Gray Level Co-Occurrence* yaitu: *Homogeneity*, *Contrast*, *Correlation* dan *Dissimilarity* mendapatkan hasil keakuratan sebesar 88%.

Pada pengujian ini dilakukan perubahan parameter distance pada *Gray Level Co-Occurrence (GLCM)* terhadap sudut yang memberikan nilai akurasi paling tinggi. Berikut adalah hasil pengujian dengan menggunakan empat fitur *contrast*, *correlation*, *homogeneity* dan *dissimilarity* pada GLCM.

B. Pengujian Distance Terhadap Sudut

TABEL 2
HASIL PENGUJIAN GLCM

Sudut Orientasi	Jarak				
	1	2	3	4	5
0°	88%	90%	86%	90%	90%
45°	88%	90%	86%	90%	90%
90°	88%	90%	86%	90%	90%
135°	88%	90%	86%	90%	90%

Pengujian jarak pada *Gray Level Co-occurrence Matrix (GLCM)* menghasilkan tingkat keakuratan yang berbeda. Pada sudut 0° akurasi tertinggi berada pada jarak 2,4, dan 5 piksel yaitu dengan tingkat keakuratan sebesar 90%, pada sudut 45°,90°,135° tingkat keakuratan tertinggi pada jarak 2,4 dan 5 piksel sebesar 90%.

C. Pengujian Parameter SVM Terhadap *Gaussian Mixture Model*

Dilakukan pengujian dua parameter kernel pada *Support Vector Machine (SVM)* terhadap *Gaussian Mixture Model*, pada parameter SVM yang akan diuji yaitu kernel rbf dan polynomial (poly). Berikut merupakan hasil dari pengujian yang dilakukan:

TABEL 3
PENGUJIAN KERNEL RBF

Parameter SVM			Parameter Gaussian			Akurasi
C	Gamma	Kernel	n_components	n_int	max_iter	
0.1	1	rbf	5	10000	100	33%
1	1	rbf	5	10000	100	24%
10	1	rbf	5	10000	100	24%
100	1	rbf	5	10000	100	24%
0.1	0.1	rbf	5	10000	100	24%
0.1	0.01	rbf	5	10000	100	24%
0.1	0.001	rbf	5	10000	100	33%
0.1	0.0001	rbf	5	10000	100	37%

TABEL 4
PENGUJIAN KERNEL POLY

Parameter SVM			Parameter Gaussian			Akurasi
C	Gamma	Kernel	n_components	n_int	max_iter	
0.1	1	Poly	5	10000	100	76%
1	1	Poly	5	10000	100	76%
10	1	Poly	5	10000	100	76%
100	1	Poly	5	10000	100	76%
0.1	0.1	Poly	5	10000	100	76%
0.1	0.01	Poly	5	10000	100	76%
0.1	0.001	Poly	5	10000	100	76%
0.1	0.0001	Poly	5	10000	100	76%
100	0.0001	Poly	5	10000	100	90%

Pada percobaan diatas, pada parameter kernel poly dengan C = 100 dan Gamma= 0.0001 mendapatkan hasil keakuratan sebesar 90% dimana, dengan angka keakuratan sebesar 90% merupakan hasil

tertinggi dibandingkan dengan kernel Poly dan rbf lainnya.

D. Pengujian Hasil Komputasi

TABEL 5
HASIL KOMPUTASI

Daun	Tingkat kecocokan (%)	Waktu pemrosesan(detik)
Betadine	97.65%	4.29
Kersen	77.79%	4.16
Mengkudu	78.01%	4.22
Sirih	80.41%	4.23
Teh	62.14%	5.31
Rata Rata	79.2%	4.44

Pada pengujian yang dilakukan untuk mendapatkan hasil pemrosesan data melalui website dengan melakukan percobaan sepuluh gambar daun per jenisnya, mendapatkan hasil rata rata tingkat kecocokan sebesar 79.2% dengan rata rata waktu pemrosesan sebesar 4.44 detik.

dapat disimpulkan Gaussian Mixture Model mempunyai tingkat kecocokan yang tinggi sebesar 79.20%. Dalam prosesnya *Gaussian Mixture Model* membutuhkan waktu 4.44 detik.

V. KESIMPULAN

Berdasarkan hasil penelitian segmentasi *Gaussian Mixture Model* dalam mengidentifikasi Tanaman Herbal maka

REFERENSI

- [1] E. Iqbal, Eliyatiningih, Nurahmanto, Dwi, Sari and V. Kartika, "Pembibitan Tanaman

- Herbal di Desa Pace Kecamatan Silo Kabupaten Jember Guna Menuju Desa Sentra H," *Jurnal Pengabdian Masyarakat Berkemajuan*, pp. 264-267, 2020
- [2] Prasasti, A. L., Irawan, B., Fajri, S. E., Rendika, A., & Hadiyoso, S. (2020). Perbandingan Ekstraksi Fitur Dan Proses Matching Pada Autentikasi Sidik Jari Manusia. *Elkomika: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 8(1), 95.
- [3] Saepudin, Encang, Rusmana, Agus, Budiono and Agung, "Penciptaan Pengetahuan Tentang Tanaman Obat Herbal dan Tanaman Obat Keluarga," *Jurnal Kajian Informasi & Perpustakaan*, vol. 4, no. 1, p. 95, 2016
- [4] Prasasti, A. L., Mengko, R. K. W., & Adiprawita, W. (2015). Vein Tracking Using 880nm Near Infrared And CMOS Sensor With Maximum Curvature Points Segmentation. In *7th WACBE World Congress On Bioengineering 2015* (Pp. 206-209). Springer, Cham.
- [5] R. A. Surya, A. Fadlil and A. Yudhana, "Ekstraksi Ciri Metode *Gray Level Co-Occurrence Matrix* (GLCM) dan Filter Gabor Untuk Klasifikasi Citra Batik Pekalongan," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 2, no. 2, 2017
- [6] K. Karnadi, "Pengembangan Aplikasi Digital Image Processing dengan Microsoft Visual Basic," *Jurnal Digital Teknologi Informasi*, vol. 1, no. 1, p. 15, 2018
- [7] Putra, R. D., Purboyo, T. W., & Prasasti, L. A. (2017). A Review of Image Enhancement Methods. *International Journal of Applied Engineering Research*, 12(23), 13596-13603.
- [8] Aziz, M. N., Purboyo, T. W., & Prasasti, A. L. (2017). A survey on the implementation of image enhancement. *Int. J. Appl. Eng. Res*, 12(21), 11451-11459
- [9] R. Achmad, S. Gulo and A. Y. Gultom, "Analisis *Gray Level Co- Occurrence Matrix* (Glcm) Dalam Mengenali Citra Ekspresi Wajah," *Jurnal Mantik*, vol. 3, no. 2, pp. 31-38, 2019.
- [10] Parapat, I. Monika, Furqon, M. Tanzil and Sutrisno, "Penerapan Metode Support Vector Machine (SVM) pada Klasifikasi Penyimpangan Tumbuh Kembang Anak," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, p. 3163, 2018.
- [11] R. A. Rizal, I. S. Girsang and S. A. Prasetyo, "Klasifikasi Wajah Menggunakan Support Vector Machine (SVM)," *Riset dan E-Jurnal Manajemen Informatika Komputer*, vol. 3, 2019.