

IMPLEMENTASI *GRAY LEVEL CO-OCCURRENCE MATRIX* (GLCM) UNTUK KLASIFIKASI PENYAKIT DAUN PADI

Arif Akbarul Huda¹, Bayu Setiaji², Fajar Rosyid Hidayat³

^{1,2,3} Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta
Jl. Ringroad Utara, Condong Catur, Depok, Sleman, 55283

¹arif.akbarul@amikom.ac.id

²bayusetiaji@amikom.ac.id

³fajar22@students.amikom.ac.id

Abstrak: Penyakit pada tanaman padi merupakan salah satu faktor yang menyebabkan turunnya tingkat produksi padi. Penyakit tersebut adalah *bacterial leaf blight*, *leaf smut*, *brown spot* dan sebagainya. Upaya identifikasi sejak dini penyakit tanaman padi dilakukan dengan pemanfaatan algoritma, salah satunya GLCM dan klasifikasi KNN. Identifikasi jenis penyakit menggunakan metode klasifikasi KNN berdasarkan ekstraksi fitur GLCM dengan mengubah citra asli menjadi citra keabu-abuan (*grayscale*). Setelah citra asli tersebut diubah menjadi citra keabu-abuan (*grayscale*), kemudian diekstraksi menggunakan GLCM untuk mendapatkan ekstraksi nilai ciri. Digunakan metode KNN untuk mengelompokkan jenis kemiripan penyakit. Data yang digunakan sebanyak 240 gambar diperoleh dari *UCI Machine Learning Repository* yang terdiri atas 3 jenis penyakit padi. Sebanyak 210 gambar sebagai *data training* dan 30 gambar lainnya untuk data uji. Hasil penelitian ini setelah dilakukan 2 kali proses uji, tingkat akurasi tertinggi yang didapatkan sebesar 93,3%.

Kata Kunci: penyakit daun padi, klasifikasi, GLCM, KNN.

Abstract: One of the factors causing the decline in rice production in Indonesia is leaf disease. These diseases are *bacterial leaf blight*, *leaf smut*, *brown spot*, etc. Researchers have developed techniques to recognize rice leaf diseases by utilizing computations and algorithms, GLCM, and KNN classification. Identification of the type of disease using the KNN classification method based on GLCM feature extraction by converting the original image into a grayscale image. After the original image is transformed into a grayscale image, it is extracted using GLCM to obtain feature value extraction. The KNN method is used to classify types of disease similarity. The data used are 240 images obtained from the *UCI Machine Learning Repository*, consisting of 3 kinds of rice diseases types. There were 210 images as training data and another 30 images as test data. The results of this research after two times the test process, the highest level of accuracy obtained is 93.3%.

Keywords: leaf disease, classification, GLCM, KNN.

I. PENDAHULUAN

Terdapat penyakit tanaman padi yang paling penting di Indonesia ialah bercak coklat (bs), bercak daun (ls), dan hawar daun bakteri (blb). Tren kehilangan hasil panen akibat penyakit hawar daun bakteri (*bacterial leaf blight*) di wilayah Indonesia berkisar antara 15% hingga 24% [1].

Upaya proses identifikasi sejak dini gagal panen telah dilakukan melalui pemanfaatan teknologi komputasi. Cuplikan gambar berbagai penyakit daun padi diolah dengan rekyasa komputasi *Convolutional Neural Network* (CNN) [2] [3], *Support Vector Machine* (SVM) [4], dan *Gray level co-occurrence matrix* (GLCM) [5].

Algoritma GLCM memiliki kemampuan mengenali ciri sebuah citra dan telah digunakan pada banyak bidang. Ekstraksi ciri menggunakan

GLCM dilakukan pada daging [6], bunga, tanaman [7] [8], buah [9] [10] dan padi [2]–[5], [11], [12]. Namun demikian, perlu digunakan pembaruan kombinasi klasifikasi untuk mendapatkan pengetahuan terbaik khususnya dalam deteksi penyakit pada daun padi [5].

Pada penelitian ini dilakukan ekstraksi ciri penyakit daun padi dengan menerapkan algoritma GLCM dan melakukan pengelompokan kemiripan penyakit menggunakan algoritma KNN (*k-nearest neighbor*). Pada proses ekstraksi fitur digunakan *feature contrast, correlation, energy* dan *homogeneity*.

II. PENELITIAN YANG TERKAIT

Identifikasi dini penyakit daun padi telah dilakukan [2] menggunakan CNN dengan arsitektur MobileNet. Didapatkan nilai akurasi confusion matrix sebesar 92% dengan *epochs* 100 dan terjadi *overfitting*. Arsitektur lain digunakan oleh [3] menggunakan EfficientNetB3. Proses luaran berupa klasifikasi terhadap penyakit *brown spot* dan *bacterial leaf* dengan akurasi 79.53%.

Peneliti [11] melakukan ekstraksi ciri penyakit menggunakan GLCM dan pengelompokan jenis penyakit dengan *backpropagation*. Menggunakan 150 data uji, diperoleh akurasi sebesar 80%.

GLCM berhasil dimanfaatkan [5] untuk ekstraksi ciri tesktur penyakit pada daun padi. Proses pengelompokan penyakit kemudian diolah menggunakan *backpropagation* 3 lapis. Namun demikian, peneliti menyebutkan mekanisme pengelompokan perlu diuji dengan teknik lain untuk mendapatkan pengetahuan metode klasifikasi yang optimal khusus penyakit tanaman padi.

III. METODE PENELITIAN

Algoritma GLCM merupakan sebuah metode yang efektif untuk mengekstraksi ciri atau menganalisis tekstur. GLCM digambarkan dengan matriks yang berisi frekuensi munculnya pasangan dua piksel dengan intensitas, jarak dan arah tertentu, sehingga diperoleh lima parameter ciri yakni *contrast, correlation, energy, homogeneity*, dan *entropy* [13].

KNN merupakan metode untuk melihat tingkat kemiripan antar objek. Tingkat kemiripan diukur berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Salah satu algoritma yang digunakan untuk mengukur jarak adalah Euclidean dengan ditunjukkan seperti pada persamaan 1.

$$D(a, b) = \sqrt{\sum_{k=1}^d (a_k - b_k)^2} \dots (1)$$

$D(a,b)$ = Jarak skalar dari kedua vektor a dan vektor b dari matriks dengan ukuran d dimensi

a_k = vektor a (data uji)

b_k = vektor b (sampel data)

d = dimensi data

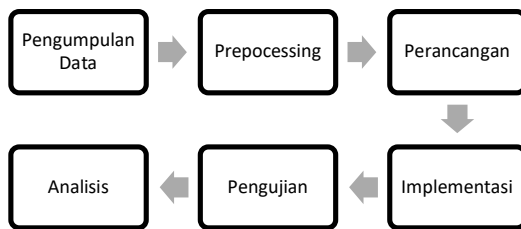
Proses penelitian ditunjukkan seperti Gambar. 1, diawali dengan pengumpulan data terdiri atas jurnal-jurnal yang relevan dan citra penyakit daun padi. Digunakan data citra sebanyak 240 yang diperoleh dari [14] dengan dibagi dalam tiga kelas yaitu *Bacterial leaf blight*, *Brown spot*, dan *Leaf smut*. Adapun koleksi citra dibagi dalam rincian sebagai berikut.

1. 70 file gambar *bacterial leaf blight* dengan format jpg.
2. 70 file gambar *brown spot* dengan format jpg.
3. 70 file gambar *leaf smut* dengan format jpg.

4. 30 file gambar yang terdiri dari 10 gambar *leaf blight*, 10 gambar *brown spot*, dan 10 gambar *leaf smut* dengan format jpg.

Data yang telah terkumpul kemudian diolah, koleksi citra dilakukan penyesuaian ukuran. Perancangan dilakukan untuk menyusun strategi pembagian data uji dan data latih. Selanjutnya memasuki tahap implementasi yakni pembuatan basis kode GLCM dan KNN. Empat sudut *co-occurrence* digunakan pada GLCM yaitu 00, 450, 900, dan 1350.

Pada tahap implementasi juga mulai dijalankan proses pengolahan data latih. Dari keseluruhan dataset yang digunakan akan dibagi menjadi 88% sebagai *data training* dan 12% sebagai data uji dan hasilnya akan dianalisis.



Gambar 1. Alur penelitian penyakit daun padi

IV. HASIL DAN PEMBAHASAN

Dilakukan pengujian menggunakan 30 data uji. Setiap data uji dilakukan pengujian sebanyak 2 kali yaitu dengan perbandingan *data training* dan data uji yang berbeda.

Pengujian pertama, perbandingan *data training* terhadap data uji sebesar 88:12 yaitu 210 gambar sebagai data latih dan 30 gambar sebagai data uji. Pengujian kedua, digunakan perbandingan 50:50 yang terdiri atas 30 gambar untuk proses latih dan 30 gambar untuk proses uji. Pada tahap ekstraksi fitur, *grayscale*, sudut dan jarak adalah ciri penting

yang digunakan untuk mendefinisikan GLCM. Sudut yang berpengaruh terhadap berubahnya nilai ekstraksi ciri dapat diketahui dengan menggunakan sudut- sudut berikut, yaitu 0, 45, 90, dan 135.

Dilakukan perhitungan diambil *sample* matriks 4 x 4 dari total matriks keseluruhan 400 x 400 *pixel*, dengan data *sample* pada Tabel 1. Proses komputasi berikutnya dapat disederhanakan dengan mengubah nilai matriks ke dalam rentang keabu-abuan yang awalnya 0 – 255 menjadi 0 – 7. Nilai setiap matriks dimasukan dalam persamaan 2 sehingga diperoleh hasil seperti pada Tabel 2.

$$Xb = \text{int} (X * (2^n - 1)/255) \dots (2)$$

Keterangan

Xb = nilai grayscale baru

X = nilai grayscale lama

2n = nilai derajat keabuan

Tabel 1. Matriks sample grayscale 4x4

Baris/kolom	0	1	2	3
0	127	122	122	125
1	128	123	122	124
2	126	127	126	123
3	124	128	129	127

Tabel 2. Matriks range baru 0-7

Baris/kolom	0	1	2	3
0	4	3	3	3
1	4	3	3	3
2	4	4	4	3
3	3	4	4	4

Kemudian diisikan hubungan spasial matriks untuk membentuk matriks kookurensi dengan d = 1 sudut 0 seperti pada Tabel 3. Matriks dinormalisasikan dengan membaginya dengan jumlah akumulasi angka yang muncul, yaitu 12. Diperoleh hasil normalisasi yang akan digunakan untuk proses ekstraksi fitur GLCM yaitu *energy*,

contrast, *correlation*, dan *homogeneity* seperti pada Tabel 4.

Tabel 3. Matriks kookurensi

Baris/kolom	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
3	0	0	0	4	1	0	0	0
4	0	0	0	3	4	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Tabel 4. Hasil normailisasi

Baris/kolom	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
3	0	0	0	0,333	0,083	0	0	0
4	0	0	0	0,25	0,333	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Dilakukan proses uji sebanyak dua kali dengan dua variasi perbandingan data latih terhadap data uji. Perbandingan pertama 88:12 ditunjukkan pada Tabel 5 dan perbandingan kedua 50:50 ditunjukkan pada Tabel 6. Adapun keterangan tabel sebagai berikut:

blb = *bacterial leaf blight*

ls = *leaf smut*

bs = *brown spot*

Tabel 5. Hasil klasifikasi uji 88:12

Data	Target	Hasil uji	Ket.
1	blb	blb	True
2	blb	blb	True
3	blb	blb	True
4	blb	blb	True
5	blb	blb	True
6	blb	blb	True
7	blb	ls	False
8	blb	blb	True
9	blb	blb	True
10	blb	blb	True
11	ls	ls	True
12	ls	ls	True

13	ls	ls	True
14	ls	ls	True
15	ls	ls	True
16	ls	ls	True
17	ls	ls	True
18	ls	blb	False
19	ls	ls	True
20	ls	ls	True
21	bs	bs	True
22	bs	bs	True
23	bs	bs	True
24	bs	bs	True
25	bs	bs	True
26	bs	bs	True
27	bs	bs	True
28	bs	bs	True
29	bs	bs	True
30	bs	bs	True

Setelah proses perhitungan GLCM, dihasilkan 16 nilai ciri yang akan digunakan sebagai data inputan untuk klasifikasi menggunakan *K-Nearest Neighbor*. Dari proses uji diperoleh hasil klasifikasi ditunjukkan seperti pada Tabel 5.

Pada proses uji pertama, diperoleh hasil dari 30 data yang diuji, 28 terhitung sesuai/benar (*True*) dan 2 data terhitung tidak sesuai/salah (*False*). Penghitungan akurasi dengan persamaan 3 ditunjukkan seperti berikut:

$$\text{akurasi} = \frac{(\text{jml data} - \text{jml salah})}{\text{jml data}} * 100\% \dots (3)$$

$$\text{akurasi} = ((30-2))/(30)*100\%$$

$$\text{akurasi} = 93,3\%$$

Pada proses uji kedua menggunakan perbandingan dataset 50:50 didapatkan hasil seperti ditunjukkan pada Tabel 6. Sejumlah 20 data dinyatakan sesuai/benar (*True*) dan 10 data dihasilkan tidak sesuai/salah (*False*). Sedangkan nilai akurasi sebesar 66,6%.

Tabel 6. Hasil klasifikasi uji 50:50

Data	Target	Hasil uji	Ket.
1	blb	ls	False
2	blb	bs	False
3	blb	ls	False
4	blb	blb	True
5	blb	ls	False
6	blb	blb	True
7	blb	ls	False
8	blb	ls	False
9	blb	ls	False
10	blb	blb	True
11	ls	ls	True
12	ls	ls	True
13	ls	ls	True
14	ls	ls	True
15	ls	blb	False
16	ls	ls	True
17	ls	ls	True
18	ls	blb	False
19	ls	ls	True
20	ls	ls	True
21	bs	bs	True
22	bs	blb	False
23	bs	bs	True
24	bs	bs	True
25	bs	bs	True
26	bs	bs	True
27	bs	bs	True
28	bs	bs	True
29	bs	bs	True
30	bs	bs	True

Perbandingan hasil seluruh pengujian sistem dengan jumlah perbandingan *dataset* yang berbeda ditunjukkan pada Tabel 7. Dari kedua *testing* diperoleh akurasi tertinggi 93,3% dengan 240 *dataset* dan akurasi terendah 66,6% dengan 60 *dataset*.

Tabel 7. Hasil seluruh pengujian

No	datalatih	datauji	true	false	akurasi
1	210	30	28	2	93,3%
2	30	30	20	10	66,6%

V. KESIMPULAN

Tekstur daun padi dapat dikenali dengan metode GLCM dengan cara menghasilkan nilai matrix dari atribut ciri berupa contrast, homogeneity, energy dan correlation pada jenis penyakit yang diteliti. Proses testing dilakukan

sebanyak 2 kali dengan jumlah dataset yang berbeda, pada proses testing menggunakan data 50:50 menghasilkan akurasi sebesar 66,6% dan testing kedua menggunakan data 88:12 menghasilkan akurasi sebesar 93,3%. Secara keseluruhan GLCM dan KNN dapat diimplementasikan ke dalam sistem identifikasi penyakit daun padi dan menghasilkan akurasi tertinggi sebesar 93,3%. Namun demikian, sistem ini belum dapat digunakan dengan mudah oleh petani, karena memerlukan perangkat lunak tertentu. Sebagai penyempurnaan berikutnya, perlu diterapkan pada mesin server dan dibuat API (Application Programming Interface) supaya dapat digunakan melalui client apps berbasis Android atau Web.

UCAPAN TERIMA KASIH

Ucapan terima kasih penulis kepada pengelola Prodi Informatika dan LPPM yang membantu ataupun memberikan dukungan terkait dengan penelitian yang dilakukan seperti bantuan fasilitas penelitian, dana hibah, dan kegiatan penelitian lain.

REFERENSI

- [1] Sutarman, "Dasar Ilmu Penyakit Tanaman," Sidoarjo, 2017.
- [2] R. A. Saputra, S. Wasyanti, A. Supriyatna, and D. F. Saefudin, "Penerapan Algoritma Convolutional Neural Network Dan Arsitektur MobileNet Pada Aplikasi Deteksi Penyakit Daun Padi," JURNAL SWABUMI, vol. 9, no. 2, 2021, [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Rice>
- [3] E. Anggiratih, S. Siswanti, S. K. Octaviani, and A. Sari, "Klasifikasi Penyakit Tanaman Padi Menggunakan Model Deep Learning Efficientnet B3 dengan Transfer Learning," Jurnal Ilmiah SINUS, vol. 19, no. 1, p. 75, Jan. 2021, doi: 10.30646/sinus.v19i1.526.
- [4] A. Purnamawati, W. Nugroho, D. Putri, and W. F. Hidayat, "Deteksi Penyakit Daun pada Tanaman Padi Menggunakan Algoritma Decision Tree, Random Forest, Naïve Bayes, SVM dan KNN," vol. 5, no. 1, 2020, doi: 10.30743/infotekjar.v5i1.2934.
- [5] N. A. Haris, H. Asgar, J. Sumah, and Kusriani, "Kombinasi Ciri Bentuk dan Ciri Tekstur Untuk Identifikasi Penyakit Pada Tanaman Padi," Jurnal Teknik Informatika dan Sistem Informasi, vol. 7, no. 2, pp. 237–249, 2020.

- [6] Y. Fernando et al., "KLASIFIKASI JENIS DAGING BERDASARKAN ANALISIS CITRA TEKSTUR GRAY LEVEL CO-OCCURRENCE MATRICES (GLCM) DAN WARNA," 2017.
- [7] D. Putra Pamungkas, "Ekstraksi Citra menggunakan Metode GLCM dan KNN untuk Identifikasi Jenis Anggrek (Orchidaceae)," vol. 1, no. 2, pp. 51–56, 2019.
- [8] Suhendri and P. Rahayu, "Metode Grayscale Co-occurrence Matrix (GLCM) Untuk Klasifikasi Jenis Daun Jambu Air Menggunakan Algoritma Neural Network," *Journal of Information Technology*, vol. 1, no. 1, pp. 15–22, 2019.
- [9] R. Widodo et al., "Pemanfaatan Ciri Gray Level Co-Occurrence Matrix (GLCM) Citra Buah Jeruk Keprok (*Citrus reticulata* Blanco) untuk Klasifikasi Mutu," 2018. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [10] M. Widyarningsih, "IDENTIFIKASI KEMATANGAN BUAH APEL DENGAN GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM)," *Jurnal SAINTEKOM*, vol. 6, no. 1, pp. 71–88, 2017, Accessed: Dec. 11, 2021. [Online]. Available: <https://ojs.stmikplk.ac.id/index.php/saintekom/article/view/7>
- [11] J. Kusanti, K. Penyakit, D. Padi, and A. Haris, "Klasifikasi Penyakit Daun Padi Berdasarkan Hasil Ekstraksi Fitur GLCM Interval 4 Sudut," *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, vol. 03, no. 01, 2018.
- [12] R. N. Whidhiasi and I. Ekawati, "IDENTIFIKASI JENIS PENYAKIT DAUN PADI MENGGUNAKAN ADAPTIF NEURO FUZZY INFERENCE SYSTEM (ANFIS) BERDASARKAN TEKSTUR," 2019.
- [13] D. Rohpandi, A. Sugiharto, M. Yoga, and S. Jati, "Klasifikasi Citra Digital Berbasis Ekstraksi Ciri Berdasarkan Tekstur Menggunakan GLCM Dengan Algoritma K-Nearest Neighbor," 2018.
- [14] H. B. Prajapati, J. P. Shah, and V. K. Dabhi, "Detection and classification of rice plant diseases," *Intelligent Decision Technologies*, vol. 11, no. 3, pp. 357–373, 2017, doi: 10.3233/IDT-170301.