

Klasifikasi Jenis Lamun Menggunakan Ekstraksi Fitur GLCM dan Algoritma K-Nearest Neighbor (KNN)

M. Mudaffarsyah¹, Muhammad Azza Al Kausar², Obi Luter Sihombing³, Halta Putra Ash Sidiq⁴,
Kirana Putri Fercia⁵, Nurul Hayaty^{6*}

^{1,2,3,4,5,6}Jurusan Teknik Informatika, FTTK, Universitas Maritim Raja Ali Haji
^{1,2,3,4,5,6}Jl. Politeknik Senggarang, Tanjungpinang 29100

*Corresponding Author: nurul.hayaty@umrah.ac.id

Abstract — Seagrass is a type of flowering plant (*Angiospermae*) that grows fully submerged in shallow coastal waters and estuaries, playing a vital role in marine ecosystems. Currently, seagrass species identification is still performed manually by experts, which is time-consuming, costly, and labor-intensive. To support more efficient conservation and ecological monitoring, an automated, fast, and accurate method is needed. This study proposes the combination of the K-Nearest Neighbors (KNN) algorithm for classification and Gray Level Co-occurrence Matrix (GLCM) for texture feature extraction. The seagrass image data was obtained from the Roboflow website, and the value of k used in KNN was set to 3. Feature extraction using GLCM was conducted at angles of 0° , 45° , 90° , and 135° . The results showed the highest accuracy at $k=3$, with 77.42% accuracy on training data and 73.33% on testing data. Therefore, the combination of KNN and GLCM has proven capable of providing fairly accurate results in identifying seagrass species.

Keywords— GLCM, KNN, Classification, Digital Image, Seagrass

Intisari— Lamun merupakan tumbuhan berbunga yang hidup terendam di perairan laut dangkal dan estuari, berperan penting dalam ekosistem maritim. Identifikasi jenis lamun hingga kini masih dilakukan secara manual oleh ahli, membutuhkan waktu, biaya, dan tenaga yang tidak sedikit. Untuk mendukung efisiensi dalam konservasi dan monitoring ekologis, diperlukan metode otomatis yang cepat dan akurat. Penelitian ini mengusulkan kombinasi metode K-Nearest Neighbors (KNN) untuk klasifikasi dan Gray Level Co-occurrence Matrix (GLCM) untuk ekstraksi fitur tekstur. Data lamun diperoleh dari situs Roboflow, dengan nilai $k=3$ pada KNN. Ekstraksi fitur GLCM dilakukan menggunakan sudut 0° , 45° , 90° , dan 135° . Hasil menunjukkan akurasi tertinggi pada $k=3$, yaitu 77,42% untuk data training dan 73,33% untuk data testing. Dengan demikian, kombinasi KNN dan GLCM terbukti mampu memberikan hasil yang cukup akurat dalam identifikasi jenis lamun.

Kata kunci— GLCM, KNN, Klasifikasi, Citra Digital, Lamun.

I. PENDAHULUAN

Lamun adalah tumbuhan berbunga (*Angiospermae*) yang hidup terendam dalam kolom air dan berkembang dengan baik di perairan laut dangkal dan estuari. Tumbuhan lamun terdiri dari daun, batang menjalar yang biasanya disebut rimpang (*rhizome*), dan akar yang tumbuh pada bagian rimpang [1]. Lamun

mempunyai peran penting dalam menjaga ekosistem pesisir sebagai habitat biota laut, penstabil sedimen, dan penyerap karbon. Namun, identifikasi jenis lamun masih mengandalkan metode manual oleh ahli lapangan, yang membutuhkan waktu, biaya, dan tenaga. Perlu ada cara otomatis yang cepat dan akurat untuk mendukung konservasi dan monitoring ekologis.

Teknologi pengolahan citra digital mampu menawarkan solusi berupa klasifikasi visual otomatis. Salah satu pendekatan yang efektif adalah analisis tekstur citra, karena tekstur mampu membedakan permukaan daun lamun meski dari segi bentuk atau warna terlihat mirip.

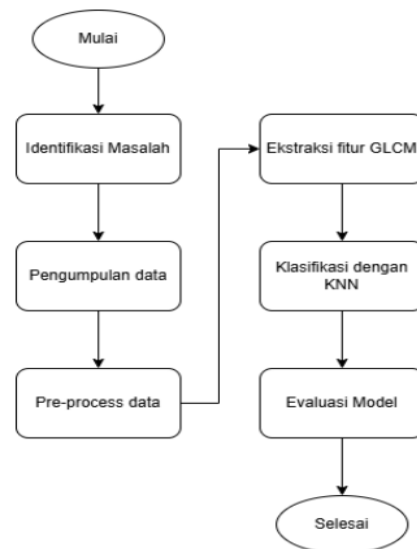
Gray Level Co-occurrence Matrix (GLCM) adalah teknik populer untuk ekstraksi fitur tekstur—menghasilkan nilai statistik seperti contrast, energy, correlation, dan homogeneity. Metode ekstraksi ini terbukti efektif dalam mengklasifikasikan jenis objek berdasarkan tekstur seperti penelitian yang telah dilakukan oleh Sarimole dan Syaeful pada tahun 2022 [2] menyatakan bahwa Klasifikasi Jenis Durian dengan ekstraksi *Gray Level Co-occurrence Matrix* (GLCM) dan *K-Nearest Neighbor* (KNN) mendapatkan hasil akurasi yang sangat baik pada nilai $k = 3$ yaitu 93%.

Setelah proses ekstraksi fitur tahap selanjutnya adalah klasifikasi. Algoritma *K-Nearest Neighbor* (KNN) digunakan karena sederhana namun efektif untuk mengklasifikasikan jenis objek. Penelitian terbaru yang dilakukan oleh Ilhami dan Sanjaya pada tahun 2022 [3] untuk mengklasifikasikan Citra Buah Pisang dengan Algoritma KNN menghasilkan akurasi sebesar 76% pada $k = 1$. Penelitian juga pernah dilakukan oleh Wijaya dan Widiarti pada tahun 2024 [4] Klasifikasi jenis batik menggunakan KNN dan GLCM menghasilkan akurasi 75%. Pada tahun 2024 Kurniati dkk [5] melakukan penelitian untuk deteksi objek menggunakan ekstraksi kombinasi KNN dan RF, KNN menghasilkan akurasi lebih tinggi dibandingkan RF yaitu 92,7% dan F1 92,5%.

Penelitian ini bertujuan untuk menerapkan metode ekstrak tekstur menggunakan GLCM kemudian diklasifikasikan menggunakan algoritma KNN. Melalui metode ini, diharapkan sistem mampu mengklasifikasikan jenis lamun dalam citra digital dengan akurat. Penelitian ini juga mengukur akurasi klasifikasi untuk mengetahui seberapa baik kombinasi metode GLCM dan KNN dalam pengenalan jenis lamun.

II. METODE PENELITIAN

Penelitian ini mengidentifikasi masalah yang ada. Setelah permasalahan dipahami dengan jelas, maka akan dilanjutkan dengan melakukan studi literatur yang berkaitan dengan masalah yang ingin diselesaikan. Gambar 1 merupakan *flowchart* metode penelitian.



Gambar 1. Flowchart Metode Penelitian

2.1 Identifikasi Masalah

Penulis melakukan studi literatur dari jurnal-jurnal terbaru untuk mengidentifikasi masalah penutupan lamun di Indonesia sudah mulai menipis untuk mengangkat judul penelitian yang berhubungan dengan lamun

2.2 Pengumpulan Data

Data yang digunakan untuk penelitian ini merupakan data yang dipublis oleh situs roboflow yang berjumlah 1004 dengan pembagian *Thalassia Hemprichii* (500), *Enhalus Acoraides* (235), *Cymodosea Rotundata* (269). Tabel 1 Menunjukkan contoh-contoh gambar yang digunakan sebagai dataset

Tabel 1. Contoh Gambar Jenis Lamun

Jenis Lamun		
<i>Thalassia Hemprichii</i>	<i>Enhalus Acoraides</i>	<i>Cymodosea Rotundata</i>
		

2.3 Pre Process Data

a. *Resize image*

Citra lamun yang diperoleh memiliki ukuran pixel yang sangat besar, ini akan memberatkan sistem dalam mengekstrak tekstur dan memungkinkan overfitting. Karena itu penulis melakukan *resize image* menjadi 128 x128 pixel.

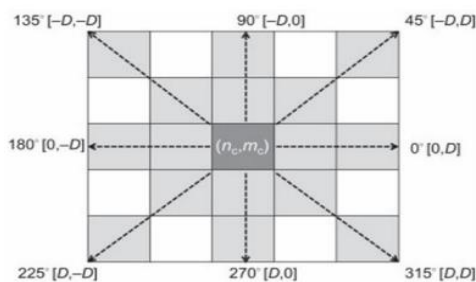
b. *Augmentasi Data*

Data gambar lamun yang diambil dari situs roboflow memiliki total gambar 1004 dengan pembagian *Thalassia Hemprichii* (500), *Enhalus Acoraides* (235), *Cymodosea Rotundata* (269) ini berarti data tersebut tidakimbang, karena itu penulis melakukan augmentasi menjadi 500 untuk jenis *Enhalus Acoraides* dan *Cymodosea Rotundata* agar data tersebut seimbang untuk persiapan ekstraksi.

2.4 Ekstraksi Fitur GLCM

Tekstur (*texture*) adalah keteraturan pola-pola tertentu yang tersusun dari beberapa piksel pada daerah tertentu dalam citra digital. Informasi tekstur digunakan untuk menentukan ciri permukaan suatu benda dalam citra yang berhubungan dengan ketajaman dan kecerahan, juga ciri spesifik dari ketajaman dan kecerahan permukaan tersebut [3]. Penelitian ini menggunakan GLCM untuk ekstraksi ciri tekstur.

GLCM adalah suatu matrik kookurensi yang berisi angka berdasarkan jumlah kemunculan piksel berdasarkan tingkat keabuan dari tiap piksel tersebut [6]. Pasangan piksel itu berada pada jarak (d) dan sudut tertentu (θ). Objek orientasi sudut yang sering digunakan yaitu sudut 0° , 45° , 90° , dan 135° (bisa dilihat pada Gambar 2).



Gambar 2. Matriks GLCM

Jarak antar piksel orientasi dengan piksel tetangga biasanya ditetapkan sebesar 1 piksel atau 2 piksel. Dalam penelitian ini menggunakan empat fitur ciri, yaitu: energi, kontras, entropi, idm dengan persamaan sebagai berikut:

a. Energi

Energi adalah fitur pengukur konsentrasi intensitas pasangan bersama matriks. Rumus yang digunakan untuk menghitung energi adalah [7] :

$$energi = \sum_{i,j}^{n-1} p_{(i,j)}^2 \quad (1)$$

b. Kontras

Perhitungan perbedaan dari intensitas antara piksel satu dan piksel lain yang berada berdekatan pada seluruh gambar [8].

$$Kontras = \sum_{n=1}^L \{ \sum_{|i-j|} GLCM_{(i,j)} \} \quad (2)$$

c. Entropi

Entropi adalah besaran skalar yang dapat didefinisikan sebagai ukuran statistic keacakan. Dan sangat membantu untuk mengidentifikasi fitur gambar masukan [9].

$$Entropi = \sum_{x,y=1}^{n-1} \log(p_{(i,j)}) * p_{(i,j)} \quad (3)$$

d. Inverse Difference Moment (IDM)

IDM digunakan untuk mengukur pengaruh homogenitas warna [10].

$$IDM = \sum_{i,j}^L \frac{(GLCM_{(i,j)})^2}{1+(GLCM_{(i-j)})^2} \quad (4)$$

Pada saat pengestrakan data akan langsung dilabeling sesuai folder jenis lamun, outputnya adalah ekstrak.csv. Proses ekstrak dilakukan dengan program menggunakan bahasa Python.

2.5 Klasifikasi dengan KNN

Algoritma K-Nearest Neighbor adalah sebuah algoritma untuk melakukan klasifikasi terhadap data baru berdasarkan data pembelajaran yang jaraknya paling dekat dengan data tersebut [4]. Setelah proses ekstrak dan labeling, data akan dibagi menjadi dua segmen yaitu data training 80% dan data testing 20%. Langkah berikutnya adalah melakukan pengurutan data dalam data training berdasarkan hasil perhitungan nilai jarak. Selanjutnya, dari hasil pengurutan tersebut diambil kategori mayoritas sejumlah K data yang memiliki nilai kedekatan. Nilai K sudah

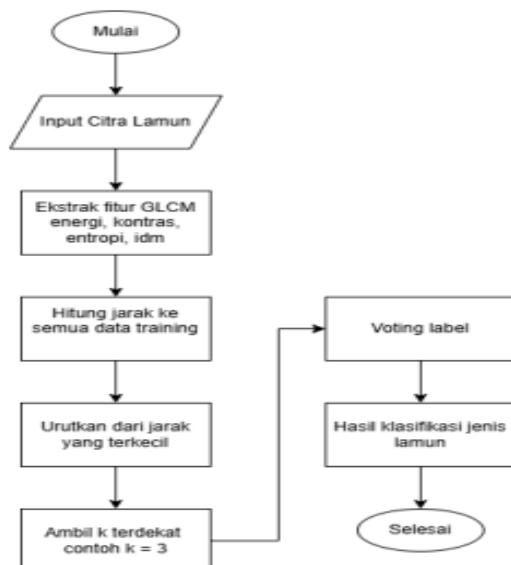
ditentukan atau diasumsikan sebelumnya sesuai dengan syarat [4] yaitu:

- Nilai K Harus lebih dari 1
- Nilai K adalah nilai ganjil (karena jika diambil K adalah bilangan genap akan ada kemungkinan hasil klasifikasi sulit ditentukan karena masing-masing kelas bernilai sama.
- Nilai K lebih dari jumlah kelas.
- Nilai K tidak melebihi jumlah data training

Untuk mencari kategori mayoritas sebanyak nilai K, pada tahap ini akan dilakukan pengecekan apabila terdapat kategori mayoritas maka kategori akan langsung terklasifikasi. Namun apabila tidak terdapat kategori mayoritas maka akan dipilih kategori dengan nilai kemiripan (cosine similarity) yang paling tinggi [4]. Untuk melihat alur dari sistem klasifikasi jenis lamun menggunakan algoritma KNN ini dapat dilihat pada Gambar 3.

2.6 Evaluasi Model

Pengujian kualitas dilakukan untuk mengetahui kinerja dari algoritma klasifikasi yang telah diterapkan. Ada beberapa cara untuk mengukur kinerja algoritma klasifikasi tiga diantaranya adalah precision, recall dan fmeasure [4]. Evaluasi model pada penelitian ini dilakukan sebanyak 4 kali yaitu evaluasi dengan data training dan testing masing-masing menggunakan 4 nilai K yaitu 3, 5, 7, 9. Hasil akurasi tertinggi didapatkan pada nilai K = 3.



Gambar 3. Flowchart Alur Sistem

III. HASIL DAN PEMBAHASAN

3.1 Pengujian Model Prediksi

Pengujian pada model prediksi dibagi menjadi dua segmen dengan menggunakan data training sebanyak 1200 dan data testing sebanyak 300. Pengujian dilakukan dengan Confusion Matrix Pada Tabel. 2 merupakan hasil pengujian model dari data training dan Tabel. 3 merupakan evaluasi model dari data testing pada sistem klasifikasi jenis lamun.

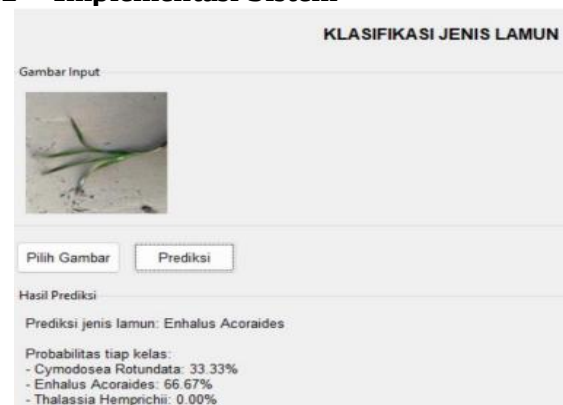
Tabel 2. Pengujian Model Data Training

CM	Jenis Lamun		
	<i>Thalassia Hemprichii</i>	<i>Enhalus Acoraides</i>	<i>Cymodosea Rotundata</i>
Precision	77%	85%	71%
Recall	68%	82%	82%
F1- Score	72%	83%	76%
Total	77,42%		

Tabel 3. Pengujian Model Data Testing

CM	Jenis Lamun		
	<i>Thalassia Hemprichii</i>	<i>Enhalus Acoraides</i>	<i>Cymodosea Rotundata</i>
Precision	80%	79%	62%
Recall	71%	74%	76%
F1- Score	75%	76%	68%
Total	73,33%		

3.2 Implementasi Sistem



Gambar 4. Percobaan Klasifikasi Jenis Lamun

Hasil klasifikasi jenis lamun menggunakan sistem dijelaskan dalam Tabel 4.

Tabel 4. Hasil Percobaan

Input	Jenis Lamun		
	<i>Thalassia Hemprichii</i>	<i>Enhalus Acoraides</i>	<i>Cymodosea Rotundata</i>
TH	66.67%	00.00%	33.33%
EA	00.00%	66.67%	33.33%
CR	33.33%	00.00%	66.67%

Hasil klasifikasi dari jenis lamun *Thalassia Hemprichii*, *Enhalus Acoraides*, *Cymodosea*

Rotundata berhasil dikenali dengan hasil 66.67% dan jenis lain 33.33%. Ini membuktikan bahwa sistem sudah bisa digunakan untuk klasifikasi jenis lamun.

Dari Tabel 2 terlihat bahwa evaluasi model dengan data training memiliki performa berbeda untuk setiap jenis lamun. *Thalassia hemprichii* memiliki presisi tertinggi (77%) tetapi *recall* rendah (68%), artinya model cukup akurat saat memprediksinya sebagai *Thalassia*, tetapi masih sering melewatkan kasus yang sebenarnya. *Enhalus acoroides* unggul di semua metrik (*presisi* 85%, *recall* 82%, *F1-score* 83%), menunjukkan model sangat baik mengenali jenis ini. Sementara *Cymodocea rotundata* memiliki *recall* tinggi (82%) tetapi presisi rendah (71%), berarti model sering salah mengidentifikasi jenis lain sebagai *Cymodocea*. Secara keseluruhan, akurasi total 77.42% menunjukkan model cukup baik, tetapi perlu peningkatan khususnya untuk *Thalassia* (*presisi* vs *recall* tidak seimbang) dan *Cymodocea* (banyak false positive).

Dari Tabel 3 terlihat bahwa evaluasi model dengan data testing menghasilkan akurasi 73.33%, di mana *Thalassia hemprichii* memiliki presisi tertinggi (80%) namun *recall* rendah (71%), *Enhalus acoroides* menunjukkan performa seimbang (*presisi* 79% dan *recall* 74%), sedangkan *Cymodocea rotundata* memiliki presisi terendah (62%) meskipun *recall*-nya cukup baik (76%). Hasil ini mengindikasikan bahwa model sudah cukup baik dalam mengklasifikasikan namun masih perlu perbaikan, terutama dalam meningkatkan ketepatan identifikasi *Cymodocea* dan kemampuan menemukan semua kasus *Thalassia*.

3.3 Diskusi

Hasil akurasi 77,42% (training) dan 73,33% (testing) mengindikasikan bahwa: (1) ukuran gambar 128x128 pixel terlalu kecil sehingga mengurangi kualitas ekstraksi tekstur GLCM, (2) selisih akurasi yang kecil (4%) membuktikan augmentasi berhasil mencegah *overfitting*, namun (3) latar belakang yang tidak dihapus (seperti pasir/batu) menambah *noise*. Untuk meningkatkan akurasi, disarankan: (a) melakukan *cropping*/segmentasi untuk menghilangkan

background, (b) meningkatkan resolusi gambar menjadi 256x256 pixel atau naikan lagi, dan (c) menambahkan fitur geometris seperti *eccentricity* (mengukur tingkat keeksentrikan bentuk) dan *elongation* (rasio panjang/lebar) sebagai pelengkap analisis tekstur GLCM untuk akurasi yang lebih baik.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan dapat disimpulkan bahwa penelitian ini berhasil membuktikan bahwa kombinasi metode KNN dan ekstraksi ciri GLCM mampu mengidentifikasi jenis lamun dengan cukup akurat. Hasil terbaik diperoleh saat menggunakan $k=3$, dengan tingkat keberhasilan 77,42% untuk data latih dan 73,33% untuk data uji.

REFERENSI

- [1] M. Frisca, A. Tana, and I. N. Kailola, "Keanekaragaman Jenis-Jenis Lamun Di Indonesia," *Journal of International Multidisciplinary Research*, [Online]. Available: <https://journal.banjaresepacific.com/index.php/jimr>
- [2] F. Matheos Sarimole and A. Syaeful, "Classification Of Durian Types Using Features Extraction Gray Level CoOccurrence Matrix (GlcM) And K-Nearest Neighbors (KNN)."
- [3] R. Syahrul Ilhamy and U. P. Sanjaya, "Algoritma K-Nearest Neighbors (KNN) untuk Klasifikasi Citra Buah Pisang dengan Ekstraksi Ciri Gray Level CoOccurrence," *Jurnal Telematika*, vol. 17, no. 2.
- [4] D. Afreyna Fauziah, A. Maududie, and I. Nuritha, "Klasifikasi Berita Politik Menggunakan Algoritma K-nearest Neighbor (Classification of Political News Content using K-Nearest Neighbor)."
- [5] F. T. Kurniati, D. H. F. Manongga, I. Sembiring, S. Wijono, and R. R. Huizen, "The object detection model uses combined extraction with KNN and RF classification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 35, no. 1, pp. 436–445, Jul. 2024, doi: 10.11591/ijeecs.v35.i1.pp436-445.
- [6] B. Jabber, K. Rajesh, D. Haritha, C. Z. Basha, dan S. N. Parveen, "An intelligent system for classification of brain tumours with GLCM and back propagation neural network," dalam *Proc. 4th Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2020*, hlm. 21-25.

- [7] Wibowo, S. A., Hidayat, B., & Sunarya, U. (2016). Simulasi dan Analisis Pengenalan Citra Daging Sapi dan Daging Babi dengan Metode GLCM dan KNN. Seminar Nasional Inovasi Dan Aplikasi Teknologi Di Industri (Seniati) 2016, 338–343.
- [8] Aferi, F. A., Purboy. T. W., & Saputra, R. E., (2018). Cotton texture segmentation based on image texture analysis using gray level run length and Euclidean distance. *International Journal of Applied Engineering Research*, 13(1), 6915–6923.
- [9] Usha, R., & Perumal, K. (2019). SVM classification of brain images from MRI scans using morphological transformation and GLCM texture features. *International Journal of Computational Systems Engineering*, 5(1), 18. <https://doi.org/10.1504/ijcsyse.2019.098415>
- [10] Sari, Y., Alkaff, M., & Pramunendar, R. A. (2018). Classification of coastal and Inland Batik using GLCM and Canberra Distance. *AIP Conference Proceedings*, 1977(June). <https://doi.org/10.1063/1.5042901>.