

Pengenalan Rumput Laut Menggunakan Euclidean Distance Berbasis Ekstraksi Fitur

Resty Wulanningrum, S.Kom¹, Aeri Rachmad, ST.MT²

¹Jurusan Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri
Jl. K.H. Achmad Dahlan No. 76
Telp. (0354) 776706 Kediri

²Jurusan Teknik Informatika, Fakultas Teknik, Universitas Trunojoyo Madura
Jl. Raya Telang, Kecamatan Kamal, Bangkalan, Madura 16912 Indonesia
Telp. (031) 3011146, Fax. (031) 3011506
E-mail: re_resty@yahoo.com, aery_r@yahoo.com

ABSTRAK

Rumput laut merupakan salah satu komoditas unggulan di Indonesia. Diantaranya mempunyai keunggulan komparatif karena tersedia dalam jumlah yang besar dan beraneka ragam. Semakin banyaknya jenis rumput laut, maka dibutuhkan suatu sistem yang digunakan untuk pengenalan rumput laut dengan proses pelatihan citra rumput laut, penggunaan pixel dari citra secara langsung dapat mengakibatkan banyaknya fitur-fitur citra rumput laut yang tidak dapat terekstraksi dengan baik. Sehingga diperlukan suatu pemrosesan awal yang dapat mengekstraksi fitur-fitur citra dengan baik. Dimana pada penelitian ini digunakan PCA (Principal Component Analysis) dan Euclidean Distance. Dari ujicoba kelas A (Sargassum), B (Gracilaria) dan C (Eucheuma Cottoni) diperoleh akurasi sebesar 96.6%, pada dataset 40 data testing 10 masing-masing kelas dan menggunakan dimensi 120.

Kata kunci: Pengenalan rumput laut, PCA, Euclidean Distance

1. PENDAHULUAN

1.1 Latar Belakang

Seiring perkembangan teknologi yang sedemikian pesat merupakan akibat dari usaha manusia dalam memenuhi kebutuhannya. Hampir dalam segala bidang, teknologi merasuk hingga di sela-sela fenomena kehidupan, namun sekecil apapun peran teknologi pada dasarnya sebagai cermin eksistensi dalam usaha manusia, dimana eksistensi tersebut merupakan bagian dari aktualisasi. Begitu pula mengenai sumber daya kelautan, hingga saat ini dari usaha-usaha manusia tidak pernah berhenti menciptakan berbagai cara untuk mengembangkan dan memanfaatkan teknologi sebagai titik temu dalam menyelesaikan tingkat waktu yang dibutuhkan dalam proses budidaya.

Secara harfiah biota laut yang pada dasarnya adalah sebagai sumber hayati laut sangat berperan penting dalam mendukung berbagai aspek kehidupan yang secara nyata dapat dirasakan manfaatnya secara nasional jika dikelola dan dimanfaatkan secara arif dan bijaksana. Salah satu komoditas marikultuer yang saat ini sedang berkembang adalah rumput laut, yang pada umumnya rumput laut digunakan sebagai bahan makanan dan minuman, namun seiring dengan berkembangnya IPTEK dewasa ini rumput laut dapat di kembangkan dan dimanfaatkan dalam berbagai macam industri misalnya tekstil, kosmetik, dan industri kefarmasian.

Klasifikasi rumput laut menggunakan citra merupakan salah satu cara yang dapat digunakan sebagai identifikasi pengenalan rumput laut. Namun dalam proses pengenalan citra rumput laut akan terjadi ketidak-efisienan jika *pixel* dalam citra langsung digunakan ke dalam proses pengenalan dan identifikasi citra rumput laut, sehingga diperlukan sebuah model komputasi untuk mengubah *pixel* dalam citra rumput laut menjadi suatu ciri rumput laut, sehingga dapat digunakan dalam skala dan orientasi rumput laut yang berbeda-beda.

Pada penelitian ini digunakan PCA (Principal Component Analysis) dalam ekstraksi ciri. Hasil PCA (Principal Component Analysis) hanya berupa real tanpa ada imajiner. Hal ini banyak membantu karena dapat mengurangi perhitungan. Sehingga dapat mempercepat dalam proses ekstraksi cirinya. Pada penelitian sebelumnya digunakan metode Jaringan Syaraf Tiruan (JST) Quick Propagation dengan pemrosesan awal Principal Components Analysis (PCA).

1.2 Perumusan Masalah

Perumusan masalah pada penelitian ini adalah:

- Bagaimana mengkonversi dan menormalisasi suatu citra sehingga citra tersebut dapat diproses ekstraksi?
- Bagaimana mengimplementasikan PCA (Principal Component Analysis) sebagai metode dalam proses ekstraksi ciri?

- c. Bagaimana menentukan fitur yang optimal untuk klasifikasi rumput laut?

1.3 Tujuan

Tujuan dari penelitian ini adalah untuk membuat sistem pengenalan klasifikasi rumput laut dengan masukan berupa citra rumput laut yang akan dikenali kemudian sistem ini akan menampilkan klasifikasi rumput laut tersebut dan tidak akan mengenali yang tidak ada dalam kasus.

1.4 Batasan Masalah

Batasan masalah dari penelitian ini adalah:

- a. Klasifikasi rumput laut yang di gunakan adalah jenis rumput laut yang bisa di buat obat (*Gracilaria*), makanan (*Eucheuma Cottoni*) dan kosmetik (*Sargassum*).
- b. Data yang diinputkan dalam sistem ini adalah file gambar berformat BMP.
- c. Ekstraksi ciri menggunakan PCA (*Pincipal Component Analysis*).
- d. Data yang dipakai untuk learning 40, 30, 20, 10 dan 10, 20, 30, 40 untuk testing.
- e. Data gambar yang digunakan berukuran 100x100 pixel.

2. LANDASAN TEORI

2.1 Rumput Laut

Rumput laut merupakan ganggang yang hidup di laut dan tergolong dalam divisio *thallophyta*. Keseluruhan dari tanaman ini merupakan batang yang dikenal dengan sebutan *thallus*, bentuk *thallus* rumput laut ada bermacam-macam ada yang bulat seperti tabung, pipih, gepeng, bulat seperti kantong, rambut dan lain sebagainya. *Thallus* ini ada yang tersusun hanya oleh satu sel (uniseluler) atau banyak sel (multiseluler). Percabangan *thallus* ada yang *thallus dichotomus* (dua-dua terus menerus), *pinate* (dua-dua berlawanan sepanjang *thallus* utama), *pectinate* (berderet searah pada satu sisi *thallus* utama) dan ada juga yang sederhana tidak bercabang. Sifat substansi *thallus* juga beraneka ragam ada yang lunak seperti gelatin (*gelatinous*), keras diliputi atau mengandung zat kapur (*calcareous*), lunak bagaikan tulang rawan (*cartilagenous*), berserabut (*spongy*) dan sebagainya.

2.2 Sargassum

Ekstraksi alginat dari rumput laut *Sargassum* dan *Turbinaria* telah dicoba untuk industri tekstil dalam proses pewarnaan yang memberikan hasil dapat memenuhi standart yang di perlukan untuk proses pewarnaan tekstil (Anggaeni et al.1997). tampaknya masih banyak kendala dari kalangan indudtri untuk dapat menggunakan produksi yang sudah dihasilkan dari proses ekstraksi rumput laut *Sargassum* hasil dari panen alami yang merupakan satu jenis rumput laut coklat alginofit di Indonesia, antara lain kualitas yang dihasilkan belum stabil

dan masih di bawah standarta yang dikualifikasikan untuk keperluan industri tertentu (cat, tekstil, kosmetik).



Gambar 1. Sargassum

2.3 Gracilaria

Gracilaria yang dalam literatur tumbuhan obat China disebut Long Xu Cai memiliki rasa asin, manis, bersifat dingin dan tidak beracun. Khasiat dari *Gracilaria* antara lain menurunkan demam, panas dalam, peluruh kencing, peluruh dahak, pencahar dalam proses pencernaan, merawat lambung, anti disentri, pelembut, peluruh gumpalan keras atau benjolan terutama bagian leher, anti tumor dll.



Gambar 2. Gracilaria

2.4 Eucheuma Cottoni

Eucheuma cottonii merupakan spesies rumput laut yang banyak dibudidayakan di perairan Indonesia. Hal tersebut dikarenakan manfaat pikokoloidnya yang besar yaitu karaginan dan agar serta teknik budidayanya yang relatif mudah dan murah. *Eucheuma cottonii* merupakan rumput laut merah (*Rhodophyta*) yang kaya akan pigmen fotosintesis dan pigmen aksesoris lainnya, yaitu klorofil a, α -karoten, β -karoten, fikobilin, neozantin dan zeaxanthin.



Gambar 3. Eucheuma Cottoni

2.5 Citra Digital

Citra (image) adalah gambar pada dua dimensi. Citra digital merupakan suatu matrik yang terdiri dari baris dan kolom, dimana setiap indeks dari matrik tersebut menyatakan suatu titik pada suatu citra. Nilai dalam matrik menyatakan tingkat kecerahan titik tersebut. Titik-titik dari citra dinamakan sebagai elemen citra atau disebut sebagai pixel (picture element). Citra juga dapat didefinisikan sebagai suatu representasi, kemiripan,

atau imitasi dari suatu objek atau benda yang didefinisikan sebagai fungsi dua dimensi $f(x, y)$, dimana x dan y adalah nilai koordinat, dan amplitudo f dari pasangan koordinat (x, y) disebut dengan nilai intensitas atau gray level dari citra pada titik tersebut (Gonzales, 2002).

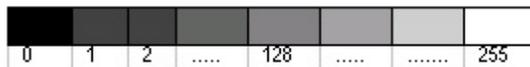
2.6 Citra Greyscale

Citra *greyscale* adalah citra berwarna keabu-abuan dengan memiliki variasi warna yaitu 8 bit ($2^8 = 256$) kemungkinan nilai.

Format citra ini disebut dengan skala keabuan karena pada umumnya warna yang digunakan adalah antara warna hitam dan putih dimana, hitam sebagai warna minimal dan putih sebagai warna maksimalnya, sehingga warna diantaranya adalah abu-abu. Untuk mendapatkan nilai keabuan digunakan sistem penghitungan tersendiri dengan mengambil nilai warna RGB citra awal, yaitu :

$$L = R * 0.114 + G * 0.587 + B * 0.299 \quad 2.1$$

Untuk format citra keabuan nilai masing-masing RGB-nya sama, sehingga nilai RGB yang baru berturut-turut adalah (L,L,L).



Gambar 4. Komposisi Warna Greyscale

2.7 PCA (Principal Component Analysis)

Pengenalan pola (*pattern recognition*) merupakan teknik yang bertujuan untuk mengklasifikasikan citra yang telah diolah sebelumnya. berdasarkan kesamaan atau kemiripan ciri yang dimilikinya.

Berdasarkan pendekatan yang digunakan, metode pengenalan pola dibagi menjadi tiga kelompok, yaitu:

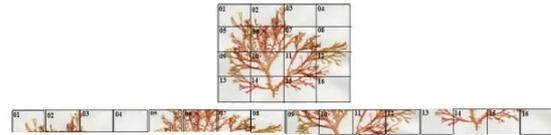
1. Statistik (*Statistical*)
2. Sintatik (*Syntatic*)
3. Jaringan Syaraf Tiruan (*Artificial Neural Network*)

PCA termasuk dalam kelompok bidang statistika yang tujuannya adalah untuk mengekstraksi ciri dari data yang akan dianalisa dan mentransformasikannya pada ruang yang baru.

2.8 Pengenalan CITRA MENGGUNAKAN METODE PCA

Dalam penelitian pada tugas akhir ini PCA juga digunakan untuk mengenali citra dari klasifikasi rumput laut.

Sebuah klasifikasi rumput laut, yang juga merupakan sebuah gambar, dapat dilihat sebagai sebuah vektor. Jika panjang dan lebar dari gambar tersebut adalah w dan h *pixel* maka jumlah komponen dari vektor ini adalah $w * h$. Setiap *pixel* dikodekan oleh satu komponen vektor.



Gambar 5. Representasi ruang vektor gambar

Principal Components Analysis (PCA) digunakan untuk mereduksi dimensi dari sekumpulan atau ruang gambar sehingga basis atau sistem koordinat yang baru dapat menggambarkan model yang khas dari kumpulan tersebut dengan lebih baik. Dalam hal ini model yang diinginkan merupakan sekumpulan gambar klasifikasi rumput laut. Basis yang baru tersebut akan dibentuk melalui kombinasi linear.

2.9 Euclidean Distance

Metode *Euclidean* yaitu metode klasifikasi tetangga terdekatnya dengan menghitung jarak antara dua buah obyek, metode ini disebut juga jarak *Euclidean* (Pablo Navarrete and Javier Ruizdel-Solar, 2003: 6-7). Rumus penghitungan jarak ditulis sebagai berikut:

$$d_e = \sqrt{\sum_{k=1}^m (fd_{i,k} - k_j)^2} \quad 2.12$$

Keterangan :

- d_e : jarak euclidean
- $.fd_i$: bobot wajah pelatihan
- k_j : data bobot wajah test
- m : jumlah data pelatihan

$$d_e = \sqrt{\left(\begin{pmatrix} fd_{11} & fd_{12} & \dots & fd_{1m} \\ \dots & \dots & \dots & \dots \\ fd_{m1} & fd_{m2} & \dots & fd_{mm} \end{pmatrix} - \begin{pmatrix} k_1 & k_1 & \dots & k_1 \\ k_2 & k_2 & \dots & k_2 \\ \dots & \dots & \dots & \dots \\ k_m & k_m & \dots & k_m \end{pmatrix} \right)^2} \quad 2.13$$

$$d_i = \sqrt{(fd_{1i} - k_1)^2 + (fd_{2i} - k_2)^2 + \dots + (fd_{mi} - k_m)^2} \quad 2.14$$

dimana $i = 1, 2, 3, \dots, M$.

$$d_{ei} = [d_1, d_2, \dots, \epsilon] \quad 2.15$$

Keterangan :

- d_e : jarak euclidean
- $.fd_i$: bobot wajah pelatihan
- k_j : data bobot wajah test
- m : jumlah data pelatihan

Dari bobot *Euclidean* (d_e) diatas yang mempunyai bobot paling kecil akan digunakan sebagai nilai fitness dari sebuah citra rumput laut yang digunakan untuk pengenalan.

3. HASIL DAN PEMBAHASAN

Dalam pengujian sistem klasifikasi rumput laut di lakukan pada data *learning* 10,20,30,40 data *testing* 10,20,30,40 dan nilai *threshold* 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1. Di bawah ini adalah tabel hasil pengujian terhadap citra rumput laut dengan ukuran gambar 100x100 *pixel*. Ujicoba dilakukan pada 3 kelas dan 2 kelas untuk ujicoba di luar kelas. Data yang digunakan untuk ujicoba secara urut, tidak random.

Tabel 1. Hasil pengujian Aplikasi dengan dataset 10 data testing 40

Dimensi	Dataset	Threshold	Kelas A	Kelas B	Kelas C	Jumlah	Benar	Salah	Akurasi
30	10	0.1	40	40	40	120	5	115	4.16
	10	0.2	40	40	40	120	10	110	9.164
	10	0.3	40	40	40	120	18	102	15
	10	0.4	40	40	40	120	20	100	16.6
	10	0.5	40	40	40	120	25	95	20.83
	10	0.6	40	40	40	120	30	90	25
	10	0.7	40	40	40	120	33	87	27.5
	10	0.8	40	40	40	120	34	86	28.3
	10	0.9	40	40	40	120	37	83	30.8
	10	1	40	40	40	120	40	80	33.3

Dari tabel di atas dapat dilihat pada saat dataset 10 dan data testing 40, akurasi paling tinggi pada nilai *threshold* 1 dengan akurasi sebesar 33,3%.

Tabel 2. Hasil pengujian Aplikasi dengan dataset 20 data testing 30

Dimensi	Dataset	Threshold	Kelas A	Kelas B	Kelas C	Jumlah	Salah	Benar	Akurasi
60	20	0.1	30	30	30	90	39	51	56.6
	20	0.2	30	30	30	90	38	52	57.7
	20	0.3	30	30	30	90	35	55	61.1
	20	0.4	30	30	30	90	30	60	66.6
	20	0.5	30	30	30	90	27	63	70
	20	0.6	30	30	30	90	27	63	70
	20	0.7	30	30	30	90	25	65	72.2
	20	0.8	30	30	30	90	23	67	74.4
	20	0.9	30	30	30	90	20	70	77.7
	20	1	30	30	30	90	20	70	77.7

Dari tabel di atas dapat dilihat pada saat dataset 20 dan data testing 30 dan dimensi 60, akurasi paling tinggi pada nilai *threshold* 1 dengan akurasi sebesar 77.7%.

Tabel 3. Hasil pengujian Aplikasi dengan dataset 30 data testing 20

Dimensi	Dataset	Threshold	Kelas A	Kelas B	Kelas C	Jumlah	Salah	Benar	Akurasi
90	30	0.1	20	20	20	60	17	43	71.6
	30	0.2	20	20	20	60	16	44	73.3
	30	0.3	20	20	20	60	16	44	73.3
	30	0.4	20	20	20	60	14	46	76.6
	30	0.5	20	20	20	60	13	47	78.3
	30	0.6	20	20	20	60	12	48	80
	30	0.7	20	20	20	60	11	49	81.6
	30	0.8	20	20	20	60	11	49	81.6
	30	0.9	20	20	20	60	11	49	81.6
	30	1	20	20	20	60	11	49	81.6

Dari tabel di atas dapat dilihat pada saat dataset 30 dan data testing 20 dan dimensi 90, akurasi paling tinggi pada nilai *threshold* 1 dengan akurasi sebesar 81.6%.

Tabel 4. Hasil pengujian Aplikasi dengan dataset 30 data testing 20

Dimensi	Dataset	threshold	Kelas A	kelas B	kelas C	jumlah	salah	benar	akurasi
120	40	0.1	10	10	10	30	6	24	80
	40	0.2	10	10	10	30	5	25	83.3
	40	0.3	10	10	10	30	4	26	86.6
	40	0.4	10	10	10	30	3	27	90
	40	0.5	10	10	10	30	3	27	90
	40	0.6	10	10	10	30	3	27	90
	40	0.7	10	10	10	30	2	28	93.3
	40	0.8	10	10	10	30	2	28	93.3
	40	0.9	10	10	10	30	1	29	96.6
	40	1	10	10	10	30	1	29	96.6

Dari tabel di atas dapat dilihat pada saat dataset 40 dan data testing 10 dan dimensi 120, akurasi paling tinggi pada nilai *threshold* 1 dengan akurasi sebesar 96.6%.

Dari gambar dibawah ini dapat lihat bahwa ujicoba 3 kelas pada dataset 40 dan data testing 10 dan menggunakan dimensi 120 didapatkan nilai akurasinya yang paling tinggi, yaitu 96.6 %. Dimana pada kelas B (*Gracilaria*) mengenali 100% dan pada kelas C (*Euclidean Cottonii*) ada 1 kesalahan yang kemudian di kenali sebagai kelas A (*Sargassum*).



Gambar 6. Ujicoba dataset 40 nilai *threshold* 1

Pada ujicoba dengan dataset 40 dan data testing terjadi hanya ada 1 kesalahan pengenalan yaitu pada kelas A (*Sargassum*) yang dikenali ke dalam kelas C (*Euclidean Cottonii*). Ini terjadi kemungkinan dari data citra rumput laut A (*Sargassum*) hampir sama dengan citra rumput laut C (*Euclidean Cottonii*). Sehingga citra rumput laut A (*Sargassum*) dikenali ke dalam kelas C (*Euclidean Cottonii*).

Ujicoba pada dataset 30 dan 20 data testing dan dimensi 90 dengan nilai *threshold* 1 ditunjukkan pada gambar di bawah ini.



Gambar 7. Ujicoba dataset 30 nilai threshold 1

Pada ujicoba diatas dengan nilai threshold 1 dataset 30 dan data testing 20 dimensi 90 didapatkan pada kelas C (*Eucheuma Cottonii*) ada 1 yang dikenali ke dalam kelas B (*Gracilaria*) dan 10 data dikenali kelas C (*Eucheuma Cottoni*). Dari ujicoba itu didapatkan kebenaran ujicoba 49 dan 11 kesalahan identifikasi.

4. KESIMPULAN

Setelah menyelesaikan perancangan dan pembuatan sistem atau aplikasi *Pengenalan Rumput Laut menggunakan Euclidean Distance berbasis Ekstaksi Fitur* serta melakukan uji coba dan evaluasi, maka dapat ditarik kesimpulan sebagai berikut :

- Pada sistem pengenalan yang berbasis pada *Principal Component Analysis* (PCA) dan *Euclidean Distance* dapat digunakan untuk mentransformasi dimensi gambar dataset citra rumput laut dengan akurasi terbaik 96.6 %. Ujicoba dilakukan pada 3 kelas yaitu kelas A (*Sargassum*), B (*Gracilaria*) dan C (*Eucheuma Cottonii*).
- Nilai *threshold* saat proses pengenalan citra rumput laut sangat berpengaruh pada tingkat akurasinya. Dataset terbaik pada dataset 40 dan data testing 10 dengan nilai dimensi 120.
- Semakin banyak data yang digunakan untuk pelatihan, maka semakin tinggi pula tingkat akurasinya.
- Penentuan nilai *Threshold* juga sangat mempengaruhi proses pencocokan/pengenalan karena nilai *threshold* digunakan sebagai batas jarak *Euclidean* dalam penentuan identitas citra rumput laut.

PUSTAKA

- Wahyu, Eko Wibowo. *Pengenalan Wajah Menggunakan Jaringan Syaraf Tiruan Quick Propagation Dengan Pemrosesan Awal Principal Components Analysis*. Tugas

Akhir. Bangkalan : Universitas Trunojoyo Madura. 2009.

- Kurniawan, Dedy. *Perbandingan Metode Principal Component Analysis (PCA) Dengan Fast Fourier Transform (FFT) Untuk Pengenalan Pola Suara Chord Nada Dasar Gitar*. Tugas Akhir. Bangkalan: Universitas Trunojoyo Madura. 2009.
- MS, Ir.Sulistijo. *Penelitian Budidaya Rumput Laut (Algae Makro/ Seaweed) di Indonesia*. Jakarta: Pusat Penelitian Oseanografi. 2002.
- Yusuf. 2005. *Budidaya Rumput Laut di Wilayah Pesisir*. <URL:http://www.damandiri.or.id/file/yusuf_kamlasiipbbab2.pdf>. Diakses pada 12 Januari 2010
- Sulaeman, Dr.Ir. Suhendar. *Pengembangan agribisnis komoditi rumput laut Melalui model klaster bisnis*.< http://www.smecda.com/deputi7/file_Infokop/EDISI%2028/komoditi_rumput_laut.pdf>.Diakses pada 13 Januari 2010.
- Gonzalez, R.C. and Woods, R.E. *Digital Image Processing, Prentice-Hall,inc*.New Jersey.2002.
- Kusumadewi, S. *Artificial Intelligence: Teknik dan Aplikasinya*, Yogyakarta: Graha Ilmu. 2003.
- Putra, D. *Sistem Biometrika Konsep Dasar, Teknik analisis Citra, dan Tahapan Membangun Aplikasi Sistem Biometrika*. Yogyakarta : Andi. 2009.
- Achmad, Ir. Balza, *Teknik Pengolahan Citra Digital menggunakan DELPHI*. Yogyakarta: Ardi Publishing. 2005.
- Shlens ,Jonathon, A *Tutorial on Principal Component Analysis*, San Diego University of California. 2005
- Agushinta, Dewi dan Karmilasari."Mosaik Citra Wajah Panramik untuk sistem Pengenalan Wajah".23-24 Agustus.Depok:Universitas Gunadarma.2006
- PhD, Obuchowski, Nancy A."Receiver Operating Characteristic Curves and Their Use in Radiology". *Diagnostic radiology Statistical analysis*.228:3-8.2003.

- [13] Gönen, Mithat."Receiver Operating Characteristic (ROC) Curves".SUGI 31 Proceedings.26-29 Maret.San Francisco.2006
- [14] JR, Tanner P. Wilson, Theodore E. Cohn, Daniel G Green, "Receiver Operating Characteristic Analysis". The Journal of General Physiology 66, 583-614.1975.
- [15] Purnomo, M.H, Arif Muntasa. "Konsep Pengolahan Citra Digital dan Ekstraksi Fitur". Yogyakarta: 2010.