

PERBAIKAN SISTEM PENGENAL JENIS POHON MANGGA MENGUNAKAN SVM DAN FK-NNC

Eko Prasetyo

*Program Studi Teknik Informatika, Fakultas Teknik Universitas Bhayangkara Surabaya
E-mail : eko1979@yahoo.com*

Abstrak. *Pohon mangga kini menjadi primadona masyarakat untuk ditanam dipekarangan rumah. Selain buah yang manis rasanya, pohonnya sendiri merupakan aset penghijauan alam. Seringkali masyarakat kecewa saat pohon mangganya berbuah ketika tahu bahwa jenis pohon mangga yang ditanamnya tidak sesuai dengan yang diharapkan saat menanamnya. Hal ini bisa dimaklumi mengingat pohon mangga cangkokan membutuhkan waktu yang cukup lama untuk tumbuh sebelum berbuah. Maka akan lebih baik jika bisa diketahui sejak awal jenis pohon mangga tersebut berdasarkan komponen pohon yang mudah diamati yaitu tekstur daun. Penelitian sebelumnya menggunakan pendekatan K-NN (K Nearest Neighbor) dan JST (Jaringan Syaraf Tiruan) Backpropagation untuk pekerjaan klasifikasi jenis pohon mangga gadung dan curut. Ternyata pendekatan tersebut memberikan akurasi prediksi sampai dengan 65.19%. Fitur yang diekstrak untuk diolah adalah : rata-rata intensitas, smoothness, entropy, 5 moment invariant, energy, dan kontras. Dalam penelitian ini, digunakan pendekatan SVM (Support Vector Machine) dengan parameter kernel Radial Basis Function (RBF) dan Fuzzy K-Nearest Neighbor in every Class (FK-NNC), akurasi prediksi yang didapatkan 86.67% untuk SVM, dan 88.89% untuk FK-NNC. Diharapkan dengan akurasi yang lebih tinggi maka sistem dapat memberikan penilaian terhadap jenis mangga secara tepat.*

Kata kunci: *klasifikasi, daun mangga, gadung, curut, support vector machine, fuzzy k-nearest neighbor in every class.*

Mangga termasuk ke dalam marga Mangifera, yang terdiri dari 35-40 anggota, dan suku Anacardiaceae [1]. Nama ilmiahnya adalah Mangifera indica. Pohon mangga termasuk tumbuhan tingkat tinggi yang struktur batangnya (habitus) termasuk kelompok arboreus, yaitu tumbuhan berkayu yang mempunyai tinggi batang lebih dari 5 m, bahkan mencapai tinggi 10-40 m.

Salah satu bagian penting dalam sistem tumbuhan adalah daun. Tumbuhan dapat mempunyai jumlah daun yang banyak. Daun hanya terdapat pada batang saja dan tidak pernah terdapat pada bagian lain pada tanaman. Bagian batang tempat duduknya atau melekatnya daun disebut buku-buku (*nodus*) batang dan tempat di atas daun yang merupakan sudut antara batang dan ketiak (*axilla*). Daun biasanya berwarna hijau yang disebut klorofil [2].

Pohon mangga kini menjadi primadona masyarakat untuk ditanam dipekarangan rumah. Selain buah yang manis rasanya, pohonnya

sendiri merupakan aset penghijauan alam. Seringkali masyarakat kecewa saat pohon mangganya berbuah ketika tahu bahwa jenis pohon mangga yang ditanamnya tidak sesuai dengan yang diharapkan saat menanamnya. Hal ini bisa dimaklumi mengingat pohon mangga cangkokan membutuhkan waktu yang cukup lama untuk tumbuh sebelum berbuah. Maka akan lebih baik jika bisa diketahui sejak awal jenis pohon mangga tersebut berdasarkan komponen pohon yang mudah diamati yaitu tekstur daun.

Atas dasar masalah tersebut, penulis melakukan penelitian pengenalan jenis pohon mangga berdasarkan tekstur daun. Dengan sistem yang dikembangkan ini, diharapkan dapat membantu masyarakat dalam mengenali jenis mangga yang akan ditanamnya hanya berdasarkan tektur daunnya saja, sehingga tidak merasa tertipu atau kecewa pada jenis mangga yang ditanamnya.

Sistematika dalam paper ini terbagi menjadi 6 bagian : bagian 1 memuat pendahuluan,

bagian 2 membahas penelitian sebelumnya yang dilakukan terhadap daun, bagian 3 membahas teori yang melandasi dalam melakukan penelitian, bagian 4 adalah desain sistem yang digunakan, bagian 5 membahas skenario pengujian, dan analisis hasil pengujian, sedangkan bagian 6 membahas kesimpulan yang didapat dari penelitian.

Penelitian sebelumnya

Proposal penelitian yang diajukan Valerina et al [3] adalah sistem untuk mengidentifikasi jenis daun tanaman obat dengan basis ciri morfologi, bentuk, dan tekstur daun tanaman obat. Meskipun masih dalam tahap pengajuan, tetapi target akurasi pengenalan yang diharapkan adalah mendekati 100%. Belum ditemukan hasil yang dipublikasikan, sehingga penulis mencoba untuk menerapkan hal yang serupa pada jenis tanaman mangga dengan berfokus pada tekstur daun. Dari proposal penelitian Valerina et al [3] maka dapat dikatakan bahwa identifikasi jenis pohon mangga dengan klasifikasi berdasarkan tekstur daun memungkinkan untuk dilakukan.

Agustin dan Prasetyo [4] melakukan penelitian untuk mengenali jenis pohon mangga gadung dan curut berdasarkan tekstur daun. Komponen warna daun dalam skema RGB yang digunakan adalah komponen warna green dengan asumsi bahwa daun berwarna hijau, sehingga mengabaikan warna apapun selain warna hijau. Analisis citra yang digunakan untuk mengekstrak fitur adalah pendekatan statistik, moment invariant, dan matrik co-occurrence. Sedangkan metode klasifikasi yang digunakan adalah K-Nearest Neighbor (K-NN) dan Jaringan Syaraf Tiruan (JST) Backpropagation, hasil kedua metode tersebut dibandingkan. Hasilnya, dengan menggunakan K-NN, akurasi tertinggi yang didapat saat uji coba adalah 52.24%, sedangkan JST Backpropagation 65.19%. Hasil ini menunjukkan bahwa sistem klasifikasi masih belum memberikan hasil yang optimal saat proses prediksi.

Karakter daun mangga

Daun pohon mangga umumnya tunggal, dengan letak tersebar, tanpa daun penumpu [5][6]. Panjang tangkai daun bervariasi dari 1,25-12,5 cm, bagian pangkalnya membesar

dan pada sisi sebelah atas ada alurnya. Aturan letak daun pada batang biasanya 3/8, tetapi makin mendekati ujung, letaknya makin berdekatan sehingga nampaknya seperti dalam lingkaran (roset). Helai daun bervariasi namun kebanyakan berbentuk jorong sampai lanset, 2-10 × 8-40 cm, agak liat seperti kulit, hijau tua berkilap, berpangkal melancip dengan tepi daun bergelombang dan ujung meluncip, dengan 12-30 tulang daun sekunder [6]. Beberapa variasi bentuk daun mangga [5]:

1. Lonjong dan ujungnya seperti mata tombak.
2. Berbentuk bulat telur, ujungnya runcing seperti mata tombak.
3. Berbentuk segi empat, tetapi ujungnya runcing.
4. Berbentuk segi empat, ujungnya membulat.

Daun yang masih muda biasanya berwarna kemerahan, keunguan atau kekuningan [5]; yang di kemudian hari akan berubah pada bagian permukaan sebelah atas menjadi hijau mengkilat, sedangkan bagian permukaan bawah berwarna hijau muda. Umur daun bisa mencapai 1 tahun atau lebih.

ANALISIS TEKSTUR CITRA

Pendekatan Statistik

Pendekatan ini sering digunakan untuk analisis tekstur yang didasarkan pada properti statistik histogram intensitas [7]. Satu kelas pengukuran didasarkan pada moment statistik. Untuk menghitung moment n th terhadap mean diberikan oleh:

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i) \quad (1)$$

di mana z_i adalah variabel random yang mengindikasikan intensitas, $p(z)$ adalah histogram level intensitas dalam region, L adalah jumlah level intensitas yang tersedia, mean (rata-rata) intensitas dihitung dengan formula:

$$m = \sum_{i=0}^{L-1} z_i p(z_i) \quad (2)$$

Ukuran smoothness relatif dari intensitas dalam region. R bernilai 0 untuk region dalam intensitas konstan dan mendekati 1 untuk

region dengan ekskursi yang besar dalam nilai level intensitas. Dalam prakteknya, varian digunakan dalam ukuran ini yang dinormalisasikan dalam range [0,1] oleh pembagian dengan $(L-1)^2$. Smoothness dihitung dengan formula :

$$R = 1 - 1 / (1 + \sigma^2) \tag{3}$$

Dimana σ adalah ukuran standart deviasi yang diukur dengan formula:

$$\sigma = \sqrt{\mu_2(z)} \tag{4}$$

Entropy digunakan untuk mengukur keacakan nilai intensitas citra, dihitung dengan formula:

$$e = \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) \tag{5}$$

Moment Invariants

Moment 2-D dari order (p + q) pada citra digital $f(x,y)$ didefinisikan sebagai [7]:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \tag{6}$$

Untuk p,q = 0, 1, 2, ..., di mana penjumlahan lebih dari nilai koordinat spasial x dan y yang merentangkan citra. *Central moment* yang berhubungan didefinisikan sebagai:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \tag{7}$$

di mana

$$\bar{x} = \frac{m_{10}}{m_{00}} \tag{8}$$

$$\text{dan } \bar{y} = \frac{m_{01}}{m_{00}} \tag{9}$$

Normalized central moment dari order (p + q) didefinisikan sebagai:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \tag{10}$$

Untuk p, q = 0, 1, 2, ..., di mana:

$$\gamma = \frac{p + q}{2} + 1 \tag{11}$$

Untuk p+q = 2, 3, ...

Sejumlah tujuh moment invariant yang tidak sensitif terhadap translasi, perubahan

skala, pencerminan, dan rotasi dapat diturunkan dari persamaan berikut:

$$\phi_1 = \eta_{20} + \eta_{02} \tag{12}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{13}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{14}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \tag{15}$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} + \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{16}$$

$$\phi_6 = (\eta_{20} + \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \tag{17}$$

$$\phi_7 = (3\eta_{21} + \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} + \eta_{30})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{17}$$

Matrik Co-occurrence

Matriks intensitas co-occurrence adalah suatu matriks yang menggambarkan frekuensi munculnya pasangan dua piksel dengan intensitas tertentu dalam jarak dan arah tertentu dalam citra [7][8]. Matriks intensitas co-occurrence $p(i_1, i_2)$ didefinisikan dengan dua langkah sederhana sebagai berikut. Langkah pertama adalah menentukan lebih dulu jarak antara dua titik dalam arah vertikal dan horizontal (vektor $d=(dx,dy)$), di mana besaran dx dan dy dinyatakan dalam piksel sebagai unit terkecil dalam citra digital. Langkah kedua adalah menghitung pasangan piksel-piksel yang mempunyai nilai intensitas i_1 dan i_2 dan berjarak di piksel dalam citra. Kemudian hasil setiap pasangan nilai intensitas diletakkan pada matriks sesuai dengan koordinatnya, di mana absis untuk nilai intensitas i_1 dan ordinat untuk nilai intensitas i_2 .

Energi dalam matrik co-occurrence yaitu fitur untuk mengukur konsentrasi pasangan intensitas pada matriks co-occurrence, dan didefinisikan dengan:

$$Energi = \sum_{i_1} \sum_{i_2} p^2(i_1, i_2) \tag{18}$$

Kontras yang digunakan untuk mengukur kekuatan perbedaan intensitas dalam citra dan dinyatakan dengan:

$$\text{Kontras} = \sum_{i_1} \sum_{i_2} (i_1 - i_2)^2 p(i_1, i_2) \quad (19)$$

DESAIN SISTEM

Perbaikan sistem klasifikasi jenis pohon mangga berdasarkan tekstur daun ini menggunakan sistem kerja seperti pada gambar 1, sebagai berikut :

1. Preprocessing

Pada bagian ini dilakukan pekerjaan awal sebelum pemrosesan citra lebih lanjut, seperti: cropping, resizing, dan pengurangan noise pada citra.

2. Segmentasi

Segmentasi menggunakan metode K-means clustering pada spektrum warna RGB dengan jumlah k mulai 3 sampai 4 sesuai dengan kondisi terbaik hasil segmentasi yang diberikan.

3. Ekstraksi fitur

Pada bagian ini, dilakukan pengambilan komponen warna hijau pada citra daun yang sudah disegmentasi. Kemudian mengekstrak fitur yaitu: rata-rata intensitas, smoothness, dan entropy dari pendekatan statistik; 5 dari 7 moment invariants; energy dan kontras dari pendekatan matrik co-occurrence.

4. Pemisahan data

Ada 60 data citra, dibagi menjadi 2 kelompok: data training dan data uji, komposisi yang digunakan menyesuaikan hasil analisis dengan teknik K-fold Cross Validation yaitu nilai K-fold bernilai 5. Jadi ada 48 citra latih (24 citra daun mangga gadung, 24 citra daun mangga curut) dan 12 citra uji (6 citra daun mangga gadung, 6 citra daun mangga curut).

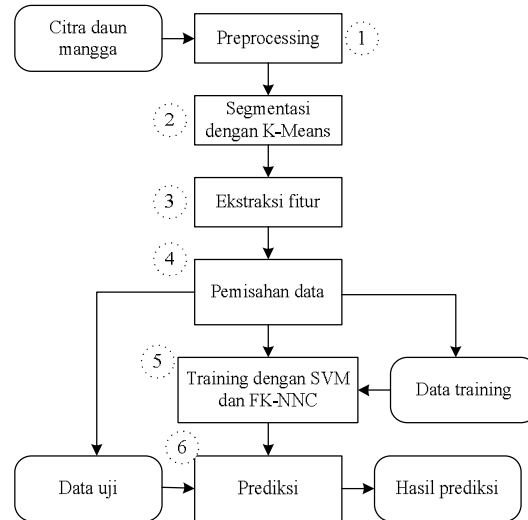
5. Training dengan SVM dan FK-NNC

Training dengan SVM [9] dan FK-NNC [10] dilakukan pada data training dengan label kelas yang sudah diberikan pada setiap data training.

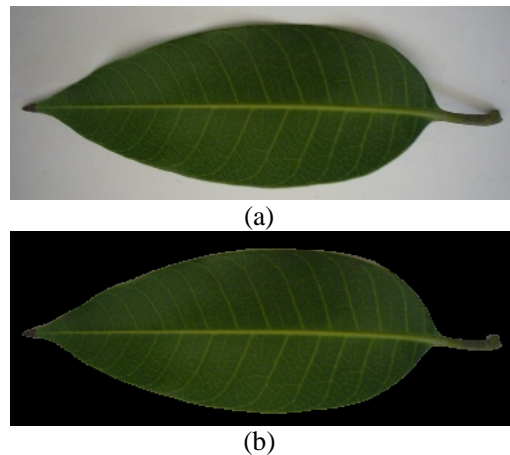
6. Klasifikasi

Proses klasifikasi dilakukan dengan memproses satu persatu data uji untuk diketahui keluaran kelas yang diberikan oleh sistem. Pada FK-NNC, masing-masing data uji dilakukan pengujian 3 kali yaitu: 3-NN, 5-NN, dan 7-NN, sedangkan SVM (Support Vector Machine) dengan parameter kernel RBF (Radial Basis

Function) dilakukan pengujian 3 kali percobaan. Kemudian hasilnya dilakukan pencocokan dengan kelas yang sesungguhnya sehingga diketahui akurasi sistem dalam melakukan klasifikasi.



Gambar 1. Desain perbaikan sistem klasifikasi jenis pohon mangga



Gambar 2. Citra daun mangga gadung; (a) belum di segmentasi; (b) sudah di segmentasi

PENGUJIAN DAN ANALISIS

Penelitian ini bertujuan untuk memperbaiki kinerja akurasi sistem pengenalan jenis mangga berdasarkan tekstur daun, dengan mengimplementasikan penggunaan SVM dan FK-NNC dalam proses klasifikasi sistem pengenalan jenis mangga. Maka, untuk fitur yang diekstrak dari komponen warna hijau dari citra daunnya juga menggunakan fitur yang sama, yaitu: rata-rata, smoothness, dan entropy dari

pendekatan statistik, 5 moment dan 7 moment invariant, sedangkan dari pendekatan matrik co-occurrence menggunakan fitur energy dan kontras.

Citra daun mangga yang diproses lebih lanjut dalam sistem adalah yang sudah melalui preprocessing dan sudah di segmentasi untuk memisahkan obyek daun dari latar belakang dalam citra, contoh daun yang belum dan sudah disegmentasi ditunjukkan pada gambar 2.

Skenario pengujian yang dilakukan adalah dengan menggunakan data training sebagai pelatihan terhadap sistem untuk mendapatkan model sebagai *black box* prediksi, kemudian menggunakan data uji untuk melakukan prediksi. Hasil prediksi data uji oleh sistem dibandingkan dengan klasifikasi data uji yang sesungguhnya (manual), dan hasil prediksi pada penelitian sebelumnya untuk diketahui tingkat perbaikannya.

Tabel 1. Fitur 12 data uji citra mangga gadung dan curut untuk K-fold = 5

DATA		FITUR									
UJI	RTA	SMO	ENT	M1	M2	M4	M6	M7	ENE	KON	
1	29.1729	0.0155	3.4826	5.5427	11.6998	22.9311	29.0055	44.8459	0.2786	67.7212	
2	33.9193	0.0230	3.3328	5.7097	11.9975	21.7963	28.5323	44.7507	0.3116	84.3237	
3	29.0943	0.0156	3.5126	5.5885	11.8344	24.5504	30.7731	48.3301	0.2784	69.3236	
4	29.0359	0.0148	3.4111	5.5716	11.8278	24.3760	30.9883	47.6693	0.2721	78.6404	
5	30.0713	0.0158	3.5678	5.5993	11.8586	24.2964	30.4569	48.0587	0.2659	69.6455	
6	30.6568	0.0159	3.5851	5.6038	11.8700	23.9468	30.0445	47.6500	0.2584	68.5670	
7	35.2274	0.0153	4.2650	5.4728	11.5021	23.2160	29.9284	45.8194	0.1738	83.7451	
8	34.4043	0.0144	4.2666	5.4251	11.3855	23.2463	29.6159	46.1905	0.1704	86.0139	
9	27.2064	0.0103	3.7283	5.3438	11.2981	22.0147	29.7404	43.1424	0.2050	68.9388	
10	26.4059	0.0102	3.6640	5.3492	11.3047	21.3572	27.5919	42.5218	0.2204	57.0943	
11	27.9246	0.0100	3.9025	5.3477	11.3233	21.2033	27.3976	42.3380	0.1850	63.5500	
12	30.2173	0.0111	4.0013	5.4143	11.4539	21.3308	27.4956	42.5375	0.1755	59.8338	

Keterangan kolom tabel 1:

RTA : Rata-rata intensitas ENT : Entropy KON : Kontras M2 : Moment 2 M6 : Moment 6
 SMO : Smoothness ENE : Energy M1 : Moment 1 M3 : Moment 4 M7 : Moment 7

Tabel 2. Hasil pengujian prediksi citra daun mangga gadung dan curut

METODE	PENGUJIAN	AKURASI	RATA-RATA AKURASI
SVM	Percobaan 1	80%	86.67%
	Percobaan 2	90%	
	Percobaan 3	90%	
FK-NNC	3-NN	91.67%	88.89%
	5-NN	83.33%	
	7-NN	91.67%	

Sebanyak 10 fitur dari masing-masing data training digunakan sebagai data training dalam SVM dan FK-NNC. Selanjutnya 10 fitur dari masing-masing data uji satu persatu dilakukan

prediksi data uji dengan model yang sudah dibangun saat pelatihan metode SVM dan FK-NNC, 10 fitur dari masing-masing 12 citra uji ditampilkan pada tabel 1.

Berdasarkan hasil penelitian sebelumnya [4], yang menyebutkan bahwa K-fold terbaik untuk pengujian tekstur daun mangga adalah 5, maka pengujian dalam penelitian ini digunakan K-fold bernilai 5 juga, yaitu 20% dari data set digunakan sebagai data uji, sedangkan 80% untuk data latih. Pada metode SVM, dilakukan percobaan pelatihan dan prediksi sebanyak 3 kali untuk diketahui akurasi kinerja sistem. Sedangkan FK-NNC dilakukan prediksi untuk 3-NN, 5-NN, dan 7-NN. Hasil pengujian pada saat proses prediksi disajikan pada tabel 2. Dari tabel tersebut dapat diamati bahwa hasil prediksi untuk SVM rata-rata 86.67%, sedangkan FK-NNC rata-rata 88.89%. Hasil

penelitian sebelumnya memberikan akurasi prediksi tertinggi 65.19%, artinya nilai akurasi yang didapat oleh SVM secara signifikan lebih tinggi 21.48%, sedangkan FK-NNC lebih tinggi 23.70%.

DAFTAR RUJUKAN

- [1] Wikipedia Indonesia – ensiklopedia bebas, *Mangga*, [online] (Updated 13 Juli 2011) Available at: <http://id.wikipedia.org/wiki/Mangga> [Accessed 5 Agustus 2011]
- [2] Tjitrosoepomo G. *Morfologi Tumbuhan*, Gajah Mada University Press: Yogyakarta, 1989.
- [3] Valerina, F., Ratu, D.A., Nuryunita, K. *Sistem Identifikasi Daun Tanaman Obat dengan Penggabungan Ciri Morfologi, Bentuk, dan Tesktur Menggunakan Probabilistic Neural Network pada Perangkat Mobile*, PM-GT, Institut Pertanian Bogor, Bogor, 2011.
- [4] Agustin, S. dan Prasetyo, E. “Klasifikasi Jenis Pohon Mangga Gadung dan Curut Berdasarkan Tesktur Daun” in *Seminar Nasional Sistem Informasi Indonesia*, Surabaya, 2011, pp.58-64.
- [5] Ahmad, U. *Pangan*, Vol. 19 No. 1, 2010, pp.71-80
- [6] Rukmana, R. *Mangga Budidaya dan Pasca panen*, Kanisius: Yogyakarta, 1997.
- [7] Gonzalez, R.C, Wood, R.E. *Digital Image Processing*, 3rd Edition, Pearson Prentice Hall: New Jersey, 2008.
- [8] Ahmad, U. *Pengolahan Citra Digital & Teknik Pemrogramannya*, Edisi 1, Graha Ilmu: Yogyakarta, 2005
- [9] Tan, P., Steinbach, M., Kumar, V. *Introduction to Data Mining*, 1st Ed, Pearson Education: Boston San Fransisco New York, 2006
- [10] Prasetyo, E. “Fuzzy K-Nearest Neighbor in Every Class untuk Klasifikasi Data” in *Seminar Nasional Teknik Informatika*, 2012, pp.57-60.