

# ANALISIS FITUR TEKSTUR DAUN MANGGA DENGAN FISHER'S DISCRIMINANT RATIO UNTUK PENCAPAIAN FITUR YANG INFORMATIF

Eko Prasetyo

Program Studi teknik Informatika Fakultas Teknik, Univ. Bhayangkara Surabaya  
email: eko@ubhara.ac.id

## ABSTRACT

*Mango tree species recognition system based on the texture of leaves on the previous system gives accuracy up to 88.89%. This indicates that feature selection of research on the mango tree species recognition need to be taken into account. In this research, analysis of mango leaf texture features are used. There are 3 types of features that is: statistics, invariant moment, and the co-occurrence matrix. Methods for analyzing the feature is Fisher's Discriminant Ration (FDR). This method obtained from a number of informative features, that is: energy (co-occurrence), uniformity (statistics), the third moment (statistics), entropy (co-occurrence), entropy (statistical), and homogeneity (co-occurrence). Performance testing is done by comparing the use of old and new features on the K-Nearest Neighbor method with the value K is 3, 5, 7, and 11. The results showed that the accuracy of the K-NN method with new features reach 0.90, and tend to be better than the old features.*

**Keywords:** features, analysis, texture, leaf mango, Fisher's Discriminant Ratio

## PENDAHULUAN

Mangga merupakan salah satu komoditas ekspor potensial Indonesia. Mangga juga berperan penting bagi kehidupan petani karena usaha tani mangga dapat meningkatkan kesejahteraan mereka. Produksi dan luas panen mangga di Indonesia meningkat sampai tahun 2009 dengan luas panen 215,387 ha, dan tahun

2013 menjadi 247,239 ha, dengan peningkatan hingga 12.55% dari tahun

2012 ke 2013 (Departemen Pertanian, 2015). Ini menunjukkan bahwa pertumbuhan pohon mangga terus meningkat di tengah menyempitnya lahan perkebunan yang tergerus oleh arus industrialisasi di berbagai sektor.

Bagi masyarakat awam yang menanam pohon mangga di pekarangan rumah, dengan berbagai jenis pohon yang ditanam, ternyata secara tidak langsung membantu pertumbuhan jumlah lahan pohon mangga di negara ini. Penanaman bibit saat ini umumnya dalam bentuk hasil cangkokan. Dengan masa tunggu 4-5 tahun sejak ditanam hingga berbuah, membuat banyak masyarakat menanamnya. Jenis mangga seperti gadung, mana lagi, golek, menjadi pilihan banyak masyarakat. Tetapi ketika berbuah, kemudian ternyata jenis buah mangga yang didapatkan berbeda dari yang diketahui di awal ketika menanam maka hal ini tentu mengecewakan.

Penelitian yang dilakukan oleh Agustin dan Prasetyo (2011) melakukan penelitian untuk menjawab masalah tersebut. Sistem yang dibuat adalah melakukan pengenalan jenis pohon mangga berdasarkan tekstur daun. Penelitian dilakukan dengan membandingkan penggunaan metode ANN *Error-Backpropagation* dan *K-Nearest Neighbor* (K-NN). Hasilnya, akurasi prediksi yang didapatkan pada metode K-NN adalah 54.24%, sedangkan ANN *Error-Backpropagation* 65.19%. Hasil ini tentu jauh dari harapan karena sistem belum dapat memberikan hasil prediksi secara baik. Prasetyo (2012) juga melakukan percobaan dengan membandingkan penerapan metode *Support Vector Machine* (SVM) dan *Fuzzy K-Nearest Neighbor in every Class* (FK-NNC) untuk mengetahui kinerja prediksinya, hasil yang didapat untuk SVM adalah 86.67%, sedangkan FK-NNC adalah 88.89%. Makalah ini memaparkan hasil penelitian tentang analisis terhadap fitur-fitur tekstur daun yang digunakan dalam melakukan klasifikasi. Pada penelitian sebelumnya fitur yang digunakan adalah: fitur statistik (rata-rata intensitas, smoothness, entropy; fitur moment invariant (moment 1, 2, 4, 6, dan 7; fitur matrik *co-occurrence* (energy, kontras). Penelitian yang dipaparkan dalam makalah ini dilakukan dengan menganalisis semua fitur dari statistik: rata-rata intensitas,

kontras, *smoothness*, *third moment*, *uniformity*, dan *entropy*. Dari *moment invariant* dilakukan analisis kembali terhadap 7 moment invariant. Sedangkan dari matrik *co-occurrence* dilakukan analisis terhadap: *entropy*, energi, kontras, dan homogenitas. Analisis dilakukan menggunakan metode *Fisher's Discriminant Ratio* untuk memilih fitur informatif secara terpisah ((Theodoridis dan Koutroumbas,

2009), (Prasetyo, 2014)), dilanjutkan dengan pengujian menggunakan metode klasifikasi K-NN dengan perbandingan

terhadap penggunaan fitur lama.

Pemaparan makalah ini dibagi menjadi 5 bagian. Bagian 1 menyajikan pendahuluan yang melatarbelakangi penulis melakukan penelitian. Bagian 2 menyajikan penelitian-penelitian terkait yang menjadi dasar bagi penulis untuk

melakukan penelitian. Bagian 3 menyajikan kerangka kerja penelitian dalam melakukan

analisa fitur tekstur daun mangga. Bagian 4

menyajikan pengujian dan analisis yang dilakukan untuk mendapatkan fitur

informatif beserta akurasi yang didapat. Dan bagian 5 menyajikan simpulan dari hasil penelitian dan saran untuk penelitian

berikutnya.

## 1. Penelitian Sebelumnya

Penelitian yang dilakukan Agustin dan Prasetyo (2011) adalah melakukan prediksi terhadap jenis pohon mangga belum berbuah berdasarkan tekstur daun. Penelitian tersebut memilih fitur statistik (rata-rata intensitas, *smoothness*, *entropy*; fitur moment invariant (moment 1, 2, 4, 6, dan 7; fitur matrik *co-occurrence* (energy, kontras). Sistem untuk implementasinya

menggunakan metode K-NN dan ANN Back-propagation. Hasil prediksi yang dilakukan

kurang memuaskan karena mendapatkan akurasi metode K-NN adalah 54.24%, sedangkan ANN Error-Backpropagation 65.19%. Penelitian yang dilakukan oleh Prasetyo (2012) melakukan percobaan dengan membandingkan penerapan

## 2. Analisis Tekstur Citra

### Pendekatan Statistik

Pendekatan yang sering digunakan untuk analisis tekstur didasarkan pada properti statistik histogram intensitas (Gonzales, 2008). 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:

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

*Smoothness* dihitung dengan formula:

$$R = 1 - 1 / (1 + \sigma^2) \quad (3)$$

Dimana  $\sigma$  adalah ukuran standart deviasi untuk mengukur kontras, nilainya didapatkan dengan formula:

$$\sigma = \mu_2(z) \quad (4)$$

Dimana  $\mu_2(z)$  adalah varian  $\sigma^2$ , yang

disebut juga momen dengan ordo 2.

Ukuran Third Moment (skewness/kecondongan) histogram, nilainya 0 untuk histogram yang simetris, positif untuk histogram yang condong ke kanan (terhadap mean) dan negatif untuk histogram yang condong ke kiri. Formula yang digunakan:

$$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i) \quad (5)$$

penggunaan metode Support Vector Machine (SVM) dan Fuzzy K-Nearest Neighbor in every Class (FK-NNC) dalam melakukan prediksi jenis pohon mangga. Hasil yang didapat untuk SVM adalah 86.67%, sedangkan FK-NNC adalah 88.89%.

Ukuran keseragaman 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) . . Ukuran ini  
maksimum ketika semua gray level sama  
(keseragaman maksimal) dan menurun dari  
sana. Formula yang digunakan:

$$U = \sum_{i=0}^{L-1} p^2(z_i) \quad (6)$$

$i=0$

2

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

$$e = \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) \quad (7)$$

**Moment Invariants**

Moment 2-D dari order (p + q) pada citra digital f(x,y) didefinisikan sebagai (Gonzales, 2008):

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (8)$$

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:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (9)$$

di mana  $\bar{x} = \frac{m_{10}}{m_{00}}$  (10)

dan  $\bar{y} = \frac{m_{01}}{m_{00}}$  (11)

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

$$\eta = \frac{\mu_{pq}}{m_{00}^{\frac{p+q}{2}}} \quad (12)$$

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

$$\gamma = \frac{p+q}{2} + 1 \quad (13)$$

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

Sejumlah tujuh moment invariant yang tidak sensitif terhadap translasi, perubahan skala, pencerminan, dan rotasi dapat diturunkan dari persamaan berikut:

$$\phi_1 = \eta_{20} + \eta_{02} \quad (14)$$

$$\phi_2 = (\eta_{20} - \eta_{02}) + 4\eta_{11} \quad (15)$$

$$\phi_3 = \frac{(\eta_{30} - 3\eta_{12})^2}{2} + \frac{(3\eta_{21} - \eta_{03})^2}{2} \quad (16)$$

$$\phi_4 = (\eta_{30} + \eta_{12}) + (\eta_{21} - \eta_{03}) \quad (17)$$

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

**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 (Ahmad, 2006). Matriks intensitas co-occurrence p(i<sub>1</sub>, i<sub>2</sub>)

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<sub>1</sub> dan i<sub>2</sub> 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<sub>1</sub> dan ordinat untuk nilai intensitas i<sub>2</sub>.

Fitur untuk mengukur keteracakan

dari distribusi intensitas disebut entropi dan didefinisikan sebagai berikut:

$$Entropi = - \sum_{i_1} \sum_{i_2} p(i_1, i_2) \log p(i_1, i_2) \quad (21)$$

Nilai entropi maksimum jika semua elemen p(i<sub>1</sub>, i<sub>2</sub>) sama, yaitu matriks yang

berhubungan dengan citra di mana tidak terdapat susunan tertentu dalam pasangan intensitas dengan jarak vektor tertentu (daerah konstan).

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

$$Energi = \sum_{i_1} \sum_{i_2} p(i_1, i_2) \quad (22)$$

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

$$Kontras = \sum_{i_1} \sum_{i_2} (i_1 - i_2)^2 p(i_1, i_2) \quad (23)$$

Kebalikan dari kontras adalah homogenitas, yaitu untuk mengukur kehomogenan variasi intensitas dalam

(18)

$$\phi_6 = (\eta_{20} + \eta_{02})(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2$$

$$+ 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (19)$$

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

$$- 3(\eta_{21} + \eta_{03})] + (3\eta_{12} + \eta_{30})(\eta_{21} + \eta_{03})$$

$$[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (20)$$

citra, dan didefinisikan dengan:

$$Homogenitas = \sum_{i_1} \sum_{i_2} \frac{p(i_1, i_2)}{1 + |i_1 - i_2|} \quad (24)$$

Nilai homogenitas membesar jika variasi intensitas dalam citra mengecil dan sebaliknya mengecil jika variasi intensitas dalam citra membesar.

### Fisher's Discriminant Ratio

Fisher's Discriminant Ratio (FDR) umumnya dipakai untuk mengukur kekuatan diskriminasi fitur individu dalam memisahkan dua kelas berdasarkan nilai yang dimilikinya. Untuk  $\mu_1$  dan  $\mu_2$  masing-masing adalah nilai rata-rata dari dua kelas,

$\sigma_1^2$  dan  $\sigma_2^2$  masing-masing adalah varian

dari dua kelas dalam fitur yang akan diukur. FDR dirumuskan seperti pada persamaan berikut.

$$FDR = \frac{(\mu_1 - \mu_2)^2}{(\sigma_1^2 + \sigma_2^2)} \quad (25)$$

Hasil yang diberikan oleh FDR adalah untuk fitur yang mempunyai perbedaan yang besar pada rata-rata dari kelas dan varian kecil dari tiap kelas, maka nilai FDR yang tinggi akan didapatkan. Jika dua fitur mempunyai perbedaan absolut rata-rata yang sama tapi berbeda dalam jumlah

varian ( $\sigma_1^2 + \sigma_2^2$ ), maka fitur dengan jumlah varian yang lebih kecil akan mendapatkan nilai FDR yang lebih tinggi. Di sisi lain, jika dua fitur mempunyai jumlah varian yang sama tetapi perbedaan absolut rata-rata berbeda maka fitur dengan perbedaan absolut rata-rata yang lebih besar akan mendapatkan nilai FDR yang lebih tinggi (Prasetyo, 2014).

### METODOLOGI PENELITIAN

Penelitian ini melakukan analisis terhadap 3 jenis fitur: statistik, moment invariant, dan matrik co-occurrence. Dari fitur statistik, fitur yang diproses adalah rata-rata intensitas, kontras, smoothness, third moment, uniformity, dan entropy. Dari moment invariant, fitur yang diproses adalah 7 moment invariant. Sedangkan dari matrik co-occurrence fitur yang diproses adalah entropy, energi, kontras, dan homogenitas. Analisis fitur dilakukan dengan metode Fisher's Discriminant Ratio (FDR) untuk mendapatkan fitur yang informatif terhadap klasifikasi.

Kerangka kerja penelitian yang dilakukan seperti yang disajikan dalam makalah ini sebagai berikut:

1. Melakukan studi literatur penelitian pengenalan jenis pohon mangga dan seleksi fitur

Pada tahap ini penulis melakukan studi

terhadap sejumlah penelitian, baik mengenai pengenalan jenis pohon

mangga, maupun pengenalan obyek yang lain. Studi pustaka terhadap metode untuk pemilihan fitur juga

perlu dilakukan untuk mengetahui

metode yang cocok digunakan. Untuk

analisis fitur tekstur daun mangga ini

penulis menggunakan Fisher's Discriminant Ratio.

2. Melakukan perhitungan analisis pemilihan fitur menggunakan FDR

Pada tahap ini penulis melakukan perhitungan nilai FDR pada semua fitur. Nilai FDR dihitung secara terpisah

pada setiap fitur, dimana setiap fitur

dengan sejumlah data akan dikorelasikan dengan label kelasnya.

3. Pemilihan fitur informatif berdasarkan nilai FDR

Pada penelitian ini, penulis melakukan pemilihan sejumlah fitur informatif berdasarkan nilai FDR semua fitur yang nilainya lebih besar atau sama dengan dari batas (*threshold*) yang ditetapkan. Jika nilai FDR sebuah fitur lebih besar atau sama dengan nilai *threshold* yang ditetapkan maka fitur tersebut dinyatakan lolos seleksi sebagai fitur yang informatif saat klasifikasi.

4. Uji perbandingan dan analisis

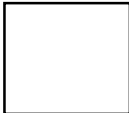
Pada tahap ini, penulis melakukan uji banding penggunaan fitur yang lolos seleksi FDR dengan fitur lama yang sudah digunakan pada metode klasifikasi. Metode klasifikasi yang digunakan adalah K-NN. Setelah dilakukan pengujian kinerja prediksi, selanjutnya dilakukan analisis terhadap hasil yang didapat pada kedua jenis fitur yang digunakan.

5. Menarik kesimpulan fitur informatif yang didapatkan.

Tabel 1 Set data pengenalan jenis mangga

Statistik						Moment invariant							Matrik co-occurrence				kelas
rata-rata	kontras	smoothness	third moment	uniformity	entropy	momen 1	momen 2	momen 3	momen 4	momen 5	momen 6	momen 7	homo genitas	entropy	energy	kontras	
20.033	22.887	0.008	0.082	0.300	3.462	4.878	10.017	18.109	20.256	39.876	25.592	39.944	1.638	0.289	40.988	0.696	1
19.090	22.556	0.008	0.092	0.316	3.387	4.875	10.016	18.112	20.188	40.133	25.670	39.614	1.604	0.306	40.240	0.703	1
21.536	24.779	0.009	0.123	0.299	3.506	5.058	10.472	20.386	21.009	41.778	26.722	42.711	1.633	0.288	58.770	0.703	1
20.133	24.382	0.009	0.120	0.337	3.229	5.111	10.586	21.109	21.992	43.648	28.548	44.371	1.437	0.327	34.883	0.762	1
18.425	22.617	0.008	0.109	0.335	3.309	4.951	10.211	19.345	21.265	42.220	26.998	41.848	1.461	0.325	30.841	0.763	1
22.406	26.869	0.011	0.189	0.313	3.573	5.139	10.634	19.741	22.087	43.224	27.585	43.817	1.584	0.304	59.876	0.740	1
25.857	29.639	0.013	0.168	0.309	3.441	5.315	11.033	20.017	20.545	40.827	26.139	45.117	1.590	0.298	67.740	0.720	1
23.551	29.328	0.013	0.224	0.357	3.179	5.308	11.014	20.107	20.789	41.239	26.407	44.145	1.435	0.347	66.871	0.756	1
24.153	26.723	0.011	0.110	0.291	3.442	5.339	11.283	19.532	21.999	43.890	28.029	42.835	1.596	0.279	60.180	0.714	1
30.203	35.069	0.019	0.264	0.322	3.351	5.577	11.728	20.050	21.126	41.718	27.570	44.129	1.565	0.312	77.286	0.714	1
27.827	27.979	0.012	0.043	0.246	3.672	5.267	10.951	19.930	21.883	45.044	28.832	42.796	1.653	0.234	46.472	0.716	1
29.001	28.874	0.013	0.040	0.241	3.714	5.307	11.032	20.326	22.368	44.714	28.768	43.808	1.700	0.229	47.353	0.701	1
27.058	27.160	0.011	0.040	0.246	3.632	5.236	10.882	20.240	22.471	43.981	28.174	45.084	1.666	0.233	46.068	0.704	1
26.560	27.301	0.011	0.053	0.258	3.586	5.240	10.889	20.238	22.437	44.054	28.083	44.515	1.643	0.245	43.702	0.708	1
23.694	26.683	0.011	0.090	0.312	3.153	5.278	11.021	19.880	20.767	41.094	26.278	43.529	1.436	0.299	39.811	0.747	1
25.421	27.755	0.012	0.114	0.285	3.422	5.276	11.029	20.741	21.871	43.193	27.386	44.911	1.576	0.272	66.678	0.716	1
25.673	28.265	0.012	0.095	0.299	3.265	5.323	11.100	20.285	21.532	42.441	27.134	45.670	1.509	0.286	40.175	0.728	1
25.561	27.496	0.011	0.078	0.284	3.383	5.333	11.140	20.296	21.614	42.569	27.214	47.297	1.563	0.271	42.452	0.720	1
30.402	32.047	0.016	0.101	0.273	3.532	5.623	11.908	21.300	22.681	44.672	28.663	48.278	1.585	0.261	53.245	0.733	1
30.996	32.387	0.016	0.092	0.270	3.526	5.634	11.930	21.450	23.230	45.588	29.269	47.348	1.578	0.257	53.730	0.737	1
29.810	32.566	0.016	0.135	0.293	3.417	5.643	11.950	21.391	23.126	45.389	29.131	47.763	1.510	0.281	52.634	0.753	1
28.287	31.414	0.015	0.143	0.301	3.346	5.584	11.790	21.036	22.448	44.236	28.564	45.504	1.524	0.290	47.783	0.734	1
25.366	27.924	0.012	0.103	0.293	3.424	5.273	10.929	20.373	21.141	41.937	26.829	43.191	1.636	0.281	43.391	0.688	1
25.116	28.244	0.012	0.121	0.305	3.343	5.276	10.930	20.495	21.508	42.602	27.360	43.395	1.523	0.292	44.803	0.730	1
26.598	29.934	0.014	0.141	0.306	3.362	5.337	11.051	21.000	22.273	44.152	28.459	44.386	1.536	0.294	46.865	0.729	1
30.475	31.091	0.015	0.095	0.248	3.787	5.349	11.086	20.852	22.415	44.666	29.648	44.221	1.811	0.236	72.226	0.655	1
24.367	27.105	0.011	0.130	0.291	3.431	5.398	11.446	22.172	24.381	48.279	30.351	47.930	1.598	0.279	61.675	0.709	1
24.229	26.181	0.010	0.068	0.288	3.321	5.381	11.439	21.565	24.278	49.235	30.851	47.312	1.547	0.272	61.996	0.719	1
25.241	27.385	0.011	0.117	0.279	3.485	5.412	11.475	22.104	24.071	47.489	29.976	47.626	1.643	0.266	61.728	0.696	1
25.734	27.460	0.011	0.104	0.272	3.498	5.416	11.486	21.943	23.564	46.495	29.409	47.026	1.632	0.259	60.875	0.701	1
22.054	23.938	0.009	0.069	0.280	3.470	5.105	10.622	18.295	19.079	37.792	24.529	39.363	1.567	0.268	38.007	0.732	2
28.205	31.724	0.015	0.145	0.311	3.241	5.279	10.887	18.960	20.686	40.908	27.110	40.929	1.528	0.299	49.663	0.714	2
25.455	25.145	0.010	0.028	0.238	3.611	5.196	10.909	19.383	20.499	40.466	26.134	41.946	1.642	0.224	39.279	0.714	2
24.725	24.338	0.009	0.024	0.236	3.600	5.181	10.899	19.523	20.555	40.642	26.215	41.798	1.633	0.222	38.325	0.714	2
25.545	24.578	0.009	0.010	0.231	3.541	5.202	10.953	18.717	19.746	38.988	25.287	40.909	1.613	0.215	37.500	0.714	2
25.356	24.441	0.009	0.011	0.232	3.548	5.216	11.003	18.861	19.832	39.189	25.386	41.120	1.615	0.216	35.054	0.714	2
26.725	25.765	0.010	0.034	0.220	3.836	5.149	10.740	19.616	21.351	42.181	27.151	42.298	1.724	0.207	59.354	0.702	2
27.328	26.418	0.011	0.035	0.222	3.843	5.193	10.838	19.956	21.584	42.640	27.298	42.884	1.725	0.209	57.935	0.702	2
30.375	27.908	0.012	0.041	0.185	4.262	5.313	11.185	20.642	22.747	44.552	28.648	48.329	1.922	0.174	78.212	0.674	2
29.690	27.173	0.011	0.041	0.182	4.264	5.266	11.069	20.589	22.738	44.574	28.484	45.476	1.888	0.170	80.370	0.689	2
22.742	23.547	0.008	0.056	0.255	3.566	5.141	10.773	20.296	22.815	44.803	29.178	44.755	1.630	0.242	52.105	0.712	2
24.564	25.750	0.010	0.079	0.258	3.651	5.226	10.938	21.028	21.872	43.492	27.833	43.945	1.656	0.247	55.541	0.711	2

202	25.020	25.465	0.010	0.058	0.249	3.599	5.225	10.956	20.555	23.538	45.741	29.243	46.881	1.633	0.235	55.726	0.711	2
	24.267	25.449	0.010	0.069	0.264	3.541	5.245	11.005	20.959	23.597	46.475	29.741	46.161	1.619	0.251	57.091	0.713	2
	27.578	27.216	0.011	0.043	0.235	3.746	5.334	11.193	20.142	21.294	42.101	27.114	42.918	1.721	0.222	61.600	0.691	2
	28.028	26.838	0.011	0.025	0.221	3.799	5.339	11.220	20.250	21.369	42.255	27.154	43.152	1.764	0.207	57.847	0.680	2
	28.405	27.170	0.011	0.025	0.220	3.814	5.347	11.230	20.201	21.307	42.125	27.103	43.120	1.779	0.207	61.098	0.675	2
	28.077	27.209	0.011	0.032	0.226	3.799	5.350	11.237	20.266	21.391	42.301	27.209	43.170	1.762	0.213	60.938	0.679	2
	27.728	27.859	0.012	0.057	0.247	3.562	5.229	10.846	19.465	20.705	40.924	26.606	41.631	1.660	0.232	66.678	0.690	2
	26.421	25.654	0.010	0.028	0.231	3.619	5.167	10.726	19.376	20.549	40.623	26.338	41.433	1.729	0.215	62.025	0.667	2
	28.112	27.839	0.012	0.051	0.238	3.652	5.242	10.881	19.638	21.033	41.603	27.126	41.977	1.725	0.224	69.182	0.673	2
	27.987	27.143	0.011	0.027	0.231	3.664	5.233	10.860	19.777	21.133	41.820	27.194	42.200	1.728	0.216	63.006	0.676	2
	23.500	22.377	0.008	0.019	0.226	3.468	5.059	10.558	18.542	19.754	38.948	25.215	40.240	1.649	0.206	54.399	0.684	2
	25.666	23.145	0.008	0.001	0.203	3.622	5.125	10.716	19.120	20.803	40.973	26.579	41.421	1.700	0.182	63.091	0.684	2
	25.671	22.780	0.008	-0.003	0.194	3.703	5.118	10.709	19.239	21.149	41.779	27.167	41.732	1.721	0.174	62.849	0.687	2
	26.095	22.275	0.008	-0.014	0.180	3.781	5.109	10.696	19.144	20.874	41.136	26.668	41.463	1.805	0.158	61.181	0.654	2
	23.131	22.435	0.008	0.039	0.221	3.700	5.171	10.952	19.601	21.694	44.477	28.653	42.351	1.738	0.205	62.206	0.680	2
	22.359	22.172	0.008	0.033	0.235	3.625	5.174	10.954	19.437	20.984	41.642	27.436	41.456	1.676	0.220	51.472	0.700	2
	23.576	21.961	0.007	0.024	0.201	3.861	5.169	10.963	19.348	20.782	41.187	27.095	41.199	1.854	0.185	56.019	0.650	2
	25.451	23.051	0.008	0.011	0.190	3.960	5.233	11.087	19.407	20.806	41.217	26.945	41.306	1.859	0.176	52.523	0.662	2





**HASIL DAN PEMBAHASAN**

**1. Set data yang diolah**

Set data yang diolah dalam penelitian ini adalah set data yang telah digunakan oleh Agustin dan Prasetyo (2011) dan Prasetyo (2012) ditambah dengan fitur lain dalam ketiga jenis yang tidak digunakan, seperti pada tabel 1. Jumlah data yang digunakan ada 60 data, terdiri dari 30 jenis mangga gadung dan 30 jenis mangga curut (=jawa). Jumlah fitur yang dianalisis ada 17 fitur, terdiri dari 6 fitur statistik, 7 moment invariant, dan 4 fitur matrik co-occurrence.

**2 Nilai FDR fitur dan seleksi**

Selanjutnya dilakukan perhitungan nilai FDR dari setiap fitur dengan menghitung pada setiap fitur dan setiap

kelompok kelas berupa rata-rata dan varian kemudian digunakan formula (25)

untuk mendapatkan nilai FDR. Hasilnya seperti disajikan pada tabel 2.

Tabel 2 Nilai FDR semua fitur

rata-kontr	smooth	third	uni	entro	mome	mome		
rata	s	ness	moment	formity	py	n 1	momen	n 3
0.019	0.521	0.5156	1.7236	2.4902	1.004	0.192	0.0877	0.433
tidak	tidak	tidak	lolos	lolos	lolos	tidak	tidak	tidak
mome	mome	momen	ener	kontr	homo			
n 4	n 5	6	momen	entrop	y	s	genitas	
0.290	0.331	0.2207	0.4836	1.0548	2.519	0.071	0.7977	
tidak	tidak	tidak	tidak	lolos	lolos	tidak	lolos	

Selanjutnya dilakukan pemilihan fitur yang lolos seleksi, fitur yang lolos seleksi dianggap sebagai fitur yang informatif dalam klasifikasi. Dengan mengamati nilai FDR pada semua fitur dimana nilai tertinggi adalah 2.5193 dan terendah adalah 0.0197, rata-ratanya adalah 0.7505. Pada penelitian ini digunakan threshold adalah rata-rata tersebut, maka berdasarkan nilai rata-rata=0.7505

sebagai threshold, maka fitur yang lolos

seleksi sebagai fitur informatif ada 6. Secara berurutan fitur terpilih dari yang paling informatif adalah: energi (co- occurrence), uniformity (statistik), third moment (statistik), entropy (co- occurrence), entropy (statistik), dan homogenitas (co-occurrence). Tidak ada satupun dari moment invariant yang lolos seleksi, ini disebabkan nilai FDR yang memang jauh di bawah threshold. Hal ini

juga berarti bahwa fitur tekstur moment invariant kurang cocok digunakan sebagai fitur dalam pengenalan jenis pohon mangga berdasarkan tekstur daun.

**3 Uji perbandingan dan analisis**

Selanjutnya dilakukan perbandingan antara fitur terpilih tersebut dengan fitur yang digunakan dalam penelitian sebelumnya. Perbandingan fitur yang digunakan disajikan pada tabel 3. Kedua kelompok fitur tersebut kemudian digunakan untuk prediksi menggunakan metode K-NN dengan pilihan sejumlah nilai K tetangga terdekat.

Tabel 3 Fitur yang digunakan sebagai perbandingan kinerja klasifikasi

Fitur Lama	Fitur Baru
Rata-rata(statistik)	third moment (statistik)
Smoothness (statistik)	uniformity (statistik)
Entropy (statistik)	entropy (statistik)
Moment 1	entropy (co-occurrence)
Moment 2	energi (co-occurrence)
Moment 4	homogenitas (co-occurrence)
Moment 6	
Moment 7	
Energi (co-occurrence)	
Kontras (co-occurrence)	

Pengujian perbandingan dilakukan

dengan teknik K-Fold Cross Validation dengan K=5, pada setiap sesi pengujian ada 12 data yang menjadi data uji, sisanya menjadi data latih. Prediksi dengan K-NN dilakukan dengan jarak Euclidean, K untuk jumlah tetangga terdekat adalah 3, 5, 7, 9, 11. Untuk fitur yang lama, hasil prediksi disajikan pada tabel 4. Sedangkan fitur baru, hasil prediksi disajikan pada tabel 5.

Tabel 4 Hasil prediksi dengan K-NN pada fitur lama

Akurasi pada setiap sesi Cross

KK	Validation rata					rata-
	1	2	3	4	5	
3	0.92	0.58	0.75	0.67	0.58	0.70
5	0.92	0.92	1.00	1.00	1.00	0.97
7	0.75	0.75	0.75	0.58	0.50	0.67
9	0.83	0.75	0.75	0.75	0.75	0.77
11	0.92	0.75	0.58	0.58	0.58	0.68



Tabel 5 Hasil prediksi dengan K-NN pada fitur baru

KK rata	Akurasi pada setiap sesi Cross					rata-
	Validation					
	1	2	3	4	5	
3	0.75	0.75	0.75	0.75	0.75	0.75
5	0.75	0.75	0.75	0.75	0.75	0.75
7	0.92	0.92	0.92	0.83	0.92	0.90
9	1.00	0.92	0.83	0.75	0.75	0.85
11	0.75	0.75	0.92	0.92	0.92	0.85

Dari informasi yang disajikan pada tabel 4 dan 5, dapat diamati bahwa rata-rata akurasi yang didapatkan dengan fitur baru hasil analisis cenderung lebih tinggi dan stabil dengan akurasi terendah 0.75, sedangkan pada sistem dengan fitur lama mempunyai akurasi dengan perbedaan yang tinggi, tertinggi 0.97 tetapi terendah mencapai 0.67. Meskipun akurasi tertinggi fitur baru 0.90 lebih rendah dari pada fitur lama 0.97 tetapi fitur baru cenderung lebih tinggi dari pada yang lama, dimana untuk K=3, K=7, K=9, dan K=11 rata-rata akurasi sistem dengan fitur baru lebih tinggi dari pada sistem dengan fitur lama. Kelebihan lain yang dimiliki oleh sistem yang baru nantinya adalah jumlah fitur yang digunakan lebih sedikit yaitu 6, sedangkan yang lama menggunakan 10 fitur. Di sisi lain, dibandingkan dengan penelitian sebelumnya yang mempunyai akurasi

fitur baru ini juga lebih baik yaitu 0.90 atau 90%.

#### KESIMPULAN DAN SARAN

Dari penelitian yang dilakukan, maka dapat disimpulkan sebagai berikut:

1. Analisis dengan FDR didapatkan bahwa fitur tekstur yang informatif dalam melakukan pengenalan jenis mangga ada 6, meliputi: energi (co- occurrence), uniformity (statistik), third moment (statistik), entropy (co- occurrence), entropy (statistik), dan homogenitas (co-occurrence).
2. Uji kinerja prediksi yang dilakukan pada fitur lama dan fitur baru hasil analisis memberikan jawaban bahwa fitur baru hasil analisis cenderung lebih informatif dan memberikan hasil

akurasi prediksi yang lebih baik daripada fitur lama.

Saran yang dapat penulis berikan dari penelitian ini sebagai petunjuk untuk penelitian berikutnya adalah sebagai berikut:

1. Perlu kajian lebih lanjut untuk mengetahui kinerja akurasi fitur baru hasil analisis dengan menerapkan metode klasifikasi yang lain.

2. Pengujian fitur-fitur yang mungkin mempunyai korelasi yang tinggi juga perlu dilakukan.

3. Penelitian ini hanya berfokus pada fitur tekstur, masih ada jenis fitur lain seperti bentuk, garis, atau jenis-jenis fitur yang lain.

#### DAFTAR PUSTAKA

- [1] Ahmad, U., 2005. *Pengolahan Citra Digital & Teknik Pemrogramannya*, Edisi 1, Yogyakarta: Graha Ilmu
- [2] Departemen Pertanian. 2015, *Luas Panen Buah-Buahan di Indonesia 2009-2013*, <http://www.pertanian.go.id>
- [3] Gonzalez, R.C, Wood, R.E. 2008, *Digital Image Processing*, 3rd Edition, Pearson New Jersey: Prentice Hall
- [4] Prasetyo, E. 2014, *Data Mining-Mengolah Data Menjadi Informasi Menggunakan Matlab*, Yogyakarta: Andi Offset
- [5] Prasetyo, E. 2012, *Perbaikan Sistem Pengenal Jenis Pohon Mangga Menggunakan SVM dan FK-NNC*, Jurnal SCAN, Vol. VII. No. 3, Oktober 2012
- [6] Prasetyo, E., Agustin, S. 2011, *Klasifikasi Jenis Pohon Mangga Gadung Dan Curut Berdasarkan Tesktur Daun, in Proceeding of Seminar Nasional Sistem Informasi Indonesia*, Surabaya: Institut Teknologi Sepuluh Nopember,
- [7] Theodoridis, S., Koutroumbas, K. (2009), *Pattern Recognition-4th edition*, Academic Press: Burlington, MA, USA.

