

Klasterisasi Kualitas Beras Berdasarkan Citra Pecahan Bulir Dan Sebaran Warna

Eko Supriyadi^{*1}, Achmad Basuki², Riyanto Sigit³

^{1,2,3}Pasca Sarjana Teknik Informatika dan Komputer Politeknik Elektronika Negeri Surabaya

E-mail: ¹ekalaya@pasca.student.pens.ac.id, ²basuki@pens.ac.id,

³riyanto@pens.ac.id

Abstrack – Rice is a staple food for Indonesian people, because rice contains complex carbohydrates and can provide various other nutrients that are important to the body. There are still many people who assume that if rice is clean, indifferent, and has a higher price, the rice is of good quality, even though it does not mean that the rice is of good quality. For this reason, we conduct an investigation into the quality of rice which can help the community to differentiate between good and bad qualities. Many factors affect rice quality such as grain fragments, non-uniform color, odor and other factors. The study used the proportion of broken rice grains and color uniformity to determine the quality of rice. The study used the broken grain parameters and color distribution which then led to clustering with the K-Means method where the cluster values were used as value classes to label the type of rice according to the quality of the medium 1,2, and 3. Furthermore, the data labeled cluster was carried out by the classification process to obtain accuracy the best, and the best classification method is neural network at 99.85%

Keywords – rice quality, rice image, K-Mean, Neural Network.

Intisari – Beras sebagai bahan pokok makanan masyarakat Indonesia, karena di dalam beras terdapat kandungan karbohidrat kompleks, serta dapat memberikan berbagai zat gizi lain yang penting bagi tubuh. Masih banyak masyarakat yang berasumsi jika beras dalam keadaan bersih, tak berbau, dan memiliki harga lebih mahal, beras tersebut berkualitas baik, padahal belum tentu beras yang dimaksud tersebut mempunyai kualitas baik. Untuk itu dalam melakukan penelitian mendeteksi kualitas beras yang nantinya dapat membantu masyarakat untuk dapat membedakan kualitas baik dan buruk. Banyak faktor yang mempengaruhi kualitas beras seperti fragmen biji-bijian, warna yang tidak seragam, bau dan faktor lainnya. Penelitian ini menggunakan persentase buliran beras pecah dan keseragaman warna untuk menentukan kualitas beras. Penelitian ini menggunakan parameter buliran pecah dan distribusi warna yang selanjutnya diproses klasterisasi dengan metode *K-Means* di mana nilai klaster dijadikan sebagai nilai *class* untuk melabeli jenis beras sesuai mutu dari medium 1,2, dan 3. Selanjutnya data yang terlabel klaster dilakukan proses klasifikasi untuk mendapatkan akurasi yang terbaik, dan metode klasifikasi yang terbaik adalah neural network sebesar 99,85%.

Kata kunci – kualitas beras, citra beras, K-Mean, Neural Network.

I. PENDAHULUAN

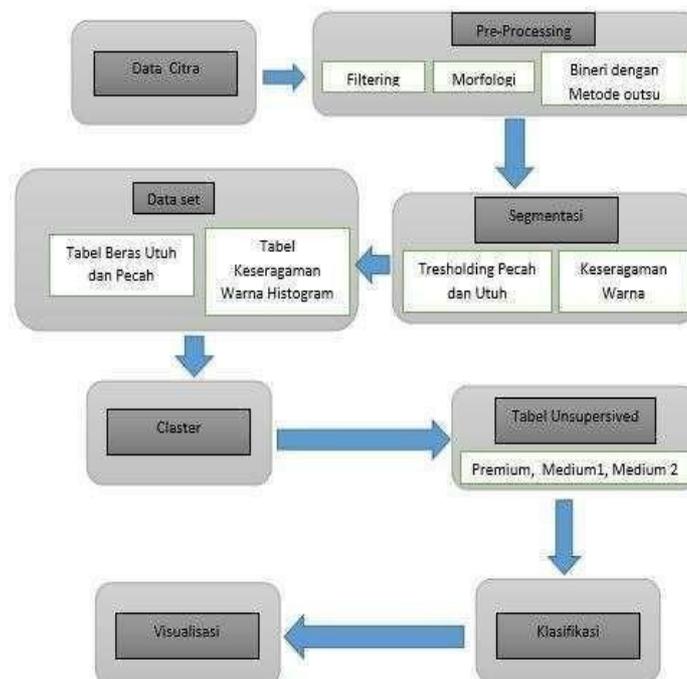
Konsumsi makanan sehari – hari penduduk Indonesia adalah nasi yang berasal dari beras. Untuk menentukan mutu beras yang dikonsumsi, maka pemerintah melalui Badan Standardisasi Nasional (BSN), Standar Nasional Indonesia (SNI) membagi 4 klasifikasi kualitas beras, yakni premium, medium I, medium II, dan medium III. Adapun syarat beras yang ditentukan, yakni bebas penyakit dan bebas hama, tidak berbau apek, asam ataupun bau lainnya, terbebas dari dedak atau bekatul, terlebih utama terbebas dari bahan kimia. Untuk menentukan kualitas dan membedakan jenis beras tersebut maka proses pemeriksaan kualitas beras selalu dilakukan oleh para ahli dibidang pertanian dengan pengawasan dan pemeriksaan yang rutin [1].

Kelemahan dalam pemeriksaan di atas memerlukan waktu yang lama dan menghasilkan produk dengan kualitas yang tidak merata karena keterbatasan visual, faktor kelelahan, dan perbedaan persepsi masing-masing pengamat. Salah satu alternatif yang dibutuhkan adalah pengolahan citra agar pendeteksian lebih cepat dan akurat. Proses pengolahan citra ini memungkinkan digunakan secara luas terlebih dengan berkembangnya teknologi *smartphone* yang menawarkan fitur kamera yang semakin baik dan murah. Beberapa penelitian terdahulu mengembangkan algoritma pemrosesan citra dengan menilai menilai beras berdasarkan panjang, lebar, luas dan juga bekerja pada deteksi warna pada bulir beras untuk menentukan kualitas beras [2][3]. Hasilnya diklasifikasikan setiap jenis beras untuk membantu menentukan evaluasi harga. Hasil eksperimen menunjukkan ada bagian signifikan dari butir beras bernilai tinggi yang tercampur dalam butiran beras yang pecah [4]. Penelitian lainnya melakukan ekstraksi citra kuantitatif dari segmen informasi [5].

Perlindungan konsumen mengenai kualitas beras sangatlah perlu karena beras adalah komoditas makanan utama masyarakat Indonesia yang dikonsumsi setiap hari. Banyak terjadi penipuan dan pemalsuan kualitas beras oleh sebagian orang. Dengan menyampur atau mengoplos beras berkualitas baik dengan beras berkualitas rendah untuk dapat keuntungan yang banyak. Untuk melindungi masyarakat dari pemalsuan tersebut maka diperlukan pendeteksi kualitas beras yang nantinya bisa membantu masyarakat dan juga membantu pemerintah untuk mengontrol kualitas beras yang baik dan yang buruk. Yang sesuai Badan Standardisasi Nasional (BSN).

II. SIGNIFIKANSI STUDI

Sistem desain yang dibuat pada penelitian ini adalah terdiri dari pengambilan citra dengan menggunakan kamera *hand phone* untuk menentukan citra beras yang digunakan, *pre-processing*, *Segmentasi fitur* dan distribusi warna untuk mendapatkan hasil kualitas beras berdasarkan jumlah bulir yang tidak normal (pecah dan tidak utuh) dan warna.



Gambar 1. Desain Sistem

A. *Data citra*

Data citra diambil menggunakan kamera *handphone* sesuai dengan jenis beras dan kualitasnya yang akan dijadikan label dari setiap data. Pengambilan citra dilakukan dengan latar belakang hitam, berjarak 15cm dan memperbolehkan posisi beras bertumpuk. Percobaan ini menggunakan 40 jenis beras dengan masing-masing antar 10 pada setiap jenis beras, jadi total ada 400 gambar.

B. *Pre-Processing*

Tahap *Pre-Processing* ini dilakukan untuk mempersiapkan gambar beras dengan tujuan meningkatkan gambar sebagai langkah proses sebelum analisis. mengoreksi penerangan latar belakang tidak seragam dan mengubah gambar menjadi gambar biner untuk memudahkan mengidentifikasi objek latar depan (butiran beras individu). kemudian dapat menganalisis objek, seperti menemukan area setiap butir beras, dan dapat menghitung statistik untuk semua objek digambar. proses selanjutnya dengan mengurangi gambar asli dengan gambar hasil proses *filtering*. Dengan memperhatikan dua langkah sebelumnya dapat diganti dengan satu langkah menggunakan *imtophat* yang pertama menghitung pembukaan morfologis dan kemudian mengurangkannya dari gambar asli. Buat versi biner dari gambar yang diproses dapat menggunakan fungsi kotak alat untuk analisis. Gunakan fungsi *imbinarize* untuk mengubah gambar skala abu-abu menjadi gambar biner. Hapus noise latar belakang dari gambar dengan menggunakan metode *Otsu*[6].

Pada langkah ini, citra yang rusak dipisahkan dari latar belakangnya. Nilai intensitas citra \geq nilai *threshold* dirubah menjadi 1 (berwarna putih) untuk nilai intensitas citra yang kurang dari nilai *threshold* dirubah menjadi 0 (berwana hitam). citra yang dihasilkan dari hasil *thresholding* adalah citra *biner* dengan menggunakan *otsu*.

C. *Segmentasi*

Pada proses *segmentasi* , kami menggunakan 3 parameter penentu deteksi kualitas beras, tahap deteksi beras utuh dan pecah, warna pada bulir beras.

1) *Thresholding* Nilai Beras Pecah Dan Utuh

Persamaan yang digunakan untuk memisahkan nilai beras yang utuh dan nilai beras yang pecah adalah:

$$g(x, y) = \begin{cases} 0, & \text{jika } f(x, y) < T \\ 1, & \text{jika } f(x, y) \geq T \end{cases} \tag{1}$$

Di mana :

$f(x,y)$ adalah citra

rayscale $g(x,y)$ adalah citra *biner*

T adalah nilai *threshold*

2) *Penyeragaman warna (Filter Warna)*

Pada langkah ini, gambar asli dipisahkan dari latar belakangnya, agar mendapatkan nilai keseragaman warna yang nantinya akan diuji dan juga merubah latar belakang gambar dengan warna biru. Kualitas gambar yang dihasilkan tergantung pada metode perkiraan yang gunakan, kisaran warna pada gambar input, dan menggunakan *dithering* atau tidak. Mengurangi jumlah warna dalam gambar melibatkan kuantisasi. Fungsi *rgb2ind* menggunakan kuantisasi sebagai bagian dari algoritma pengurangan warnanya. *rgb2ind* mendukung dua metode kuantisasi: kuantisasi seragam dan kuantisasi *varians minimum*.

TABEL I
CITRA KESERAGAMAN WARNA

penyeragaman warna	Histogram penyeragaman warna
	

D. *Data Set*

Dari proses *thresholding* pemisahan bulir beras pecah, utuh dan histogram keseragaman warna, Proses parameter tersebut dilakukan untuk membangun dataset dari hasil gambar jenis jenis beras. Proses ini menghasilkan dataset sebanyak 400 dataset dari nilai proses jenis beras tersebut untuk mendapatkan nilai mean dan standart *deviasi*.

E. *Klaster*

Dari data nilai mean, standart deviasi dan keseragaman warna maka dilakukan proses klastering, dalam penentuan klaster pecah beras dibagi menjadi tiga yaitu kualitas 1, kualitas 2 dan klaster 3 dengan formula :

$$J(V) = \sum_{i=1}^{C_i} \sum_{j=1}^{C_i} (||x_i - v_j||)^2 \tag{2}$$

Di mana:

'|| xi - vj ||' adalah jarak *Euclidean* antara xi dan vj.

'Ci' adalah jumlah titik data di *cluster* ke-i.

'C' adalah jumlah pusat *cluster*.

Langkah algoritmik untuk pengelompokan *k-means*

Misalkan $X = \{x_1, x_2, x_3, \dots, x_n\}$ menjadi himpunan titik data dan $V = \{v_1, v_2, \dots, v_c\}$ menjadi himpunan pusat.

- 1) Secara acak pilih pusat *cluster* 'c'.
- 2) Hitung jarak setiap titik data dengan pusat *cluster*.
- 3) Tetapkan titik data yang jaraknya dari pusat *cluster* minimum dari semua pusat cluster.
- 4) Penghitung ulang pusat *cluster* baru menggunakan:

$$v_i = \left(\frac{1}{c_i}\right) \sum_{j=1}^{c_i} x_j \tag{3}$$

Di mana, 'ci' merepresentasikan jumlah titik data dalam klaster ke-i.

- 5) Hitung ulang jarak antara setiap titik data dan pusat cluster baru yang diperoleh.
- 6) Jika tidak ada titik data yang ditetapkan ulang, hentikan, jika tidak ulangi dari langkah 3).

F. *Tabel data Unsupervised*

Tabel ini dihasilkan dari proses klaster sesuai masing jenis beras dan ditentukan nilai dari masing-masing *class* dari jenis beras.

TABEL III
CONTOH DATA NILAI CLASS BERAS PECAH, UTUH DAN NILAI WARNA

jenis beras	good rice (%)	bad rice (%)	nilai warna kuning (%)	nilai warna putih (%)	class
arwana	76.20.00	1,006944444	43	13	1
...
banyuwangi1	45.31.00	53.12.00	32	58	2
...
bintang	68.18.00	31.81	56	25	1
...
bintang 5	54.65	45.34.00	54	28	1
...
bintang ijo	56.92	43.07.00	35	52	2
...
ir64	64.64	35.35.00	51	31	1
...

G. *Klasifikasi*

Proses klasifikasi ini menggunakan *tool* weka untuk mendapatkan nilai akurasi terbaik dari masing masing metode yang digunakan. Sebelum masuk proses klasifikasi, data *supervised* dinormalisasi terlebih dahulu supaya data yang diproses nantinya menghasilkan data yang tepat dan akurat. Klasifikasi data *supervised* dengan 7 metode *K-Nearest Neighbour*, *Neural Network*, *Naïve Bayes*, *Fuzzy NN*, *SVM*, *C4.5*, *Random forest* dalam validasi silang, buat setiap rekaman sebagai data uji berurutan dalam penelitian ini. Weka digunakan untuk memilih fitur.

1. *K-Nearest Neighbour*

KNN merupakan algoritma *supervised learning* dimana hasil dari *query instance* yang baru diklasifikasi berdasarkan kelas terbanyak data pada algoritma KNN. Hasil dari klasifikasi ditentukan dari kelas terbanyak. Definisi kedekatan jarak metrik, seperti jarak *Euclidean*. Rumus persamaan 1 Jarak *Euclidean* berikut ini:

$$D_{xy} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{4}$$

keterangan :

D : jarak kedekatan

S : data *training*

E : data *testing*

G : jumlah antar atribut individu 1 s/d n.

f : fungsi *similitary* atribut i antara kasus *X* dan kasus *Y*

G = Atribut individu antara 1 sampai dengan n Langkah dalam menghitung metode *K-Nearest Neighbor* antara lain :

1. Menentukan parameter *K* (jumlah tetangga terdekat).
2. Dengan menggunakan persamaan 1, penghitungan kuadrat jarak *Euclid* (*query instance*) dari objek terhadap data sampel yang diberikan.
3. Kemudian mengurutkan objek tersebut ke dalam kelompok dengan jarak *Euclid* terkecil.
4. Mengumpulkan kategori *Y* (Klasifikasi *Nearest Neighbor*).
5. Prediksi perhitungan nilai *query instance* menggunakan kategori *Nearest Neighbor* yang paling mayoritas.

2. *Neural Network (NN)*

Neural Network atau *Artificial Neural Networks (ANN)* adalah kinerja yang meniru jaringan sel-sel saraf (*neuron*) di otak manusia. komputasi metode ini adalah dalam hubungan antar *neuron*. Lapis ini terdiri dari *layer* masukan dan *layer* keluaran sedangkan pada *Multilayer Neural Network*, *layer* masukan dan *layer* keluaran ada *hidden layer* yang berisi nilai-nilai *node* keluaran yang digunakan sebagai nilai *node* masukan lain (Tahir dkk, 2018). *Neural Network* dapat dirumuskan sebagai berikut:

$$Z = a_1w_1 + a_2w_2 + \dots + akwk + b \tag{5}$$

Keterangan :

a : nilai masukan

w : nilai bobot

b : variabel mewakili karakteristik arah untuk klasifikasi

Karakteristik kontinu dan monoton menurun ini merupakan fungsi aktivasi ,yang memiliki *range* (0,1) dalam Fungsi sigmoid.

3. *Naïve Bayes (NB)*

Naive Bayes merupakan metode *supervised learning* untuk klasifikasi serta memecahkan masalah diagnostik dan prediksi menggunakan probabilitas dan statistik berdasarkan *teorema Bayes* dengan asumsi independensi yang kuat antara fitur yang dimiliki sebuah objek. *Teorema Naïve Bayes*, dirumuskan sebagai berikut :

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)} \tag{6}$$

Keterangan :

P = probabilitas setiap kelas

C = variabel mewakili kelas

F1...Fn = variabel mewakili karakteristik arah untuk klasifikasi

4. *Fuzzy NN*

Model FNN lebih sederhana karena tidak menggunakan lapisan ‘ouput_fuzzy’. Dari fuzzy rule pada conjunction layer yang digunakan adalah fungsi keanggotaan keluaran adalah singletons yang menyertai bobot koneksi pada lapisan terakhir, sehingga FNN ini hanya terdiri terdiri dari empat lapisan. (Kasabov, 1998) .

Rumus yang digunakan berikut:

$$z = \sum_{i=1}^m \pi_i \omega_i / \sum_{i=1}^m \pi_i \tag{7}$$

dimana level aktivasi dari unit output, μ_i dan ω_i berturut-turut adalah fungsi keanggotaan dan jalur terbobot unit ke i pada lapisan antara output fuzzy ke output.

5. *Support Vector Machine (SVM)*

Mesin *Vektor* memiliki fungsi penilaian yang menghitung skor untuk input baru. Mesin *Vektor* adalah pengklasifikasi biner (dua kelas); jika keluaran dari fungsi skoring negatif maka

masukannya tersebut tergolong kelas $y = -1$. Jika skorannya positif maka input tersebut tergolong kelas $y = 1$.

Persamaan untuk fungsi penilaian, digunakan untuk menghitung skor untuk input vektor x .

$$\sum_{i=1}^m \alpha_i y^{(i)} K(x^{(i)}, x) + b \tag{8}$$

Fungsi ini beroperasi pada setiap titik data dalam set pelatihan ($i = 1$ hingga m).

Di mana :

$x^{(i)}$, $y^{(i)}$ mewakili contoh pelatihan ke- i

- a. $x^{(i)}$ adalah vektor input yang dapat berupa dimensi apapun.
- b. $y^{(i)}$ adalah label kelas, yang hanya memiliki satu dari dua nilai, baik -1 atau 1.
- c. α_i adalah koefisien yang terkait dengan contoh pelatihan ke- i .

x : adalah vektor input klasifikasi

K : fungsi kernel.

- a. beroperasi pada dua vektor dan hasilnya adalah skala.
- b. Ada berbagai kemungkinan pilihan fungsi kernel.
- c. b : nilai skala

6. Algoritma C4.5

Algoritma C4.5 merupakan kelompok algoritma decision tree. Algoritma ini mempunyai input berupa training samples dan samples. Training samples merupakan data contoh yang digunakan untuk membangun sebuah tree yang telah diuji kebenarannya. Sedangkan samples merupakan field-field data yang digunakan sebagai parameter dalam klasifikasi data (Sujana, 2010). Secara umum alur proses algoritma C4.5 guna membangun pohon keputusan dalam data mining adalah sebagai berikut (Mahmud, et al, 2014):

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk tiap-tiap nilai
- c. Bagi kasus dalam cabang.
- d. Pengulangan proses pada setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Information gain adalah atribut selection measure yang digunakan untuk memilih test atribut tiap node dalam tree (Sujana, 2010). Dalam memilih atribut yang digunakan sebagai simpul, simpul akar (root) ataupun simpul internal berdasarkan nilai gain tertinggi dari atribut yang ada. Perhitungan nilai gain digunakan rumus sebagai berikut (Mahmud, et al, 2014):

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \tag{9}$$

Keterangan :

S : himpunan kasus

A : atribut

N : jumlah partisi atribut A

$[S_i]$: jumlah kasus pada partisi ke- i

$[S]$: jumlah kasus pada S

Untuk menghitung Entropy menggunakan rumus :

$$Entropy(S) = \sum_{i=1}^n -pi * log_2 pi \tag{10}$$

Keterangan :

S : himpunan kasus

A : fitur

N : jumlah partisi S

pi : proporsi dari Si terhadap S

7. Random Forest

Random forest adalah kombinasi dari masing – masing *tree* yang baik kemudian dikombinasikan ke dalam sebuah model. *Random Forest* bergantung pada sebuah nilai vector random dengan distribusi yang sama pada semua pohon yang masing masing *decision tree* memiliki kedalaman yang maksimal. *Random forest* adalah *classifier* yang terdiri dari *classifier* yang berbentuk pohon $\{h(x, \theta k), k = 1, . . . \}$ dimana θk adalah random vector yang didistribusikan secara independen dan masing masing *tree* pada sebuah unit kan memilih class yang paling populer pada input x. Berikut ini karakteristik akurasi pada *random forest*.

1. Memusatkan *random forest*

Terdapat *classifier* $h_1(x), h_2(x), . . . , h_k(x)$ dan dengan *training set* dari distribusi *random vector* Y,X, Berikut fungsi yang terbentuk

$$mg(X, Y) = avkI (h_k(X) = Y) - \max_{j \neq Y} avkI (h_k(X) = j) \tag{11}$$

Fungsi error yang digunakan

$$PE^* = P_{X,Y} (mg(X, Y) < 0) \tag{12}$$

Hasil dari penggabungan fungsi

$$P_{X,Y} (P_{\theta}(h(X, \theta) = Y) - \max_{j \neq Y} (P_{\theta}(h(X, \theta) = j) < 0) \tag{13}$$

Pada hasil tersebut menjelaskan mengapa *random forest* tidak *overfit* saat *tree* di tambahkan, tetapi menghasilkan nilai yang terbatas pada error.

2. Kekuatan dan Korelasi

Fungsi yang dihasilkan adalah

$$PE^* \leq \sum var(P_{\theta}(h(X, \theta) = Y) - (P_{\theta}(h(X, \theta) = j))s_j^2 \tag{14}$$

Pada fungsi tersebut kekuatan tidak bergantung pada *forest*. Berikut dijelaskan macam fitur-fitur yang digunakan untuk *random forest*.

1. Random Forest menggunakan seleksi input yang random

Bagging digunakan untuk pemilihan fitur secara *random*. Masing – masing *training set* diambil dengan penggantinya dari *training set* asli. Kemudian sebuah *tree* di tanam pada sebuah *training set* menggunakan seleksi fitur *random*. Ada dua alasan penggunaan

bagging yaitu yang pertama penggunaan *bagging* untuk meningkatkan akurasi ketika fitur *random* digunakan. Yang kedua *bagging* digunakan untuk memberikan perkiraan dari kesalahan generalisasi (PE*) dari gabungan *tree*, untuk memperkirakan kekuatan dan korelasi. *Random Forest* yang paling sederhana dengan fitur random dibentuk dengan seleksi secara random, pada masing – masing node, sebuah grup kecil dari input variable yang terbagi. Membentuk *tree* menggunakan metodologi CART ke ukuran yang maksimum.

2. *Random Forest menggunakan kombinasi input yang linear*
 Misalkan terdapat beberapa input, M, F mengambil fraksi pada M yang akan memimpin dalam meningkatkan kekuatan tetapi pada korelasi yang tinggi. Pendekatan yang lain terbentuk dengan mendefinisikan lebih banyak fitur dengan mengambil kombinasi *random linear* dari sejumlah variable input. Fitur tersebut variabel L yaitu jumlah dari variable yang dikombinasikan. Variabel L secara random diseleksi dan ditambahkan bersama dengan koefisien yang memiliki nomor *random* [-1,1]. Kombinasi linear F dihasilkan. Prosedur ini di sebut *Forest-RC*.

III. HASIL DAN PEMBAHASAN

Ada beberapa pengujian yang dilakukan untuk mengetahui hasil dari metode yang diusulkan. Untuk menentukan kualitas Premium ,Medium 1 ,Medium 2, dilihat dari hasil presentase dari tiap – tiap pengujian.

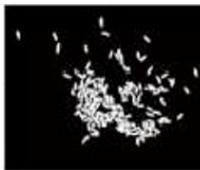
Pengujian ini dilakukan secara bertahap yaitu:

- a. Ekstraksi *parameter* dari jumlah pecahan bulirberasnya.
- b. Ekstraksi *parameter* beras dari keseragaman warnanya.
- c. Melihat nilai kualitas beras dari gabungan fitur: jumlah pecahan beras dan keseragaman warna dalam tabel *means* dan *standart deviasi*.
- d. Klasterisasi.
- e. Klasifikasi.

A. Ekstraksi parameter Beras dari Jumlah Pecahan Bulir Beras

Nilai pecah bulir beras yang dihasilkan nantinya dipakai pada proses berikutnya untuk menentukan kualitas bulir beras yang diinginkan. Dengan melihat jumlah bulir utuh nilai presentase kualitas beras yang utuh dan yang pecah pada tabel nilai Kualitas Beras yang utuh dan yang pecah.

TABEL III
BULIR PECAH DAN UTUH

Jenis Beras	Beras Pecah	Beras Utuh
Beras Arwana Beras utuh: 76.20% Beras pecah: 23.12%		
Beras Banyuwangi Beras utuh: 45.31% Beras pecah: 53.12%		

B. Ekstraksi parameter Beras Dari Keseragaman Warna

Nilai keseragaman warna beras yang dihasilkan nantinya juga dipakai sebagai penentu kualitas beras pada proses berikutnya untuk menentukan kualitas bulir beras yang diinginkan. Dari keseragaman warna nilai presentase kualitas beras pada tabel nilai histogram keseragaman warna di bawah.

TABEL IV
PENYERAGAMAN WARNA DAN HISTOGRAM

Jenis Beras	Penyeragaman Warna	Histogram
Beras Arwana Nilai warna kuning beras: 43% Nilia warna putih beras: 13% Nilia warna kuning & putih beras: 44%		
Beras Banyuwangi Nilai warna kuning beras: 30% Nilia warna putih beras: 48% Nilia warna kuning & putih beras: 22%		

C. Melihat Kualitas Beras kombinasi nilai dari Jumlah Pecahan Bulir Beras dan Keseragaman Warna dalam tabel means dan standart deviasi.

Dari proses diatas menghasilkan data dari masing-masing image beras yang diambil. data tersebut ditabelkan sesuai dari hasil pemrosesan yang dilakukan. terbagi menjadi dua table yaitu data Segmentasi dan data Penyeragaman warna. Summary dari keseluruhan data yang dihasilkan, ditampilkan data nilai means dan standart deviasi, nilai min, nilai max, nilai medium dan nilai means keseragaman warna untuk mencari kluster kualitas beras. Summary dari keseluruhan data yang dihasilkan, ditampilkan data nilai means dan standart deviasi, nilai min, nilai max, nilai medium dan nilai means keseragaman warna untuk mencari kluster kualitas beras bisa dilihat pada tabel data supervised di bawah

TABEL V
TABEL SUPERVISED

jenis beras	good rice (%)	bad rice (%)	nilai warna kuning (%)	nilai warna putih (%)	class	jenis beras	good rice (%)	bad rice (%)	nilai warna kuning (%)	nilai warna putih (%)	class
Arwana	76,2	23,70	43,00	13,00	3	bunga pandan	55,64	44,35	51,22	23,12	3
Arwana	3	bunga pandan	55,64	44,35	51,22	23,12	3
banyuwangi	52,97	47,02	41,00	42,00	1	bunga pandan	3

jenis beras	good rice (%)	bad rice (%)	nilai warna kuning (%)	nilai warna putih (%)	class	jenis beras	good rice (%)	bad rice (%)	nilai warna kuning (%)	nilai warna putih (%)	class
banyuwangi	1	Cap Nyonya	68,18	31,81	43,57	9,37	3
Bintang	68,18	31,81	56,00	25,00	3	Cap Nyonya	3
Bintang	3	coco	65,11	34,88	41,59	40,41	1
bintang 5	67,46	32,53	41,00	45,00	3	coco	1
bintang 5	3	guci	54,42	45,57	26,13	46,63	3
bintang ijo	56,92	43,07	35,00	52,00	1	guci	3
bintang ijo	1	koki merah	78,2	21,79	46,27	22,63	3
ir64	62,5	37,50	34,00	44,00	3	koki merah	3
ir64	3	kuda	59,85	40,14	40,08	19,15	3
Kelapa	69,38	30,61	50,00	34,00	3	kuda	3
Kelapa	3	ladang padi	63,46	36,53	43,7	33,91	3
kelapa muda	89,9	10,09	51,00	24,00	3	ladang padi	3
kelapa muda	3	lele	61,9	38,09	42,12	26,52	3
Lumbung	83,06	16,93	36,00	46,00	3	lele	3
Lumbung	3	lele rejeki	36,12	63,87	23,38	65,68	1
nasi uduk	83,33	16,66	37,00	46,00	3	lele rejeki	1
nasi uduk	3	lumbung padi	52,33	47,66	42,08	12,08	3
pin2	91,2	8,79	50,00	34,00	3	lumbung padi	3
pin2	3	lumbung	83,06	16,93	46,69	23,74	3
tys	43,45	56,54	44,00	41,00	1	lumbung	3
tys	1	menco	52,97	47,02	35,18	48,05	3
555	17,88	82,11	41,33	11,75	1	menco	3
555	1	Raja Lele	94,38	5,61	35,44	28,39	3
naga	51,86	48,13	50,52	16,84	3	Raja Lele	94,38	5,61	35,44	28,39	3
naga	3	Raja Lele	3
Pandan Wangi	94,11	5,88	45,89	31,49	3	rosita	79,2	20,79	42,88	25,56	3

jenis beras	good rice (%)	bad rice (%)	nilai warna kuning (%)	nilai warna putih (%)	class	jenis beras	good rice (%)	bad rice (%)	nilai warna kuning (%)	nilai warna putih (%)	class
Pandan Wangi	3	rosita sedap	3
Pandawa	58,66	41,33	33,46	52,78	1	Sedap wangi	71,42	28,57	9,25	118	2
Pandawa	1	Sedap wangi	2
Pari anom	40,76	59,23	24,16	59,25	1	Sentra ramos	83,33	16,66	18,13	120,03	2
Pari anom	1	Sentra ramos	2
Polif	45,49	54,5	47,69	25,34	1	Kuda hitam	73,78	26,21	56,23	16,63	3
Polif	1	Kuda hitam	3
Raja bandeng	41,7	58,29	37,66	14,93	1	Beras super	70,04	29,95	44,87	10,78	3
Raja bandeng	1	Beras super	3

D. Klasterisasi

Tujuan dari proses klasterisasi ini memproses setiap data unsupervised dari hasil ekstraksi menggunakan parameter warna, tekstur pada bulir beras, dengan menggunakan metode k-means data yang dihasilkan lebih akurat untuk menentukan 3 klaster yaitu klaster 1, klaster 2, dan klaster 3. Yang mana setiap hasil klaster jenis beras akan terseleksi sesuai dengan class yang telah ditentukan Pengklaster dari nilai mean , nilai deviasi pecah beras dengan keseragaman warna dapat dilihat dari gambar 2 berikut :

```

Jenis beras ke 1 termasuk dalam klaster 3
Jenis beras ke 2 termasuk dalam klaster 1
Jenis beras ke 3 termasuk dalam klaster 3
Jenis beras ke 4 termasuk dalam klaster 3
Jenis beras ke 5 termasuk dalam klaster 1
Jenis beras ke 6 termasuk dalam klaster 3
Jenis beras ke 7 termasuk dalam klaster 3
Jenis beras ke 8 termasuk dalam klaster 3
Jenis beras ke 9 termasuk dalam klaster 3
Jenis beras ke 10 termasuk dalam klaster 3
Jenis beras ke 11 termasuk dalam klaster 3
Jenis beras ke 12 termasuk dalam klaster 1
Jenis beras ke 13 termasuk dalam klaster 1
Jenis beras ke 14 termasuk dalam klaster 1
Jenis beras ke 15 termasuk dalam klaster 1
Jenis beras ke 16 termasuk dalam klaster 3
Jenis beras ke 17 termasuk dalam klaster 3
Jenis beras ke 18 termasuk dalam klaster 1
Jenis beras ke 19 termasuk dalam klaster 3
Jenis beras ke 20 termasuk dalam klaster 3
Jenis beras ke 21 termasuk dalam klaster 3
Jenis beras ke 22 termasuk dalam klaster 3
Jenis beras ke 23 termasuk dalam klaster 3
Jenis beras ke 24 termasuk dalam klaster 3
Jenis beras ke 25 termasuk dalam klaster 1
Jenis beras ke 26 termasuk dalam klaster 3
Jenis beras ke 27 termasuk dalam klaster 3
Jenis beras ke 28 termasuk dalam klaster 3
Jenis beras ke 29 termasuk dalam klaster 3
Jenis beras ke 30 termasuk dalam klaster 3
Jenis beras ke 31 termasuk dalam klaster 1
Jenis beras ke 32 termasuk dalam klaster 1
Jenis beras ke 33 termasuk dalam klaster 1
Jenis beras ke 34 termasuk dalam klaster 1
Jenis beras ke 35 termasuk dalam klaster 3
Jenis beras ke 36 termasuk dalam klaster 3
Jenis beras ke 37 termasuk dalam klaster 2
Jenis beras ke 38 termasuk dalam klaster 2
Jenis beras ke 39 termasuk dalam klaster 3
Jenis beras ke 40 termasuk dalam klaster 3
    
```

Gambar 2. Hasil Klaster

Dari hasil klaster di atas, dapat ditentukan class dari jenis beras yang nantinya dibuat tabel *data supervised*

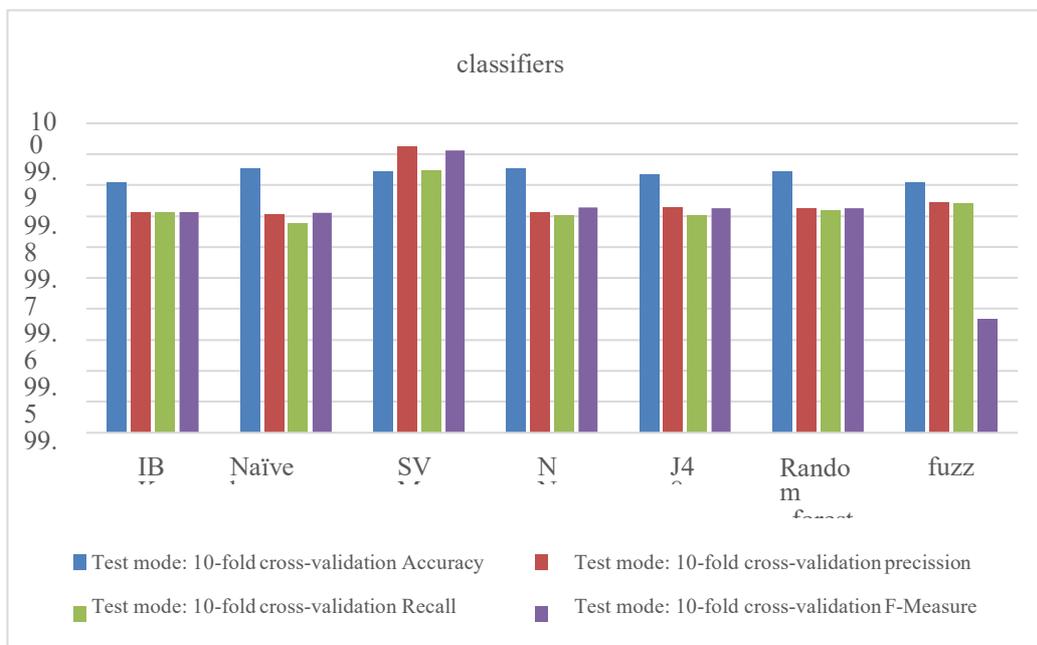
1. *Klasifikasi*

Hasil klasifikasi fitur 7 metode *K-Nearest Neighbour, Neural Network, Naïve Bayes, Fuzzy NN, SVM, C4.5, Random forest* dalam validasi silang, buat setiap rekaman sebagai data uji berurutan dalam penelitian ini.

Tabel di bawah ini adalah hasil pemilihan fitur di mana operator yang digunakan adalah *K-Nearest Neighbour, Neural Network, Naïve Bayes, Fuzzy NN, SVM, C4.5, Random forest* dalam k = 10, dalam 100 iterasi dan sejumlah node dalam lapisan tersembunyi banyak dalam jumlah node output serta jumlah Pembelajaran yang mendalam Konfigurasi yang tersisa adalah menggunakan parameter standar pada Weka.

TABEL VI
DATA PERBANDINGAN ANTAR METODE

classifiers	Test mode: 10-fold cross-validation			
	Accuracy	preccision	Recall	F-Measure
KNN	99,8097	99,71	99,71	99,713
Naïve bayes	99,8523	99,707	99,676	99,712
SVM	99,8459	99,925	99,848	99,911
NN	99,8553	99,712	99,704	99,727
C45	99,8354	99,731	99,704	99,724
Random forest	99,8459	99,724	99,718	99,724
fuzzy	99,8108	99,743	99,742	99,369



Gambar 3. Perbandingan antar metode menggunakan weka

IV. KESIMPULAN

Dari penulisan ini menghasilkan analisa dari setiap eksperimen bahwa metode yang digunakan menghasilkan 3 segmen di dalam satu image beras yang terdiri dari bulir beras utuh, bulir beras pecah, dan keseragaman warna. Masing-masing segmen memiliki nilai sebagai parameter dalam pencarian *class* masing – masing kualitas diproses klustering. Setelah proses tersebut didapat data *supersived* dan dilakukan proses klasifikasi menggunakan algoritma *Neural Network* menghasilkan akurasi 99,85%. Dengan hasil ini ini membuktikan bahwa *hipotesa* yang digunakan dapat menyelesaikan permasalahan dan menjadi parameter untuk mendeteksi kualitas beras yang diinginkan.

REFERENSI

- [1] E. Z. Parveen, "Assessment of Quality of Rice Grain using Optical and Image Processing Technique," pp. 265–270, 2017.
- [2] I. Shape *et al.*, "Pengolahan Citra Digital Menggunakan Metode Susan Detection Dan Neurofuzzy Untuk Identifikasi Komponen," vol. 4, no. 6, pp. 30–45, 2017.
- [3] S. F. Ali, H. Jamil, R. Jamil, I. Torij, and S. Naz, "Low cost solution for rice quality analysis using morphological parameters and its comparison with standard measurements," *Proc. 2017 Int. Multi-Topic Conf. INMIC 2017*, vol. 2018–January, pp. 1–6, 2018.
- [4] D. Ngampak and P. Piamsa-Nga, "Image analysis of broken rice grains of Khao Dawk Mali rice," *Proc. 2015-7th Int. Conf. Knowl. Smart Technol. KST 2015*, pp. 115–120, 2015.
- [5] N. Pratibha, M. Hemlata, M. Krunali, and P. S. T. Khot, "Analysis and Identification of Rice Granules Using Image Processing and Neural Network," vol. 10, no. 1, pp. 25– 33, 2017.
- [6] Indra Abdam Muwakhid, Dewi Nurdiyah . "Otsu Method For Image Finish Segmentation With Components of Hue Saturation Value "TRANSFORMATIKA, Vol.15, No.2, January 2018, pp. 67- 73 ISSN: 1693-3656, journals.usm.ac.id/index.php/transformatika.
- [7] S. Shantaiya and U. Ansari, "Identification Of Food Grains And Its Quality Using Pattern Classification," *Image (Rochester, N.Y.)*, vol. 2, no. 2, pp. 70– 74, 2010.
- [8] G. Van Dalen, "Determination of the size distribution and percentage of broken kernels of rice using flatbed scanning and image analysis," *Food Res. Int.*, vol. 37, no. 1, pp. 51–58, 2004.
- [9] H. S. Gujjar, "A Method for Identification of Basmati Rice grain of Indiaand Its Quality Using Pattern Classification," vol. 3, no. 1, pp. 268–273, 2013.
- [10] D. Sharma and S. D. Sawant, "Grain quality detection by using image processing for public distribution," *Proc. 2017 Int. Conf. Intell. Comput. Control Syst. ICICCS 2017*, vol. 2018– January, pp. 1118–1122, 2018.
- [11] P. Kongsawat, "Quality Assessment of Thai Rice Kernels Using Low Cost Digital Image Processing System," 2018.
- [12] Nagoda, N., & Ranathunga, L. (2018), "Rice Sample Segmentation and Classification Using Image Processing and Support Vector Machine," *2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS)*, (978), 179–184.

- <https://doi.org/10.1109/ICIINFS.2018.8721312>
- [13] Asif, M. J., Shahbaz, T., Tahir Hussain Rizvi, S., & Iqbal, S. (2019), "Rice Grain Identification and Quality Analysis using Image Processing based on Principal Component Analysis," *RAEE 2018 - International Symposium on Recent Advances in Electrical Engineering*, 1–6. <https://doi.org/10.1109/RAEE.2018.8706891>
- [14] Parveen, E. Z. (2017), "Assessment of Quality of Rice Grain using Optical and Image Processing Technique," 265–270.
- [15] Ruslan, R., & Ibrahim, M. F. (2018), "Effect of Background Color on Rice Seed Image Segmentation using Machine Vision," *2018 International Conference on Computational Approach in Smart Systems Design and Applications(ICASSDA)*,1–4.

UCAPAN TERIMA KASIH

- a. Penulis mengucapkan terima kasih terutama kepada Allah SWT,
- b. Achmad Basuki sebagai pembimbing pertama yang selalu memberi arahan dan masukan dalam penulisan journal ini.
- c. Riyanto Sigit sebagai pembimbing pendamping yang selalu memberi ide dan masukan dalam penulisan journal ini.
- d. PDDIKTI yang telah memberikan dana penelitian PTM kepada penulis
- e. Kepada Tim Jurnal Informatika Polbeng yang telah meluangkan waktu untuk membuat template ini sehingga penulis dapat mempublikasikan hasil dari penelitian yang telah dilakukan. Besar harapan semoga karya ilmiah ini dapat bermanfaat bagi pembaca.