KLASIFIKASI NAIVE BAYES UNTUK PENDETEKSIAN STATUS HARAM PADA MAKANAN KEMASAN JEPANG

Teguh Syam Mardiana[1], Ken Kinanti Purnamasari. S.Kom., M.T.[2]

[1,2]Teknik Informatika - Universitas Komputer Indonesia
Jl. Dipatiukur 112-114 Bandung
E-mail : teguhsyamm@gmail.com [1], kinantiken@gmail.com[2]

ABSTRAK

Komposisi merupakan sebuah label yang di dalamnya terdapat tulisan, gambar, atau kombinasi keduanya yang disertakan pada makanan kemasan dengan cara ditempel atau dicetak. Komposisi merupakan bagian yang sangat penting untuk memberikan informasi bahan yang terkandung dalam suatu makanan kemasan sehingga konsumen dapat memperoleh produk yang diharapkan, sehat, serta aman untuk dikonsumsi. Penggunaan bahasa dalam komposisi pada makanan kemasan itu berbeda-beda, salah satunya yaitu yang terdapat pada makanan kemasan Jepang yang menggunakan aksara Jepang sebagai informasi bahan kandungan yang terdapat didalamnya. Dalam penelitian ini dibuat sebuah sistem OCR dengan memanfaatkan citra komposisi sebagai media identifikasinya. Metode klasifikasi yang digunakan adalah metode Naive Bayes yang dikombinasikan dengan metode ekstraksi fitur gabungan Zoning (ICZ dan ZCZ) untuk melakukan pengenalan karakter pada citra komposisi makanan kemasan Jepang yang selanjutnya dapat dilakukan pendeteksian status haram pada makanan kemasan Jepang tersebut. Aksara Jepang yang diklasifikasi hanya karakter ze, ra, chi, dan n yang membentuk bahan haram zerachin (gelatin). Data latih yang digunakan sebanyak 75 karakter yang didalamnya terdapat 5 karakter ze, 11 karakter ra, 18 karakter chi, dan 41 karakter n. Sedangkan data uji yang digunakan sebanyak 20 citra komposisi yaitu 5 citra komposisi halal, 8 citra komposisi haram yang mengandung zerachin, dan 7 citra komposisi haram yang mengandung bahan lain. Berdasarkan beberapa hasil pengujian didapatkan hasil akurasi terbaik sebesar 33,33% pada pengujian yang menggunakan parameter pembagian 6 zona pada ekstraksi fitur Zoning (ICZ dan ZCZ)nya. Didapatkan kesimpulan bahwa metode Naive Bayes dapat mengklasifikasi citra komposisi makanan kemasan Jepang, hanya saja hasil akurasinya masih kurang baik. Hasil akurasi tersebut dipengaruhi dari proses pengujian pada tahap klasifikasi yang juga akurasinya masih rendah.

kata kunci : Kecerdasan Buatan, Pengolahan Citra, OCR, Zoning, Naive Bayes.

1. PENDAHULUAN

Sebelumnya, ada beberapa aplikasi yang sudah dibuat untuk mendeteksi status haram pada makanan kemasan Jepang. Beberapa aplikasi tersebut masih menggunakan barcode sebagai media identifikasinya. Kekurangan penggunaan barcode yaitu harus dilakukan pembaharuan secara berkala seiring dengan munculnya makanan kemasan baru. Pembaharuan tersebut akan memakan banyak kapasitas pada memori sehingga dirasa masih kurang efektif untuk saat ini. Berbeda hal jika ada sebuah sistem Optical Character Recognition (OCR) yang dapat mendeteksi status haram pada makanan kemasan Jepang dengan menggunakan citra komposisi sebagai identifikasinya. Sistem akan melakukan pelatihan terhadap karakter-karakter pembentuk bahan haram, sehingga pada proses pengujian, sistem secara otomatis dapat melakukan pengenalan karakter pada citra komposisi, yang selanjutnya sistem dapat melakukan pendeteksian status haram pada makanan kemasan tersebut.

Naive Bayes merupakan salah satu metode klasifikasi klasik yang menggunakan konsep probabilistik. Sebelumnya metode Naive Bayes pernah diimplementasikan pada penelitian terkait dengan OCR, salah satunya yaitu penelitian untuk pengidentifikasian huruf alphabet tulisan tangan[1]. Hasil dari penelitian tersebut mendapatkan akurasi rata-rata mencapai 80%. Selain itu, metode Naive Bayes juga pernah digunakan untuk mengklasifikasi beberapa kasus lainnya seperti : Pendeteksian penyakit lidah manusia dengan akurasi hampir 85%[2], klasifikasi status gizi dengan akurasi 93,2%[3] dan akurasi paling tinggi 96% pada pengenalan pola untuk pengidentifikasian lima biji-bijian[4]. Dari beberapa fakta didapatkan bahwa metode Naive Bayes masih sangat optimal dalam penggunaannya karena dapat mengklasifikasi beberapa kasus dengan baik.

Zoning adalah salah satu metode ekstraksi fitur yang paling populer untuk sistem OCR. Metode ekstraksi fitur ini sudah pernah diimplementasikan pada penelitian untuk pengenalan aksara sunda huruf cetak dengan akurasi lebih dari 90%[5]. Pada penelitian tersebut menggunakan metode ekstraksi fitur gabungan Zoning ICZ dan ZCZ dengan menghasilkan akurasi yang tinggi dibandingkan

hanya menggunakan metode ICZ atau ZCZ saja[5]. Dengan sama-sama mempunyai tingkat akurasi yang tinggi pada masing-masing penelitian tersebut, maka metode klasifikasi Naive Bayes dan metode ekstraksi fitur Zoning (ICZ dan ZCZ) akan dikombinasikan dan diimplementasikan untuk sistem pendeteksian status haram pada makanan kemasan Jepang di penelitian ini. Penggunaan metode klasifikasi Naive Bayes dan metode ekstraksi fitur Zoning (ICZ dan ZCZ) belum diketahui hasil akurasinya jika dikombinasikan sehingga diharapkan dapat memberikan akurasi yang tinggi di kasus pada penelitian ini.

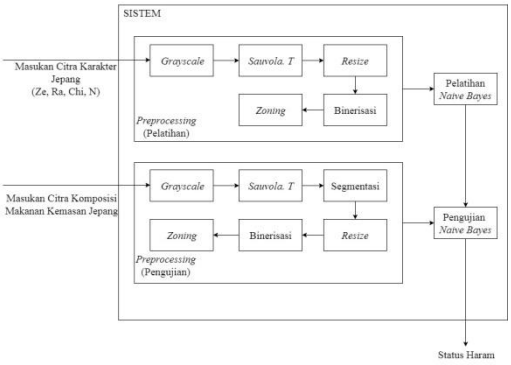
2. ISI PENELITIAN

2.1 Analisis Masalah

Berdasarkan latar belakang dan identifikasi masalah, maka akan dijelaskan tentang analisis masalah yang ditemukan dalam penelitian ini. Adapun masalah di dalam penelitian ini yaitu belum diketahui akurasi metode klasifikasi Naïve Bayes yang dikombinasikan dengan metode ekstraksi fitur gabungan Zoning (ICZ dan ZCZ) yang diimplementasikan pada sebuah sistem OCR pendeteksian status haram pada makanan kemasan Jepang.

2.2 Analisis Proses

Analisis solusi merupakan suatu tahapan dimana Proses yang dilakukan dalam sistem ini dibagi menjadi beberapa bagian, dimana setiap proses memiliki peranan masing-masing untuk pendeteksian status haram untuk makana kemasan Jepang ini.



Gambar 1 Arsitektur Sistem

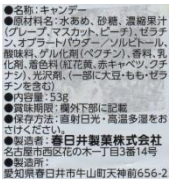
Pada blok diagram di atas, tahapan awal yaitu proses pelatihan dimana pertama kali yang dilakukan adalah memasukan citra aksara Jepang yaitu karakter ze, ra, chi, dan n ke dalam sistem. Langkah selanjutnya adalah melakukan proses preprocessing meliputi grayscale, sauvola threshold, resize, binerisasi dan ekstraksi fitur menggunakan Zoning ICZ & ZCZ sehingga didapatkan suatu nilai array desimal. Selain itu dalam tahapan pengujian yang dilakukan adalah memasukan citra komposisi makanan kemasan Jepang. Selanjutnya adalah melakukan proses preprocessing seperti pada proses pelatihan yang juga nantinya didapatkan suatu nilai array desimal. Nilai array desimal tersebut kemudian akan diolah kembali pada proses klasifikasi dengan metode Naive Bayes baik dalam pelatihan maupun pengujian. Hasil dari klasifikasi akan digunakan untuk mengetahui status haram pada makanan kemasan Jepang tersebut.

2.2.1 Analisis Data Masukan

Contoh data masukan yang digunakan pada proses pelatihan yaitu citra aksara Jepang meliputi karakter ze, ra, chi, dan n seperti pada gambar 2. Sedangkan sebagai data masukan yang digunakan pada proses pengujian yaitu citra komposisi makanan kemasan Jepang seperti pada gambar 3. Kedua data masukan tersebut menggunakan format gambar .jpg, Citra tersebut diperoleh dari pengambilan secara langsung menggunakan kamera ponsel dengan resolusi 8 megapixel dengan jarak 15cm. Pengambilan citra pun tidak terlalu gelap agar aplikasi dapat lebih maksimal dalam melakukan preprocessing.



Gambar 2 Citra Masukan Data Latih



Gambar 3 Citra Masukan Data Uji

2.2.1 Grayscale

Grayscale merupakan proses untuk mengubah citra berwarna (RGB) menjadi citra keabuan yang berfungsi untuk mengecilkan range warna menjadi 0 sampai dengan 255. Berikut adalah rumus konversi citra berwana (RGB) menjadi grayscale.

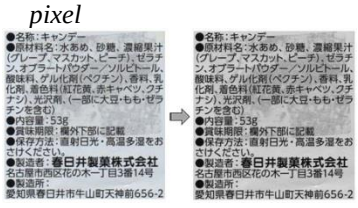
$$I = (0.2989* R)+(0.5870*G)+(0.1141*B) \quad (1)$$

Keterangan :

R : komponen nilai merah (*Red*) dari suatu titik *pixel*

G : komponen nilai hijau (*Green*) dari suatu titik *pixel*

B : komponen nilai biru (*Blue*) dari suatu titik *pixel*



Gambar 4 Citra Hasil Grayscale

2.2.3 Sauvola Threshold

Sauvola threshold merupakan metode threshold yang mampu memberikan peningkatan dengan cara menghitung ambang menggunakan rentang nilai dinamis dari nilai standar deviasi citra grayscale. Sauvola threshold memiliki keunggulan

dibandingkan dengan metode thresholding lainnya seperti metode bernsen, eikvil, parker, niblack dan otsu[20][21]. Berikut adalah rumus sauvola threshold.

$$m(x,y) = \frac{\sum_{i=min}^{i=max}\sum_{j=min}^{j=max} img(i,j)}{(max\ i - min\ i + 1) * (max\ i - min\ i + 1)} \quad (2)$$

$$s(x,y) = \sqrt{\frac{\sum_{i=min}^{i=max}\sum_{j=min}^{j=max}(img(i,j) - m(x,y))^2}{((max\ i - min\ i + 1) * (max\ i - min\ i + 1)) - 1}} \quad (3)$$

$$T(x,y) = m(x,y) * \left(1 + k * \left(\frac{s(x,y)}{R} - 1\right)\right) \quad (4)$$

Keterangan :

R : Nilai maksimum dari standar deviasi (128 untuk citra *grayscale*)

k : *Kernel* dengan nilai antara 0.2 – 0.5.

m: Fungsi yang menghasilkan nilai rata-rata dari sejumlah *pixel* citra.

s : Fungsi yang menghasilkan nilai standar deviasi dari sejumlah *pixel* citra

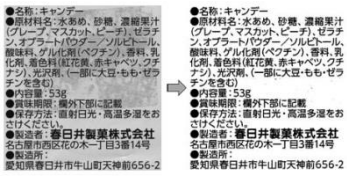
T : Fungsi yang menghasilkan nilai *threshold* (ambang).

x : Nilai koordinat lebar citra.

y : Nilai koordinat tinggi citra.

Setelah nilai ambang T(x,y) sudah didapatkan, selanjutnya masukkan ke persamaan berikut.

$$f(x,y) = \begin{cases} 0, & \text{img}(x,y) < \text{T}(x,y) \\ 255, & \text{img}(x,y) \geq \text{T}(x,y) \end{cases} \quad (5)$$



Gambar 5 Citra Hasil Threshold

2.2.4 Segmentasi

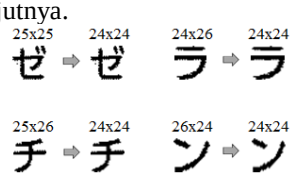
Segmentasi merupakan teknik untuk membagi suatu citra menjadi beberapa daerah (region) dimana setiap daerah memiliki kemiripan atribut. Dalam penelitian ini segmentasi dilakukan dengan pemotongan baris dan kolom citra. Pemotongan pertama dilakukan untuk seluruh baris citra tulisan tangan, kemudian dilanjutkan dengan pemotongan kolom untuk setiap citra hasil pemotongan baris.

Tabel 1 Proses Segmentasi

Pemotongan Baris	1	(グレープ、マスカット、ピーチ)、ゼラチ
	2	ン、オブラートパウダー／ソルビトール、
Pemotongan Kolom	1	(グレープ、マスカット、ピーチ)、ゼラチ
	2	ン、オブラートパウダー／ソルビトール、

2.2.5 Resize

Resize citra artinya mengubah besarnya ukuran citra digital dalam pixel. Pada proses ini, setiap karakter yang ada akan dipaksa berada dalam ukuran yang sudah ditentukan sehingga bentuknya tidak proporsional. Tetapi jika hal ini dilakukan secara konsisten dan menyeluruh pada setiap karakter maka tidak akan merusak informasi yang diperlukan untuk proses selanjutnya.

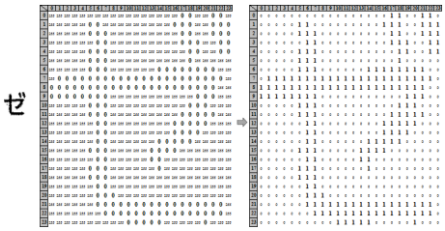


Gambar 6 Citra Hasil Resize

2.2.6 Binerisasi

Citra biner adalah citra dengan setiap pixel hanya dinyatakan dengan sebuah nilai dari dua kemungkinan yaitu nilai 0 dan 1. Citra yang dikonversi menjadi biner biasanya dari citra dengan warna hitam putih. Dalam penelitian ini citra biner digunakan sebagai data masukan pada tahap klasifikasi. Berikut adalah persamamaan binerisasi citra.

$$f(x) = \begin{cases} 0, & x = 0 \\ 1, & x = 255 \end{cases} \quad (6)$$



Gambar 7 Citra Karakter Hasil Binerisasi

2.2.7 Ekstraksi Fitur Zoning

Ekstraksi ciri Zoning merupakan, ekstraksi ciri dengan konsep utamanya adalah pembagian citra kedalam sejumlah zona yang sama agar dapat dikenali ciri dari setiap karakter huruf dan angka [12], yang nantinya akan menghasilkan sebuah nilai yang akan diproses pada tahapan klasifikasi. Langkah-langkah dalam algoritma Zoning ICZ (Image Centroid Zone).

1. Hitung nilai centroid citra masukan.

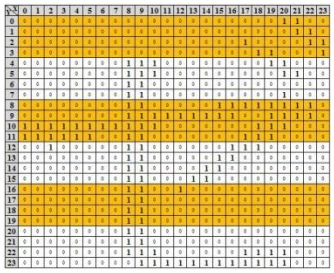
$$xc = \frac{\sum_{y=0}^{y=max}\sum_{x=0}^{x=max} img(x,y) * Bx}{t}$$

$$yc = \frac{\sum_{x=0}^{x=max}\sum_{y=0}^{y=max} img(x,y) * By}{t}$$



Gambar 8 Contoh Karakter Yang Akan Diekstraksi

2. Bagi ke dalam jumlah zona yang sama. Dalam penelitian ini menggunakan beberapa pembagian zona, yaitu 6, 8, dan 12 zona. Sebagai gambaran berikut contoh pembagian 6 zona dan perhitungannya.



Gambar 9 Matrik Zoning 24x24 Karakter Ze

3. Hitung jarak piksel hitam dari centroid pada setiap zona.

$$d(p,c) = \sqrt{(px - xc)^2 + (py - yc)^2}$$

atau bisa diturunkan sebagai

$$d((px,py),(xc,yc)) = \sqrt{(px - xc)^2 + (py - yc)^2}$$

Perhitungan jarak piksel terhadap centroid pada zona 1 dapat dilihat di bawah. Lakukan pada semua zona.

Perhitungan pada Zona 1

$d((px,py),(xc,yc)) = \sqrt{(px - xc)^2 + (py - yc)^2}$

$Jarak_{1.1}((20,0),(12,12)) = \sqrt{(20-12)^2 + (0-12)^2} = 14{,}4222051$

$Jarak_{2.1}((21,0),(12,12)) = \sqrt{(21-12)^2 + (0-12)^2} = 15$

…

$Jarak_{13.1}((23,23),(12,12)) = \sqrt{(23-12)^2 + (3-12)^2} = 14{,}2126704$

4. Hitung jarak rata-rata antara titik

$$\sum_{m=1}^{m=6} RJ_m = \frac{Jarak_{1.m} + Jarak_{2.m} + Jarak_{3.m} + \cdots Jarak_{n.m}}{t_m}$$

Perhitungan jarak rata-rata pada zona 1 dapat dilihat di bawah. Lakukan pada semua zona.

Perhitungan pada Zona 1

$$\sum_{m=1}^{m=6} RJ_m = \frac{Jarak_{1.1} + Jarak_{2.1} + Jarak_{3.1} + \cdots Jarak_{n.1}}{t_m}$$

$$\text{Rataan Jarak}_1 = \frac{Jarak_{1.1} + Jarak_{2.1} + Jarak_{3.1} + \cdots Jarak_{12.1}}{t_1}$$

$$= \frac{164{,}7518879}{12} = 12{,}67322214734602$$

5. Didapatkan 6 nilai sebagai berikut.

Ze1|ICZZona 1 = 12,67322214734602
Ze1|ICZZona 2 = 12,012497207679937
Ze1|ICZZona 3 = 12,447487267849914
Ze1|ICZZona 4 = 11,409830874191181
Ze1|ICZZona 5 = 11,185689842746477
Ze1|ICZZona 6 = 10,709418276009616

Langkah-langkah untuk ekstraksi fitur Zoning ZCZ (Zone Centroid Zone).

1. Langkah yang dilakukan sama dengan pada Zoning ICZ, Hanya saja nilai fitur yang harus dilakukan adalah pembagian zona terlebih dahulu sebelum dilakukan perhitungan centroid. Sehingga pada Zoning ZCZ didapatkan beberapa nilai centroid pada setiap zona.
2. Nilai yang dihasilkan sebagai berikut.

Ze1|ZCZZona 1 = 4,721785887569821
Ze1|ZCZZona 2 = 5,2598512621445055
Ze1|ZCZZona 3 = 6,266306046689112
Ze1|ZCZZona 4 = 4,025892101809964
Ze1|ZCZZona 5 = 1,46964141804604
Ze1|ZCZZona 6 = 4,095808930800819

Tabel 2 Nilai Fitur Hasil Ekstraksi

Karakter	Nilai Fitur ICZ	Nilai Fitur ZCZ
ゼ	Zona 1 = 12,67322214734602 Zona 2 = 12,012497207679937 Zona 3 = 12,447487267849914 Zona 4 = 11,409830874191181 Zona 5 = 11,185689842746477 Zona 6 = 10,709418276009616	Zona 1 = 4,721785887569821 Zona 2 = 5,2598512621445055 Zona 3 = 6,266306046689112 Zona 4 = 4,025892101809964 Zona 5 = 1,46964141804604 Zona 6 = 4,095808930800819
ゼ	Zona 1 = 12,1520920346993395 Zona 2 = 11,517637496096641 Zona 3 = 12,9180933756966833 Zona 4 = 11,774304565748105 Zona 5 = 11,71740464721132 Zona 6 = 11,204265885334451	Zona 1 = 6,585186118788576 5 Zona 2 = 5,343669893271543 Zona 3 = 6,028100403797702 Zona 4 = 5,386656689061503 Zona 5 = 2,033548153373633 Zona 6 = 3,7892566232574176
ラ	Zona 1 = 9,815704917563945 Zona 2 = 9,2975211332706211 Zona 3 = 11,375360590190143 Zona 4 = 9,147618075919464 Zona 5 = 7,456429991169783 Zona 6 = 9,39874519725485	Zona 1 = 5,27003864952598 Zona 2 = 6,4039242156787 49 Zona 3 = 7,673956413475698 Zona 4 = 1,4128990204491 96 Zona 5 = 1,729126786466033 8 Zona 6 = 2,1459737457608 09
ラ	Zona 1 = 10,685100190679 76 Zona 2 = 0 Zona 3 = 12,344423773295137 Zona 4 = 12,218913863708 77 Zona 5 = 10,562231499771158 Zona 6 = 10,456940163557647	Zona 1 = 4,338280731159644 Zona 2 = 0 Zona 3 = 6,258517038284154 Zona 4 = 1,330056307974577 Zona 5 = 1,627277986846677 6 Zona 6 = 3,133835219296107
チ	Zona 1 = 8,012749103327335 Zona 2 = 9,41841302575218 Zona 3 = 10,158762579985678 Zona 4 = 7,592153617140983 Zona 5 = 7,656575506990917 Zona 6 = 8,965168581122711	Zona 1 = 2,616887652671708 Zona 2 = 4,465719115888621 Zona 3 = 5,759863980887081 Zona 4 = 1,258061887780747 5 Zona 5 = 1,412335645338953 7 Zona 6 = 1,984631787785905 4
チ	Zona 1 = 9,141663910276941 Zona 2 = 8,798054549379625 Zona 3 = 10,0899165195443 89 Zona 4 = 8,740215188717475 Zona 5 = 8,623828443755032 Zona 6 = 10,363979509874575	Zona 1 = 4,171640436000164 Zona 2 = 1,275412073846563 7 Zona 3 = 4,920798851643807 Zona 4 = 1,412696783230135 Zona 5 = 1,530421515165979 Zona 6 = 2,098183374044004 3
ン	Zona 1 = 13,439618456275456 Zona 2 = 14,967811074186626 Zona 3 = 14,609756273433398 Zona 4 = 12,38956051600307 Zona 5 = 11,948957391616256 Zona 6 = 13,439973808660172	Zona 1 = 3,395414262140956 Zona 2 = 8,593759175857413 Zona 3 = 1,516123775561495 Zona 4 = 1,725456785586119 8 Zona 5 = 2,754679910652 39 Zona 6 = 1,868587419059603 5
ン	Zona 1 = 12,979078303553663 Zona 2 = 13,685677959551846 Zona 3 = 13,321328228344745 Zona 4 = 11,950069732803595 Zona 5 = 10,614544829711189 Zona 6 = 12,604500897567021	Zona 1 = 4,801682613375976 Zona 2 = 7,104436990221722 Zona 3 = 1,521482441874218 8 Zona 4 = 1,662769960302215 Zona 5 = 2,085941674157796 Zona 6 = 2,987641033811849

2.2.8 Klasifikasi Naive Bayes

Naive Bayes adalah suatu metode pengklasifikasian paling sederhana dengan menggunakan peluang yang ada, dimana diasumsikan bahwa setiap variable x bersifat bebas (independence). Data yang digunakan dapat bersifat kategorial (diskrit) maupun kontinyu. Namun, pada penelitian ini akan digunakan data kontinyu, karena hasil ekstraksi fitur citra merupakan data kontinyu berupa angka hasil penghitungannya. Maka untuk data kontinyu dapat diselesaikan dengan menggunakan langkah-langkah sebagai berikut.

2.2.8.1 Pelatihan

1. Menghitung nilai mean

$$\mu = \frac{\sum x_i}{\text{n}} \qquad (7)$$

Berikut adalah contoh perhitungannya. Lakukan untuk semua nilai fitur pada setiap kelas yang ada.

$$\mu = \frac{\sum \text{Nilai Fitur Ze|ICZ}_{\text{Zona 1}}}{\text{n}}$$

$$\mu = \frac{\text{Ze}_1\text{|ICZ Zona 1} + \text{Ze}_2\text{|ICZ Zona 1}}{\text{n}}$$

$$\mu = \frac{12{,}67322214734602 + 12{,}152092034699395}{2}$$

$$\mu = \frac{24{,}82531418}{2} = 12{,}41265709$$

2. Menghitung variance

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2 \quad (8)$$

Berikut adalah contoh perhitungannya. Lakukan untuk semua nilai fitur pada setiap kelas yang ada.

$$\sigma^2 = \frac{1}{2-1}\sum_{i=1}^{n}(\text{Nilai Fitur Ze}|\text{ICZ}_{\text{Zona 1}} - \mu)^2$$

$$\sigma^2 = \frac{1}{2-1}.(12{,}67322214734602 - 12{,}41265709)^2 + (12{,}15209203469939 5 - 12{,}41265709)^2 = 0{,}135788297$$

Sehingga didapatkan rekapitulasi nilai mean dan variance sebagai berikut.

Tabel 3 Nilai Mean & Variance

Kelas Karakter	Fitur	μ	σ^2	Kelas Karakter	Fitur	μ	σ^2
Ze	$ICZ_{Zona\ 1}$	12,41265709	0,135788297	Ra	$ICZ_{Zona\ 1}$	10,25040255	0,37792407
	$ICZ_{Zona\ 2}$	11,76506735	0,122443067		$ICZ_{Zona\ 2}$	4,648760567	43,22194961
	$ICZ_{Zona\ 3}$	12,68279032	0,110735054		$ICZ_{Zona\ 3}$	11,85989218	0,469541726
	$ICZ_{Zona\ 4}$	11,59206772	0,066420536		$ICZ_{Zona\ 4}$	10,68326597	4,716428908
	$ICZ_{Zona\ 5}$	11,45154724	0,141360317		$ICZ_{Zona\ 5}$	9,009330745	4,823001505
	$ICZ_{Zona\ 6}$	10,95684208	0,122437078		$ICZ_{Zona\ 6}$	9,92784268	0,559888293
	$ZCZ_{Zona\ 1}$	5,653486003	1,736130211		$ZCZ_{Zona\ 1}$	4,80415969	0,434086409
	$ZCZ_{Zona\ 2}$	5,301760578	0,005512781		$ZCZ_{Zona\ 2}$	3,201962108	20,50512268
	$ZCZ_{Zona\ 3}$	6,147203225	0,028370964		$ZCZ_{Zona\ 3}$	6,966236726	1,001734312
	$ZCZ_{Zona\ 4}$	4,706274395	0,925840131		$ZCZ_{Zona\ 4}$	1,371477664	0,003431458
	$ZCZ_{Zona\ 5}$	1,751594786	0,158995403		$ZCZ_{Zona\ 5}$	1,678202387	0,005186589
	$ZCZ_{Zona\ 6}$	3,942532777	0,046987159		$ZCZ_{Zona\ 6}$	2,639904483	0,487935145
Chi	$ICZ_{Zona\ 1}$	8,577206507	0,637224321	N	$ICZ_{Zona\ 1}$	13,20934838	0,106048616
	$ICZ_{Zona\ 2}$	9,108233788	0,19242232		$ICZ_{Zona\ 2}$	14,32674452	0,821932662
	$ICZ_{Zona\ 3}$	10,12433955	0,00236989		$ICZ_{Zona\ 3}$	13,96554225	0,830023414
	$ICZ_{Zona\ 4}$	8,166184403	0,659022686		$ICZ_{Zona\ 4}$	12,16981512	0,096576074
	$ICZ_{Zona\ 5}$	8,140201975	0,467789122		$ICZ_{Zona\ 5}$	11,28175111	0,890328443
	$ICZ_{Zona\ 6}$	9,664574045	0,978336007		$ICZ_{Zona\ 6}$	13,02223735	0,349007493
	$ZCZ_{Zona\ 1}$	3,394264044	1,208628109		$ZCZ_{Zona\ 1}$	4,098548438	0,988795338
	$ZCZ_{Zona\ 2}$	2,870565595	5,089029511		$ZCZ_{Zona\ 2}$	7,849098083	1,109040286
	$ZCZ_{Zona\ 3}$	5,340331416	0,352015146		$ZCZ_{Zona\ 3}$	1,518803109	1,43577E
	$ZCZ_{Zona\ 4}$	1,335379336	0,011955975		$ZCZ_{Zona\ 4}$	1,694113373	0,001964819
	$ZCZ_{Zona\ 5}$	1,47137858	0,006972136		$ZCZ_{Zona\ 5}$	2,420310792	0,223605414
	$ZCZ_{Zona\ 6}$	2,041407581	0,006446981		$ZCZ_{Zona\ 6}$	2,428114226	0,626140496

2.2.8.2 Pengujian

Berikut adalah potongan karakter citra kompsisi makanan kemasan Jepang yang dilakukan pengujian.

Tabel 4 Contoh Karakter Yang Akan Diuji

No	Karakter	Nilai Fitur ICZ	Nilai Fitur ZCZ
1.		Zona 1 = 13,025233469603139 Zona 2 = 11,410754257617823 Zona 3 = 13,791873159443282 Zona 4 = 11,705892525941424 Zona 5 = 12,284985913444595 Zona 6 = 11,432536236765731	Zona 1 = 4,559097203215948 Zona 2 = 5,897394017768255 Zona 3 = 4,994586374032392 Zona 4 = 5,094496037888883 Zona 5 = 1,082842712474619 Zona 6 = 3,885957798383947
2.		Zona 1 = 11,482920100648125 Zona 2 = 9,739332964963 53 Zona 3 = 11,966315386442417 Zona 4 = 10,812547792676513 Zona 5 = 9,019370143936545 Zona 6 = 10,420798224556306	Zona 1 = 5,305410615550341 5 Zona 2 = 5,952380952380952 6 Zona 3 = 6,245963343061068 Zona 4 = 1,337706967465351 4 Zona 5 = 2,024604141377068 Zona 6 = 3,282619190207 08

Untuk mengetahui termasuk kelas apa karakter citra uji tersebut. Maka langkah pengujian dalam algortima Naive Bayes harus dilakukan, berikut tahapannya adalah sebagai berikut.

1. Mencari probabilitas (*prior*) untuk masing-masing kelas pada data latih. Caranya yaitu, menghitung jumlah data tiap kelas dibagi jumlah total data secara keseluruhan

   Prior karakter Ze = 5/75=0,0666
   Prior karakter Ra = 11/75=0,1466
   Prior karakter Chi = 18/75=0,24
   Prior karakter N = 41/75=0,5466

2. Menghitung densitas probabilitasnya

$$\varphi_{\mu\sigma}(x) = \frac{1}{\sqrt{2\pi(\sigma^2)}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (9)$$

Berikut adalah contoh perhitungannya. Lakukan untuk semua nilai fitur pada setiap kelas yang ada.

Perhitungan pada kelas karakter Ze

$$= \frac{1}{\sqrt{2*3{,}14*(0{,}135788297)}}\, e^{-\frac{(13{,}025233469603139 - 12{,}41265709)^2}{2*0{,}135788297}}$$

$$= \frac{1}{\sqrt{6{,}28*0{,}135788297}}\, e^{-\frac{0{,}61257639^2}{0{,}271576594}}$$

$$= \frac{1}{\sqrt{0{,}852750505}}\, e^{-\frac{0{,}375249821}{0{,}271576594}}$$

$$= \frac{1}{0{,}923444912}\, e^{-\frac{0{,}375249821}{0{,}271576594}}$$

$$= 1{,}08290163\ e^{-1{,}38174581}$$

$$= 1{,}08290163 * 0{,}251139728$$

$$= 0{,}271959621$$

Sehingga didapatkan rekapitulasi nilai densitas fitur setiap kelas.

Tabel 5 Nilai Densitas Setiap Kelas

Kelas Karakter	Fitur	φ	Kelas Karakter	Fitur	φ
Ze	$ICZ_{Zona\ 1}$	0,271959621	Ra	$ICZ_{Zona\ 1}$	2,44478E-5
	$ICZ_{Zona\ 2}$	0,682995003		$ICZ_{Zona\ 2}$	0,035764128
	$ICZ_{Zona\ 3}$	0,004642665		$ICZ_{Zona\ 3}$	0,010939668
	$ICZ_{Zona\ 4}$	0,404468274		$ICZ_{Zona\ 4}$	0,164462117
	$ICZ_{Zona\ 5}$	0,090956268		$ICZ_{Zona\ 5}$	0,059740062
	$ICZ_{Zona\ 6}$	0,452624419		$ICZ_{Zona\ 6}$	0,070608864
	$ZCZ_{Zona\ 1}$	0,214500263		$ZCZ_{Zona\ 1}$	0,565184027
	$ZCZ_{Zona\ 2}$	7,88943E-22		$ZCZ_{Zona\ 2}$	0,073815848
	$ZCZ_{Zona\ 3}$	1,60783E-10		$ZCZ_{Zona\ 3}$	0,057275129
	$ZCZ_{Zona\ 4}$	0,382298966		$ZCZ_{Zona\ 4}$	0
	$ZCZ_{Zona\ 5}$	0,245203398		$ZCZ_{Zona\ 5}$	8,01065E-15
	$ZCZ_{Zona\ 6}$	0,779257011		$ZCZ_{Zona\ 6}$	0,116374853
Chi	$ICZ_{Zona\ 1}$	9,05209E-8	N	$ICZ_{Zona\ 1}$	0,044375912
	$ICZ_{Zona\ 2}$	9,46397E-7		$ICZ_{Zona\ 2}$	0,002495663
	$ICZ_{Zona\ 3}$	0		$ICZ_{Zona\ 3}$	0,430114837
	$ICZ_{Zona\ 4}$	3,65681E-5		$ICZ_{Zona\ 4}$	0,421369311
	$ICZ_{Zona\ 5}$	6,18603E-9		$ICZ_{Zona\ 5}$	0,240308391
	$ICZ_{Zona\ 6}$	0,081660384		$ICZ_{Zona\ 6}$	0,018082068
	$ZCZ_{Zona\ 1}$	0,207061275		$ZCZ_{Zona\ 1}$	0,360484843
	$ZCZ_{Zona\ 2}$	0,071907855		$ZCZ_{Zona\ 2}$	0,068033838
	$ZCZ_{Zona\ 3}$	0,567543567		$ZCZ_{Zona\ 3}$	0
	$ZCZ_{Zona\ 4}$	8,1738E-257		$ZCZ_{Zona\ 4}$	0
	$ZCZ_{Zona\ 5}$	9,49911E-5		$ZCZ_{Zona\ 5}$	0,015456918
	$ZCZ_{Zona\ 6}$	1,2527E-114		$ZCZ_{Zona\ 6}$	0,092389301

3. Setelah mendapatkan nilai densitas setiap karakter, selanjutnya yaitu mencari nilai *posterior* untuk masing-masing kelas.

$$Posterior = \frac{Prior * Likehood}{Evidence} \quad (10)$$

Sebelumnya dilakukan pencarian nilai *evidence* terlebih dahulu seperti dibawah ini.

Evidence =
(*Prior* Ze) * φ(Ze|$ICZ_{Zona\ 1}$) * φ(Ze|$ICZ_{Zona\ 2}$) * φ(Ze|$ICZ_{Zona\ 3}$) * φ(Ze|$ICZ_{Zona\ 4}$) * φ(Ze|$ICZ_{Zona\ 5}$) * φ(Ze|$ICZ_{Zona\ 6}$) * φ(Ze|$ZCZ_{Zona\ 1}$) * φ(Ze|$ZCZ_{Zona\ 2}$) * φ(Ze|$ZCZ_{Zona\ 3}$) * φ(Ze|$ZCZ_{Zona\ 4}$) * φ(Ze|$ZCZ_{Zona\ 5}$) * φ(Ze|$ZCZ_{Zona\ 6}$) +
(*Prior* Ra) * φ(Ra|$ICZ_{Zona\ 1}$) * φ(Ra|$ICZ_{Zona\ 2}$) * φ(Ra|$ICZ_{Zona\ 3}$) * φ(Ra|$ICZ_{Zona\ 4}$) * φ(Ra|$ICZ_{Zona\ 5}$) * φ(Ra|$ICZ_{Zona\ 6}$) * φ(Ra|$ZCZ_{Zona\ 1}$) * φ(Ra|$ZCZ_{Zona\ 2}$) * φ(Ra|$ZCZ_{Zona\ 3}$) * φ(Ra|$ZCZ_{Zona\ 4}$) * φ(Ra|$ZCZ_{Zona\ 5}$) * φ(Ra|$ZCZ_{Zona\ 6}$) +
(*Prior* Chi) * φ(Chi|$ICZ_{Zona\ 1}$) * φ(Chi|$ICZ_{Zona\ 2}$) * φ(Chi|$ICZ_{Zona\ 3}$) * φ(Chi|$ICZ_{Zona\ 4}$) * φ(Chi|$ICZ_{Zona\ 5}$) * φ(Chi|$ICZ_{Zona\ 6}$) * φ(Chi|$ZCZ_{Zona\ 1}$) * φ(Chi|$ZCZ_{Zona\ 2}$) * φ(Chi|$ZCZ_{Zona\ 3}$) * φ(Chi|$ZCZ_{Zona\ 4}$) * φ(Chi|$ZCZ_{Zona\ 5}$) * φ(Chi|$ZCZ_{Zona\ 6}$) +
(*Prior* N) * φ(N|$ICZ_{Zona\ 1}$) * φ(N|$ICZ_{Zona\ 2}$) * φ(N|$ICZ_{Zona\ 3}$) * φ(N|$ICZ_{Zona\ 4}$) * φ(N|$ICZ_{Zona\ 5}$) * φ(N|$ICZ_{Zona\ 6}$) * φ(N|$ZCZ_{Zona\ 1}$) * φ(N|$ZCZ_{Zona\ 2}$) * φ(N|$ZCZ_{Zona\ 3}$) * φ(N|$ZCZ_{Zona\ 4}$) * φ(N|$ZCZ_{Zona\ 5}$) * φ(N|$ZCZ_{Zona\ 6}$) =

Didapatkan hasil.

1,90083E-39 + 0 + 0 + 0 = 1,90083E-39

Nilai tersebut akan digunakan pada selanjutnya yaitu menghitung *posterior*. Berikut contoh perhtiungan pada karakter ze.

***Posterior* Karakter Ze =**

$$\frac{(Prior\ Ze) * \varphi(Ze|ICZZona\ 1) * \varphi(Ze|ICZZona\ 2) * \varphi(Ze|ICZZona\ 3) * \varphi(Ze|ICZZona\ 4) * \varphi(Ze|ICZZona\ 5) * \varphi(Ze|ICZZona\ 6) * \varphi(Ze|ZCZZona\ 1) * \varphi(Ze|ZCZZona\ 2) * \varphi(Ze|ZCZZona\ 3) * \varphi(Ze|ZCZZona\ 4) * \varphi(Ze|ZCZZona\ 5) * \varphi(Ze|ZCZZona\ 6)}{Evidence}$$

$$\frac{0{,}25 * 0{,}271959621 * 0{,}682995003 * 0{,}004642665 * 1{,}404468274 * 0{,}090956268 * 0{,}452624419 * 0{,}214500263 * 7{,}88943E-22 * 1{,}60783E-10 * 0{,}382298966 * 0{,}245203398 * 1{,}779257011}{5{,}6571E-36}$$

$$\frac{1{,}90083E-39}{1{,}90083E-39} = 1$$

Lakukan perhitungan *posterior* untuk semua kemungkinan karakter.

Tabel 6 Hasil Perhitungan Posterior

Posterior Karakter	Hasil
Ze	1
Ra	0
Chi	0
N	0

Dari hasil perhitungan posterior pada setiap kelas di atas, carilah nilai terbesar, dimana nilai posterior terbesar menunjukkan hasil pengenalan kelas untuk data uji tersebut yaitu karakter ze. Lakukan untuk semua karakter uji untuk dilakukan klasifikasi pengujian, sehingga didapatkan hasil pengenalan pada bebeberapa karakter uji pada citra komposisi makanan kemasan tersebut.

Tabel 7 Hasil Pengenalan

No	Citra Karakter	Kelas	Hasil Pengenalan	Keterangan	No	Citra Kar2akter	Kelas	Hasil Pengenalan	Keterangan
1.	ゼラチン	Ze	Ze	Benar	11.	パウダー	X	N	Salah
2.	ラ	Ra	Ra	Benar	12.	ー	X	N	Salah
3.	チ	Chi	Chi	Benar	13.	ー	X	Chi	Salah
4.	ン	N	N	Benar	14.	ー	X	Chi	Salah
5.	ー	X	Ze	Salah	15.	ー	X	N	Salah
6.	ー	X	N	Salah	16.	ー	X	N	Salah
7.	ー	X	N	Salah	17.	ー	X	Ra	Salah
8.	ー	Ra	Ra	Benar	18.	ー	X	N	Salah
9.	ー	X	N	Salah	19.	ー	X	N	Salah
10.	ー	X	Chi	Salah	20.	ー	X	N	Salah

Dari tabel analisis rencana pengujian di atas, didapatkan kesimpulan sebagai berikut.

a. Jumlah karakter benar = 5
b. Jumlah karakter salah = 15
c. Jumlah keseluruhan karakter = 20

Kemudian dihitung akurasi pengenalannya sehingga didapatkan hasil sebagai berikut.

$$\text{Akurasi} = \frac{\text{Jumlah Karakter Benar}}{\text{Jumlah Keseluruhan Karakter}} * 100\%$$

$$\text{Akurasi} = \frac{5}{20} * 100\% = 25\%$$

2.2.9 Pendeteksian Status Haram

Pada proses ini, sistem akan melakukan penelusuran dan pencocokan bahan haram yang tedapat pada listHaram terhadap hasil pengenalan karakter pada citra komposisi makanan kemasan Jepang yang sudah dilakukan sebelumnya. Ada beberapa kondisi dimana sistem dapat melakukan pendeteksian status haram, seperti terlihat pada tabel berikut.

Tabel 8 Kondisi Pendeteksian Status

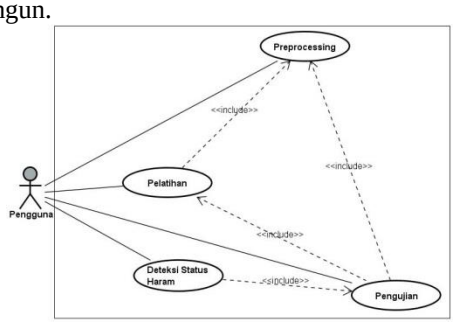
No	Hasil Pengenalan	Status
1.	Ze-Ra-Chi-N	Haram
2.	Ze-Ra-Chi-X	Halal
3.	Ze-Ra-X-N	Halal
4.	Ze-X-Chi-N	Halal
5.	X-Ra-Chi-N	Halal
6.	X-X-X-X	Halal

Tabel 3.23 di atas menjelaskan bahwa jika hasil pengenalan karakter sesuai dan berurutan seperti pada kondisi no. 1 yaitu Ze-Ra-Chi-N, maka status dari makanan kemasan Jepang tersebut adalah haram. Selain kondisi tersebut, maka status yang didapat adalah halal.

Dari hasil pengenalan pada kasus yang sudah dilakukan sebelumnya, maka status yang didapatkan adalah haram pada citra masukan komposisi makanan kemasan tersebut.

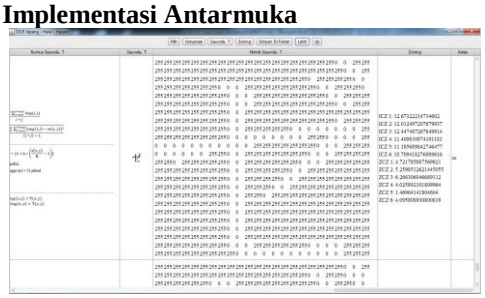
2.3 Use Case Diagram

Use case diagram merupakan pemodelan yang menunjukkan interaksi antara sistem dengan pengguna (aktor). Dalam *use case* yang dibuat menggunakan lima *use case*, dimana ada *use case* yang berelasi dengan *use case* lainnya. Berikut adalah *use case diagram* untuk aplikasi yang dibangun.

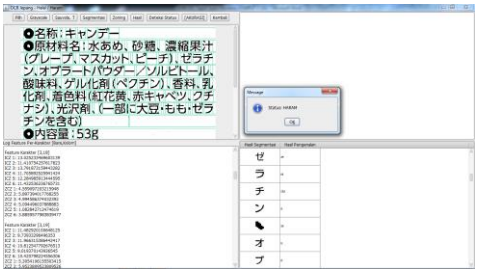


Gambar 8 Use Case Diagram

2.4 Implementasi Antarmuka



Gambar 9 Antarmuka Menu Pelatihan



Gambar 10 Antarmuka Menu Pengujian

2.5 Pengujian Akurasi

Dalam penelitian ini pengujian akurasi yaitu melakukan proses pengujian dengan banyaknya citra komposisi makanan kemasan Jepang yang dikenali sebagai zerachin di dalam sistem terhadap banyaknya citra komposisi haram keseluruhan. Dalam penelitian ini pengujian akurasi digunakan beberapa parameter pada proses ekstraksi fitur Zoning-nya, yaitu pembagian dengan beberapa zona seperti pembagian 6 zona, 8 zona, dan 12 zona. Tabel 4.11 merupakan beberapa parameter yang digunakan.

Tabel 8 Hasil Pengujian

Pengujian ke-	Akurasi
1. Pembagian 6 zona	33,33%
2. Pembagian 8 zona	20%
3. Pembagian 12 zona	0%

Dari tabel hasil pengujian diatas, didapatkan Pengujian terbaik pada parameter ke-1, dimana didapatkan akurasi sebesar 33,33%.

3. PENUTUP

3.1 Kesimpulan

Berdasarkan hasil pengujian dari sistem pendeteksian status haram makanan kemasan Jepang menggunakan metode ekstraksi fitur Zoning (ICZ dan ZCZ) yang dikombinasikan dengan metode klasifikasi Naive Bayes memperoleh akurasi tertinggi mencapai 33,33% pada hasil pengujian yang menggunakan parameter pembagian 6 zona pada ekstraksi fiturnya. Didapatkan kesimpulan bahwa metode Naive Bayes dapat mengklasifikasi citra komposisi makanan kemasan Jepang, hanya saja hasil akurasinya masih kurang baik. Hasil akurasi tersebut dipengaruhi dari proses pengujian pada tahap klasifikasi yang juga akurasinya masih rendah.

3.2 Saran

Agar penelitian berikutnya mengenai pendeteksi status haram makanan kemasan Jepang mendapatkan akurasi yang lebih tinggi, maka berikut adalah saran yang dapat dijadikan pertimbangan, yaitu :

1. Menggunakan metode segmentasi lainnya seperti metode deteksi tepi modern Canny atau metode segmentasi lainnya yang dapat memisahkan semua karakter dengan lebih baik pada citra uji.
2. Dapat menggunakan metode klasifikasi lainnya seperti Support Vector Machine (SVM) atau Deep Learning sebagai perbandingan.

DAFTAR PUSTAKA

[1] Akhmad R. M. Rizky B. S. Akhmad F. H. Ach Dwi A. Nuri N. A. U, "Optical Character Recognition Dengan Metode Naive Bayes".

[2] Fahmi Riekal, "Analisis Gray Level Difference Method Dan Metode Naive Bayes Mengidentifikasi Penyakit Lidah Manusia", Universitas Komputer Indonesia, Bandung, 2016.

[3] Sri K, "Klasifikasi Status Gizi Menggunakan Naive Bayesian Classification" Universitas Islam Indonesia, Yogyakarta.

[4] Kusmian A, "Implementasi Algoritma Naive Bayes Untuk Klasifikasi Citra Berdasarkan Ekstraksi Ciri Tekstur Dengan Metode Matriks Kookurensi", Universitas Komputer Indonesia, Bandung, 2015.

[5] I. Mulia, "Pengenalan Aksara Sunda Menggunakan Ekstraksi Ciri Zoning Dan Klasifikasi Support Vector Machine", Institut Pertanian Bogor, Bogor, 2012.

[6] Sugiyono, "Metode Penelitian Kuantitatif, Kualitatif, dan Kombinasi (Mixed Methods), Bandung : Alfabeta, 2014.

[7] Sutoyo T. dkk, "Teori Pengolahan Citra Digital", Yogyakarta : Andi.

[8] U. Ahmad, "Pengolahan Citra Digital Dan Teknik Pemrogramannya", Yogyakarta : Graha Ilmu.

[9] Purnama. Agus, "Operasi Pengolahan Citra Digital", Tersedia di http://elektronika-dasar.web.id/operasi-pengolahan-citra-digital/, [Di akses 28 April 2017].

[10] A. Kadir. A. Susanto, "Teori Dan Aplikasi Pengolahan Citra", Yogyakarta : ANDI, 2013.

[11] C. Solomon. T. Breckon, "Fundamental Of Digital Image Processing", West Sussex : John Wiley & Sons. Ltd, 2011.

[12] K. Khurshid. I. Siddiqi. C. Faure. dan N. Vincent, "Comparison Of Niblack Inspired Binarization Methods For Ancient Document", Document Recognition And Retrieval XVI, San Jose, 2009.

[13] S. V. Rajashekararadhya. P. Vanaja R, "Zone Based Feature Extraction Algorithm For Handwritten Numeral Recognition Of South Indian Scripts", Digital Technology Journal, vol. 2, pp.41-51, 2009.

[14] Prasetyo. E, "Data Mining Konsep Dan Aplikasi Menggunakan Matlab", Yogyakarta : ANDI.

[15] B. Hariyanto, "Rekayasa Sistem Berorientasi Objek", Bandung : Informatika, 2004.

[16] S. R. A. Shalahudin. M, "Rekayasa Perangkat Lunak Terstruktur Dan Berorientasi Objek", Bandung : Informatika, 2013.

[17] R. Sianipar, "Teori Dan Implementasi Java", Bandung : Informatika, 2013.

[18] B. Hariyanto, "Rekayasa Sistem Berorientasi Objek", Bandung : Informatika, 2004.

[19] Kim Hamilton. Russel Miles, "Learning UML 2.0", O'Reilly : Sebastopol, 2006.

[20] J. Sauvola. T. Seppanen. S. Haapakoski. M. Pietikainen, "Adaptive Document Binarization", 4th Int. Conf. On Document Analysis And Recognition, Ulm, Germany, pp.147-152, 1997.

[21] Z. A. Agus. Y. W. Arya. Laili, “Algoritma Thresholding Adaptif Berdasarkan Deteksi Blok Terhadap Citra Dokumen Terdegradasi”, Institut Teknologi Sepuluh Nopember, Surabaya.