IMPLEMENTASI METODE NAÏVE BAYESIAN UNTUK MENERJEMAHKAN BAHASA ISYARAT INDONESIA

Novrizal Eko Sihbudi[1], Galih Hermawan, S.Kom., M.T.[2]

Teknik Informatika – Universitas Komputer Indonesia
Jl. Dipatiukur 112-114 Bandung
E-mail : zhekho@gmail.com[1], galih.hermawan@yahoo.co.id[2]

ABSTRAK

Bahasa isyarat adalah bahasa yang mengutamakan bahasa tubuh, dan gerak bibir, bukannya suara untuk berkomunikasi. Penyandang tunarungu adalah kelompok utama yang menggunakan bahasa ini, biasanya dengan mengkombinasikan bentuk tangan, orientasi dan gerak tangan, lengan, dan tubuh, serta ekspresi wajah untuk mengungkapkan pikiran mereka. Cara membedakan suatu bahasa isyarat dengan yang lainnya yaitu dengan membedakan pola citra bahasa isyarat tersebut. Metode Principal Component Analysis adalah suatu metode yang dapat mengekstrak ciri dari kumpulan frame berdimensi sama. Hasil dari ekstraksi ciri citra tersebut kemudian akan digunakan sebagai nilai masukan untuk menentukan hasil klasifikasi citra menggunakan metode naive bayesian. Metode naive bayesian merupakan sebuah metode pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan.

Berdasarkan hasil penelitian yang telah dilakukan dapat diketahui identifikasi citra bahasa isyarat dapat dilakukan dengan klasifikasi Naïve Bayesian menggunakan metode PCA sebagai metode untuk ekstraksi kumpulan citra bahasa isyarat dengan dimensi sama dan metode naive bayesian sebagai metode untuk klasifikasi. Dari hasil pengujian yang telah dilakukan identifikasi citra bahasa isyarat mata berdasarkan pola menggunakan metode PCA dan naive bayesian didapat tingkat akurasi sebesar 69%.

Kata Kunci: Identifikasi Bahasa isyarat, PCA, Naive Bayesian,Kecerdasan Buatan

1. PENDAHULUAN

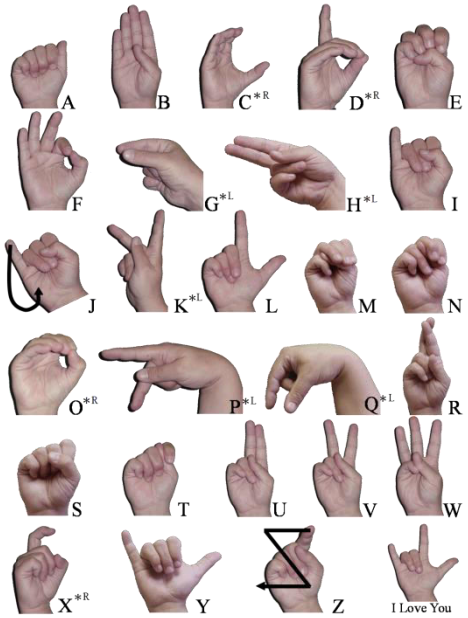
Bahasa isyarat adalah bahasa yang mengutamakan bahasa tubuh, dan gerak bibir, bukannya suara untuk berkomunikasi. Penyandang tunarungu adalah kelompok utama yang menggunakan bahasa ini, biasanya dengan mengkombinasikan bentuk tangan, orientasi dan gerak tangan, lengan, dan tubuh, serta ekspresi wajah untuk mengungkapkan pikiran mereka. Untuk Indonesia, ada dua sistem bahasa isyarat yang berlaku, Bahasa Isyarat Indonesia (BISINDO) dan Sistem Isyarat Bahasa Indonesia (SIBI). BISINDO merupakan bahasa isyarat yang dikembangkan oleh para penyandang tunarungu, sedangkan SIBI merupakan upaya pemerintah untuk menyesuaikan bahasa isyarat dengan ejaan yang disempurnakan (EYD) [1].

Penelitian yang pernah dilakukan untuk identifikasi bahasa isyarat indonesia menggunakan webcam sebagai media input salah satunya adalah Handgesture To Text Dengan Metode Artificia Intelligence KNN (K-Nearest Neighbour) dimana disebutkan bahwa untuk identifikasi huruf yang sesuai dengan Sistem Bahasa Isyarat Indonesia (SIBI) menggunakan webcam sebagai media input dengan metode KNN didapatkan akurasi sebesar 91,8% tetapi huruf J dan Z belum bisa di identifikasi karena terdiri lebih dari 1 gerakan [6]. Sedangkan penelitian untuk mengklasifikasi gerakan yang telah dilakukan adalah pada penelitian Analisis dan Implementasi Aplikasi Pengenalan Gerak Bibir Menjadi Teks Menggunakan Metode Support Vector Machine disebutkan bahwa PCA digunakan sebagai metode ekstraksi untuk mendapatkan ciri dari kumpulan citra dengan dimensi yang sama [4], kemudian pada penelitian Appearance Based Recognition of Words in American Sign Language disebutkan tracking tidak digunakan untuk menghindari kemungkinan error yang terjadi ketika segmentasi [2]. Untuk perbandingan tingkat akurasi dari beberapa metode untuk mengklasifikasi bahasa isyarat adalah Tamil Sign Language to Speech Translation dimana disebutkan bahwa untuk pengklasifikasian bahasa isyarat dengan inputan citra huruf bahasa isyarat dengan metode Naïve Bayesian mempunyai tingkat akurasi yang paling tinggi sebesar 93% dibandingkan dengan metode Proximal Support Vector (PSVM) sebesar 91% dan metode K-Nearest Neighbor (KNN) sebesar 78% [5].

Dari permasalahan dan solusi yang telah dijelaskan, maka penelitian ini akan mengklasifikasikan bahasa isyarat yang terdiri lebih dari satu gerakan sesuai dengan Sistem Isyarat Bahasa Indonesia (SIBI) ke dalam text dengan menerapkan metode PCA untuk proses ekstraksi dan metode Naïve Bayesian untuk klasifikasi,

1.1 Bahasa Isyarat

Bahasa isyarat adalah salah satu media untuk berkomunikasi yang mengutamakan komunikasi manual, bahasa tubuh dan gerak bibir. Di dunia penggunaan bahasa isyarat dalam setiap negara sangat unik dan bisa saja berbeda – beda, sebahai contoh Inggris dan Amerika menggunakan bahasa isyarat yang berbeda meskipun mempunyai bahasa tertulis yang sama, lain lagi dengan Inggris dan Spanyol meskipun memiliki bahasa tertulis yang berbeda tetapi kedua Negara tersebut menggunakan bahasa isyarat yang sama, karena sampai saat ini belum ada bahasa isyarat internasional yang sukses untuk diterapkan. Di Indonesia, ada dua sistem bahasa isyarat yang berlaku, Bahasa Isyarat Indonesia (BISINDO) dan Sistem Isyarat Bahasa Indonesia (SIBI).



Gambar 1. Abjad Sistem Isyarat Bahasa Indonesia

1.2 Kecerdasan Buatan

Kecerdasan buatan (Artificial inteligence) adalah salah satu cabang ilmu pengetahuan yang berhubungan dengan pemanfaatan mesin untuk memecahkan persoalan yang rumit dengan cara yang lebih manusiawi. Hal ini biasanya dilakukan dengan mengikuti karakteristik dan analogi berpikir dari kecerdasan manusia, dan menerapkannya sebagai algoritma yang dikenal oleh komputer.

Semakin pesatnya perkembangan teknologi menyebabkan adanya perkembangan dan perluasan lingkup yang membutuhkan kehadiran kecerdasan buatan. Karakteristik cerdas sudah mulai dibutuhkan di berbagai disiplin ilmu dan teknologi. Kecerdasan Buatan tidak hanya merambah di berbagai disiplin ilmu yang lain. Irisan antara psikologi dan kecerdasan buatan melahirkan sebuah area yang dikenal dengan nama cognition & psycolinguistics. Irisan antara teknik elektro dengan kecerdasan buatan melahirkan berbagai ilmu seperti pengolahan citra, teori kendali, pengenalan pola dan robotika.

Kecerdasan buatan digunakan untuk menganalisis pemandangan dalam citra dengan perhitungan simbol-simbol yang mewakili isi pemandangan tersebut setelah citra diolah untuk memperoleh ciri khas. Kecerdasan buatan bisa dilihat sebagai tiga kesatuan yang terpadu yaitu persepsi, pengertian dan aksi. Persepsi menerjemahkan sinyal dari dunia nyata dalam citra menjadi simbol-simbol yang lebih sederhana, pengertian memanipulasi simbol-simbol tersebut untuk memudahkan penggalian suatu informasi tertentu, dan aksi menerjemahkan simbol-simbol yang telah dimanipulasi menjadi sinyal lain yang dapat merupakan hasil akhir.

1.3 Pengolahan Citra (Image Processing)

Untuk melakukan pengenalan suatu objek berdasarkan tekstur yang dimiliki objek tersebut maka dibutuhkanlah suatu teknik untuk pengolahan citra. Pengolahan citra adalah pemrosesan citra dengan menggunakan komputer untuk menghasilkan citra yang kualitasnya lebih baik. Umumnya, operasi-operasi pada pengolahan citra diterapkan pada citra bila :

1. Perbaikan atau memodifikasi citra perlu dilakukan untuk meningkatkan kualitas penampakan atau untuk menonjolkan beberapa aspek informasi yang terkandung di dalam citra.
2. Elemen di dalam citra perlu dikelompokkan, dicocokkan, atau diukur.
3. Sebagian citra perlu digabung dengan bagian citra yang lain.

Pengolahan Citra bertujuan memperbaiki kualitas citra agar mudah diinterpretasi oleh manusia atau mesin (dalam hal ini komputer).

1.4 Preprocessing

Preprocessing adalah proses untuk memudahkan dalam menghasilkan nilai dari ekstraksi ciri. Pada tahap ini citra inputan yang telah didapat akan dilakukan proses resize dan threshold.

1.5 Resize

Pada proses preprocessing dilakukan proses resize citra. Resize adalah proses mengubah resolusi atau ukuran horizontal dan vertikal pada suatu citra, hal ini perlu dilakukan untuk menghasilkan ukuran yang sama pada semua citra yang dimasukkan. Tujuan dari resize citra adalah untuk mempercepat proses ekstraksi fitur warna dan fitur tekstur pada gambar yang memiliki dimensi (ukuran) yang besar.

1.6 Threshold

Threshold atau citra biner pada sebuah image digital merupakan proses untuk mengubah warna dimana setiap pikselnya hanya berisikan informasi intensitas warna hitam dan putih. Untuk mendapatkan nilai biner pada citra maka dilakukan proses threshold. Threshold berfungsi untuk memudahkan proses klasifikasi citra agar lebih terfokus terhadap masing-masing piksel.

1.7 Metode Principal Component Analysis (PCA)

Proses ekstraksi PCA adalah (Principal Component Analysis), dimana hasil proyeksi tersebut diektraksi dengan perhitungan PCA untuk mendapatkan feature dari image. Feature adalah komponen–komponen penting dari image- image training yang didapatkan dari proses training. Feature inilah yang nanti akan digunakan untuk mengidentifikasikan image yang akan dikenali.

Tahapan awal dari ekstraksi PCA adalah pembentukan matriks yang datanya diambil dari pixel setiap citra. Data citra direpresentasikan dalam ukuran matrik m x n, dimana m adalah jumlah citra yang dilatih dan n merupakan dimensi dari citra tersebut, kemudian dilakukan proses ekstraksi menjadi citra dengan dimensi yang lebih kecil yang hasilnya diproyeksikan menjadi sebuah matriks seperti pada persamaan berikut [6].

$$x = \begin{bmatrix} x_{1,1}x_{1,2}\ x_{1,3} \dots & x_{1,n} \\ x_{2,1}x_{2,2}x_{2,3} & x_{2,n} \\ x_{3,1}\ x_{3,2}x_{3,3} & x_{3,n} \\ \dots & \dots \\ x_{m,1}x_{m,2}x_{m,3} \dots & x_{m,n} \end{bmatrix} \quad (1)$$

Dimana
x = Matriks citra
Setelah matriks data citra terbentuk, maka proses berikutnya adalah proses perhitungan untuk mencari rata-rata hasil seluruh citra. Pencarian nilai rata-rata ini tujuannya untuk mengetahui noise atau persamaan tiap vector yang dapat menganggu keakuratan perhitungan pada PCA, yang dapat dihitung dengan menggunakan persamaan :

$$\alpha = \frac{(x_{1,i} + x_{2,i} + x_{3,i} + \dots + x_{m,i})}{m} = \frac{\sum_{j=1}^{m} x_{j,i}}{m} = [\tau_1, \tau_2, \dots \dots \dots, \tau_n] \quad (2)$$

Dimana
α = Rata-rata matriks seluruh citra
$x_{1,i\dots}$ $x_{m,I}$ = Matrik citra ke-1 sampai citra ke-m
m = Jumlah matriks citra
Terlihat bahwa hasil perhitungan persamaan diatas adalah merupakan sebuah *array* yang memuat nilai n, (τ1, τ2, … … …, τn). *Array* tersebut nantinya akan digunakan untuk mencari nilai rata-rata nol (*zero mean*), setelah dibentuk menjadi matriks 1xn terlebih dahulu. Namun, ukuran matriks nilai rata- rata tidak berkesesuaian dengan matriks data citra sampel yang memiliki ukuran mxn, maka perlu dilakukan penggandaan nilai rata-rata.
Hasil perhitungan nilai rata-rata nol digunakan untuk mendapatkan nilai matriks kovarian. Matriks kovarian merupakan matriks reduksi dari proses ekstraksi PCA, yang memiliki dimensi lebih kecil dibandingkan dimensi matriks citra asli. Berikut ini adalah persamaan matriks kovarian.

Y * YI = C …(3)
Dimana :
Y = Matriks Zero Means
YI = Transpose Matriks Zero Means

Setelah matriks kovarian didapatkan maka langkah selanjutnya adalah menentukan matriks *eigenvektor.* Matriks *eigen* dapat dicari menggunakan persamaan:

$$C - Z = |C - \lambda I|$$

$$|C - \lambda I| = 0 \quad (4)$$

Dimana *C* adalah matriks kovarian. Z adalah matriks *eigenvektor* dengan λsebagai *scalar* pembentuknya dan *I* sebagai matriks identitas, persamaan tersebut jika diimplementasikan ke bentuk matriks menjadi persamaan sebagai berikut:

$$\begin{bmatrix} C_{11}C_{12} \dots C_{1,n} \\ C_{21}C_{2,2} \dots C_{2,n} \\ \dots \dots \dots \dots \\ C_{m,1}C_{m,2} \dots C_{m,n} \end{bmatrix} - \lambda \begin{bmatrix} 1\ 0 \dots 0 \\ 0\ 1 \dots 0 \\ \dots \dots \dots \dots \\ 0\ 0 \dots 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} C_{11}C_{12} \dots C_{1,n} \\ C_{21}C_{2,2} \dots C_{2,n} \\ \dots \dots \dots \dots \\ C_{m,1}C_{m,2} \dots C_{m,n} \end{bmatrix} - \begin{bmatrix} \lambda\ 0 \dots 0 \\ 0\ \lambda \dots 0 \\ \dots \dots \dots \dots \\ 0\ 0 \dots \lambda \end{bmatrix} = 0$$

$$\begin{bmatrix} C_{11} - \lambda & C_{12} & \dots & C_{1,n} \\ C_{21} & C_{2,2} - \lambda & \dots & C_{2,n} \\ \dots & \dots & \dots & \dots \\ C_{m,1} & C_{m,2} & \dots & C_{m,n} - \lambda \end{bmatrix} = 0 \quad (5)$$

1.8 Klasifikasi Naïve Bayesian

Naïve bayesian adalah suatu metode pengklasifikasian paling sederhana dengan menggunakan peluang yang ada, dimana diasumsikan bahwa setiap variable X bersifat bebas (independence) [4]. Karena asumsi variabel tidak saling terikat, maka didapatkan :

$$P(X \mid Y = y) = \prod_{i=1}^{q} P(X_i \mid Y = y) \quad (6)$$

Terdapat beberapa langkah dalam pengklasifikasian menggunakan metode naive bayesian, berikut adalah langkah - langkahnya :

Training :

1. Hitung rata-rata (*mean*) tiap fitur dalam dataset training dengan.
   $\mu = \frac{\sum x_i}{n}$ (7)
   Dimana:
   μ = mean eigenvektor dataset
   n = banyaknya data
   $\sum x_i$ = jumlah nilai data

2. Kemudian hitung nilai varian dari dataset training tersebut seperti pada.

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

Dimana:
σ^2= varians
μ= mean
x_i = nilai data
n = banyaknya data

Testing :

1. Hitung probabilitas (*Prior*) tiap kelas yang ada dengan cara menghitung jumlah data tiap kelas dibagi jumlah total data secara keseluruhan.
2. Selanjutnya menghitung densitas probabilitasnya. Fungsi densitas mengekspresikan probabilitas relatif. Data dengan mean μ dan standar deviasi σ, fungsi densitas probabilitasnya adalah :

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

Dimana :
x = data masukan
π = 3,14
σ = standar deviasi
μ = mean

3. Setelah didapatkan nilai densitas probabilitasnya, selanjutnya menghitung posterior masing-masing kelas dengan menggunakan persamaan.

$$posterior = \frac{prior\ x\ likehood}{evidence}$$

Atau bisa ditulis

$$Posterior = \frac{P(C_i)x\ P(X1 \ldots\ldots\ldots\ldots\ldots Xn|C_i)}{P(X1 \ldots\ldots\ldots\ldots\ldots Xn|C_i)}$$

4. Setelah didapat nilai posterior masing-masing kelas, maka kelas yang sesuai untuk data masukan adalah kelas yang memiliki nilai posterior terbesar.

1.9 Pengujian *Confusion Matrix*

Pengujian yang dilakukan pada metode klasifikasi terdapat pada bagian akurasi dari hasil klasifikasi. Akurasi sebuah klasifikasi berpengaruh terhadap performa dari suatu metode klasifikasi. Untuk melakukan pengujian akurasi dapat digunakan *confusion matrix* yaitu sebuah matriks dari prediksi yang akan dibandingkan dengan kelas yang asli dari data masukan. Setiap kolom dari matriks berkorespondensi kepada hasil klasifikasi dan setiap baris pada masukan. Akurasi sebuah klasifikasi dimana i=j menerangkan akurasi dari klasifikasi pada setiap kelas [9]. Berikut contoh *Confusion Matrix* dapat dilihat pada tabel 2.

Tabel 2 *Confusion Matrix*

Kelas		Hasil Klasifikasi	
		0	1
Target	0	00	01
	1	10	11

Untuk menghitung akurasinya digunakan formula :

$$\frac{jumlah\ akurasi}{total\ data} X\ 100\% = Akurasi\%$$

2. ISI PENELITIAN

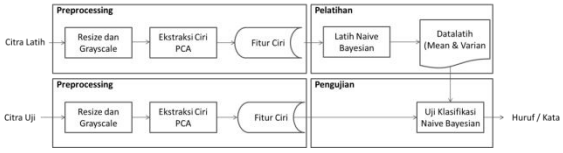
2.1 Analisis Masalah

Bahasa isyarat adalah bahasa yang mengutamakan bahasa tubuh, dan gerak bibir, bukannya suara untuk berkomunikasi. Untuk mengidentifikasi bahasa isyarat menggunakan webcam sebagai perangkat input diperlukan metode yang dapat mengekstraksi ciri pada gerakan. Pada penelitian sebelumnya telah dilakukan sistem identifikasi bahasa isyarat yaitu Tamil Sign Language to Speech Translation dimana disebutkan bahwa untuk pengklasifikasian bahasa isyarat tamil dengan inputan citra menggunakan metode Naïve Bayesian mempunyai tingkat akurasi yang paling tinggi sebesar 93% dibandingkan dengan metode Proximal Support Vector (PSVM) sebesar 91% dan metode K-Nearest Neighbor (KNN) sebesar 78% [4]. Sedangkan penelitian untuk identifikasi bahasa isyarat menggunakan webcam sebagai media input yaitu Handgesture To Text Dengan Metode Artificia Intelligence KNN (K-Nearest Neighbour) disebutkan bahwa untuk identifikasi huruf yang sesuai dengan Sistem Bahasa Isyarat Indonesia (SIBI) menggunakan webcam sebagai media input dengan metode KNN didapatkan akurasi sebesar 91,8% tetapi huruf J dan Z belum bisa teridentifikasi karena terdiri lebih dari 1 gerakan [6]. Maka perlu dilakukan penelitian lebih lanjut untuk meningkatkan akurasi pada sistem identifikasi bahasa isyarat.

Naive Bayesian merupakan salah satu algoritma yang dapat mengklasifikasi citra berdasarkan pelatihan yang diberikan. Sebelum proses klasifikasi, akan dilakukan ekstraksi citra terlebih dahulu untuk mendapatkan ciri dari citra tersebut. Metode yang akan digunakan untuk ekstraksi adalah PCA.

2.2 Analisis Sistem

Analisis Sistem pada penelitian ini akan menjelaskan proses yang ada pada pada sistem identifikasi bahasa isyarat secara langsung menggunakan webcam sebagai media input. Berikut adalah alur proses sistem yang akan dibangun dapat dilihat pada gambar 2 :



Gambar 2 Blok Diagram Sistem

2.3 Analisis Citra Latih

Berdasarkan penelitian sebelumnya disebutkan dengan pengujian menggunakan jumlah frame 10, 20, 30, pada saat jumlah frame yng digunakan adalah 20 frame, tingkat akurasi akan meningkat dari ketika menggunakan 10 frame, tetapi

ketika jumlah frame yang digunakan adalah 30 frame, maka tingkat akurasi menjadi konstan [4]. Berdasarkan hal tersebut pada penelitian ini frame yang akan digunakan berjumlah 30 frame. Data berupa 30 frame yang diekstrak menggunakan software irfanview dari video yang sebelumnya direkam menggunakan webcam dari bentuk isyarat huruf A-Z, dan beberapa kata dengan kombinasi dua bentuk yang sesuai dengan Kamus Isyarat Bahasa Indonesia edisi kedua 2000. Jakarta : Departemen Pendidikan Nasional. [9]. Jarak antara pengisyarat dengan webcam saat pengambilan data sample adalah 80-90cm, dan waktu pengambilan adalah siang hari pukul 11.00 WIB. Sample yang akan dianalisis adalah 6 frame huruf I dengan bentuk gerakan yang serupa dan huruf J yang terdiri lebih dari satu bentuk gerakan untuk mengetahui hasil identifikasi citra bahasa isyarat yang terdiri lebih dari satu bentuk gerakan.

Tabel 3 Huruf I dan J SIBI

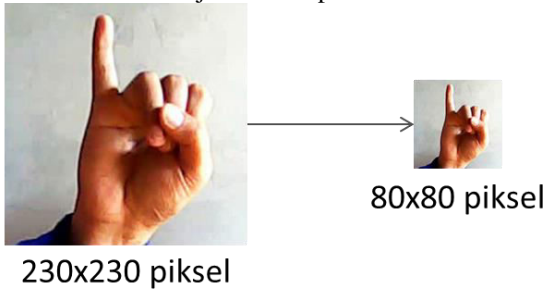
Huruf	Bentuk Isyarat					
	Frame-1	Frame-2	Frame-3	Frame-4	Frame -5	Frame-6
I						
J						

2.4 Analisis Proses Preprocessing

Preprocessing bertujuan untuk mendapatkan nilai ciri fitur yang akan dipakai sebagai masukan dalam pelatihan maupun pengujian naive bayesian, pada penelitian ini preprocessing terdiri dari tiga tahap yaitu Resize, grayscale dan ekstraksi PCA.

2.5 Resize

Resize merupakan proses mengubah ukuran citra, proses ini dilakukan untuk menyamakan ukuran citra bahasa isyarat yang menjadi masukan baik pada saat pelatihan maupun saat pengujian. Berdasarkan penelitian Analisis dan Implementasi Aplikasi Pengenalan Gerak Bibir Menjadi Teks Menggunakan Metode Support Vector Machine disebutkan saat pengujian pengaruh dimensi pada tingkat akurasi menggunakan dimensi 40x40, 60x60, dan 80x80, akurasi tertinggi diperoleh pada dimensi 80x80 sebesar 75% [4]. Pada penelitian ini citra akan di resize menjadi 80x80 piksel.



Gambar 3 Proses Resize

2.4 Threshold

Threshold adalah proses mengubah citra RGB menjadi citra biner, proses ini bertujuan untuk menyederhanakan nilai piksel yang sebelumnya citra RGB memiliki 3 nilai menjadi satu nilai pada setiap pikselnya. Konversi informasi suatu citra warna ke dalam citra biner dapat dilakukan dengan operasi ambang batas tunggal, yaitu nilai pixel dikelompkan menjadi dua kelompok seperti ditunjukan pada rumus berikut:

$$f_0(x,y) = \begin{matrix} 0, & f_0(x,y) < 100 \\ 255, & f_0(x,y) \geq 100 \end{matrix}$$

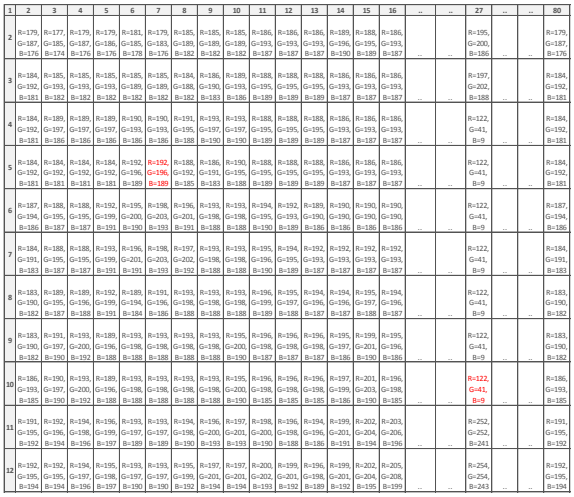
Piksel-piksel yang nilainya intensitasnya dibawah 100 diubah menjadi hitam(nilai intensitas 0), sedangkan piksel-piksel yang nilai intensitanya diatas 100 diubah menjadi warna putih(nilai intensitas = 255).

Contoh perhitungan threshold :

Citra masukan yang digunakan pada contoh perhitungan threshold ini adalah citra yang dimensianya telah di resize seperti gambar 4.



Gambar 4 Citra RGB Huruf I frame ke-1 dengan ukuran 80x80



Gambar 5 Matriks Nilai RGB

Contoh nilai RGB pada piksel (5,5) yaitu :

R : 188
G : 192
B : 196

Perhitungan threshold piksel (5,5) :

Threshold = (188 + 192 + 196)/3
Threshold = 188
188 > 100
Nilai Pixel = 255

Contoh nilai RGB pada piksel (10,27) yaitu :

R : 122
G : 41
B : 9

Perhitungan threshold piksel (5,5) :

Threshold = (122+41+9)/3

Threshold = 57
57 < 100
Nilai Pixel = 0

Hasil citra masukan yang telah di *threshold* dapat dilihat pada gambar 6.



Gambar 6 Citra 80x80 Hasil *Trehsold*

Setelah didapatkan nilai piksel 0 dan 255, selanjutnya nilai 255 bisa dirubah menjadi nilai 1, sehingga matriks terdiri dari dua nilai yaitu 0 dan 1.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	..	..	80
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	..	..	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	..	..	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	..	..	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	..	..	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	..	..	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	..	..	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	..	..	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	..	..	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
80	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1

Gambar 7 Matriks *threshold* 80x80

2.5 Ekstraksi Ciri Menggunakan Metode PCA

Menggunakan citra yang telah melalui tahap resize dan threshold, langkah selanjutnya adalah ekstraksi ciri menggunakan PCA Berikut adalah tahap-tahap ekstraksi ciri citra tangan menggunakan PCA , didapatkan eigenvektor dari tiap-tiap eigenvalue sebagai berikut :

Eigenvalues I:{-0.0034909006025451994;0.0034979702072709005; 1.0729490167158229 ; 4.427050983125396}
Eigenvektor :
untuk Eigenvalue -0.0034909006025451994:
[0.9956553327487252 ; 1 ; 0 ; 0 ; 0 ; 0]
untuk Eigenvalue -0.000003534722971764206-0.00349441486790229 55î:
[0.999976502646367-0.004367896648869726î ; 1 ; 0 ; 0 ; 0 ; 0]
untuk Eigenvalue -0.000003534722971764206+0.00349441486790229 55î:
[0.999976502646367+0.004367896648869726î ; 1 ; 0 ; 0 ; 0 ; 0]
untuk Eigenvalue 0.0034979702072709005:
[1.0043916651514138 ; 1 ; 0 ; 0 ; 0 ; 0]
untuk Eigenvalue 1.0729490167158229:
[-1 ; 1 ; -5.854101991377212 ; -1 ; 1 ; 1]
untuk Eigenvalue 4.427050983125396:
[-1 ; 1 ; 0.8541019662490332 ; -1 ; 1 ; 1]

Langkah selanjutnya adalah menghitung rata-rata dari eigenvector setiap eigenvalue. Dengan menggunakan langkah yang sama pada huruf J maka akan didapatkan tabel eigen vektor sebagai berikut

Tabel 4 Eigenvektor huruf I dan J

Huruf	EigenVektor					
I	0.332609	0.166667	0.166667	0.334065	-0.80902	0.309017
J	0.665218	0.324559	0.233768	0.668131	-1.61803	0.618034

2.6 Analisis Proses Pelatihan Naive Bayesian

Pelatihan naive bayesian dilakukan untuk mendapatkan nilai mean dan varian. Berikut penjelasan dari proses pelatihan naive bayesian :

2.7 Latih data

Latih data merupakan tahap untuk pelatihan data dengan menggunakan metode naive bayesian. Dari fitur ciri yang ada pada tabel 3.4, kemudian akan dihitung mean dan varians setiap kelasnya, nilai mean dan varian tersebut nantinya akan digunakan untuk pencarian probabillitas likehood, dengan rumus densitas. Nilai mean dapat dihitung dengan menggunakan persamaan 2.8 dan nilai varians dapat dihitung dengan menggunakan persamaan :

$$\mu = \frac{\sum x_i}{n}$$

Dimana:
μ = mean
n = banyaknya data
$\sum x_i$ = jumlah nilai data

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

Dimana:
σ^2= varians
μ= mean
x_i = nilai data
n = banyaknya data
Berikut perhitungan nilai mean dan varians berdasarkan eigen vektor pada table 4:

Kelas Huruf I
μ Eigenvektor Huruf I

$$= \frac{\begin{matrix}0.332609222 + 0.166666667 \\ +0.166666667 + 0.334065278 \\ + 0.809016999 \\ +0.309016994\end{matrix}}{6}$$

$$= 0.22165$$

Tabel 5 Mean dan Varian huruf I dan J

Huruf	Nilai	Fitur
		Eigenvektor
I	Mean (μ)	0.22165
	Varian (σ^2)	0.15708429686
J	Mean (μ)	0.14861333333
	Varian (σ^2)	0.78327432239

2.8 Simpan Data Latih

Simpan data merupakan tahap untuk menyimpan mean dan varian dari proses latih data kedalam bentuk txt.

2.9Analisis Proses Pengujian Naive Bayesian

Pengujian dilakukan untuk mengklasifikasi citra bahasa isyarat yang dimasukan dengan data yang sudah dilatih, sehingga dapat diketahui hasil klasifikasinya.

2.9 Citra Uji

Citra uji merupakan citra yang akan melalui tahap pengujian terhadap data latih, sebelumnya dilakukan tahap yang sama dengan citra latih untuk mendapatkan nilai Fitur Ciri. Contoh citra yang akan diuji dapat dilihat pada tabel 6.

Tabel 6 Contoh Citra Uji

Bentuk Isyarat					
Frame-1	Frame-2	Frame-3	Frame-4	Frame -5	Frame-6
					

Berikut nilai fitur ekstraksi ciri citra uji dapat dilihat pada tabel 7.

Tabel 7 Eigenvektor citra uji

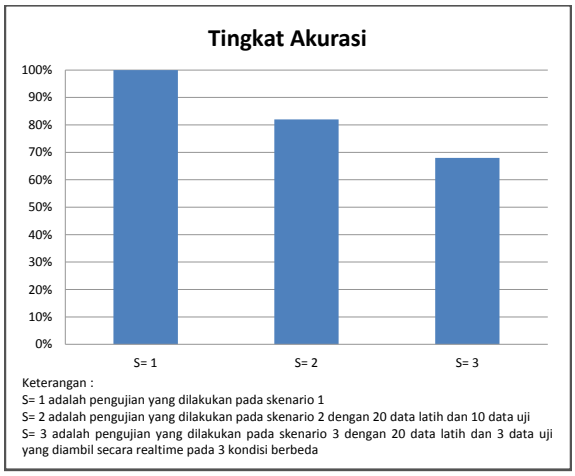
EigenVektor					
0.332609	0.196667	0.166667	0.334065	-0.80902	0.309017

Setelah dilakukan proses pengujian naive bayesian didapat nilai posterior bahasa isyarat huruf J lebih besar dari bahasa isyarat huruf I, maka kelas yang cocok untuk citra uji diatas adalah bahasa isyarat huruf J.

3. PENGUJIAN DAN HASIL

Pengujian metode bertujuan untuk mengetahui hasil dari metode PCAdan naive bayesian dalam mengidentifikasi bahasa isyarat. Pada penelitian ini, rencana pengujian dilakukan dengan beberapa skenario. Berikut adalah skenario yang akan dilakukan.

1. Menguji citra bahasa isyarat yang telah dijadikan data latih.
2. Menguji citra bahasa isyarat yang belum dilatih sebelumnya dan menguji pengaruh jumlah data latih terhadap tingkat akurasi.



Berdasarkan pengujian yang telah dilakukan metode PCA dan metode naïve Bayesian dapat digunakan untuk mengidentifikasi citra bahasa isyarat. dan berdasarkan pengujian diatas didapat kesimpulan bahwa konsistensi data latih berpengaruh terhadap tingkat akurasi dimana saat pengujian secara real time menggunakan kondisi lokasi yang berbeda tingkat akurasi yang dihasilkan.cenderung turun.

4. PENUTUP

4.1 Kesimpulan

Hasil yang didapat dari penelitian yang telah dilakukan dalam penyusunan skripsi ini serta mengacu pada tujuan penelitian, maka dapat disimpulkan.

1. Metode PCA dan naive bayesian dapat digunakan untuk mengidentifikasi bahasa isyarat berdasarkan kumpulan citra.
2. Tingkat akurasi metode PCA dan naive bayesian dalam mengidentifikasi bahasa isyarat berdasarkan kumpulan citra menggunakan webcam adalah sebesar 69%.

4.2 Saran

Berdasarkan hasil dari penelitian yang telah tercapai saat ini, terdapat beberapa saran yang mungkin bermanfaat jika ada yang akan melakukan penelitian yang sejenis, yaitu :

1. Dataset citra yang digunakan sebaiknya memiliki kualitas yang sama.
2. Untuk mendapat tingkat keakurasian yang tinggi dalam mengklasifikasikan berbagai citra, sebaiknya menggunakan data latih yang banyak.
3. Menambahkan metode tracking agar bounding area bisa mengikuti posisi tangan.

DAFTAR PUSTAKA

[1] Angga Nikola Fortuna. 2011. “Persepsi Siswa Tunarungu Terhadap Penggunaan Sistem Isyarat

Bahasa Indonesia Dalam Komunikasis". Jurnal. Volume 3, No. 3 September 2011

[2] Morteza Zahedi, Daniel Keysers, dan Hermann Ney. 2003. "Appearance Based Recognition of Words in American Sign Language". Paper. Computer Science Department, RWTH, Aachen University.

[3] Kumud Tripathi, Neha Baranwal dan G. C. Nandi. 2015. "Continuous Indian Sign Language Gesture Recognition and Sentence Formation" Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)

[4] Revi Febriana S. 2015 "Analisis dan Implementasi Aplikasi Pengenalan Gerak Bibir Menjadi Teks Menggunakan Metode Support Vector Machine". *Skripsi*. Fakultas Teknik Elektro, Univeristas Telkom

[5] Putu Wijaya Merta, I Made Gede Sunarya, I Ketut Resika Arthana. 2015. " Handgesture To Text Dengan Metode Artificia Intelligence KNN (K-Nearest Neighbour)". Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika (KARMAPATI). Volume 4, No. 1, Januari 2015.

[6] Juniar Prima Rakhman. 2015. "Translasi Bahasa Isyarat". *Tesis*. Pascasarjana Teknik Informatika ITS Sukalio.

[7] Kamus Isyarat Bahasa Indonesia. Edisi Kedua. 2000. Jakarta : Departemen Pendidikan Nasional.

[8] *Ibnu Subiyanto.* 1993. Seri Diktat Kuliah *"Metode Penelitian"*. Gunadarma

[9] *Umi Proboyekti. 2009. Agile Software Development. Bahan Ajar Rekayasa Perangkat Lunak. Yogyakarta : Universitast Kristen Duta Wacana*