Analisis dan Implementasi Pengenalan Wajah berbasis Video dengan Menggunakan Algoritma Viola Jones dan SVM (*Support Vector Machine)*

Analysis and Implementation of Video Base Face Recognition using Viola Jones Algorithm and Support Vector Machine

Marliani Harahap[1], Bedy Purnama, S.Si., M.T.[2], Mahmud Dwi Sulistyo, S.T., M.T[3]
[1, 2, 3] Program Studi S1 Teknik Informatika, Fakultas Informatika, Universitas Telkom
[1]hrplia@hotmail.com, [2]bedy.purnama@gmail.com, [3]mahmuddwis@gmail.com

Abstrak - Saat ini telah banyak aplikasi yang memanfaatkan pengolahan citra yang digunakan dalam bidang biometrik. Biasanya teknik pengolahan citra ini digunakan untuk proses deteksi dan pengenalan, melalui karateristik-karakteristik alami yang terdapat pada tubuh manusia, contohnya pengenalan wajah. Berbagai penelitian pengenalan wajah telah dilakukan sebelumnya salah satunya dengan menggunakan Algoritma Viola Jones dan SVM berbasis gambar dengan akurasi 84,28%.Pada tugas akhir ini dibangun suatu sistem yang mampu mengenal wajah dan menghitung jumlah wajah yang dikenali, dengan inputan berbasis video. Data yang diperoleh dari hasil perekaman diproses sehingga mampu mengklasifikasi nama setiap wajah dengan memanfaatkan metode Viola Jones sebagai deteksi dan SVM (*Support Vector Machine)* sebagai pengenalan. Dengan implementasi metode tersebut, sistem dapat mengenali nama wajah yang dikelompokkan kedalam kelas-kelas. Dari hasil pengujian sistem tersebut didapat akurasi 21.79% pada kernel linier.

Kata Kunci : video, wajah,Viola Jones, SVM

Abstract – *Nowadays, there are many applications using image processing in the field of biometric. As a usual, these image processing techniques use for the detection and recognition process through characteristics that only exist naturally in the human body, for example, face recognition. Various studies about face recognation have been done before. One of them used the Viola Jones Algorithm and SVM-based image with 84.28% of accuracy. In this final project, It will be built a system that can recognize and count the number of recognizable faces, with a video-based input. The data obtained from the recording process so as to classify the name of each face by utilizing Viola Jones as a detection method and SVM (Support Vector Machine) as a recognation method. With the implementation of those methods, the system can recognize the face names, which are grouped into classes. From the test results of the system obtained 21.79% in linier kernel,*

Keywords: *Video, Face, SVM, Viola Jones*

I. PENDAHULUAN

Sekarang ini telah banyak aplikasi yang memanfaatkan pengolahan citra. Hal ini ditunjukkan dengan semakin banyaknya metode yang dikembangkan untuk mendukung berbagai macam aplikasi komputer yang digunakan pada berbagai bidang, salah satu contohnya pada bidang *biometric*. Bidang ini telah banyak diaplikasikan pada kehidupan sehari-hari. Teknik yang dipakai biasanya berfungsi untuk melakukan proses identifikasi dengan menggunakan karakteristik alami yang terdapat pada manusia, seperti contohnya wajah, sidik jari, iris, alis mata dan retina mata. Wajah merupakan salah satu fitur yang sangat popular dipakai untuk autentikasi sistem. Karena wajah merupakan bagian dari tubuh yang tidak dapat diduplikasi.

Untuk mampu mengenali objek, dilakukan pendeteksian dengan menggunakan pendekatan Algoritma *Viola Jones*[1]. Pendekatan Algoritma *Viola Jones* untuk mendeteksi objek bergerak dalam inputran citra dengan menggabungkan empat metode,yaitu *Integral Image, AdaBoost Algorithm, Cascade Classifier, Haar Features*[2]. *Haar Features* memproses gambar dalam kotak-kotak, dimana dalam satu kotak terdapat beberapa pixel. Per kotak itu pun kemudian diproses dan didapatkan perbedaan nilai yang menandakan daerah gelap dan terang. Dalam penelitian ini menggunakan gambar bergerak (video) sebagai inputan, perhitungan dan penjumlahan pixel terjadi secara terus-menerus dan membutuhkan waktu yang lama. Oleh karena itu,penjumlahan diganti dengan perhitungan secara integral sehingga didapatkan hasil lebih cepat. Hasil deteksi dari *Haar Features* kurang akurat jika menggunakan satu fungsi maka dibutuhkan beberapa fungsi yang banyak.

Pemroses *haar features* yang banyak diatur dalam *cascade classifier*. Untuk mendapatkan hasil deteksi yang akurat dibutuhkan *Haar Features* yang banyak untuk diproses atau diatur di dalam *cascade classifer*.

Salah satu proses dari pengenalan wajah adalah ekstraksi fitur. Tujuan utama dari ekstraksi fitur digunakan untuk mengambil ciri penting dari suatu objek, pada penelitian ini menggunakan PCA (*Principal Component Analysis)* sebagai metode ekstraksi ciri. Selanjutnya untuk proses ekstraksi ciri menggunakan PCA, untuk mereduksi dimensi ciri dari suatu objek[3]. Kemudian dilakukan latihan dan pengklasifikasi menggunakan klasifikasi *SVM* dimana pada penelitian sebelumnya akurasi klasifikasi ini cukup baik dalam hal melatih dan mengujicoba pengenalan. [4]

II. *PREPROCESSING*

Principal Component Analysis (PCA), disebut juga *Karhunen-Loeve Transform* (KLT) atau merupakan teknik ekstraksi fitur dari representasi data yang banyak digunakan pada wilayah pengenalan pola dan *Computer Vision* [9]. PCA bertujuan untuk mengetahui nilai variansi total pada data dan merepresentasikannya dalam variabel yang berjumlah lebih sedikit. Dengan kata lain, PCA mengurangi ukuran dimensi data asli namun masih menyimpan ciri-ciri paling representatif dari data tersebut. Hal ini bermanfaat untuk kompresi dan penyimpanan data, serta mengurangi waktu komputasi pemprosesan data tersebut.

Secara matematis PCA didefinisikan sebagai berikut [10]. Dimisalkan terdapat vektor kolom citra $Xi = [X_1, X_2...X_M]^T$ di mana M merupakan jumlah citra dan merupakan jumlah kolom dalam matriks *Xi*. Maka μ, yaitu mean dari matriks *Xi* dapat dihitung sebagai berikut.

$$\mu = \frac{1}{M} \sum_{i}^{M} Xi \quad (1.1)$$

Pengurangan data matriks Xi dengan rata-rata citra μ akan menghasilkan data dengan nilai rata-rata nol. Proses ini disebut juga dengan normalisasi data yang akan menghasilkan matriks data φ_t.

$$\varphi_t = Xi - \mu;\ i = 1,2.....M \quad (1.2)$$

Hasil dari matriks data φ_t ini kemudian digunakan untuk membangun matriks kovariansi L sebagai berikut.

$$L = \frac{1}{M} \sum_{t=1}^{M} \varphi_t \varphi_t^T \quad (1.3)$$

Dimisalkan A merupakan matriks yang terdiri dari gabungan matriks data φ_t. A = $[\varphi_1, \varphi_{2....} \varphi_M]$ maka, persamaan di atas dapat diubah menjadi :

$$L = AA^T \quad (1.4)$$

Nilai vektor eigen u_t dan nilai eigen λ_t untuk matriks kovariansi *L* dapat dihitung dengan persamaan berikut :

$$Lu_t = \lambda_t u_t \quad (1.5)$$

PCA dilakukan dengan mempertahankan hanya sejumlah *k* vektor eigen tidak nol yang terbesar, yaitu $u_t' = [u_1, u_{2....Uk}]$. Teknik ini disebut *Eigen Value Decomposition*.

III. LEARNING

Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan salah satu teknik yang masuk ke dalam kelas *supervised* sehingga dalam pengimplementasiannya membutuhkan kelas target. SVM ini dapat menemukan solusi yang *global optimal*, yaitu SVM selalu mencapai solusi yang sama untuk setiap kali dilakukan running[13].

Tujuan dalam SVM ini adalah berusaha untuk menemukan *hyperplane* atau fungsi pemisah (klasifier) yang optimal yang dapat memisahkan dua set data dari dua kelas yang berbeda, dengan menggunakan ruang hipotesis berupa fungsi- fungsi linier dalam sebuah ruang fitur (*feature space*) berdimenasi tinggi, dimana dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementastikan *learning bias* yang berasal dari teori pembelajaran statistic[14].

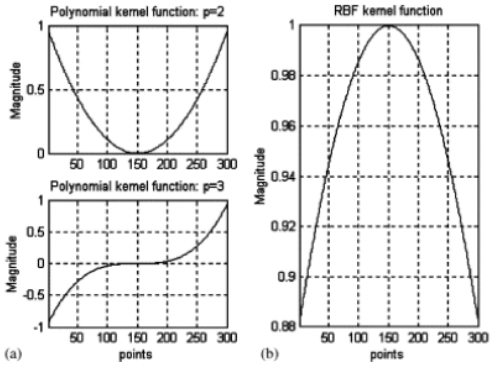
SVM Pada Non Linierly Separable Data

Merupakan data yang tidak bisa diklasifikasikan secara sempurna menggunakan SVM dasar secara linier. Dalam hal ini bisa digunakan *soft marging*, yaitu dengan melunakan konstrain yang ada dengan cara memberikan toleransi terhadap data yang tidak bisa diklasifikasikan secara sempurna. Dalam hal ini diperkenalkan (*slack variabel*) yang dihitung dari jarak data yang terklasifikasi salah terhadap *hyperplane* kelas data tersebut. Dan C (parameter pengontrol *trade off*) antara margin dan error klasifikasi (*slack variabel*).

Untuk mengubah data inputan dipetakan ke dimensi yang elbih tinggi, pada kasus nonlinierly memerlukan adanya kernel. Beberapa kernel untuk kasus *non linierly* yaitu :

Tabel Error! No text of specified style in document..**1 Jenis Kernel[15]**

Jenis Kernel	Defenisi
Polynomial	$\mathrm{K}(x_i, x_j) = (x_i x_j + 1)^p$
Gaussian	$\mathrm{K}(x_i, x_j) = exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$

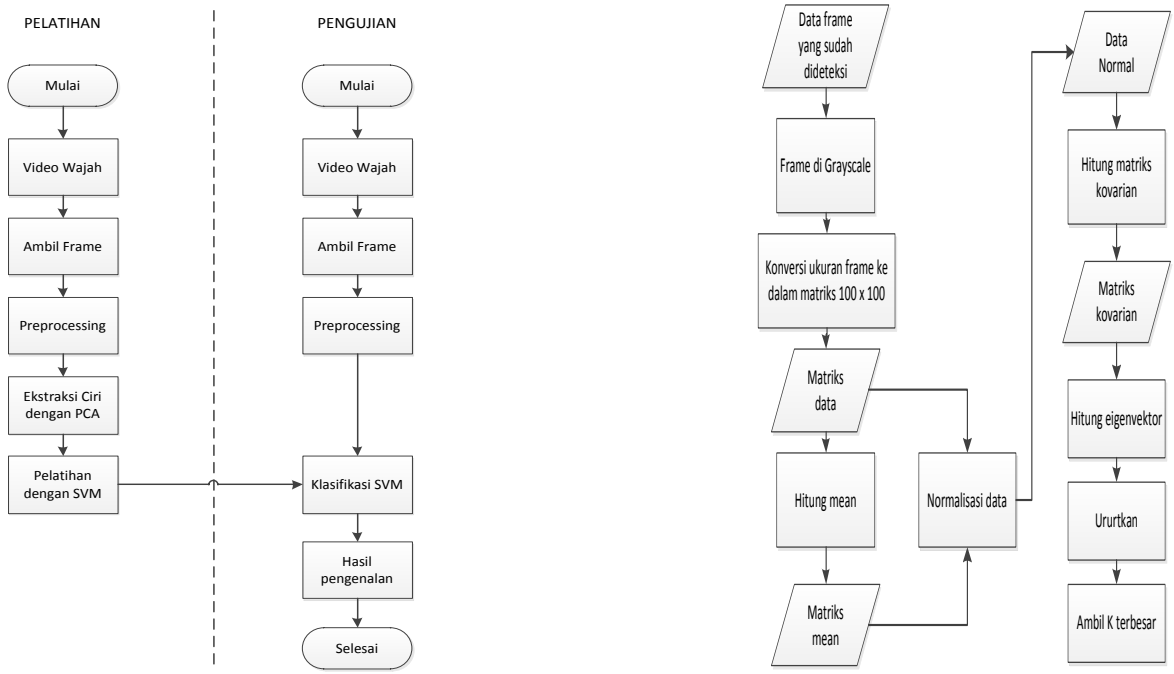


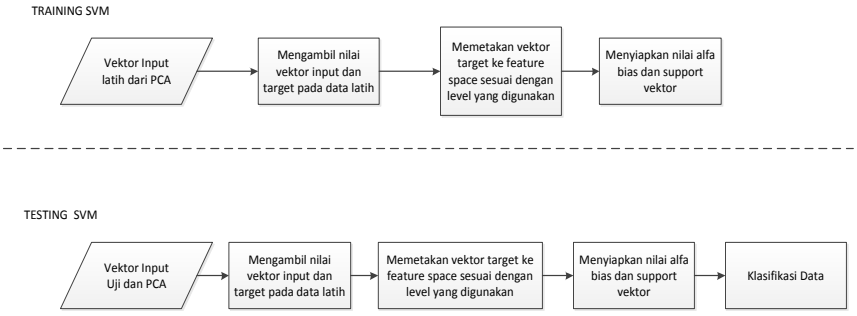
Gambar Error! No text of

SVM untuk Multi Kelas

Ada dua pendekatan utama untuk SVM multi kelas ini, pertama menggabungkan beberapa SVM biner, sedangkan yang kedua adalah secara langsung menggunakan semua data yang terdiri dari beberapa kelas ke dalam sebuah bentuk permasalahan optimasi. Pada pendekatan yang pertama, terdapat dua metode yaitu Metode Satu-Lawan-Semua (One-against-all, OAA) dan metode Satu-Lawan-Satu (One-Against-One, OAO)[13][14

I. PERANCANGAN SISTEM





II. HASIL DAN ANALISIS

Implementasi Perangkat Lunak

Implementasi pada tugas akhir ini perangkat lunak yang digunakan adalah sebagai berikut:

1. Tools : Visual Studio 2013 dengan bahasa C#
2. Sistem operasi : Microsoft Windows 7
3. Framework : Microsoft .NET *Framework* 4
4. Library
   a. *Intel Open Source Computer Vision Wrapper Library* untuk C# dikenal sebagai EmguCV.
   b. *Haarcascade_frontalface_default.xml* yang didapat dari sourceforge EmguCV sebagai cascade classifier yang telah ditraining dengan memiliki 15 fitur haar dengan ukuran training data sebesar 20x20 piksel.
   c. *PrincipalComponentAnalysis* yang didapat dari Accord.NET Framework
   d. *MulticlassSupportVectorMachine* tool yang dari sourceforge EmguCV bagian *machine learning.*

Video Dataset

Untuk data set yang berupa video terdapat 4 jenis video yang akan digunakan secara berbeda pada disetiap skenario. 4 jenis video terdiri dari video yang berjumlah 20 orang, video 7 orang, video 4 orang dan video 3 orang. Video yang jumlah orang nya lebih dari 5 akan digunakan sebagai dataset untuk pendeteksian. Sedangkan video lainya akan digunakan sebagai dataset untuk pengenalan. Setiap jenis video mempunyai waktu berkisar 3 detik hingga 5 detik yang akan dijadikan sebagai data training dan testing dipilih manual (dipilih mana yang baik) yang artinya video yang dilihat dari pergerakan wajah orang-orang secara teratur.

Tujuan Pengujian

Pengujian pembangunan sistem dalam tugas akhir ini memiliki beberapa tujuan yaitu:

1. Menganalisis pengaruh perbandingan komposisi data pelatihan dan data pengujian terhadap performa sistem;
2. Menganalisis pengaruh parameter perubahan nilai kernel terhadap proses *training* dalam klasifikasi kelas.

Skenario Pengujian

Beberapa skenario pengujian dilakukan untuk mencapai tujuan yang telah didefinisikan. Adapun skenario pengujian yang akan dilakukan pada sistem, diantaranya:

- **Skenario 1: Pengujian Viola Jones dengan Haar Cascade**

Pada skenario ini mendeteksi wajah tepat pada kamera. Pada skenario ini akan diuji dengan parameter scale Factor dan minimum Neighbours. Kedua parameter tersebut merupakan parameter keberhasilan deteksi adaboost [5]. Scale Factor adalah parameter yang menspesifikasikan seberapa besar image yang direduksi setiap skala image. Parameter ini mempunyai efek terhadap area wajah yang terdeteksi dengan penskalaan dari classifier. Minimal Neighbors adalah parameter yang menspesifikasikan setiap kandidat objek yang terdeteksi untuk dihindari. Parameter ini mempunyai efek menghindarkan wajah yang berdekatan antara satu terhadap wajah lainnya. Parameter ini menghitung jarak antara area wajah terdeteksi yang berdekatan. Serta juga dapat mengukur kefektifan sistem ini ketika menghitung orang.

D alam skenario ini yang pertama kali dilakukan mencari *scaleFactor* wajah pada data video yang berbeda untuk klasifikasi dengan nilai *minNeigbors* nya tetap. Penulis memiliki beberapa video, salah satunya video dengan jumlah orang yang banyak yaitu 20 orang. Video ini yang akan digunakan dalam menentukan nilai parameter dari *scaleFactor* dan *minNeigbors* yang terbaik. Hasil deteksi wajah terbaik diprioritaskan bukan banyaknya wajah yang terdeteksi, tapi hasil mana yang tidak ada error. Error tersebut seperti salah deteksi wajah, dimana jika *background* pada video terdapat banyak *noise* dan *noise* itu terdeteksi sebagai wajah.

- **Skenario 2: Pengujian SVM dengan Kernel**

Pada skenario ini menggunakan kernel diberbagai skenario untuk mencapai klasfikasi secara optimum. Kernel SVM yang diuji meliputi kernel *Linier, Polynomial* dan *Gaussian.*Dimana pengujian SVM dengan kernel hanya menggunakan variable tetap yang sudah ada dari library.

Hasil Pengujian dan Analisis

Beberapa skenario pengujian dilakukan untuk mencapai tujuan yang telah didefinisikan. Adapun skenario pengujian yang akan dilakukan pada sistem, diantaranya:

- **Skenario 1: Pengujian Viola Jones dengan Haar Cascade**

Pengujian ini mencari nilai *scaleFactor* terbaik menggunakan nilai 1.1, 12, 1.3 dimana nilai minimum dari scaleFactor adalah 1.1. Ketika menggunakan nilai dibawah dari 1.1 maka program tidak dapat memproses training pada *haarcascade_frontalface.xml.* Berikut ini adalah tabel hasil mencari *scaleFactor* yang optimal:

Tabel 2.2 *scaleFactor* yang optimal

scaleFactor	***MinNeighbor***	**Jumlah yang terdeteksi**	
		20 Orang	7 Orang
1.1	3	14 orang	6 orang
1.2	3	10 orang	6 orang
1.3	3	9 orang	6 orang

Dari tabel diatas dapat disimpulkan *scaleFactor* yang optimum adalah 1.1 dengan *minNeighbor* 3 karena dapat mendeteksi wajah yang paling banyak. Untuk video dengan jumlah wajah yang sedikit nilai *scaleFactor* tidak mengalami perubahan nilai jumlah yang terdeteksi. Pengujian ini menggunakan nilai 3, 10, 20 untuk mendeteksi jumlah wajah secara tepat dan tidak ada redundant pada wajah yang sama. Jika menggunakan nilai lebih kecil dari 3 maka akan banyak sekali terdapat wajah yang terdeteksi dan terdeteksi di bagian yang lain (bukan area wajah).

Gambar 2.1 Penggunaan nilai minNeigbor lebih kecil dari 3

Berikut ini adalah tabel hasil mencari *minNeighbor* yang optimal:

Tabel 2.3 *minNeighbor* yang optimal

scaleFactor	***MinNeighbor***	**Jumlah yang terdeteksi**	
		20 Orang	7 Orang
1.1	3	14 orang	6 orang
1.1	10	7 orang	6 orang
1.1	20	3 orang	6 orang

Dari tabel diatas dapat disimpulkan *scaleFactor* yang optimum adalah 1.1 dengan *minNeighbor* 3. Untuk video dengan jumlah wajah yang sedikit nilai *scaleFactor* tidak mengalami perubahan nilai. Jadi untuk skenario selanjutnya yaitu SVM akan digunakan nilai scale Factor sebesar 1.1 dan minNeigbors sebesar 3.

Gambar 2.2 Jumlah wajah yang terdeteksi dengan menggunakan minNeighbor 3 dan scaleFactor 1.1

1.1.2 Skenario 2 : Pengujian SVM dengan Kernel

Skenario kedua ini akan dilakukan uji coba berbagai macam kernel SVM yang diharapkan dapat memperoleh kernel optimum . Terdapat 3 kernel yang akan diujicobakan meliputi kernel linier, polynomial, dan Radial. Berikut ini adalah hasil dari ujicoba penggunaan kernel yang berbeda dalam pengklasfikasian untuk pengenalan nama.

Tabel 2.4 Pengujian SVM dengan Kernel

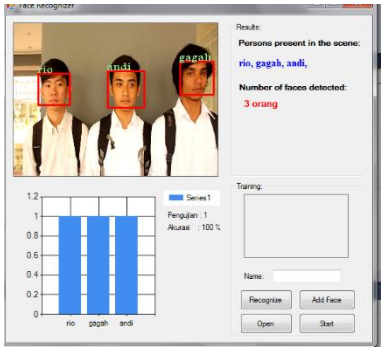
Data Uji berupa video yang dapat dikenali	**Kernel**		
	Linier	Gaussian	Polynomial
3 wajah	21.79%	2.56%	1.28%

Dari tabel diatas dapat disimpulkan kernel yang baik dalam mengenali kelas adalah linier sedangkan untuk Gaussian dan polynomial hanya satu kali atau dua kali dapat mengenali kelas dan cenderung salah. Dalam skenario ini menggunakan nilai parameter *default* dari setiap kernel karena hasil klasifikasi jika menginput nilai parameter yang berbeda-beda akan didapat hasil 0%.

Dalam kasus ini terdapat beberapa kemungkinan skenario ini dianggap gagal dalam pengklasfikasian yaitu sederhananyaya proses *preprocessing* yang dilakukan sehingga dapat berpengaruh pada ekstraksi ciri, dan persebaran data pada tiap kelas dari vektor ciri yang dihasilkan oleh PCA belum baik

Skenario ini menggunakan jumlah video hanya dengan 3 orang karena saat menggunakan video dengan jumlah wajah 7 orang atau 5 orang klasifikasi tidak berjalan dan cenderung hanya dapat mengklasifikasikan wajah yang posisi nya *frontal face*.

Berdasarkan tabel diatas dapat disimpulkan klasifikasi dengan linear mampu mengenali nama dari 3 wajah tersebut. Didapat akurasi sebesar 21.79% untuk 17 frame dari total 78 frame yang terdeteksi sebagai 3 nama dari setiap wajah. Sedangkan untuk kasus *Gaussian* dan *polynomial* dapat mendeteksi 3 nama dari setiap wajah tetapi masih salah kelas yang artinya kelas yang bukan A dikenali sebagai kelas A.



Gambar 2.3 Video yang dapat dikenali dengan benar

Pada gambar diatas terlihat akurasi pengenalan sebesar 100% dikarenakan wajah dari setiap orang dapat dikenalin secara tepat oleh program. Tetapi perlu dilakukan pengecekan secara manual yaitu dengan mengecek seluruh frame yang ada dan melihat apakah program dapat mengenali kelas secara tepat. Dari seluruh frame hanya 17 frame yang terdeteksi secara benar, ini disebabkan oleh karena pergerakan wajah yang bergerak dari kanan ke kiri.

III. KESIMPULAN DAN SARAN

Kesimpulan dari hasil penelitian diatas dapat disimpulkan sebagai berikut :

1. Pendeteksian wajah menggunakan Viola Jones dengan Cascade Classifier menghasilkan bounding box adalah persegi.
2. Akurasi yang dihasilkan adalah 100% dengan ekstraksi ciri menggunakan PCA dan klasifikasi SVM dengan parameter kernel linear.

Saran dari hasil penelitian diatas adalah sebagai berikut :

1. Untuk penelitian selanjutnya algoritma SVM bisa diganti dengan algoritma klasifikasi lainnya
2. Selain penggunaan nilai Kernel pada SVM, bisa juga digunakan nilai w,b,α untuk mendapatkan performansi yang lebih baik
3. Untuk PCA bisa ditambahkan eigen value dan eigen face sebagai parameter untuk ekstraksi ciri.

REFERENSI

[1] Wulandari, "Deteksi Wajah menggunakan Algoritma Viola-Jones", http://purpalacious.arieflatu.net/2010/03/deteksi-wajah-menggunakan-algoritma-viola-jones/#more-188, September 2012

[2] Viola, Jones, "Robust Real Time Face Detection", *International Journal of Computer Vision,* 2004Bihan, J., Michael F, V., & Pantic, M. (2011, March). Action Unit detection unit sparse appearance descriptors in space-time video volumes. *Automatic Face & Gesture Recognition and Workshops (FG 2011)*, 314-321.

[3] Smith, Lindsay. 2002. A Tutorial on Principal Component Analysis. Sumber : http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf. Diakses 2012

[4] I.Frolov, R.Sadykhov, "Experimental system for face identification based on support vector machines", in *Conference on Information system and technologies,* 2008, Minsk

[5] Barczak, C.H Messom, "Real Time Computation of Haar-like feature atgeneric angles for detection algorithms" in Institute of Information & Mathematical Sciences Massey University, New Zealand, 2002

[6] P. I. Wilson dan D. J. Fernandez, "Facial Feature Detection Using Haar Classifiers," *Journal of Computing Sciences in Colleges*, pp. 127-133, 2006

[7] Cho, Junguk, S. Mirzaei, J. Oberg dan R. Kastner, "FPGA-Based Face Detection System Using Haar Classifiers," dalam *International Symposium on Field Programmable Gate Arrays*, ACM, 2009

[8] Anto Satriyo Nugroho, "Bioinformatika dan Pattern Recognition", http://www.ilmukomputer.com, 2013

[9] P. Viola, M. J. Jones dan D. Snow, "Detecting Pedestrians Using Patterns of Motion and Appearance," dalam *Ninth IEEE International Conference*, IEEE, 2003

[10] Zhang dan Yang. 2004. Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition. Berada di IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 26, No. 1, January 2004

[11] Mushtofa, Annas. 2008. *Analisis dan Implementasi Pengenalan Telinga Menggunakan Principal Component Analysis (PCA) dan Linear Discriminant Analysis (LDA)*. Tugas Akhir, Jurusan Teknik Informatika IT Telkom, Bandung

[12] Luthfi. 2012. *Identifikasi Individu melalui Citra Telapak Tangan Menggunakan Countourlet Transform – Principal Component Analysis*. Tugas Akhir, Jurusan Teknik Informatika IT Telkom, Bandung

[13] Smith, Lindsay. 2002. A Tutorial on Principal Component Analysis. Sumber : http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf. Diakses 2012.

[14] Santosa, Budi, "*Data mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis*", Graha Ilmu : Bandung, 2007.

[15] Sembiring, Krisantus. 2007. Penerapan Teknik *Support Vector Machine* untuk Pendeteksian Intrusi pada Jaringan. Institut Teknologi Bandung.