

UJI IMPLEMENTASI ALGORITMA VIOLA-JONES DALAM PENGENALAN WAJAH

Cahaya Jatmoko¹, De Rosal Ignatius Moses Setiadi²,
Danu Hartanto³, Alvin Faiz Kurniawan⁴

^{1,2,3,4}Program Studi Teknik Informatika, Fakultas Ilmu Komputer,
Universitas Dian Nuswantoro

e-mail: ¹cahaya.jatmoko@dsn.dinus.ac.id, ²moses@dsn.dinus.ac.id,
³danu.hartanto@gmail.com, ⁴alvin@gmail.com

Abstrak

Salah satu algoritma yang sering digunakan untuk melakukan deteksi pada wajah yaitu Viola-Jones. Metode ini merupakan gabungan dari 3 buah fitur yaitu integral image, adaboost dan cascade classifier. Masing-masing fitur mempunyai fungsi tersendiri dan saling melengkapi. Integral image digunakan dalam penentuan ada dan tidaknya gambar, adaboost untuk memilih dan mengatur nilai threshold, sedangkan cascade classifier untuk mengklasifikasi daerah yang akan di deteksi. Untuk memudahkan deteksi, terutama pada bagian mata maka digunakan Haar like feature. Proses pengenalan wajah telah dilakukan pada gambar dengan satu objek dan beberapa objek. Hasil implementasi juga dapat mengenali objek foto lukisan dan foto tampak samping. Dari seluruh percobaan di dapatkan nilai rata-rata sebesar 65% dengan sebaran nilai akurasi tertinggi 70%, sensitivitas 55% dan spesifitas 71%.

Kata Kunci: deteksi wajah, viola-jones, Haar Like-Feature, AdaBoost Machine Learning, Cascade Classifier

1. PENDAHULUAN

Deteksi wajah sering digunakan dan dimanfaatkan di berbagai perangkat teknologi digital pada zaman sekarang dengan memanfaatkan perkembangan teknologi komputer. Saat ini perkembangan aplikasi komersial yang sering menggunakan algoritma pendeteksi wajah maupun pengenalan wajah sangat pesat dalam perkembangannya. Maka dari itu, kita perlu belajar untuk mengikuti perkembangan teknologi agar mampu mengembangkan teknologi pengenalan wajah agar lebih optimal. Pengembangan harus didasarkan pada kecepatan dan akurasi sistem pendeteksi wajah yang tepat dan cepat. Deteksi wajah pada manusia menjadi sangat penting untuk perkembangan teknologi pengolahan citra digital, seperti perkembangan teknologi pada kamera digital, *smartphone*, *gadget*, dan lain-lain. Pengenalan deteksi pada wajah banyak informasi yang dapat diperoleh, seperti mata, hidung, dan mulut. Tujuan tersebut untuk mengenali posisi fitur wajah manusia [1]. Pada kasus penentuan lokasi wajah pada pendeteksi wajah untuk menentukan lokasi posisi mata, hidung, dan mulut. Pada penilitan ini, kami berusaha menyajikan suatu sistem pendeteksi wajah yang dapat mendeteksi banyak wajah secara bersamaan dengan algoritma viola-jones yang diimplementasikan pada *smartphone*. Beberapa masalah dalam proses deteksi wajah disebabkan oleh adanya perbedaan posisi wajah, komponen pada wajah seperti jenggot atau kumis, ekspresi wajah, adanya objek yang menghalangi, dan kondisi ada saat citra terekognisi.

Berdasarkan literature mengenai pentingnya melakukan deteksi wajah dan tantangan proses deteksi wajah maka perlu adanya pemilihan algoritma yang tepat [2]. Metode Viola-Jones merupakan metode pendeteksian obyek yang memiliki tingkat keakuratan yang cukup tinggi yaitu sekitar 93,7 % dengan kecepatan 15 kali lebih cepat daripada detektor Rowley Baluja-Kanade dan kurang lebih 600 kali lebih cepat daripada detektor Schneiderman Kanade. Metode ini, diusulkan oleh Paul Viola dan Michael Jones pada tahun 2001.

Akurasi yang tinggi ditunjukkan oleh metode Viola Jones [2] apabila dibandingkan dengan metode perhitungan jarak seperti Euclidean distance, Neural Network, Segmentasi, dan Template Matching, sedangkan kekurangan Viola Jones yaitu lebih sulit melakukan deteksi dengan posisi objek tidak menghadap kamera secara keseluruhan. Kerangka deteksi wajah mampu memproses gambar yang sangat cepat saat mencapai tingkat deteksi tinggi.

Pada proses ini membuahkan tiga kontribusi utama. Pertama adalah representasi gambar baru yang disebut “Gambar Integral” yang memungkinkan fitur yang digunakan sebagai detector untuk dikomputasi dengan sangat cepat. Kedua adalah classifier sederhana dan efisien yang dibangun dengan algoritma AdaBoost untuk memilih sejumlah visual kritis yang kecil dari satu set fitur potensial yang besar [3]. Kontribusi ketiga adalah sebuah metode untuk menggabungkan klasifikasi ke dalam *cascade* yang memungkinkan daerah latar belakang dapat dibuang saat melakukan komputasi dalam menentukan daerah wajah [4][5].

Beberapa penelitian telah dilakukan untuk mendeteksi wajah, seperti yang telah dilakukan oleh Kovac, dkk [6] yang telah mengimplementasikan algoritma Viola Jones untuk mengklasifikasi warna kulit dan mendeteksi wajah. Viola Jones dikombinasikan dengan fitur ekstraksi warna RGB dan YCbCr. Hasil deteksi masih menunjukkan akurasi yang rendah dan masih terdapat banyak proses salah deteksi. Penelitian lainnya dalam proses deteksi wajah telah dilakukan oleh Zhao, dkk [7] pada tahun 2012 menggunakan AdaBoost untuk mengklasifikasi objek warna kulit dan sudah dapat digunakan untuk mendeteksi wajah dengan baik. menurut Wardoyo dkk [8], deteksi wajah telah dilakukan menggunakan algoritma Principal Component Analysis (PCA). Proses pengenalan wajah membutuhkan waktu yang cukup lama namun menghasilkan sensitivitas 100%. Pada beberapa objek masih terjadi kesalahan proses deteksi sehingga hanya menghasilkan 69,33%. Penelitian lain yang dilakukan oleh Denny dkk [2] melakukan kombinasi fitur warna pada algoritma Multi Layer Perceptron (MLP). Nilai akurasi dan spesifitas menjadi lebih meningkat dibanding hanya menggunakan MLP saja.

2. TINJAUAN PUSTAKA

2.1 Algoritma Viola-Jones

Deteksi wajah sering digunakan dan dimanfaatkan di berbagai perangkat teknologi digital pada zaman sekarang dengan memanfaatkan perkembangan teknologi komputer. Saat ini perkembangan aplikasi komersial yang sering menggunakan algoritma pendeteksi wajah maupun pengenalan wajah sangat pesat dalam perkembangannya. Maka dari itu, kita perlu belajar untuk mengikuti perkembangan teknologi agar mampu mengembangkan teknologi pengenalan wajah agar lebih optimal. Pengembangan harus didasarkan pada kecepatan dan akurasi sistem pendeteksi wajah yang tepat dan cepat [3]. Banyak dari sistem pendeteksi wajah tersebut menggunakan metode Viola Jones sebagai pendeteksi objek wajah manusia.

Viola-Jones menggunakan banyak pengklasifikasi yang berbeda, masing-masing melihat bagian gambar yang berbeda. Setiap *classifier* individu yang lemah (kurang akurat, menghasilkan lebih banyak false positive, dll) dari pada classifier akhir karena mengambil lebih sedikit informasi. Ketika hasil dari setiap *classifier* digabungkan, maka akan menghasilkan *classifier* yang kuat seperti pada cuplikan coding berikut:

```
Def train (self, training):  
    Training_data = [ ]  
    For x in range(len(training))  
    Training_data.append((integral_image(training[x][0]), training[x][1]))
```

Tahapan algoritma Viola-Jones:

- a. Inisialisasi bobot
- b. Menormalkan bobot
- c. Pilih classifier lemah terbaik (berdasarkan kesalahan tertimbang)

- d. Perbarui bobot berdasarkan kesalahan pengklasifikasi yang dipilih
- e. Ulangi langkah 2–4 T kali di mana T adalah jumlah pengklasifikasi lemah yang diinginkan
- f. Tambahkan data training metode ke kelas ViolaJones di mana dapat mengimplementasikan pelatihan.

2.2 Haar Like Feature

Semua wajah manusia memiliki beberapa sifat yang serupa. Keteraturan ini dapat dicocokkan dengan menggunakan Fitur Haar. Beberapa sifat yang umum pada wajah manusia:

- a. Daerah mata lebih gelap dari pipi atas.
- b. Daerah jembatan hidung lebih terang dari mata.

Komposisi sifat yang membentuk fitur wajah yang cocok:

- a. Lokasi dan ukuran: mata, mulut, pangkal hidung
- b. Nilai: gradien berorientasi intensitas piksel

Keempat fitur yang cocok dengan algoritma ini kemudian dicari dalam gambar wajah (ditampilkan di kanan). Fitur persegi panjang:

Nilai = Σ (piksel di area hitam) - Σ (piksel di area putih)

- a. Tiga jenis: dua, tiga, empat persegi panjang, Viola & Jones menggunakan fitur dua persegi panjang
- b. Misalnya: perbedaan kecerahan antara persegi panjang putih & hitam di area tertentu
- c. Setiap fitur terkait dengan lokasi khusus di jendela kecil

Apabila nilai selisih antara daerah terang dengan daerah gelap di atas nilai ambang (*threshold*), maka daerah tersebut dinyatakan memiliki fitur seperti pada cuplikasi coding berikut.

```

Input: a 24 x 24 image with zero mean and unit variance
Output: a d x 1 scalar vector with its feature index f ranging from 1 to d
Set the feature index f = 0
Compute feature type (a)
for all (i,j) such that 1 ≤ i ≤ 24 and 1 ≤ j ≤ 24 do
    for all (w,h) such that i + h - 1 ≤ 24 and j + 2w - 1 ≤ 24 do
        compute the sum S1 of the pixels in [i,i + h - 1] x [j,j + w - 1]
        compute the sum S2 of the pixels in [1,i + h - 1] x [j + w,j + 2w - 1]
        record this feature parametrized by (l,i,j,w,h): S1 - S2
        f = f + 1
    end for
end for
    
```

Nilai Haar merupakan perhitungan selisih piksel antara daerah yang gelap dan terang sesuai persamaan (1) berikut:

$$F(Haar) = \sum F White - \sum F Black$$

Dimana:

- F(Haar) = Nilai fitur total
- $\sum F White$ = Nilai fitur objek terang
- $\sum F Black$ = Nilai fitur objek gelap

Kotak *rectangular Haar* dapat dihitung secara cepat dengan menggunakan Integral *Image*. Pemrosesan dengan fungsi Haar dinilai masih kurang apabila menggunakan satu buah fungsi saja, sehingga perlu digunakan beberapa fungsi dengan tujuan menambah akurasi. Fungsi-fungsi Haar diolah dan di organisir dalam classifier cascade. berikut ini merupakan cuplikan coding pada *integral image* dalam *haar feature*:

```

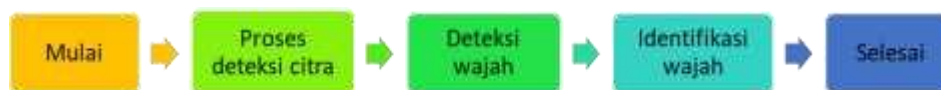
Input: an image I of size N x M
Output: its integral image II of the same size
Set II(1,1) = I(1,1)
for i = 1 to N do
    for j = 1 to M do
        II(i,j) = I(i,j) + II(i,j - 1) + II(i - 1,j) - II(i - 1,j - 1) and II
        is defined to be zero whenever its argument (i,j) ventures out of I's
        domain.
    end for
end for
    
```

2.3 Cascade Classifier

Untuk meningkatkan efisiensi komputasi dan juga mengurangi tingkat *false positive*, Viola dan Jones menggunakan serangkaian klasifikasi yang semakin kompleks yang disebut *cascade* [1]. Sebuah jendela input dievaluasi pada klasifikasi pertama dari *cascade* dan jika class tersebut mengembalikan false maka perhitungan pada jendela itu berakhir dan detektor mengembalikan false. Jika *classifier* mengembalikan *true* maka jendela dilewatkan ke *classifier* berikutnya dalam *cascade*. Klasifikasi berikutnya mengevaluasi jendela dengan cara yang sama. Jika jendela melewati setiap *classifier* dengan semua mengembalikan true maka detektor mengembalikan *true* untuk jendela itu. Semakin banyak jendela yang terlihat seperti wajah, semakin banyak klasifikasi dievaluasi dan semakin lama waktu yang dibutuhkan untuk mengklasifikasikan jendela itu. Karena sebagian besar jendela dalam gambar tidak terlihat seperti wajah, sebagian besar dengan cepat dibuang sebagai bukan wajah.

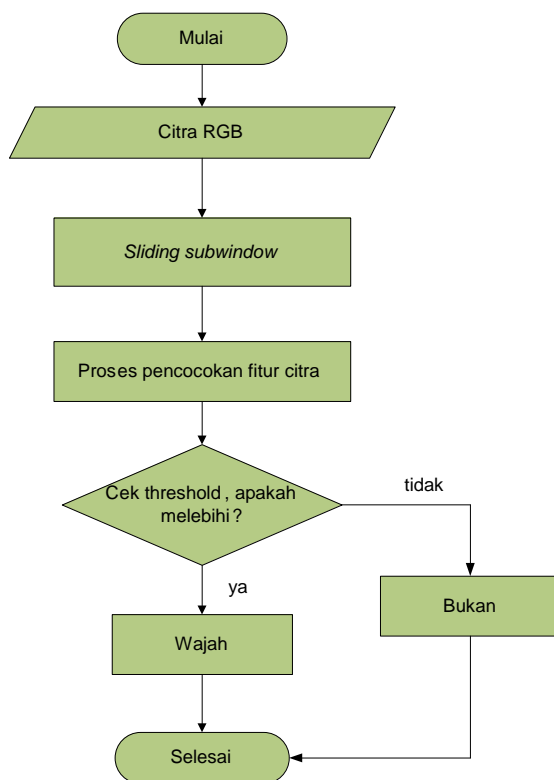
3. METODE PENELITIAN

Pada metode penelitan kali ini, kami menganalisis berbagai teknologi digital seperti kamera, kamera pada *smartphone*, dan cctv. Bagaimana proses penangkapan objek gambar yang terekam di dalam perangkat tersebut. Dengan pengamatan ini kami menganalisis bahwa untuk menangkap sebuah objek gambar diperlukan informasi kejelasan suatu objek. Dalam metode penelitian pendeteksi wajah kali ini kami menggunakan berbagai macam cara untuk merekam gambar wajah individu dari penelitian tersebut untuk menangkap dan mengenali objek gambar wajah individu, kamera memerlukan informasi pengenalan wajah yaitu mata, hidung, dan mulut. Hal ini untuk pengenalan pendeteksi wajah kami menggunakan metode algoritma Viola-Jones. Tahapan operasi algoritma yang digunakan pada makalah ini telah dijabarkan pada Gambar 1. Input citra berupa citra wajah yang diambil secara langsung dari *smartphone*. Proses selanjutnya akan dijelaskan secara bertahap sebagai berikut :



Gambar 1. Tahapan Penelitian

Pada penelitian ini terdapat 3 proses utama yaitu deteksi citra, deteksi wajah dan identifikasi wajah. Tahap pertama yaitu proses deteksi citra. Dalam makalah ini, digunakan 10 buah foto yang terdiri dari 4 foto yang mengandung konten wajah dan 6 buah foto tanpa konten wajah. Pada tahap pelatihan sistem, foto di segmentasi dengan *cascade classifier* seperti terlihat pada Gambar 2 berikut.



Gambar 2. Pola deteksi wajah dengan Cascade Classifier

Gambar 2 merupakan alur deteksi foto untuk mengidentifikasi ada atau tidaknya konten wajah. Citra wajah asli kemudian di RGB, selanjutnya di sliding untuk memperjelas konten wajah. Apabila sliding selesai, maka akan dilakukan pencocokan citra dan dilanjutkan proses cek threshold sehingga konten wajah menjadi lebih jelas. Threshold lebih memudahkan dalam proses pencocokan.

Tahap kedua adalah melakukan pemotongan ROI. Pada tahap ini, proses cropping digunakan untuk memotong konten wajah saja tanpa rambut atau komponen lainnya yang berada di sekeliling wajah. Hasil cropping belum dapat dikatakan sebagai citra wajah karena pada beberapa citra inputan masih dikenali sebagai citra bukan wajah. Objek yang tidak terdeteksi tersebut merupakan background citra yang mempunyai struktur mirip wajah.

Tahap ketiga yaitu melakukan ekstraksi fitur. Citra hasil tahap kedua yang masih salah deteksi, pada tahapan ini akan diperjelas. Ekstraksi dilakukan dengan menghitung nilai intensitas piksel melalui 12 atribut fitur. Beberapa fitur yang sangat penting untuk di hitung yaitu kurtosis, mean, standar deviasi, dan skewness. Fitur tersebut digunakan untuk menghitung setiap layer RGB.

Setelah tahap ketiga, citra kemudian di analisa untuk mengetahui konten wajah dengan akurat. Dalam makalah ini, digunakan perhitungan akurasi, sensitivitas, dan spesifitas. Nilai terbaik dari ketiganya adalah 100%, namun hingga saat ini tidak ada system yang mampu mendeteksi pada ketiga buah perhitungan tersebut dengan nilai sempurna. Akurasi digunakan untuk mengklasifikasi kebenaran objek wajah. Sensitivitas digunakan untuk menghitung gambar yang terdeteksi benar yang terdapat object of interest. Spesifitas digunakan untuk menghitung jumlah objek selain wajah yang tidak terdeteksi. Perhitungan prosentase sesuai persamaan (6) sampai persamaan (8)

$$Sensitivitas = \frac{TP}{TP + FN} \times 100\% \tag{6}$$

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \tag{7}$$

$$Spesifitas = \frac{TN}{FP + TN} \times 100\% \tag{8}$$

Dimana :

TP = citra yang terkategori benar dengan hasil yang benar secara akurat

TN = citra yang terkategori benar dengan hasil yang salah secara akurat

FP = citra yang terkategori salah dengan hasil yang benar secara akurat

FN = citra yang terkategori salah dengan hasil yang salah secara akurat








Apabila nilai dari hasil di atas berkisar 0% hingga 10% maka, sistem tersebut tidak dapat mendeteksi citra dengan baik, nilai 11% hingga 50%, sistem dapat mendeteksi dengan cukup baik, sedangkan 51% sampai 75%, sistem dapat mendeteksi citra dengan baik.

4. HASIL DAN PEMBAHASAN

4.1 Pengujian ke 1

Pengujian pertama di lakukan dengan cara membandingkan beberapa sampel citra maupun objek yang akan dideteksi apakah berupa wajah seseorang ataupun tidak pada Tabel 1.

Tabel 1. Hasil pengujian pertama berupa beberapa citra dan objek

| No. | Citra | Jenis | TP | TN | FP | FN |
|-----|---|--------------------|----|----|----|----|
| 1 |  | foto patung | 1 | 0 | 1 | 0 |
| 2 |  | foto mainan | 0 | 1 | 0 | 1 |
| 3 |  | foto wajah manusia | 1 | 1 | 0 | 0 |
| 4 |  | foto kaleng | 0 | 1 | 0 | 1 |
| 5 |  | gambar kartun | 0 | 1 | 0 | 1 |
| 6 |  | foto wajah manusia | 1 | 1 | 0 | 0 |
| 7 |  | foto binatang | 0 | 1 | 0 | 1 |

| | | | | | | |
|----|---|-----------------------------|---|---|---|---|
| 8 |  | foto lukisan | 1 | 1 | 0 | 0 |
| 9 |  | foto manusia tampak samping | 1 | 1 | 0 | 0 |
| 10 |  | foto siluet | 0 | 1 | 1 | 0 |

Berdasarkan hasil pada Tabel 1, diperoleh $Akurasi = \frac{5+9}{5+2+4+9} = \frac{14}{20} = 0,7 \times 100\% = 70\%$. Hal ini menandakan bahwa tingkat akurasi pada metode Viola Jones ini tergolong cukup akurat di mana hasil dari citra harus mencakup fitur wajah, yaitu mata, hidung, dan mulut. Apabila dilihat dari nilai $Sensitivitas = \frac{5}{5+4} = \frac{5}{9} = 0,55 \times 100\% = 55\%$ yang artinya objek dapat dideteksi oleh sistem cukup baik. Sedangkan untuk $Spesifitas = \frac{5}{2+5} = \frac{5}{7} = 0,71 \times 100\% = 71\%$. Percobaan pertama menghasilkan nilai spesifisitas sebesar 71% menandakan bahwa sistem mampu membedakan antara wajah dan bukan wajah dengan baik. Maka dari ketiga hasil parameter diatas dapat dihitung tingkat kehandalan sistem sebagai berikut :


$$\frac{Akurasi + Sensifitas + Spesifisitas}{3}$$

$$Hasil\ rata - rata = \frac{70\% + 55\% + 71\%}{3} = 65\%$$

4.2 Pengujian ke 2

Pada pengujian kedua disediakan beberapa citra untuk menguji kemampuan sistem ini untuk mendeteksi multiple object sebagai media uji, apakah sistem ini mampu mendeteksi wajah jika terdapat banyak wajah sekaligus. Berikut ini disajikan sebuah citra dari hasil proses deteksi wajah yang ditunjukkan pada Tabel 2. Citra wajah yang terdeteksi oleh sistem diberikan tanda berupa kotak berwarna merah. Tanda kotak tersebut mengidentifikasikan bahwa adanya seorang wajah manusia.

Tabel 2. Hasil pengujian kedua pada foto dengan banyak wajah

| No. | Citra | Keterangan | Total wajah asli | Total wajah terdeteksi |
|-----|---|-----------------------------|------------------|------------------------|
| 1 |  | Ukuran citra di zoom (30%) | 12 | 12 |
| 2 |  | Citra dengan kontras normal | 5 | 5 |

| | | | | |
|---|---|----------------------------|----|----|
| 3 |  | Citra dengan kontras (-20) | 16 | 3 |
| 4 |  | Citra dengan kontras (+20) | 15 | 12 |

Dari hasil pengujian di atas pada Tabel 2 dapat disimpulkan bahwa daya tangkap pendeteksi wajah pada sistem ini dipengaruhi intensitas cahaya yang diterima oleh sensor pendeteksi wajah. Kemudian kami uji kedua citra tersebut (nomor 3 dan nomor 4) dengan sampel 16 wajah yang memiliki perbedaan intensitas cahaya agar mengetahui tingkat kesuksesan sistem saat minim cahaya maupun cukup cahaya. Kedua hasil perhitungan tersebut cukup berbeda, sehingga disimpulkan bahwa citra yang memiliki cukup cahaya, keakuratan sistem dalam mendeteksi wajah akan meningkat.

5. KESIMPULAN

Pendeteksian wajah menggunakan metode Viola Jones memiliki kemampuan dan hasil yang cukup baik pada pengujian pertama, hal ini terlihat berdasarkan pengujian pertama yang dilakukan terhadap identifikasi wajah dan bukan wajah dalam kondisi yang memungkinkan dapat mempengaruhi hasil pendeteksian wajah, hasil tersebut menunjukkan tingkat akurasi sebesar 65% untuk mendeteksi wajah tersebut.

Pendeteksi wajah pada pengujian kedua, dilakukan dengan cara multiple object yang akan diuji dengan intensitas cahaya. Hasilnya pada citra yang terdapat cukup cahaya maka pendeteksian wajah dapat dilakukan dengan optimal. Namun sebaliknya, jika pada citra terdapat intensitas cahaya yang kurang maka pendeteksian wajah tidak dapat berjalan secara optimal. Maka pada pengujian kedua dapat disimpulkan bahwa intensitas cahaya juga berpengaruh ketika proses deteksi wajah.

Pada saat melakukan pengujian sistem mampu mendeteksi wajah dengan baik walaupun wajah terhalang oleh kaca atau hijab selama konturnya sama dengan kontur wajah.

DAFTAR PUSTAKA

- [1] R. Wahyusari and B. Haryoko , "Penerapan Algoritma Viola Jones untuk Deteksi Wajah," *SimetriS*, vol. 18, pp. 44-49, 2014.
- [2] D. Hardiyanto and D. A. Sartika, "Optimalisasi Metode Deteksi Wajah berbasis Pengolahan Citra untuk Aplikasi Identifikasi Wajah pada Presensi Digital," *SETRUM (Sistem Kendali-Tenaga-Elektronika-TelekomunikasiKomputer)*, vol. 7, no. 1, pp. 107-116, 2018.
- [3] P. A. Jusia, "Face Recognition Menggunakan Metode Algoritma Viola Jones Dalam Penerapan Computer Vision," *Jurnal Ilmiah Media Processor* , vol. 11, no. 1, pp. 663-675, 2016.
- [4] A. R. Syafira and G. Ariyanto, "Sistem Deteksi Wajah Dengan Modifikasi Metode Viola Jones," *Jurnal Emitor* , vol. 17, no. 1, pp. 26-33, 2017.
- [5] A. Suharso, "Pengenalan Wajah Menggunakan Metode Viola-Jones dan Eigenface Dengan Variasi Posisi Wajah Berbasis Webcam," *TechnoXplore : Jurnal Ilmu Komputer & Teknologi Informasi* , vol. 2, no. 1` , pp. 19-30, 2016.

- [6] J. Kovac, P. Peer and F. Solina, "Illumination independent color-based face detection," in *3rd International Symposium on Image and Signal Processing and Analysis, 2003. ISPA 2003*, Roma, Italia, 2003.
- [7] Q. Zhao, D. Zhang and H. Lu, "Supervised LLE in ICA space for facial expression recognition," in *2005 International Conference on Neural Networks and Brain*, 2005.
- [8] S. Wardoyo, R. Wiryadinata and R. Sagita, "Sistem Presensi Berbasis Algoritma Eigenface Dengan Metode Principal Component Analysis," *SETRUM*, vol. 3, no. 1, 2014.