

## Sistem Deteksi Wajah Pada Camera *Real Time* dengan menggunakan Metode Viola - Jones

Rizka Eka Putri<sup>1\*</sup>, Tekad Matulatan<sup>2</sup>, Nurul Hayaty<sup>3</sup>  
<sup>1,2,3</sup>Jurusan Informatika, Fakultas Teknik, Universitas Maritim Raja Ali Haji  
<sup>1,2,3</sup>Jl. Politeknik Senggarang, Tanjungpinang 29100  
\*Corresponding Author: rizkha07.rep@gmail.com

**Abstract**— In general, human are given the mind to be able to determine or be able to distinguish individuals who appear either human, animal, plant, and other objects that are known or unknown. And it is possible for human to recognize these object from their sight and from their brain memory. Especially on the human face, human can recognize whether the object is human or not human, and can recognize the object very well through his own eyes.face detection system in human becomes very important in the development of science of digital image processing. The research has been done with many advantages and disadvantages. From a face many information features that can be read, such as eyes, nose, and mouth. The detection system uses Viola Jones method as an object detection method. The Viola Jones method is known to have considerable Speed and accuracy as it combines several concepts (Haar feature, Integral image, Adaboost, Cascade classifier) into a main method for detecting objects.Based on tests conducted on face identification under conditions that may affect face detection results, the results show an accuracy of 67,6 % to detect the face.

**Keywords**— *Detection, Viola Jones, Adaboost, Cascade Classifier*

**Intisari**— Secara umum manusia diberikan akal dan fikiran untuk bisa menentukan atau dapat membedakan individu – individu yang tampak entah itu manusia, hewan, tumbuhan dan benda – benda lain yang dikenal maupun tidak dikenal. Dan sangat memungkinkan untuk manusia mengenali objek – objek tersebut dari penglihatan mereka dan dari memori otak mereka. Terutama pada wajah manusia, manusia dapat mengenali apakah objek itu manusia atau bukan manusia, dan dapat mengenali objek tersebut dengan sangat baik melalui penglihatannya sendiri. Sistem deteksi wajah pada manusia menjadi sangat penting dalam perkembangan ilmu pengolahan citra digital. Penelitian tersebut telah banyak dilakukan dengan kelebihan dan kekurangannya. Dari sebuah wajah banyak informasi fitur yang dapat dibaca, misalnya mata, hidung, dan mulut. sistem pendeteksian tersebut menggunakan metode *Viola Jones* sebagai metode pendeteksi objek. Metode *Viola Jones* dikenal memiliki kecepatan dan keakuratan yang cukup tinggi karena menggabungkan beberapa konsep (Fitur *Haar*, Citra Integral, *AdaBoost*, *Cascade Classifier*) menjadi sebuah metode utama untuk mendeteksi objek. Berdasarkan pengujian yang dilakukan terhadap identifikasi wajah dalam kondisi yang memungkinkan dapat mempengaruhi hasil pendeteksian wajah, hasil tersebut menunjukkan tingkat akurasi sebesar 67.6% untuk mendeteksi wajah tersebut.

**Kata kunci**— *Deteksi, Viola – Jones, Adaboost, Cascades Classifier*

## I. PENDAHULUAN

Sistem deteksi wajah pada manusia menjadi sangat penting dalam perkembangan ilmu pengolahan citra digital. Penelitian tersebut telah banyak dilakukan dengan kelebihan dan kekurangannya. Dari sebuah wajah banyak informasi fitur yang dapat dibaca, misalnya mata, hidung, dan mulut. Karena alasan ini, maka metode pendeteksian wajah adalah untuk menentukan "ya" atau "tidak" adanya wajah dalam sebuah gambar. Dalam kasus penentuan lokasi wajah, tujuannya adalah untuk mendeteksi letak posisi dari fitur wajah manusia misalnya mata, hidung, dan mulut dengan menggunakan metode pendeteksian wajah.

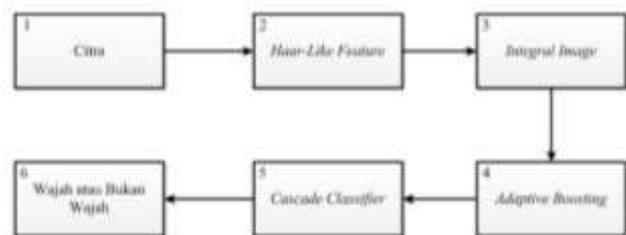
Sistem deteksi ini telah banyak diimplementasikan namun pada dasarnya sistem ini belum sempurna perlu terus dikembangkan untuk memperoleh hasil yang diinginkan. Satu hal yang dapat ditambah atau diimplementasikan dalam pengembangan teknologi pengenalan wajah yaitu dengan menambah tingkat kecepatan dan akurasi dalam pendeteksian wajah. Banyak dari sistem pendeteksian tersebut menggunakan metode Viola Jones [1] sebagai metode pendeteksi objek. Metode Viola Jones dikenal memiliki kecepatan dan keakuratan yang cukup tinggi karena menggabungkan beberapa konsep (Fitur Haar, Citra Integral, AdaBoost, *Cascade Classifier*) menjadi sebuah metode utama untuk mendeteksi objek.

## II. BAHAN DAN METODE

Metode *viola-jones* merupakan metode yang paling banyak digunakan untuk mendeteksi objek hal ini dikarenakan metode *viola-jones* memiliki algoritma yang efisien, sehingga tidak memerlukan waktu lama dalam melakukan proses pendeteksian objek. Proses pendeteksian objek dilakukan dengan mengklasifikasikan sebuah *image* setelah sebelumnya sebuah pengklasifikasi dibentuk dari data *training*.

Penelitian yang berjudul "*Rapid Object Detection Using A Boosted Cascade of Simple Features*" [1], menjelaskan kerangka deteksi wajah yang mampu memproses gambar yang sangat cepat saat mencapai tingkat deteksi tinggi. Penelitian ini menelurkan tiga kontribusi

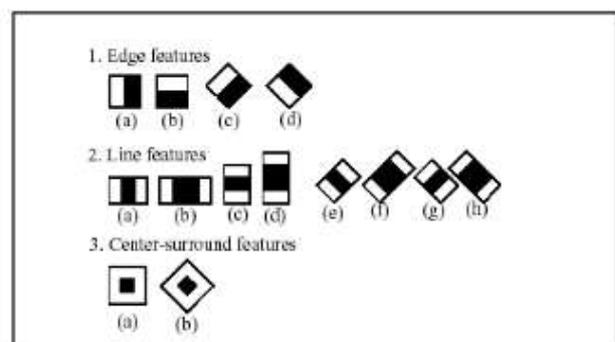
utama. Pertama adalah representasi gambar baru yang disebut "Gambar Integral" yang memungkinkan fitur yang digunakan sebagai detektor untuk dikomputasi dengan sangat cepat. Kedua adalah *classifier* sederhana dan efisien yang dibangun menggunakan algoritma pembelajaran *AdaBoost* untuk memilih sejumlah fitur visual kritis yang kecil dari satu set fitur potensial yang besar. Kontribusi ketiga adalah sebuah metode untuk menggabungkan pengklasifikasi dalam "*cascade*" yang memungkinkan daerah latar belakang gambar dengan cepat dibuang saat mengeluarkan beberapa komputasi dalam menentukan daerah wajah.



Gambar 1 Skema proses *viola jones*

### A. Fitur Haar

*Haar Like Feature* memproses gambar dalam kotak-kotak, dimana dalam satu kotak terdapat beberapa piksel. Tiap kotak itu pun kemudian diproses dan didapatkan perbedaan nilai (*threshold*) yang menandakan daerah gelap dan terang dengan mengurangi daerah yang gelap dan daerah yang putih, jika perbedaannya berada di atas *threshold* maka dikatakan fitur itu ada. Pola *Haar* dapat dilihat pada Gambar 2.



Gambar 2 Fitur *Haar* [1]

Pada proses pemilihan fitur *Haar*, fitur – fitur tersebut digunakan untuk mencari fitur wajah seperti mata, hidung, dan mulut. Pada

setiap kotak – kotak fitur tersebut terdiri dari beberapa piksel dan akan dihitung selisih antara nilai piksel pada kotak terang dengan nilai piksel pada kotak gelap. Apabila nilai selisih antara daerah terang dengan daerah gelap di atas nilai ambang (*threshold*), maka daerah tersebut dinyatakan memiliki fitur.

Nilai Haar diperoleh dari selisih jumlah nilai piksel daerah gelap dengan jumlah nilai piksel daerah terang yang ditunjukkan pada persamaan (1) :

$$F(Haar) = \sum F \text{ White} - \sum F \text{ Black} \tag{1}$$

Keterangan :

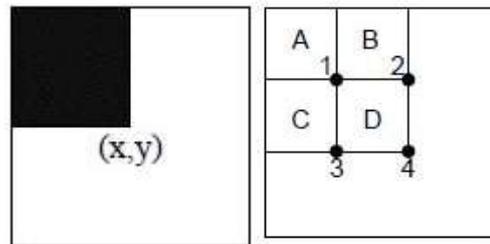
- $F(Haar)$  = Nilai fitur total
- $\sum F \text{ White}$  = Nilai fitur pada daerah terang
- $\sum F \text{ Black}$  = Nilai fitur pada daerah gelap

Kotak *rectangular Haar* dapat dihitung secara cepat dengan menggunakan Integral *Image*. Lalu untuk gambar bergerak (video), perhitungan dan penjumlahan piksel terjadi secara terus menerus dan membutuhkan waktu yang lama. Oleh karena itu, penjumlahan diganti dengan integral sehingga didapatkan hasil lebih cepat. Hasil deteksi dari *Haar* kurang akurat jika hanya menggunakan satu fungsi saja sehingga biasanya digunakan beberapa fungsi sekaligus. Semakin banyak fungsi yang digunakan maka hasilnya akan semakin akurat. Pemrosesan *Haar* yang banyak tersebut diorganisir atau diatur di dalam *classifier cascade* [2].

### B. Integral Image

Integral *Image* adalah sebuah citra yang nilai tiap pikselnya merupakan penjumlahan dari nilai piksel kiri atas hingga kanan bawah. Untuk memudahkan proses perhitungan nilai dari setiap fitur *Haar* pada setiap lokasi gambar digunakan teknik yang disebut citra integral. Secara umum integral mempunyai makna menambahkan bobot, bobot merupakan nilai – nilai piksel yang akan ditambahkan ke dalam gambar asli. Nilai integral dari setiap piksel merupakan jumlah dari semua piksel sebelah atasnya dan di sebelah kirinya. Keseluruhan gambar dapat diintegrasikan dengan operasi bilangan bulat per piksel [3].

Dengan menggunakan integral *image* dapat mengetahui nilai piksel untuk beberapa segiempat yang lain misalkan, seperti segiempat D pada Gambar 5 dapat dilakukan dengan cara menggabungkan jumlah piksel pada area segiempat A+B+C+D, dikurangi jumlah dalam segiempat A+B dan A+C, ditambah jumlah piksel di dalam A. Dengan A+B+C+D adalah nilai dari integral image pada lokasi 4, A+B adalah nilai pada lokasi 2, A+C adalah nilai pada lokasi 3, dan A pada lokasi 1 [4].



Gambar 3 Perhitungan nilai fitur dari citra integral

$$D = (A+B+C+D)-(A+B)-(A+C)+A \tag{2}$$

Sebagai contoh jika nilai integral Image titik 1 adalah A, titik 2 adalah A+B, titik 3 adalah A+C, dan titik 4 adalah A+B+C+D, maka jumlah piksel didaerah D dapat diketahui dengan cara  $4+1-(2+3)$ .

### C. Adaboost

Viola dan Jones menggunakan sebuah metode *machine learning* yang disebut *AdaBoost*. *AdaBoost* menggabungkan banyak *classifier* lemah untuk membuat sebuah *classifier* kuat. Lemah disini berarti urutan filter pada *classifier* hanya mendapatkan jawaban benar lebih sedikit. Jika keseluruhan *classifier* lemah digabungkan maka akan menjadi *classifier* yang lebih kuat. *AdaBoost* memilih sejumlah *classifier* lemah untuk disatukan dan menambahkan bobot pada setiap *classifier*, sehingga akan menjadi *classifier* yang kuat. *Viola – Jones* menggabungkan beberapa *AdaBoost classifier* sebagai rangkaian filter yang cukup efisien untuk menggolongkan daerah *image*. Masing – masing filter adalah satu *AdaBoost classifier* terpisah yang terdiri *classifier* lemah atau satu filter fitur [3]

$$h_t(x) = \begin{cases} 1, & \text{jika } p_t f_t(x) < p_t \theta_t(x) \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Keterangan :

- $h_t(x)$  = klasifikasi lemah
- $p_t$  = Parity ke t
- $\theta_t$  = *threshold* ke t
- x = adalah dimensi sub *image* misalnya  $24 \times 24$

langkah – langkah untuk mendapatkan *classifier* kuat dinyatakan dalam suatu algoritma berikut [1]:

- 1) Diberikan contoh gambar  $(x_1, y_1), \dots, (x_n, y_n)$  dimana  $y_i = 0$  untuk contoh positif dan  $y_i = 1$  untuk contoh negatif.
- 2) Inisialisasi bobot :
  - M = Jumlah negatif
  - L = Jumlah positif
  - N = M + L
- 3) Untuk  $t = 1, \dots, T$  menormalkan bobot( $D_t$ ) sehingga  $D_t$  adalah distribusi probabilitas.

Untuk setiap fitur, t melatih *classifier*  $h_t$ , untuk setiap fitur tunggal. Kesalahan ( $\epsilon_t$ ) dievaluasikan dengan bobot  $D_t$  :

$$\epsilon_t = \sum_i D_t | h_t(x_i) - y_i | \quad (4)$$

Dari persamaan diatas didapatkan sebuah *classifier* kuat dengan persamaan (5) :

$$h(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, & \text{Otherwise} \end{cases} \quad (5)$$

Dimana,

$$\alpha_t = \log \frac{1}{\beta_t} \quad (6)$$

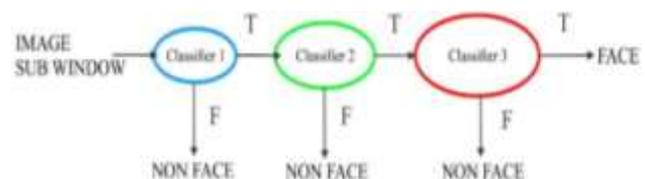
Dengan menggabungkan beberapa *AdaBoost classifier* sebagai rangkaian filter yang cukup efisien untuk menggolongkan daerah *image*. Masing – masing filter adalah satu *AdaBoost classifier* terpisah yang terdiri *classifier* lemah atau satu filter *Haar*. Selama proses pemfilteran, bila ada salah satu filter gagal untuk melewati sebuah daerah gambar, maka daerah itu langsung digolongkan sebagai

bukan wajah. Namun ketika filter melewati sebuah daerah gambar dan sampai melewati semua proses filter yang ada dalam rangkaian filter, maka daerah gambar tersebut digolongkan sebagai wajah [2].

#### D. Cascades Classifier

Tahap terakhir dalam metode *Viola – Jones* adalah *Cascade of Classifier*. Dengan mengkombinasikan pengklasifikasian dalam sebuah struktur *cascade* atau *Cascade of Classifier*, kecepatan dari proses pendeteksian dapat meningkat, yaitu dengan cara memusatkan perhatian pada daerah – daerah dalam *image* yang berpeluang saja. Hal ini dilakukan untuk menentukan di mana letak objek yang dicari pada suatu *image*.

Karakteristik dari algoritma *Viola – Jones* adalah adanya klasifikasi bertingkat. Klasifikasi pada algoritma ini terdiri dari tiga tingkatan dimana tiap tingkatan mengeluarkan subcitra yang diyakini bukan wajah. Hal ini dilakukan karena lebih mudah untuk menilai subcitra tersebut bukan wajah ketimbang menilai apakah subcitra tersebut berisi wajah. Dibawah ini merupakan proses rangkaian fitur yang dilalui oleh setiap *classifier* [3].



Gambar 4 Cascades Classifier[1]

Pada klasifikasi tingkat pertama, tiap subcitra akan diklasifikasi menggunakan satu fitur. Hasil dari klasifikasi pertama ini berupa T (*True*) untuk gambar yang memenuhi fitur *Haar* tertentu dan F (*False*) bila tidak. Klasifikasi ini kira – kira akan menyisakan 50% subcitra untuk diklasifikasi di tahap kedua. Hasil dari klasifikasi kedua berupa T (*True*) untuk gambar yang memenuhi proses integral *image* dan F (*False*) bila tidak.

Seiring dengan bertambahnya tingkatan klasifikasi, maka diperlukan syarat yang lebih spesifik sehingga fitur yang digunakan menjadi lebih banyak. Jumlah subcitra yang lolos klasifikasi pun akan berkurang hingga mencapai

jumlah sekitar 2%. Hasil dari klasifikasi terakhir berupa T (*True*) untuk gambar yang memenuhi proses *AdaBoost* dan F (*False*) bila tidak[3].

Struktur *cascade* mencerminkan fakta bahwa dalam setiap gambar tunggal, mayoritas dari *sub-window* negatif. Dengan demikian, *cascade* berguna untuk menolak negatif sebanyak mungkin dan secepat mungkin di level yang memungkinkan. Sementara contoh positif akan memicu evaluasi setiap *classifier* pada *cascade* [1].

### III. HASIL DAN PEMBAHASAN

beberapa parameter yang menjadi faktor diluar kebiasaan yang memungkinkan dapat mempengaruhi hasil pendeteksian wajah. Ini dilakukan agar dapat menguji tingkat akurasi dari aplikasi dalam mendeteksi wajah seseorang:

1) Berdasarkan umur balita



Gambar 5 Deteksi wajah balita 2 tahun pada layar ponsel



Gambar 6 Deteksi wajah balita pada gambar

2) Berdasarkan umur lansia

Terdapat deteksi wajah pada lansia yang ditampilkan di gambar 7 dan gambar 8. Pengujian ini diambil pada sampel gambar yang di hadapkan pada kamera aplikasi.



Gambar 7 Deteksi wajah lansia wanita

3) Berdasarkan gaya wajah dan jarak



Gambar 8 Deteksi wajah dengan jarak  $\pm 40$



Gambar 9 Deteksi wajah dengan jarak  $\pm 40$  cm dengan jumlah 2 orang



**Gambar 10** Deteksi wajah dengan jarak  $\pm 134$  cm dengan wajah menghadap kesamping



**Gambar 13** Deteksi wajah dengan bantuan Flash



**Gambar 11** Deteksi wajah dengan jarak  $\pm 260$  cm dengan menghadap depan

#### 4) Berdasarkan intensitas cahaya

Gambar hasil deteksi ini diambil pada saat malam hari dengan bantuan cahaya lampu ruangan.



**Gambar 12** Deteksi wajah dengan bantuan cahaya lampu ruangan

Pada Gambar 14 deteksi wajah menggunakan bantuan dari *flash* (cahaya) ponsel dengan posisi *flash* menghadap ke arah user.



**Gambar 14** Deteksi wajah dengan bantuan flash

Pada Gambar 15 deteksi wajah menggunakan bantuan dari *flash* (cahaya) ponsel dengan posisi *flash* berada pada sebelah kiri user.



**Gambar 15** Deteksi wajah dengan bantuan *flash*

Deteksi wajah pada Gambar 16 menggunakan bantuan dari *flash* (cahaya) ponsel dengan posisi flash berada di atas user.

5) Berdasarkan keadaan wajah



Gambar 16 Deteksi wajah dengan keadaan wajah

Wajah yang memiliki jenggot dan memakai kaca mata serta jubah terdeteksi oleh sistem pendeteksi, sedangkan wajah boneka bayi tidak terdeteksi pada sistem pendeteksi.

Wajah hewan tidak terdeteksi pada sistem hanya terdeteksi wajah manusia seperti yang terlihat pada Gambar 17.



Gambar 17 Deteksi wajah dengan wajah hewan

Berdasarkan hasil dari pengujian diatas, maka dapat dihitung tingkat akurasi dari pengujian pendeteksian wajah. Untuk tingkat akurasi *detection rate* diperoleh :

$$\text{Hasil} = \frac{21}{31} \times 100\%$$

Hasil = 67,6%

Dari hasil kalkulasi diatas diperoleh bahwa tingkat akurasi pendeteksian wajah berdasarkan pengujian yang penulis lakukan adalah sebesar 67.6%. sedangkan untuk tingkat *false* positif rate tidak ada, sehingga dari hasil pengujian ini dapat diartikan keakuratan pendeteksi wajah dengan cukup baik.

#### IV. KESIMPULAN

Adapun kesimpulan dari sistem deteksi wajah pada *camera real – time* dengan menggunakan metode *Viola Jones* ini adalah :

1. Pendeteksian wajah menggunakan metode *Viola Jones* memiliki kemampuan dan performansi yang cukup baik, hal ini terlihat berdasarkan pengujian yang dilakukan terhadap identifikasi wajah dalam kondisi yang memungkinkan dapat mempengaruhi hasil pendeteksian wajah, hasil tersebut menunjukkan tingkat akurasi sebesar 67.6% untuk mendeteksi wajah tersebut.
2. Sistem mampu mendeteksi wajah yang terhalang selama konturnya sama dengan kontur wajah pada template (misal, terhalang kacamata dan topi), dan sebaliknya.
3. Dalam penelitian ini juga dijelaskan sebuah *library* untuk mendeteksi bagian wajah menggunakan sebuah file xml yaitu *frontal\_face.xml* yang berfungsi sebagai *classifier* untuk mendeteksi wajah secara frontal.

#### REFERENSI

- [1] Viola, P., Jones, M., 2001, *Rapid Object Detection Using A Boosted Cascade of Simple Features*, IEEE Conference on Computer Vision and Pattern Recognition, Jauai, Hawaii.
- [2] Saputra, D, I, S., Anjar, W, S., Ramadhan, K, A, N., dan Pamungkas, R, A, 2017, Pelacakan dan deteksi wajah menggunakan video langsung pada webcam.
- [3] Putro, M, D., Adji, T, B., Winduratna, B., 2012. Sistem deteksi wajah dengan

menggunakan metode viola jones, Magister  
Instrumentasi Elektro FT UGM.

International Journal of Computer  
Applications (0975 –8887).

- [4] Ankur, S., Manjunath, T.N., Veena, N.,  
2016. A Survey on Face Recognition  
Technology – Viola Jones Algorithm.