

# Klasifikasi Citra Multi Wajah Menggunakan Domain Adaptive Faster Region Convolutional Neural Network

Ali Akbar Lubis<sup>1</sup>, Pahala Sirait<sup>2</sup>, Albert<sup>3</sup>, Andrya Tanrisono<sup>4</sup>, Andy<sup>5</sup>

Program Studi Teknik Informatika, STMIK Mikroskil

<sup>1</sup>ali.akbar@mikroskil.ac.id, <sup>2</sup>pahala@mikroskil.ac.id, <sup>3</sup>151110369@students.mikroskil.ac.id,

<sup>4</sup>151112718@students.mikroskil.ac.id, <sup>5</sup>151110733@students.mikroskil.ac.id

## Abstrak

Wajah merupakan bagian dari identifikasi biometrik karena wajah merupakan bagian langsung dari tubuh manusia yang tidak mudah untuk dicuri atau diduplikasi. Metode yang diterapkan dalam klasifikasi wajah seringkali hanya mampu mengklasifikasi citra yang di dalamnya hanya terdapat satu wajah, namun mengalami kesulitan dalam mengklasifikasi beberapa wajah dalam sebuah citra. Belakangan ini banyak dilakukan penelitian untuk mengatasi masalah pengenalan wajah ini menggunakan metode deep learning, salah satunya adalah Faster R-CNN dimana algoritma ini melakukan pencocokan menggunakan dataset kumpulan wajah manusia dengan mengekstraksi ciri-ciri yang membedakan setiap wajah dan dijadikan perbandingan untuk mengenali identitas wajah dengan akurat. Faster R-CNN tidak dilengkapi dengan pendeteksi wajah, oleh karena itu kita menggabungkannya dengan segmentasi warna kulit manusia untuk mengekstraksi bagian wajah untuk selanjutnya diidentifikasi menggunakan algoritma Faster R-CNN ini. Pengujian dilakukan dengan menggunakan citra yang di dalamnya terdapat lebih dari 1 wajah dengan berbagai posisi hadap wajah dan ekspresi wajah. Hasil dari pengujian yang telah dilakukan dengan rata-rata recall, presicion dan accuracy adalah recall sebesar 90.33%, presicion sebesar 85.33% dan accuracy sebesar 78.33%.

**Kata Kunci :** Klasifikasi Citra, Pengenalan Wajah, Faster R-CNN

## Abstract

The face is part of biometric identification because the face is a direct part of the human body that is not easy to steal or duplicate. The method applied in face classification is often only able to classify images in which there is only one face, but has difficulty classifying multiple faces in an image. Lately a lot of research has been done to overcome this face recognition problem using deep learning methods, one of them is the R-CNN Faster where this algorithm does a matching using a dataset of human face collections by extracting the characteristics that distinguish each face and used as a comparison to recognize facial identity accurately. The R-CNN Faster is not equipped with face detection, therefore we combine it with human skin color segmentation to extract part of the face for further identification using this R-CNN Faster algorithm. The tests carried out using images in which there are more than 1 face with various face to face positions and facial expressions. The result of test that has been carried out with an average of recall, presicion and accuracy are recall of 90.33%, presicion of 85.33% and accuracy of 78.33%.

**Keyword :** Image classification, Face Recognition, Faster R-CNN

## 1. PENDAHULUAN

Sistem biometrika merupakan teknologi pengenalan diri dengan menggunakan bagian tubuh atau perilaku manusia. Sistem akan mencari dan mencocokkan identitas seseorang dengan suatu basis data acuan yang telah disiapkan sebelumnya melalui proses pembelajaran. Sidik jari dan wajah merupakan contoh biometrika berdasarkan bagian tubuh. Pemanfaatan pengolahan citra pada bidang biometrika telah mengalami kemajuan yang sangat pesat. Salah satu bidang yang banyak diteliti dan dikembangkan adalah kemampuan komputer untuk dapat mengenali identitas seseorang melalui citra wajah [1].

Untuk deteksi wajah, metode yang banyak digunakan adalah algoritma *Viola-Jones*. Hal ini dikarenakan sistem deteksi wajah *Viola-Jones* dapat melakukan proses deteksi dengan akurat dan cepat yaitu sekitar 15 kali lebih cepat dari pendekatan yang sudah ada sebelumnya. Sedangkan untuk pengenalan wajah, metode *eigenface* dan *Euclidean distance* banyak digunakan karena bekerja dengan cara sederhana dan cepat. Metode ini pada dasarnya adalah mereduksi citra menjadi vektor ciri dimana vektor ciri memiliki ukuran yang sangat kecil dibandingkan ukuran aslinya, sehingga komputasi yang diperlukan sangat pendek. Namun, metode-metode tersebut hanya mampu melakukan pengenalan wajah untuk gambar wajah yang menghadap ke depan kamera dan tidak dapat digunakan untuk mendeteksi beberapa wajah manusia dalam sebuah gambar [2].

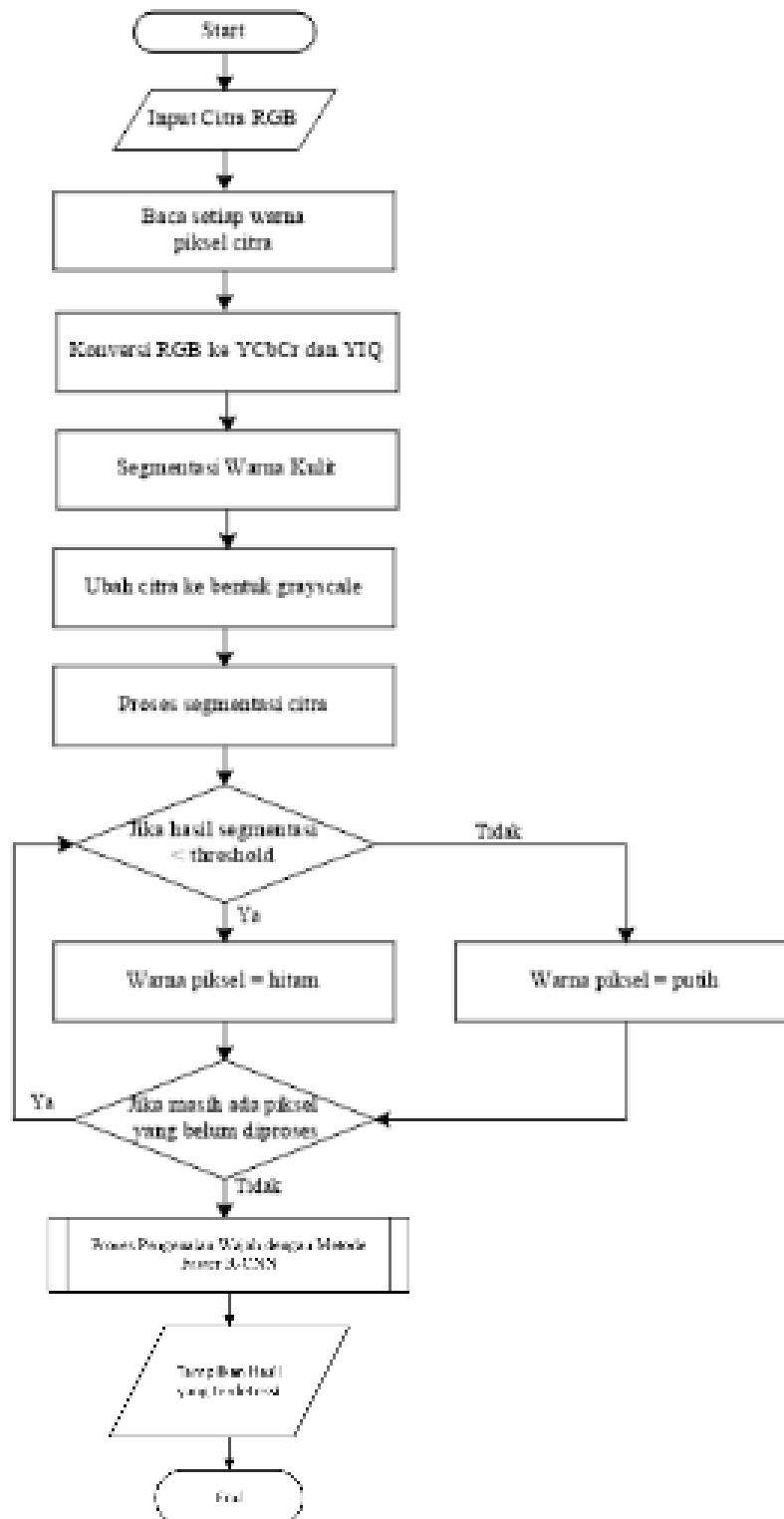
Beberapa penelitian tentang penerapan jaringan saraf tiruan seperti pengenalan wajah yang dilakukan dengan *Support Vector Machines (SVM)* dan *Jaringan Saraf Tiruan (JST)*. SVM bekerja dengan cara mendefinisikan batas antara dua kelas dengan jarak maksimal dari data yang terdekat, jarak maksimal ini didapatkan dengan menemukan *hyperplane* (garis pemisah) terbaik pada input *space* yang diperoleh dengan mengukur *margin hyperplane*, *margin* merupakan jarak antara *hyperplane* dengan titik terdekat dari masing-masing kelas, hanya saja SVM tidak dapat bekerja dengan baik untuk data yang bersifat linear [3]. Selain SVM, dalam melakukan pengenalan wajah dapat digunakan JST. JST merupakan jaringan dari sekelompok unit pemroses kecil yang dimodelkan berdasarkan jaringan saraf manusia. JST merupakan sistem adaptif yang dapat mengubah strukturnya untuk memecahkan masalah berdasarkan informasi eksternal maupun internal yang mengalir melalui jaringan tersebut dengan kata lain teknik ini memiliki kemampuan untuk belajar dari pengalaman. JST memiliki banyak metode yang dapat digunakan dalam pengenalan wajah [4].

Berdasarkan keunggulan penerapan algoritma berbasis JST pada berbagai pengenalan wajah, maka pada penelitian ini JST digunakan untuk proses pengenalan wajah [5]. Salah satu jenis algoritma yang menerapkan konsep jaringan saraf tiruan untuk melakukan proses pendeteksian objek adalah algoritma *Faster Region-Convolutional Neural Network (Faster R-CNN)*. *Faster R-CNN* adalah sebuah algoritma pendeteksi yang terdiri dari dua tahapan, yang mencakup tiga buah komponen, seperti *shared bottom convolutional layers*, *region proposal network (RPN)* dan *region-of-interest (ROI) based classifier*. Pertama, sebuah citra input akan direpresentasikan sebagai sebuah pemetaan ciri konvolusional yang dihasilkan oleh *shared bottom convolutional layers*. Berdasarkan pada peta ciri tersebut, RPN akan menghasilkan objek kandidat, dimana setelah itu *ROI-wise classifier* akan memprediksi label kategori dari sebuah vektor ciri yang dihasilkan dengan menggunakan *ROI-pooling* [6].

## 2. METODE PENELITIAN

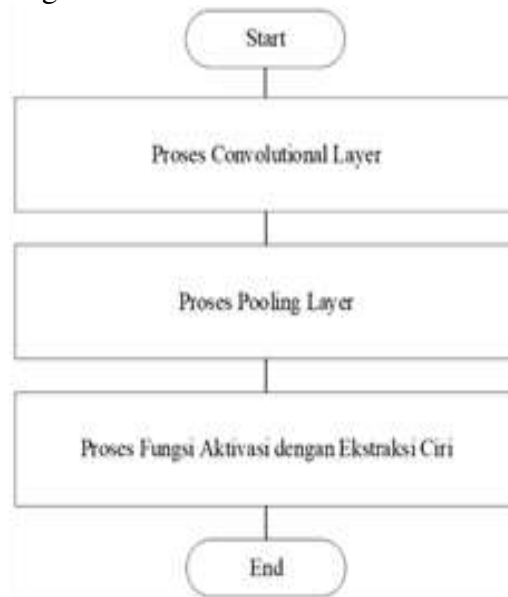
### 2.1 Analisis Proses

Analisis proses pendeteksiian objek pada penelitian ini menggunakan flowchart dapat dilihat pada Gambar 1 berikut:



Gambar 1 Flowchart Analisis Proses Pendeteksiian Objek

Salah satu tahapan pada Gambar 1 adalah proses pengenalan Wajah dengan menggunakan Metode Faster R-CNN. Dimana rincian proses yang dijalankan ada Metode Faster R-CNN dapat dilihat pada gambar 2 berikut:

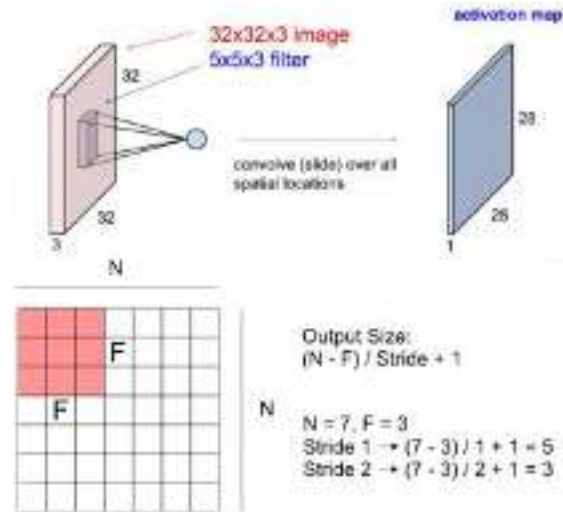


Gambar 2 Proses Pengenalan dengan Metode Faster R-CNN

Pada gambar 2 dapat dilihat terdapat 3 tahapan yaitu Convolutional layer, Pooling layer, dan Fully connected layer [6]:

1. Convolutional Layer

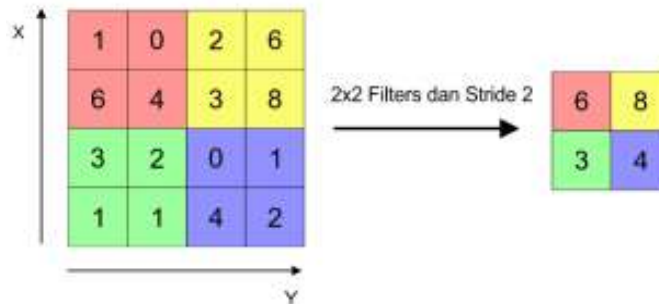
Seluruh data yang menyentuh lapisan konvolusional akan mengalami proses konvolusi. lapisan akan mengkonversi setiap filter ke seluruh bagian data masukan dan menghasilkan sebuah activation map atau feature map 2D. Filter yang terdapat pada Convolutional Layer memiliki panjang, tinggi (pixels) dan tebal sesuai dengan channel data masukan. Setiap filter akan mengalami pergeseran dan operasi “dot” antara data masukan dan nilai dari filter. Lapisan konvolusional secara signifikan mengalami kompleksitas model melalui optimalisasi outputnya. Hal ini dioptimalkan melalui tiga parameter, *depth*, *stride* dan pengaturan *zero padding*. Ilustrasi dapat dilihat pada gambar 3:



Gambar 3 Convolutional Layer [6].

## 2. Pooling Layer

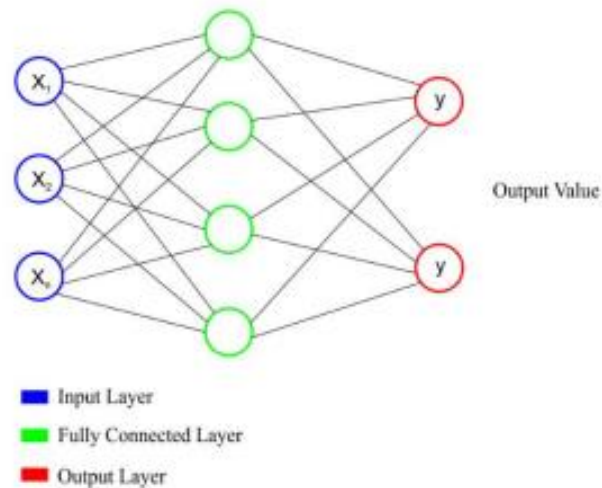
Pooling Layer merupakan tahap setelah Convolutional Layer. Pooling Layer terdiri dari sebuah filter dengan ukuran dan stride tertentu. Setiap pergeseran akan ditentukan oleh jumlah stride yang akan digeser pada seluruh area feature map atau activation map. Dalam penerapannya, pooling Layer yang biasa digunakan adalah Max Pooling dan Average Pooling. Sebagai contoh, apabila kita menggunakan Max Pooling 2x2 dengan Stride 2, maka pada setiap pergeseran filter, nilai yang diambil adalah nilai yang terbesar pada area 2x2 tersebut, Sedangkan Average Pooling akan mengambil nilai rata-rata. Ilustrasi Pooling layer dapat dilihat pada gambar 4:



Gambar 4 Pooling Layer [7]

## 3. Fully Connected Layer

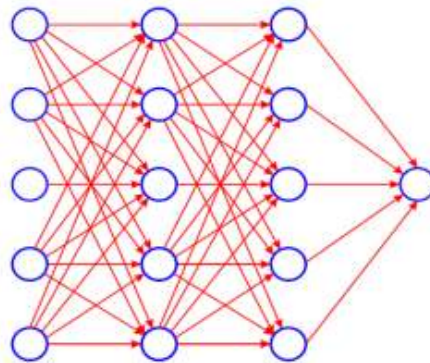
Feature map yang dihasilkan oleh tahap sebelumnya berbentuk multidimensional array. Sehingga, Sebelum masuk pada tahap Fully Connected Layer, Feature Map tersebut akan melalui proses “flatten” atau reshape. Proses flatten menghasilkan sebuah vektor yang akan digunakan sebagai input dari Fully Connected Layer. Fully Connected Layer memiliki beberapa Hidden Layer, Action Function, Output Layer dan Loss Function. Ilustrasi Fully Connected layer dapat dilihat pada gambar 5 :



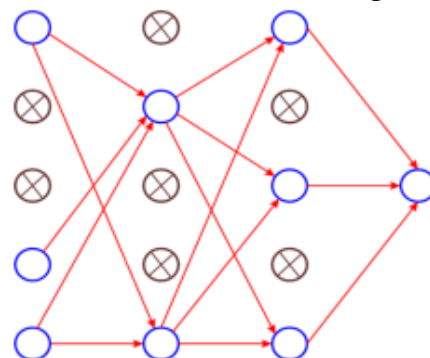
Gambar 5 Fully Connected Layer [7]

#### 4. Dropout

Dropout merupakan salah satu usaha untuk mencegah terjadinya overfitting dan juga mempercepat proses learning. Overfitting adalah kondisi dimana hampir semua data yang telah melalui proses training mencapai persentase yang baik, tetapi terjadi ketidaksesuaian pada proses prediksi. Dalam sistem kerjanya, Dropout menghilangkan sementara suatu neuron yang berupa Hidden Layer maupun Visible Layer yang berada di dalam jaringan. Ilustrasi proses dropout dapat dilihat pada gambar 6 dan gambar 7:



Gambar 6 Sebelum Dropout [6]



Gambar 7 Setelah Dropout [6]

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil

Hasil Pengenalan citra multi wajah akan dilakukan pada setiap wajah yang terdeteksi. Contoh hasil pengenalan citra multi wajah dapat dilihat pada gambar 8 berikut:



Gambar 8. Form Hasil Pendeteksian



#### 3.2 Pengujian

Pengujian akan dilakukan dengan menggunakan berbagai jenis citra multi wajah dengan posisi hadap wajah dan ekspresi wajah yang berbeda. Pengujian dilakukan dengan menggunakan *Confusion Matrix* dengan melibatkan nilai *Recall*, *Precision* dan *Accuracy*. Hasil pengujian dapat dilihat pada tabel 1 sampai dengan tabel 3.

Tabel 1 Hasil Pengujian untuk Pengenalan Multi Wajah Menghadap ke Depan

No.	Citra Uji	Jumlah Wajah	Jumlah Wajah Terdeteksi	Jumlah Objek Terdeteksi	Keterangan
1		4	3	0	Kenny = TP Albert = TP Andy = FN Andrya = TP
2		4	4	0	Kenny = TP Andy = TP Albert = TP Andrya = TP
3		4	4	0	Albert = TP Andrya = TP Kenny = TP Andy = TP



4		4	4	0	Andrya = TP Kenny = FP Albert = FP Andy = TP
5		4	4	1	Andrya = TP Andrya = FN Albert = TP Kenny = TP Andy = TP






$$\text{Recall} = 17 / (17+2) \times 100\% = 89\%$$

$$\text{Precision} = 17 / (17+2) \times 100\% = 89\%$$

$$\text{Accuracy} = 17 / (17 + 0 + 2 + 2) = 80\%$$

Berdasarkan tabel 1 menunjukkan pendeteksian multi wajah menghadap ke depan diperoleh recall, presisi dan akurasi yang sangat baik.

Tabel 2 Hasil Pengujian untuk Pengenalan Multi Wajah Menghadap berbagai arah berbeda

No.	Citra Uji	Jumlah Wajah	Jumlah Wajah Terdeteksi	Jumlah Objek Terdeteksi	Keterangan
1		4	3	0	Albert = TP Andrya = TP Andy = FN Agus = TP
2		4	4	0	Albert = TP Andrya = TP Andy = TP Agus = FP
3		4	4	0	Albert = TP Andrya = TP Andy = TP Munarso = TP
4		4	3	0	Andrya = TP Albert = TP Munarso = FP Andy = TP
5		4	2	0	Andrya = TP Albert = TP Munarso = FP Andy = TP








$$\text{Recall} = 16 / (16+1) \times 100\% = 94\%$$

$$\text{Precision} = 16 / (16+3) \times 100\% = 84\%$$

$$\text{Accuracy} = (16+0) / (16+0+3+1) \times 100\% = 80\%$$

Berdasarkan Tabel 2 pendeteksian multi wajah menghadap berbagai arah diperoleh recall, presisi dan akurasi yang sangat baik.

Tabel 3 Hasil Pengujian untuk Pengenalan Multi Wajah Berekspresi

No.	Citra Uji	Jumlah Wajah	Jumlah Wajah Terdeteksi	Jumlah Objek Terdeteksi	Keterangan
1		4	4	0	Andrya = TP Andy = TP Agus = FP Albert = TP
2		4	4	0	Andrya = TP Andy = TP Agus = FP Albert = TP
3		4	4	1	Andrya = FN Andrya = TP Agus = TP Albert = TP
4		4	4	0	Andrya = TP Albert = FP Munarso = TP Andy = TP
5		4	4	0	Andrya = TP Munarso = TP Andy = FN Albert = TP

$$\text{Recall} = 15 / (15+2) \times 100\% = 88\%$$

$$\text{Precision} = 15 / (15+3) \times 100\% = 83\%$$

$$\text{Accuracy} = (15+0) / (15+0+3+2) \times 100\% = 75\%$$

Berdasarkan Tabel 3 pendeteksian multi wajah berekspresi diperoleh recall, presisi dan akurasi yang sangat baik.

Berdasarkan hasil pengujian yang sudah dilakukan dapat dihitung rata-rata nilai recall, presisi dan akurasi untuk citra multi wajah dengan berbagai posisi hadap wajah dan ekspresi dapat dilihat pada table 4 berikut:

Tabel 4 Rekapitulasi Hasil Pengujian

Pengujian	Recall (%)	Precision (%)	Accuracy (%)
Wajah hadap ke depan	89	89	80
Wajah hadap ke berbagai arah	94	84	80

Wajah berekspresi	88	83	75
<b>Rata-Rata</b>	<b>90.33</b>	<b>85.33</b>	<b>78.33</b>

#### 4. KESIMPULAN

Berdasarkan pengujian yang sudah dilakukan dapat disimpulkan bahwa

1. Penerapan metode Faster R-CNN menghasilkan nilai rata-rata Recall, precision dan akurasi yang baik, dimana masing-masing bernilai 90%, 85% dan 78%.
2. Citra multi wajah yang menghadap kedepan dan ke berbagai arah menghasil nilai akurasi yang lebih baik daripada citra multi wajah berekspresi. Dimana nilai akurasi citra multi wajah hadap ke depan dan ke berbagai arah adalah 80%, sementara citra multi wajah berekspresi hanya 75%.
3. Citra multi wajah hadap ke depan menghasilkan precision paling baik pada penelitian ini dengan nilai 89%.

#### 5. SARAN

Adapun saran yang dapat diberikan dan membantu dalam penelitian selanjutnya adalah proses Segmentasi warna kulit manusia menyebabkan tingginya nilai False Negative karena mengidentifikasi objek selain wajah, maka dapat dikombinasikan dengan algoritma pendeteksian wajah lainnya yang dapat mendeteksi wajah berekspresi, menghadap ke berbagai arah dengan akurat.

#### DAFTAR PUSTAKA

- [1] Desiani, Anita., Arhami, Muhammad. 2006. Konsep Kecerdasan Buatan. Yogyakarta: Penerbit Andi.
- [2] Syuhada, Fahmi., Wijaya, I Gede Pasek Suta., Bimantoro, Fitri. 2018. Pengenalan Wajah Untuk Sistem Kehadiran Menggunakan Metode Eigenface dan Euclidean Distance, Prodi. Teknik Informatika, Fakultas Teknik, Universitas Mataram.
- [3] Ragul, G., Magesh, C. K., Thiyagarajan, R. 2013. Comparative study of statistical models and classifiers in face recognition. International Conference on Information Communication and Embedded Systems (ICICES), pp. 1-6.
- [4] Nur, Yanti., Fathur, Zaini Rachman., Nurwahidah, Jamal., Era, Purwanto., Fachrurozy. 2018. Jaringan Syaraf Tiruan Untuk Pengenalan Citra Sidik Jari Pada Smart Home. Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK), Vol. 5, No. 5, Oktober 2018, hlm. 597-604.
- [5] Yuhua, Chen., Wen, Li., Chirstos, Sakridis., Dengxin, Dai., Luc, Van Gool. 2018. Domain Adaptive Faster R-CNN for Object Detection in the Wild. 2018 IEEE Conference on Computer Vision and Pattern Recognition.
- [6] Santoso, Aditya., Gunawan, Ariyanto. 2018. Implementasi Deep Learning Berbasis Keras Untuk Pengenalan Wajah. Skripsi thesis, Universitas Muhammadiyah Surakarta.
- [7] Pangestu, Muftah A., Bunyamin, Hendra. 2018. Analisis Performa dan Pengembangan Sistem Deteksi Ras Anjing pada Gambar dengan Menggunakan Pre-Trained CNN Model. urnal Teknik Informatika dan Sistem Informasi (JuTISI), Volume 4, pp. 337-344.