

PENGENALAN WAJAH MENGGUNAKAN METODE PRETRAINED CONVOLUTIONAL NEURAL NETWORKS SECARA REAL-TIME PADA SISTEM ABSENSI BERBASIS WEBCAM

^aKholilul Rohman Kurniawan, ^bFajar Astuti Hermawati

^aTeknik Informatika, Universitas 17 Agustus 1945, Surabaya, Indonesia

^bTeknik Informatika, Universitas 17 Agustus 1945, Surabaya, Indonesia

E-mail: kholilulrk98@gmail.com, fajarastuti@untag-sby.ac.id

Abstrak

Sistem pencatatan kehadiran merupakan proses yang sulit jika dilakukan secara manual, bahkan dalam membuat recap data terkadang tidak sinkron atau terdapat *human error*. Sistem kehadiran cerdas dan otomatis yang berguna untuk mengelola data kehadiran dapat diimplementasikan menggunakan berbagai cara biometrik seperti pengenalan wajah salah satunya. Dengan menggunakan sistem ini, masalah yang terdapat pada pencatatan nantinya diharapkan dapat terselesaikan. Sistem absensi berbasis pengenalan wajah sebelumnya, ada beberapa kelemahan seperti masalah intensitas cahaya dan masalah pose kepala. Oleh karena itu untuk mengatasi masalah ini, ada berbagai metode pengenalan wajah yang digunakan. Di antara berbagai metode dalam sistem absensi yang terdapat pada sistem pengenalan wajah, sistem ini menggunakan metode pengenalan wajah menggunakan metode *Deep Learning*. Metode ini menggunakan *Convolutional Neural Networks* (CNN) yang memiliki ratusan layer untuk melakukan proses *learning* pada citra sehingga output menjadi lebih akurat. Dan proses *learning* tersebut digunakan untuk mengklasifikasikan citra wajah orang tersebut. Penelitian ini bertujuan untuk membangun sistem absensi secara real-time dengan trigger telapak tangan sebagai *capture* citra untuk mengenali wajah, serta menerapkan metode *Convolutional Neural Networks* (CNN) pada proses pengenalan wajah, dan metode *Viola Jones* untuk deteksi telapak tangan. Langkah utama dalam sistem ini adalah mendeteksi telapak tangan sebagai *capture* citra pada wajah dan mengenalinya. Perbandingan wajah yang terdeteksi dapat dilakukan dengan mencocokkan dengan dataset wajah siswa yang telah disimpan kedalam database dan mencatat kehadiran kedalam sistem. Sistem absensi cerdas ini diharapkan dapat menjadi cara yang efektif untuk menjaga kehadiran dan catatan siswa dengan proses learning yang cepat dan akurasi sistem pengenalan wajah yang tinggi. Rancangan sistem absensi cerdas telah berhasil diselesaikan pada penelitian ini. Evaluasi kinerja pada telapak tangan memiliki nilai *precision* sebesar 87,19 % dan *recall* sebesar 97,4 %. Hasil *training* menggunakan *pretrained* CNN memiliki akurasi sebesar 93,33 %. Pada pengujian pertama tingkat akurasi dengan posisi wajah lurus menghadap kamera yang di implementasikan pada sistem absensi ini sebesar 96 %, dan untuk pengujian kedua dengan posisi wajah bermacam macam memiliki tingkat akurasi pengenalan wajah sebesar 93%.

Kata kunci: Pengenalan Wajah, *Convolutional Neural Networks*, Sistem Pencatatan Kehadiran.

Abstract

The attendance recording system is a difficult process if it is done manually, even when making data recap is sometimes out of sync or there is human error. Intelligent and automatic systems that are useful for helping various experienced data can be implemented using biometric methods such as facial recognition, one of them. By using this system, the problems in recording are expected to be resolved. Previous face recognition based attendance systems, there were several drawbacks such as light intensity problems and head problems. Therefore, to solve this problem, various facial recognition methods are used. Among the various methods of attendance found in the facial recognition system, this system uses the facial recognition method using the Deep Learning method. This method uses Convolutional Neural Networks (CNN) which has hundreds of layers to carry out the learning process on the image so that the output becomes more accurate. And the learning process is used to classify the person's facial image. This study aims to build a real-time

attendance system with palm trigger as image capture to create faces, as well as applying the Convolutional Neural Networks (CNN) method in the face recognition process, and the Viola Jones method for detecting palms. The main step in this system is the supervision of the hand to capture images on faces and recognize them. Comparison of detected faces can be done by matching the dataset of student faces that are stored in the database and recording the presence of the system. This smart attendance system is expected to be an effective way to maintain student attendance and notes with a fast and accurate learning process with a high facial recognition system. The design of the intelligent attendance system has been successful in this study. The performance evaluation on the palms has a precision value of 87.19% and a recall of 97.4%. The results of training using CNN that have been previously trained have an accuracy of 93.33%. In the first test, the camera's accuracy by facing forward which is implemented in this attendance system is 96%, and for the second test with various facial positions with a face recognition rate of 93%.

Key words: Face Recognition, Convolutional Neural Networks, Attendance Recording System.

1. PENDAHULUAN

Seiring berkembangnya teknologi di era digital saat ini, salah satunya pada teknologi pengolahan citra yang banyak diterapkan pada proses kegiatan belajar mengajar khususnya pada sekolah menengah atas, yaitu pada sistem pencatatan kehadiran. Saat ini sudah banyak sistem absensi yang menggunakan proses pengambilan data berupa pengenalan identitas diri pada manusia secara otomatis dengan memanfaatkan sistem biometrik, salah satunya adalah pengenalan wajah, karena wajah memiliki banyak struktur sebagai sistem keamanan, contohnya seperti mata, alis, hidung, mulut, maupun struktur wajah lainnya. Dimana setiap individu memiliki ciri-ciri wajah yang berbeda atau unik [1].

Pada tiap wajah seseorang memiliki banyak variabilitas termasuk bentuk, warna dan ukuran. Faktor-faktor lain yang dapat mempengaruhi akurasi meliputi pencahayaan, obyek yang menghalangi bagian wajah meliputi topi, syal, sorban, hijab, kumis, jenggot, maupun kacamata dan ekspresi wajah seperti tersenyum atau tertawa [2].

Terdapat beberapa metode yang dapat digunakan untuk melakukan proses pengenalan wajah seperti, *Support Vector Machine* (SVM), *Principal Component Analysis* (PCA), dan *Convolutional Neural Networks* (CNN) [3]. *Principal Component Analysis* (PCA), dimana metode ini menggunakan algoritma *EigenFace* sebagai teknik untuk memotong atau mereduksi dimensi dari citra yang akan diproses. Dalam salah satu pengujian yang menggunakan metode *Principal Component Analysis* (PCA) didapatkan hasil

keakuratan sebesar 80% dalam mengenali citra dengan benar [4]. *Support Vector Machine* (SVM), untuk proses pengenalan wajah metode ini hanya dapat digunakan untuk data yang bersifat linear saja. Penggunaan metode *Support Vector Machine* (SVM) pada data non-linear juga dapat dilakukan, tetapi dengan memodifikasinya menggunakan fungsi kernel. Tetapi untuk dapat menemukan fungsi kernel yang sesuai dengan karakteristik data yang dipakai, maka harus menggunakan multi kernel. Dalam penelitian yang menggunakan metode *Support Vector Machine* (SVM) ini didapatkan hasil keakuratan yang mencapai 89% dalam mengenali wajah manusia dengan baik [5].

Diantara ketiga metode tersebut, metode *Convolutional Neural Networks* (CNN) memiliki hasil tingkat keakuratan yang cukup tinggi dibandingkan SVM dan PCA yang menggunakan algoritma *Eigenface*, yaitu sebesar 93,7% pada penelitian yang dilakukan Wagh Priyanka [6] Salah satu metode dari *Deep Learning* ini dapat memiliki ratusan layer untuk melakukan proses *learning* pada citra sehingga output menjadi lebih akurat, oleh karena itu dalam sistem pengenalan wajah yang akan dikembangkan ini digunakan metode CNN untuk melakukan proses *learning*.

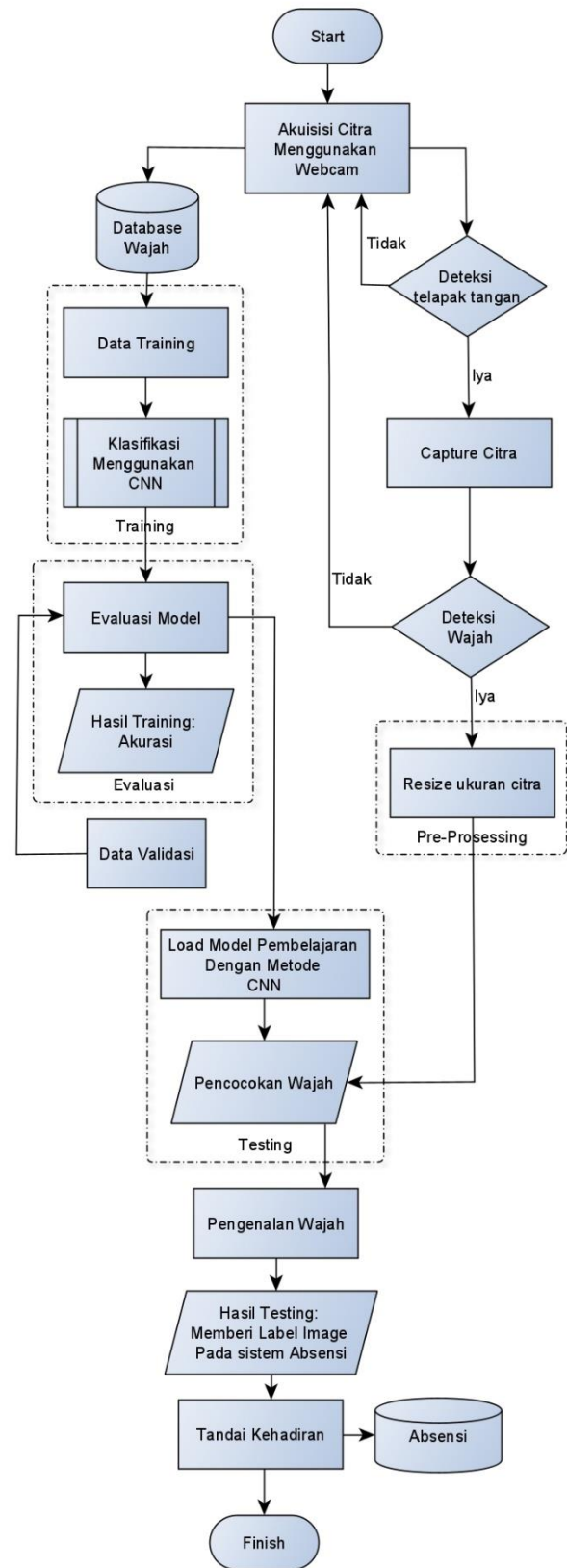
Metode pada proses pembelajaran dari *Deep Learning* ini yaitu metode CNN yang juga sebelumnya telah di terapkan system pengenalan wajah sebagai absensi yang berbasis CCTV dalam studi kasus di *FTP Polytechnic School*, dalam kasus ini CCTV tidak hanya di gunakan sebagai sistem keamanan saja, melainkan ditambahkan sedikit fungsi dari sistem keamanan tersebut sebagai pencatat kehadiran mahasiswa pada

perguruan tinggi di *FTP Polytechnic School*. Target dari sistem tersebut dapat menangani hingga 500 mahasiswa. Namun dengan adanya kendala penyimpanan terhadap data yang di dapat dari obyek yang di tangkap oleh CCTV dan juga tergantung pada *preprocessing*, maka sistem membatasi dataset yaitu 120 mahasiswa untuk dapat melakukan *preprocessing* [7].

Dalam penelitian ini akan dikembangkan sebuah sistem pengenalan wajah sebagai absensi menggunakan metode *Convolutional Neural Networks (CNN)* yang nantinya dapat mengambil citra yang didapat dari kamera dan memudahkan gambar untuk di proses, dengan penambahan konsep sebagai triger pada saat sistem memproses gambar tersebut yang menggunakan telapak tangan sebagai tanda pengenal bahwa citra nantinya akan segera di proses oleh sistem dan wajah tersebut dapat dikenali oleh sistem. Sistem agar dapat mengenali obyek yang telah tersimpan, maka langkah pertama dapat dilakukan pengumpulan dan menyimpan data berupa obyek dari siswa SMA berupa wajah kedalam database melalui input kamera, dan untuk proses pengambilan data kehadiran, target dapat mengarahkan telapak tangan terhadap input kamera sebagai aktivasi citra sehingga wajah dari target dapat di proses oleh sistem kehadiran dan telah melakukan absensi dengan pengenalan wajah.

2. METODE PENELITIAN

Secara garis besar rancangan sistem absensi cerdas pada penelitian ini digambarkan pada Skema metode yang ditunjukkan pada Gambar 1. Pertama dilakukan proses pembelajaran terhadap dataset wajah dan telapak tangan. Kerja sistem pada absensi cerdas ini dilakukan beberapa tahap yaitu, akuisisi pada citra, dilakukan deteksi telapak tangan, deteksi wajah dan pengenalan wajah, proses klasifikasi pada citra, dan pencatatan kehadiran yang nantinya akan disimpan kedalam database.



Gambar.1 Alur kerja sistem absensi cerdas

2.1 Deteksi Telapak Tangan

Telapak tangan merupakan obyek yang nantinya akan digunakan sebagai triger guna mengaktifkan *capture* citra pada sistem absensi yang berbasis pengenalan wajah. Sebelum dapat dideteksi sebagai telapak tangan, data pada telapak tangan akan dilakukan proses pembelajaran dan pelabelan sebagai telapak tangan. Dengan dilakukan beberapa tahapan agar sistem dapat mendeteksi obyek sebagai telapak tangan. Data yang dibutuhkan pada deteksi telapak tangan ini ada dua, dan pengumpulan data dilakukan secara manual, dataset telapak tangan dikumpulkan menjadi dua folder yaitu, citra *positive* dan citra *negative* pada dataset telapak tangan. Dataset pada *positive image* terdapat 500 citra telapak tangan menggunakan 10 model *sample* dari telapak tangan yang berbeda beda, yang mana didalam citra tersebut terdapat wajah, dan nantinya akan dilakukan tahap labeling pada bagian telapak tangan saja. Dataset pada *Negative Image* terdapat 612 gambar yang berupa background kelas maupun gambar real seperti citra wajah tanpa adanya tangan pada gambar tersebut.

2.2 Training Cascade Object Detector

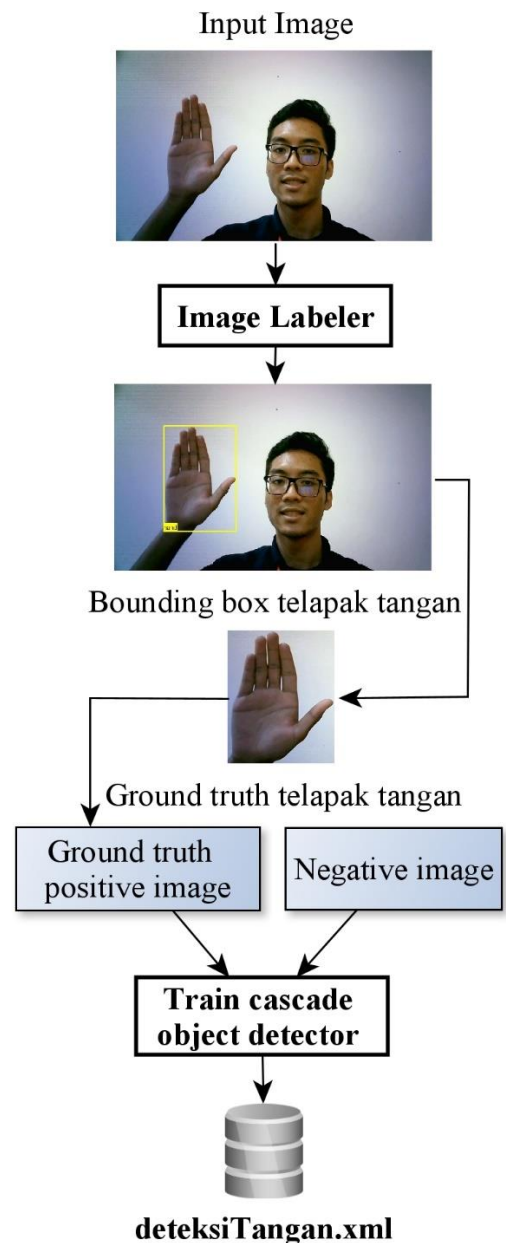
Proses training pada *Cascade of Classifiers* melibatkan dua data yang digunakan. Dalam kebanyakan kasus, pengklasifikasi dengan lebih banyak fitur akan mencapai tingkat akurasi yang lebih tinggi dan menurunkan tingkat positif palsu. Pada saat yang sama, pengklasifikasi dengan lebih banyak fitur memerlukan lebih banyak waktu untuk proses *training*. Pada prinsipnya, deteksi dapat bekerja dengan optimal yang mana pertama ditentukan *hyperparameter*: i) jumlah tahapan pengklasifikasi, ii) jumlah fitur di setiap tahapan, dan iii) ambang batas setiap tahapan, dipertukarkan untuk meminimalkan jumlah yang diharapkan dari evaluasi fitur. Sayangnya untuk menemukan hasil deteksi yang optimal, sampai saat ini masih menjadi masalah yang sangat sulit. Dalam praktiknya, proses kerja yang sangat sederhana digunakan untuk menghasilkan klasifikasi yang sangat efisien. Setiap tahapan dalam *Cascade* dilakukan pengurangan pada tingkat positif palsu dan menurunkan tingkat deteksi. Target dipilih untuk mengurangi rasio positif palsu dan menurunkan deteksi maksimum. Pada setiap tahap dilakukan proses data latih dengan menambahkan fitur hingga deteksi target

dan tingkat positif palsu terpenuhi (tingkatan ini ditentukan dengan menguji detektor pada set validasi). Selanjutnya tahapan ditambahkan sampai target keseluruhan untuk positif palsu dan tingkat deteksi terpenuhi [8]. Tahap labeling pada *Positive Image* pada telapak tangan dilakukan sebelum masuk ke proses *training*, guna mendapatkan ROI yang akan di deteksi. Pada proses *Image Labeling* ini menggunakan *tool* yang terdapat pada apps Matlab yaitu *Image Labeler*. Proses pertama adalah mendefinisikan ROI berupa “*Hand*” sesuai dengan objek yang akan dilabeli yaitu telapak tangan, kemudian proses selanjutnya adalah memasukkan semua *Positive Image* yang berupa 500 citra telapak tangan yang mana didalam citra tersebut juga terdapat wajah, proses ketiga yaitu memberi ROI label sebanyak 500 *Positive Image* pada objek yang akan dilabeli yaitu telapak tangan berupa *Bounding Box*, pada satu citra terdapat satu ROI telapak tangan. Proses terakhir adalah mengeksport ROI label yang telah di definisikan dalam bentuk tabel *groundtruth*. Hasil *Bounding Box* telapak tangan yang sudah diberi ROI label inilah yang nantinya akan digunakan dalam proses *training* menggunakan fungsi *cascade object detector* pada Matlab.

Tahap selanjutnya dilakukan proses *training* data pada *Bounding Box* telapak tangan yang sudah diberi ROI label menggunakan *train cascade object detector*. *Train cascade object detector* sendiri menggunakan algoritma *Viola-Jones* untuk mendeteksi wajah, hidung, mata, mulut, atau bagian tubuh lainnya. *Cascade object detector* juga dapat digunakan untuk melatih data berupa *object* lainnya, menggunakan *Image Labeler* pada *object* yang ingin di deteksi, contohnya seperti mendeteksi telapak tangan pada sistem ini. Sebelum dilakukan klasifikasi terhadap *object* yang akan di deteksi yaitu telapak tangan, langkah pertama adalah memuat *Bounding Box* telapak tangan yang telah diberi label sebelumnya dengan format *gTruth_hand.mat*, kemudian proses selanjutnya pilih kotak pada *Ground Truth* sebagai pertanda akan dieksekusinya tabel pada kolom yang telah dipilih, setelah itu menambahkan path pada folder *Positive Image* dan di tampung ke dalam variabel untuk diproses selanjutnya, begitupun juga pada *Negative Image* menentukan path pada folder yang sudah ditentukan, tahap

selanjutnya buat *object image data store* yang berisi gambar *negative* guna untuk memberi data sebagai citra yang tidak akan terdeteksi pada saat telapak tangan di deteksi sebagai *object* yang diinginkan. Tahap yang paling penting dalam proses ini yaitu *men-train* data menggunakan *train cascade object detector* dengan menentukan *hyperparameter* berupa *FalseAlarmRate* bernilai 0.4, *NumCascadeStages* bernilai 20 (sebagai nilai *default*), dan *FeatureType* dengan fitur HOG untuk mendeteksi objek telapak tangan. Output yang dihasilkan dari klasifikasi ini yaitu *deteksiTangan.xml*, nantinya file tersebut akan di gunakan pada parameter dari function *Viola-Jones*.

Setelah proses *training* pada dataset telapak tangan dilakukan, model yang didapat dari hasil *training* akan dilakukan deteksi telapak tangan menggunakan metode *Viola Jones* pada citra dengan kamera *webcam* sebagai input, dengan memasukkan tabel *groundtruth* berformat “xml” kedalam parameter algoritma *Viola Jones* dan ditampung kedalam suatu variabel yang mana variabel tersebut akan di lanjutkan kedalam proses deteksi wajah. Proses pengujian Deteksi telapak tangan pada sistem ini di dapat pada saat akuisisi citra menggunakan kamera webcam, dengan mengarahkan telapak tangan pada webcam berjarak 0.5 – 1 meter, selanjutnya sistem akan mendeteksi adanya telapak tangan pada input dengan memanggil fungsi *vision.Cascade Object Detector* dengan memasukkan model hasil data latih telapak tangan kedalam parameter, yang mana sistem akan otomatis mendeteksi adanya *object* yang akan digunakan sebagai triger yaitu telapak tangan. Sistem akan menampilkan *rectangle* dengan label *Hand* apabila mendeteksi telapak tangan, yang selanjutnya akan dilakukan proses deteksi pada wajah. Proses pembelajaran pada dataset telapak tangan dijelaskan pada gambar 2.



Gambar.2 Proses Pembelajaran Deteksi Telapak Tangan

2.3 Deteksi wajah

Wajah merupakan obyek pada tahap selanjutnya setelah sistem dapat mendeteksi telapak tangan sebagai triger yang nantinya citra wajah akan dideteksi dan dilanjutkan kedalam proses pengenalan wajah. Sebelum dapat dideteksi sebagai wajah. Deteksi wajah dalam sistem ini dilakukan menggunakan metode yang sama pada deteksi telapak tangan yaitu diproses

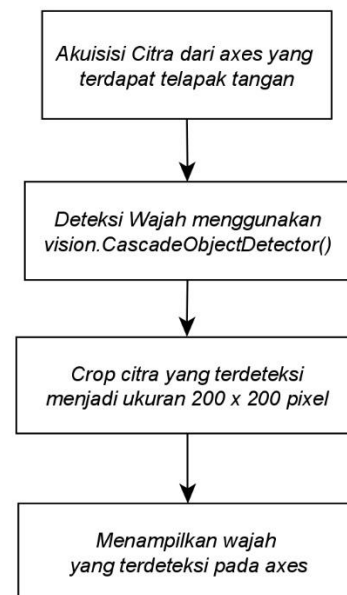
menggunakan metode *Viola Jones* dengan cascade deteksi obyek yang terdapat pada matlab.

Algoritma *Viola dan Jones* pertama kali digunakan untuk mendeteksi wajah, namun pada metode ini dapat digunakan untuk mendeteksi berbagai objek lain, tergantung dari data pembelajaran yang dipakai. Pada wajah manusia ada beberapa kesamaan yang kemungkinan besar dapat terjadi, sehingga dibutuhkan beberapa metode yang digunakan *Viola Jones* untuk mendeteksi wajah dengan tepat dan akurat pada suatu citra [9]. *Viola Jones* terdiri dari tiga metode seperti pada gambar.3, metode pertama yaitu *Haar-Like Feature* yang digunakan untuk ekstraksi fitur sebagai konsep pertama deteksi wajah pada gambar. Selanjutnya jika *Haar-Like Feature* ini ditemukan pada sebuah wajah, akan diteruskan ke metode yang kedua yaitu, metode *AdaBoost* yang digunakan untuk pemilihan pengklasifikasi - pengklasifikasi lemah yang dapat digabungkan menjadi pengklasifikasi kuat. Metode ketiga adalah metode *Cascade Classifier* digunakan untuk menyusun secara optimal kombinasi pengklasifikasi [10].



Gambar 3. Blok diagram metode *viola jones* [11]

Langkah awal untuk dilakukannya deteksi wajah pada sistem ini adalah dengan akuisisi citra sebagai input yang didapat dari axes setelah dilakukan capture citra pada proses deteksi telapak tangan, selanjutnya dilakukan deteksi wajah menggunakan *viola jones object detector*, apabila terdeteksi wajah pada citra akan dilakukan cropping citra yang terdeteksi menjadi ukuran 200x200 piksel yang bertujuan agar citra terfokuskan pada area wajah saja untuk mempermudah proses selanjutnya yaitu pengenalan wajah. Proses ini diperjelas dalam gambar.4.

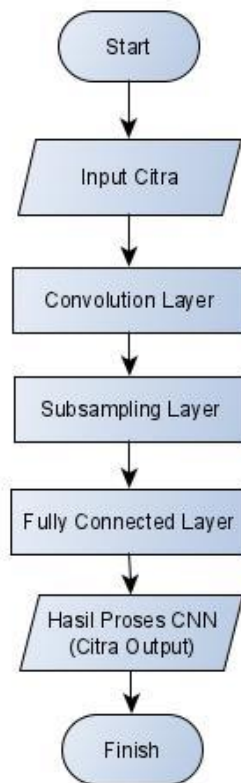


Gambar 4. Blok diagram metode *viola jones*

2.4 Pengenalan wajah

Tahapan pengenalan wajah merupakan tahapan setelah dideteksi adanya wajah dalam sistem absensi pada penelitian ini, dengan dilakukan proses pembelajaran pada dataset wajah menggunakan metode pretrained CNN sampai proses klasifikasi pada citra. Pada pengumpulan dataset wajah dilakukan pengambilan 1500 citra wajah dari 30 siswa yang mana 1 siswa dibutuhkan 50 citra wajah untuk dilakukan training nantinya. 1050 citra dimasukkan kedalam data train, dan sisanya 450 citra dimasukkan kedalam data test, untuk tiap siswa terdapat 35 citra pada tiap folder data train dan 15 citra pada tiap folder data test yang telah di beri label sesuai dengan identitas siswa tersebut. Untuk

pengambilan dataset wajah ini diambil secara *random* dengan jarak 0,5-1 meter dari kamera yang mana terdapat beberapa skenario dalam pengambilan seperti citra wajah dalam posisi yang bervariasi, ada yang menghadap ke atas, ke bawah, menghadap kesamping kanan 30 derajat maupun kesamping kiri 30 derajat.



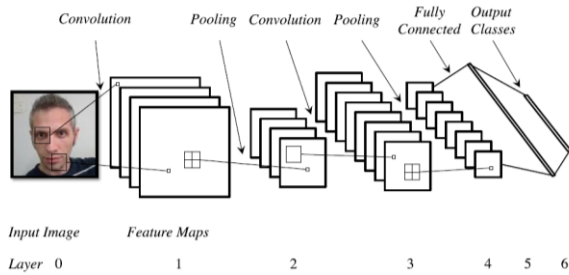
Gambar 5. Diagram Proses CNN

Tahapan pembelajaran terhadap dataset wajah pada penelitian ini menggunakan metode dari *deep learning* yaitu *Convolutional Neural Networks*. Gambar.6 menunjukkan Proses CNN terdapat ratusan layer untuk memprosesnya. Dari ratusan layer yang digunakan pada CNN, hanya terdapat tiga layer penting yang nantinya dilakukan untuk mengklasifikasi citra seperti pada gambar.5. Layer pertama yaitu *Convolution Layer*, pada lapisan convolution dilakukan proses konvolusi gambar input melalui serangkaian filter konvolusional, konvolusi dilakukan pada input data dengan menggunakan filter atau kernel seperti pada proses *Convolution Layer* pada gambar.7, kemudian menghasilkan output yang disebut *activation map* atau *feature map*. Pergeseran filter ditentukan oleh sebuah parameter yang bernama *stride* dan

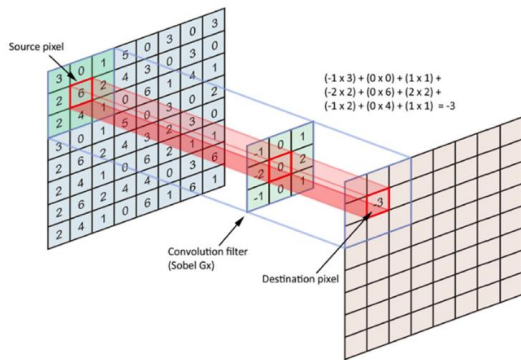
padding. *Stride* bergerak setiap kali dengan ukuran langkah biasanya 1 yang artinya filter bergeser piksel demi piksel. Dengan meningkatkan ukuran langkah (*stride*), filter akan berjalan di atas input dengan interval yang lebih besar dan dengan demikian memiliki lebih sedikit tumpang tindih antara sel. Selanjutnya menentukan parameter yang jumlah piksel (berisi nilai 0) yang akan ditambahkan di setiap sisi dari input. Hal ini digunakan dengan tujuan karena ukuran pada *Feature Map* selalu lebih kecil dari input, maka dari itu kita harus melakukan sesuatu untuk mencegah penyusutan *Feature Map* dengan menggunakan *padding*, sehingga kita bisa menggunakan *convolutional layer* yang lebih dalam dan menjaga ukuran spasial konstan setelah melakukan konvolusi [12].

Selanjutnya proses *subsampling layer* pada proses ini dilakukan metode *max pooling*. Gambar.8 menunjukkan proses *Max pooling*. *Max pooling* membagi *output* dari *convolution layer* menjadi beberapa *grid* kecil lalu mengambil nilai maksimal dari setiap *grid* untuk menyusun matriks citra yang telah direduksi, yang bertujuan untuk memungkinkan jaringan saraf convolutional mendeteksi objek ketika citra disajikan dalam berbagai posisi atau gambar dengan cara apapun [12].

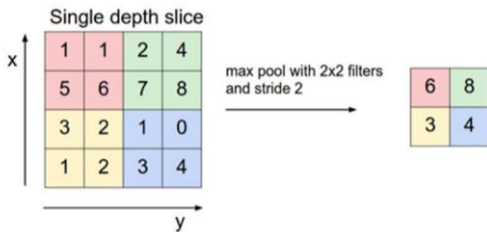
Pada gambar.9 menunjukkan Layer terakhir pada CNN yaitu, *fully connected layer*. Pada layer ini setiap *neuron* pada *convolution layer* perlu ditransformasi menjadi data satu dimensi terlebih dahulu sebelum dapat dimasukkan ke dalam sebuah *fully connected layer*, hal tersebut menyebabkan data kehilangan informasi spasialnya dan tidak *reversible*, karena pada *fully connected layer* lapisan yang sepenuhnya ini hanya dapat menerima data satu dimensi, sehingga perlu mengatur data tiga dimensi menjadi vektor satu dimensi dengan proses *flattening*. Proses terakhir yaitu fungsi aktivasi, pada tahap ini komputer akan menghitung nilai matriks dan dicocokkan dengan data latih, data latih yang memiliki nilai paling tinggi dari matriks maka itu yang akan dipilih oleh komputer untuk menentukan data apa yang ada pada matriks tersebut [12].



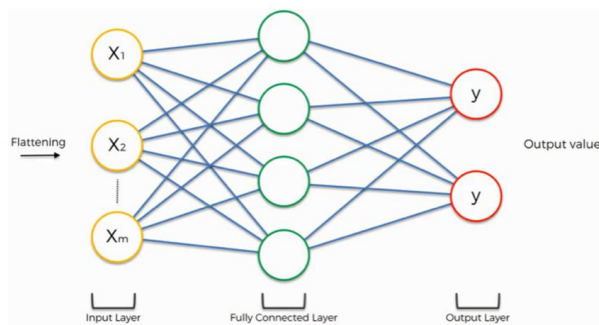
Gambar 6. Arsitektur CNN [13]



Gambar 7. Proses Convolution Layer [12]



Gambar 8. Proses Max Pooling [12]



Gambar 9. Lapisan fully connected layer [12]

2.5 Arsitektur Pretrained CNN

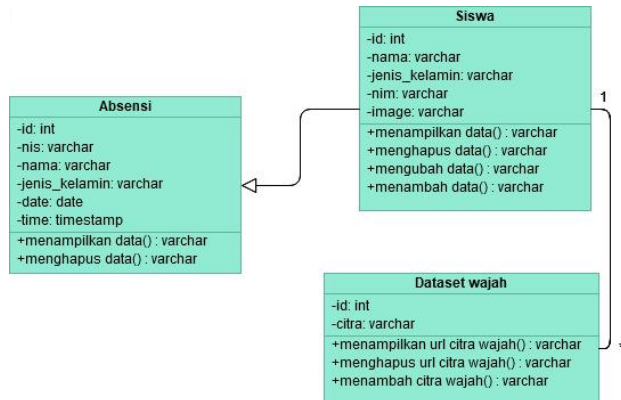
Pada perancangan sistem yang akan dikembangkan ini, metode CNN menggunakan jaringan yang sudah ditraining dengan teknik yang disebut dengan transfer learning, karena jauh lebih cepat dan lebih mudah dari pada melatih jaringan dari awal yaitu *alexnet*.

Arsitektur CNN AlexNet terdiri dari lima lapisan convolutional, tiga lapisan max-pooling, dan tiga lapisan fully-connected. Arsitektur CNN selengkapnya dapat dilihat pada Tabel 1. Pada susunan tersebut terdiri dari lima lapisan tersebut terdiri dari lima lapisan *convolutional*, tiga lapisan *max-pooling*, dan tiga lapisan *fully-connected*. Layer pertama yang bernama 'data' merupakan layer input yang ukurannya sama dengan ukuran citra. Ukuran citra pada dataset diubah menjadi ukuran citra yang digunakan pada AlexNet, yaitu 227x227x3. Lapisan tengah terdiri dari dua blok yang diulang yaitu lapisan *convolutional*, reLU, normalisasi dan lapisan *max-pooling*. Blok pertama terdiri dari lapisan *convolutional* ('conv1') dengan ukuran kernel 11x11 sebanyak 96 channel, selanjutnya yaitu reLU dan normalisasi terdapat 5 jumlah citra per elemen, kemudian diakhiri dengan *max-pooling* ('pool1') dengan ukuran kernel 3x3. Blok kedua sama dengan blok yang pertama dimana terdiri dari *convolutional layer* ('conv2') dengan ukuran kernel 5x5, reLU, normalisasi dan *max-pooling* ('pool2') dengan ukuran kernel 3x3. Dua blok yang diulang terdiri dari *convolutional layer*, yaitu ('conv3'), dan ('conv4') yang masing-masing kernel berukuran 3x3 dan reLU. Satu blok terdiri dari *convolution layer*, reLU dan *max-pooling*. Lapisan terakhir yaitu yang mencakup *fully connected layers* dan *softmax layer*. Di dalam arsitektur AlexNet ini, terdapat tiga *fully connected layers* ('fc6'), ('fc7'), ('fc8'), dan *softmax layer* ('prob').

Tabel 1. Arsitektur CNN

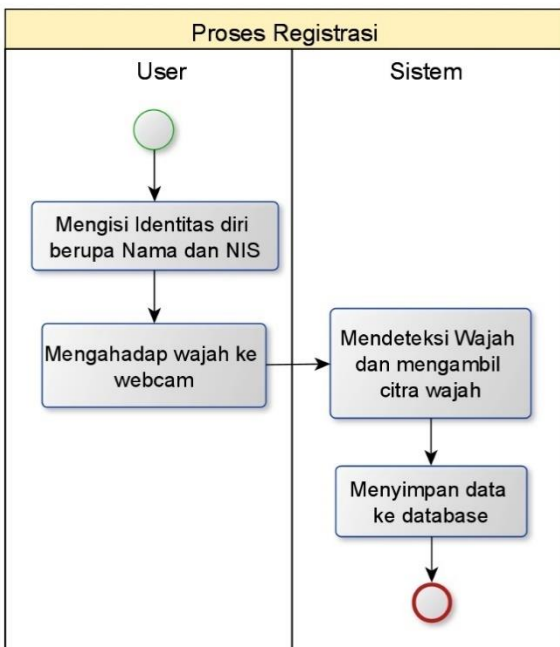
Nama	Channel	Ukuran	Kernel
'data'	3	227x227	
'conv1'	3	227x227	11x11
'pool1'	96	55x55	3x3
'conv2'	48	27x27	5x5
'pool2'	256	27x27	3x3
'conv3'	256	13x13	3x3
'conv4'	192	13x13	3x3
'conv5'	192	13x13	3x3
'pool5'	256	13x13	3x3
'fc6'	4096	6x6	6x6
'fc7'	4096	1	1x1
'fc8'	1000	1	1x1

2.6 Perancangan Sistem Absensi Cerdas



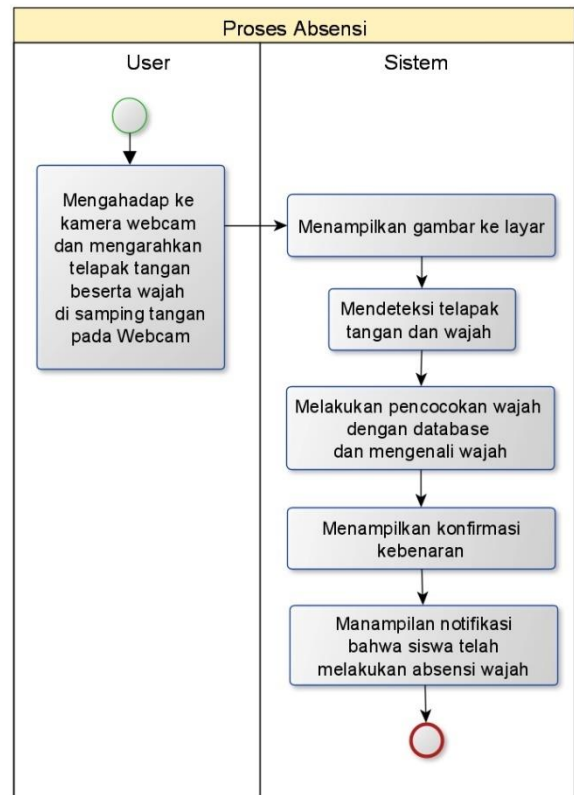
Gambar 10. Class Diagram Sistem Absensi Pengenalan Wajah

Gambar 10. menunjukkan kondisi obyek yang dinyatakan dalam *attribute/properties* pada sistem absensi pengenalan wajah yang nanti akan dirancang dalam penelitian ini. Dataset wajah menampung semua data wajah pada siswa yang nantinya akan dilakukan proses *learning*. Pada entitas siswa menyimpan semua identitas siswa berupa nomor induk, nama, jenis kelamin, dan wajah siswa itu sendiri. Pada entitas absensi menyimpan semua kehadiran dari siswa yang telah melakukan absensi seperti nomor induk, nama, jenis kelamin, dan waktu kehadiran siswa.



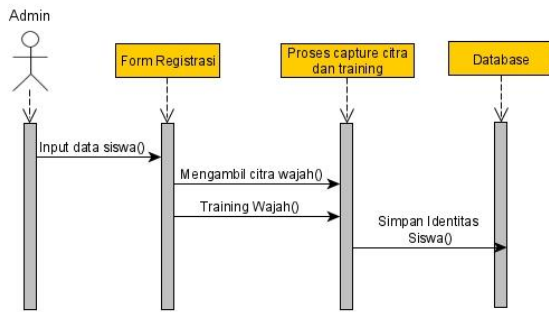
Gambar 11. Activity Diagram Proses Registrasi

Gambar 11. merupakan alur kerja atau proses registrasi yang dilakukan oleh siswa, agar identitas diri dapat dikenali oleh sistem. Dimana siswa mengisi identitas diri dan melakukan capture wajah yang nantinya data tersebut disimpan ke dalam sistem, agar siswa dapat melakukan absensi menggunakan pengenalan wajah.



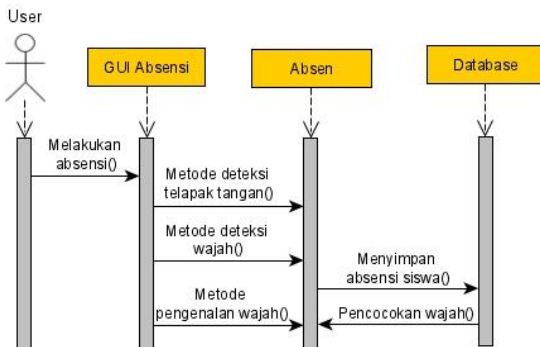
Gambar 12. Activity Diagram Proses Absensi

Gambar 12. merupakan alur kerja atau proses absensi dengan pengenalan wajah. Pertama user menghadapkan wajah ke webcam dan mengarahkan telapak tangan ke arah webcam agar terdeteksi, setelah terdeteksi adanya telapak tangan, selanjutnya sistem akan mendeteksi wajah yang terekam kedalam webcam dan melakukan proses pengenalan wajah dan mencocokkan wajah kedalam dataset wajah menggunakan metode CNN. Setelah wajah terkenali sistem akan mencatat kehadiran siswa kedalam database.



Gambar 13. *Sequence Diagram* Registrasi dan Training Wajah

Gambar 13. merupakan *sequence diagram* yang menjelaskan alur admin melakukan pengambilan identitas dan citra wajah pada siswa kemudian disimpa kedalam database, setelah itu admin melakukan proses training wajah siswa agar wajah dapat dikenali pada saat melakukan proses absensi dengan pengenalan wajah.



Gambar 14. *Sequence Diagram*

Gambar 14. merupakan *sequence diagram* yang menjelaskan alur user atau siswa saat melakukan absensi dengan pengenalan wajah. Dimana yang pertama siswa melakukan absensi dengan mengarahkan wajah beserta telapak tangan kearah webcam untuk dideteksi, setelah terdeteksi telapak tangan sistem akan mendeteksi wajah dan melakukan proses pencocokan wajah kemudian menyimpannya kedalam database.

3. HASIL DAN PEMBAHASAN

Percobaan dilakukan dengan menggunakan dua skenario. Skenario pertama adalah mengevaluasi kinerja pada deteksi telapak tangan. Yang kedua

adalah pengujian terhadap deteksi dan pengenalan wajah, dengan metode yang telah diusulkan. Pada deteksi telapak tangan Kami melakukan evaluasi kinerja dengan menghitung akurasi *precision* dan *recall*. Pada deteksi dan pengenalan wajah kami melakukan perhitungan akurasi terhadap dua pengujian yaitu, posisi wajah lurus menghadap kamera pada saat melakukan absensi dan posisi wajah bervariasi menghadap kamera pada saat melakukan absensi. Pengujian dilakukan menggunakan komputer dengan spesifikasi sebagai berikut: prosesor Intel Core i5-8250U 1.60 GHz 1.80 GHz dengan memori 12GB dan GPU NVIDIA GeForce 930MX, dan kamera webcam Full HD Pegatah 1080P.

3.1 Evaluasi kinerja deteksi telapak tangan

Pada tahap ini dilakukan evaluasi untuk mengetahui kinerja dari deteksi telapak tangan pada sistem ini, ada dua perhitungan untuk mengetahui kinerja dengan mencari nilai *precision* dan *recall*. Presisi bukanlah pembelajaran mendalam untuk konsep deteksi objek. melainkan konsep pembelajaran mesin yang sederhana. Dengan mencari nilai presisi untuk mengevaluasi kinerja pada telapak tangan. Presisi merupakan presentase dari item yang terpilih yang benar. Sedangkan *recall* merupakan presentase dari berapa banyak kasus positif yang diidentifikasi oleh model. Pada konsepnya proses ini bertujuan untuk mengevaluasi dua *bounding box* yaitu *Ground Truth Box* dan *Predicted Box* agar mengetahui kinerja dari deteksi telapak tangan. Dari hasil perhitungan *precision* dan *recall* untuk mengevaluasi kinerja pada deteksi telapak tangan pada sistem ini memiliki nilai 87,19% untuk *precision* dan 97,4% untuk *recall*.

3.2 Hasil training wajah menggunakan pretrained CNN

Pada proses *training* terhadap dataset wajah digunakan beberapa percobaan yang memakai beberapa parameter untuk hasil akurasi yang bagus, dengan merubah *MiniBatchSize* dan *MaxEpochs*. Pada tabel 2. menunjukkan perbandingan proses pelatihan dengan pretrained CNN menggunakan data latih pertama dan data latih kedua. Terlihat bahwa proses pelatihan data latih pertama memakan waktu lebih cepat dan jumlah iterasinya lebih signifikan dengan jumlah epoch yang lebih kecil. Dengan batch mini rata-

rata dari data pelatihan pertama lebih kecil daripada yang kedua, akurasi klasifikasi dari data pengujian pertama menunjukkan hasil yang lebih rendah daripada data pelatihan kedua.

Tabel 2. Perbandingan pelatihan CNN

	1 st Data Training	2 nd Data Training
Epoch	4	5
Iteration	152	155
Waktu (Detik)	289	365
Mini-batch loss average	30	36
Akurasi klasifikasi	91,94%	93,33%

3.3 Pengujian Sistem absensi

Pada tahap pengujian ini sistem diuji coba untuk mendeteksi dan mengenali objek wajah di dalam suatu ruang kelas pada saat pagi kondisi pencahayaan cukup terang sebelum kegiatan belajar mengajar berlangsung, pada pengujian ini kamera webcam diletakkan didalam kelas didekat pintu masuk dan pada saat melakukan proses absensi jarak antara webcam dengan siswa yaitu 0.5-1 meter. Data training dalam pengujian ini terdiri atas 1500 citra wajah dari 30 siswa yang berbeda dan dengan 50 pose yang berbeda, dengan kondisi wajah yang berbeda pula. Untuk pengujian deteksi dan pengenalan wajah pada sistem absensi ini dilakukan dua pengujian, untuk pengujian pertama terdapat beberapa bentuk yang diujikan, seperti:

1. Pengujian dalam bentuk pencahayaan yang cukup pada pagi hari
2. Pengujian dalam jarak wajah dengan kamera 0.5-1 meter
3. Pengujian dengan posisi wajah siswa lurus sempurna menghadap kamera.

Tabel 3 merupakan pengujian pertama pada sistem absensi ini dengan menguji sistem untuk mendeteksi dan mengenali objek wajah di dalam suatu ruang kelas pada saat pagi hari dengan pencahayaan yang cukup, pada pengujian ini webcam diletakkan didalam kelas dekat dengan pintu masuk, yang mana pada saat melakukan absensi kehadiran menggunakan pengenalan wajah, posisi wajah siswa pada pengujian satu

lurus tepat didepan kamera dengan jarak 0.5-1 meter, seperti gambar pengambilan pada tabel 3. Dari hasil pengujian didalam kelas dengan pencahayaan yang cukup pada pagi hari dan jarak antara webcam dengan siswa yaitu 0,5-1 meter, terdapat 29 dari 30 wajah yang dapat dikenali dengan tepat. Dalam waktu kurang lebih 1-3 detik sistem dapat mengenali wajah, dengan menggunakan triger telapak tangan sebagai capture pada citra.

Pada tabel 3 diperoleh jumlah data yang valid sebanyak 29 dari 30 data wajah yang diuji. Dengan persentase yang diperoleh:

$$\text{Akurasi} = \frac{\text{Jumlah Klasifikasi Benar}}{\text{Jumlah keseluruhan data uji}} \times 100$$

$$\text{Akurasi} = \frac{29}{30} \times 100$$



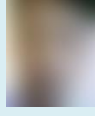

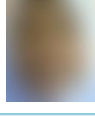
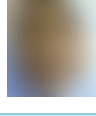
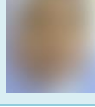
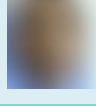
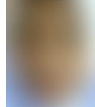

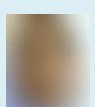
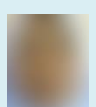








$$\text{Akurasi} = 0.96 \times 100$$








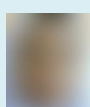

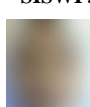
$$\text{Akurasi} = 96\%$$

Tabel 3. Hasil pengujian pertama

Wajah Yang Di deteksi	Identitas	Wajah Yang terdeteksi	Identitas	Nilai Kebenaran
	SISWA 1 200001		SISWA 1 200001	Benar
	SISWI 2 200008		SISWI 2 200008	Benar
	SISWA 3 200018		SISWA 3 200018	Benar
	SISWA 4 200021		SISWA 4 200021	Benar
	SISWA 5 200028		SISWA 5 200028	Benar

Wajah Yang Di deteksi	Identitas	Wajah Yang terdeteksi	Identitas	Nilai Kebenaran
	200038		200038	Benar
	200043		200043	Benar
	200052		200052	Benar
	200104		200104	Benar
	200106		200106	Benar
	200108		200108	Benar
	200117		200117	Benar
	200128		200128	Benar
	200157		200157	Benar
	200175		200175	Benar

Wajah Yang Di deteksi	Identitas	Wajah Yang terdeteksi	Identitas	Nilai Kebenaran
	200191		200191	Benar
	200204		200204	Benar
	200224		200224	Benar
	200234		200108	Salah
	200251		200251	Benar
	200263		200263	Benar
	200265		200265	Benar
	200283		200283	Benar
	200284		200284	Benar
	200292		200292	Benar

Wajah Yang Di deteksi	Identitas	Wajah Yang terdeteksi	Identitas	Nilai Kebenaran
	200305		200305	Benar
	200317		200317	Benar
	200320		200320	Benar
	200335		200335	Benar
	200338		200338	Benar

Pengujian kedua pada deteksi dan pengenalan wajah untuk sistem absensi ini dilakukan pengujian pada posisi wajah, untuk pengujian kedua ini terdapat beberapa bentuk yang diujikan, seperti:

1. Pengujian dalam bentuk pencahayaan yang cukup pada pagi hari
2. Pengujian dalam jarak wajah dengan kamera 0.5-1 meter
3. Pengujian dengan posisi wajah siswa yang bervariasi dengan menghadap kamera, dengan posisi wajah ada yang menghadap ke atas, ke bawah, menghadap kesamping kanan 30 derajat maupun kesamping kiri 30 derajat.

Tabel 4. merupakan pengujian kedua pada sistem absensi ini dengan menguji sistem untuk mendeteksi dan mengenali objek wajah di dalam suatu ruang kelas pada saat pagi hari dengan pencahayaan yang cukup, pada pengujian ini webcam diletakkan didalam kelas dekat dengan pintu masuk, yang mana pada saat melakukan absensi kehadiran menggunakan pengenalan

wajah pada percobaan kedua, posisi wajah siswa pada pengujian kedua menghadap kamera dengan jarak 0.5-1 meter dengan bermacam macam posisi wajah, ada yang menghadap keatas, menghadap kesamping, menghadap kebawah, memejamkan mata, dan ada pula wajah yang sedikit blur, seperti gambar pengambilan pada tabel 4.

Tujuan dilakukannya dua pengujian pada sistem ini yaitu, agar dapat mengetahui posisi wajah seperti apa yang dapat dikenali oleh sistem dengan akurasi yang cukup baik.

Dari hasil pengujian kedua didalam kelas dengan pencahayaan yang cukup pada pagi hari dan jarak antara webcam dengan siswa yaitu 0,5-1 meter, terdapat 28 dari 30 wajah yang dapat dikenali dengan tepat. Dalam waktu kurang lebih 1-3 detik sistem dapat mengenali wajah, dengan menggunakan triger telapak tangan sebagai capture pada citra.

Pada tabel 4. diperoleh jumlah data yang valid sebanyak 28 dari 30 data wajah yang diuji. Dengan persentase yang diperoleh:

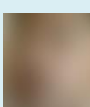
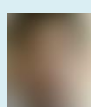
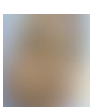
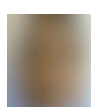
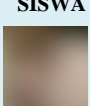
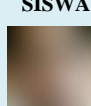
$$\text{Akurasi} = \frac{\text{Jumlah Klasifikasi Benar}}{\text{Jumlah keseluruhan data uji}} \times 100$$

$$\text{Akurasi} = \frac{28}{30} \times 100$$


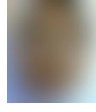


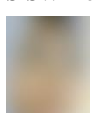



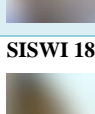
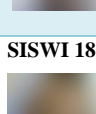

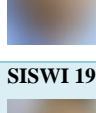

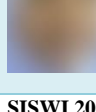

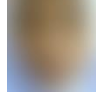

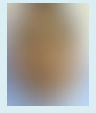
$$\text{Akurasi} = 0.93 \times 100$$










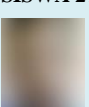

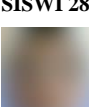
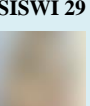



$$\text{Akurasi} = 93\%$$

Tabel 4. Hasil pengujian kedua

Wajah Yang Di deteksi	Identitas	Wajah Yang terdeteksi	Identitas	Nilai Kebenaran
	200001		200001	Benar
	200008		200008	Benar
	200018		200018	Benar

Wajah Yang Di deteksi	Identitas	Wajah Yang terdeteksi	Identitas	Nilai Kebenaran
	200021		200021	Benar
	200028		200028	Benar
	200038		200038	Benar
	200043		200043	Benar
	200052		200338	Salah
	200104		200104	Benar
	200106		200106	Benar
	200108		200108	Benar
	200117		200117	Benar
	200128		200128	Benar

Wajah Yang Di deteksi	Identitas	Wajah Yang terdeteksi	Identitas	Nilai Kebenaran
	200157		200157	Benar
	200175		200175	Benar
	200191		200191	Benar
	200204		200204	Benar
	200224		200043	Salah
	200234		200234	Benar
	200251		200251	Benar
	200263		200263	Benar
	200265		200265	Benar

Wajah Yang Di deteksi	Identitas	Wajah Yang terdeteksi	Identitas	Nilai Kebenaran
	200283		200283	Benar
	200284		200284	Benar
	200292		200292	Benar
	200305		200305	Benar
	200317		200317	Benar
	200320		200320	Benar
	200335		200335	Benar
	200338		200338	Benar

4. KESIMPULAN

Berdasarkan hasil implementasi dan evaluasi yang telah dilakukan, maka dapat diambil kesimpulan sebagai berikut:

1. Dari pengujian terhadap deteksi telapak tangan sebagai trigger capture pada sistem absensi ini terkadang masih mendeteksi obyek lain yang bukan telapak tangan seperti jendela, kaca, maupun tembok yang terdapat coretan. Tetapi dalam pengujian saat melakukan absensi kehadiran menggunakan background yang polos, sistem dapat mendeteksi telapak tangan dengan sempurna dan cepat.
2. Dataset yang dibutuhkan untuk proses training pada telapak tangan masih kurang, karena pada dataset yang di training menggunakan 500 *Positive Image* yang berisikan telapak tangan sebelah kanan saja dan 612 *Negative Image*, dikarenakan *Positive Image* yang dilakukan untuk tahap training yaitu telapak tangan bagian kanan saja, maka pada saat melakukan absensi menggunakan telapak tangan bagian kiri sistem tidak dapat mendeteksi telapak tangan tersebut.
3. Sistem dapat mendeteksi telapak tangan dengan sempurna dalam keadaan pencahayaan yang cukup dan dengan jarak 0,5-1 meter dari kamera.
4. Dari pengujian terhadap deteksi dan pengenalan wajah, sistem ini dapat mendeteksi dan mengenali wajah dengan sempurna dalam keadaan pencahayaan yang cukup dengan jarak pengambilan 0,5-1 meter dari kamera dengan posisi wajah lurus sempurna menghadap kamera. Membutuhkan waktu 1-3 detik agar sistem ini dapat mengenali wajah siswa dan menampilkan informasi kehadiran berupa identitas diri siswa dalam monitor.
5. Rancangan sistem absensi cerdas telah berhasil diselesaikan pada penelitian ini. Evaluasi kinerja pada telapak tangan memiliki nilai *precision* sebesar 87,19 % dan *recall* sebesar 97,4 %. Hasil *training* menggunakan *pretrained CNN* memiliki akurasi sebesar 93,33 %. Pada pengujian pertama tingkat akurasi dengan posisi wajah lurus menghadap kamera yang di implementasikan pada sistem absensi ini sebesar 96 %, dan untuk pengujian kedua dengan posisi wajah bermacam macam memiliki tingkat akurasi pengenalan wajah sebesar 93%.

REFERENSI

- [1] W. Setiawan *et al.*, “Multiple discriminant analysis with fukunaga koontz transform and support vector machine for image-based face detection and recognition,” vol. 7, no. 2, pp. 63–68, 2013.
- [2] M. Zufar and B. Setiyono, “Convolutional Neural Networks Untuk Pengenalan Wajah Secara Real-time,” *J. Sains dan Seni ITS*, vol. 5, no. 2, p. 128862, 2016.
- [3] F. Endrianti, W. Setiawan, and Y. Wihardi, “Sistem Pencatatan Kehadiran Otomatis di Ruang Kelas Berbasis Pengenalan Wajah Menggunakan Metode Convolutional Neural Network (CNN),” *JATIKOM - J. Apl. dan Teor. Ilmu Komput.*, vol. 1, no. 1, pp. 40–44, 2018.
- [4] A. Budi, S. Suma’inna, and H. Maulana, “Pengenalan Citra Wajah Sebagai Identifier Menggunakan Metode Principal Component Analysis (PCA),” *J. Tek. Inform.*, vol. 9, no. 2, pp. 166–175, 2018, doi: 10.15408/jti.v9i2.5608.
- [5] M. Athoillah, “Pengenalan Wajah Menggunakan SVM Multi Kernel dengan Pembelajaran yang Bertambah,” *J. Online Inform.*, vol. 2, no. 2, p. 84, 2018, doi: 10.15575/join.v2i2.109.
- [6] P. Wagh, R. Thakare, J. Chaudhari, and S. Patil, “Attendance system based on face recognition using eigen face and PCA algorithms,” in *Proceedings of the 2015 International Conference on Green Computing and Internet of Things, ICGCIoT 2015*, 2016, doi: 10.1109/ICGCIoT.2015.7380478.
- [7] N. T. Son *et al.*, “Implementing CCTV-based attendance taking support system using deep face recognition: A case study at FPT polytechnic college,” *Symmetry (Basel)*, vol. 12, no. 2, 2020, doi: 10.3390/sym12020307.
- [8] P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” 2001.
- [9] M. K. Dabhi and B. K. Pancholi, “Face Detection System Based on Viola - Jones Algorithm,” vol. 5, no. 4, pp. 2015–2017, 2016.
- [10] A. L. Pramana, “SISTEM ABSENSI BERBASIS PENGENALAN WAJAH MENGGUNAKAN METODE INDEPENDENT COMPONENT ANALYSIS MELALUI CCTV,” *J. Inf. Model.*, pp. 1–18, 2015.
- [11] Prasetya dkk, “Deteksi wajah metode viola jones pada opencv menggunakan pemrograman python,” *Simp. Nas. RAPI XI FT UMS*, pp. 18–23, 2012.
- [12] F. A. Hermawati, “Klasifikasi citra pada Deep Learning,” *J. Inf. Model.*, pp. 1–44, 2020.
- [13] M. Guarascio, G. Manco, and E. Ritacco, “Deep learning,” *Encycl. Bioinforma. Comput. Biol. ABC Bioinforma.*, vol. 1–3, no. D1, pp. 634–647, 2018, doi: 10.1016/B978-0-12-809633-8.20352-X.