Detection of Human Face Targets with Viola-Jones Method

Rahmiati[1], Ambiyar[2],Rika Melyanti[2], Eka Sabna[3,]

[1,]STMIK Amik Riau,Jl.Purwodadi Indah , Pekanbaru, Indonesia

[2] UNP, Jalan Prof. Dr. Hamka, Air Tawar , Padang, Indonesia

[3] STMIK Hangtuah , Jl. Mustafa Sari No.5 Tangkerang Selatan, Pekanbaru, Indonesia

[4] STMIK Hangtuah , Jl. Mustafa Sari No.5 Tangkerang Selatan, Pekanbaru, Indonesia

Article Info	Abstract
	This study aims to implement Viola Jones into a simple face detection system by utilizing the existing library in Open CV and using the Python programming language. In this study, there were 10 sample images consisting of 8 images of human faces and 2 images of animals. The results of the study were 9 images were successfully detected and 1 image was less precise in determining the face. The system can detect the presence of several (more than one) faces in an image. The system can also detect an object that resembles a face when the object has the same contour as a human face, for example, a face on a statue.

1. Introduction

The development of science and technology of object detection systems has developed very rapidly. This object detection system can be applied to an object or to a member of the human body. Biometric technology is a technology that utilizes human limbs as objects of detection. This technology is used as the basis for the detection system in the human body. In the medical field, it is known that some members of the human body are unique or different from other humans. One of these body parts is the face. Differences in human faces are caused by changes or differences in facial expressions such as: laughing, smiling, angry, sad, and others. Many methods have been applied to detect

or recognize human faces. Those methods are *Artificial Neural Network, Fourier Decomposition, Eigenface, Fisherface, Hidden Markov Models, Principal Component Analysis*. Yanti (2013), using the *Fourier decomposition* ordecomposition *Fourier* to recognize faces.

Social media such as *Facebook* is one of the media applications that have implemented human face detection. Where a photo is*uploaded* it will automatically detect the human face.

Today a lot of research on face detection is developing very quickly. Many methods have been applied to face target detection, one of which is by using thealgorithm *Viola-Jones*. Themethod *Viola-Jones* has a fairly high level of accuracy. This is

Rahmiati.

Email : rahmiati@sar.ac.id

because in this method several concepts are combined into a method for detecting objects (faces).

Another study conducted a performance study of popular facial recognition methods, distinguishing facial recognition methods into 3 (three) categories, namely 1) appearance-based, where the Kernel PCA method using the SVM classifier produces an accuracy of 99.05% which is the best performance compared to other methods in the same category; 2) feature-based, the best performance in this category is the SIFT & MLBP method with an accuracy of 99.73%; and 3) soft computing-based, Gabor Jet integrated with Borda Count Classification (BCC) produces 99.8% accuracy . The integration of the Gabor Jet and BCC methods has the best performance in this category. In one study also divided facial recognition methods into Holistic Matching Methods, Feature-based (structural) Methods and Hybrid Methods [36]

Based on this, this research will explain howworks *Viola Jones*, and apply it to the detection system by utilizing the libraries *Open CV* and *Python. Open CV was* chosen because *Viola Jones is* already included in the detection library, while *Python was* chosen because of its simplicity compared to other programming languages. After the system is completed, testing of facial character image samples is carried out.

Biometrics comes from the Greek words *bios* and *metron*. Bios means life and *metron* means measure. So Biometrics can be interpreted as an automatic method to identify an individual based on one or more body parts of humans or the attitudes of humans themselves that are special or special. (Wikipedia, nd)

Biometrics is a technology to identify a person's biological characteristics. The characteristics of a person are relatively different from person to person or it can be said that each individual has unique characteristics. This uniqueness can then be used to identify and verify a document, application, system, computer, and so on. This technology can be in the form of a fingerprint scanner, retinal scanner, face recognition and so on. (BINUS, nd)

Many studies have been carried out to unravel the complexity of studying the accuracy of face recognition, among others, by giving different lighting treatments to targets with several poses (boontua, 2018). The result is that the same or almost the same lighting treatment between the training data and the test data gives the result that the accuracy of face recognition is higher than the lighting treatment which is much different. In addition to lighting and poses, another uncontrollable condition is the color or tone of the face. Face hue causes different levels of lighting captured by the camera, in the experiment color images increase the accuracy of facial recognition higher than grayscale images [28].

In other studies also adopted the Development Network (DN) which has an accuracy of more than 95% in recognizing target faces with complex backgrounds (Wang, D., & Liu, L. (2015).The uncontrolled environment forced researchers to recognize multiple targets by combining the Haar Cascade and Eigenface methods to detect 55 faces in one process with an accuracy of 91.67% (Mantoro, T., Ayu, M. A., & Suhendi. (2018) . Another study revealed that the Minimum Eigen Value using the SVM classifier resulted in better facial recognition accuracy for multiple targets, which was 83% [32]. The researcher also examines real-time multiface recognition using the

Viola-Jones method as a face detection method and performs feature extraction using Speeded up Robust Features (SURF), and utilizes the M-estimator Sample Consensus (MSAC) in face matching [33]. The accuracy of face recognition by integrating the three methods is 95.9%

Biometric technology is usually applied to analyze human physical and attitude for authentication purposes.

Biometrics is a technology that can be used to identify a person based on physiological characteristics*(physiological)*that is characteristic of the body that are unique. In fact, there are some parts of the human body that are not similar to other humans, such as the face, fingerprints, and the retina of the eye. Basically the human eye has the same shape and color, but the retina is generally not the same. Likewise with voice, facial texture, and fingerprints. These body parts are then developed and used as attributes, for example in the security sector.

Application of Biometric Technology There are many applications of biometric technology that have developed rapidly to date. Although widely available, there are five biometric systems that are currently widely used:

A. Retina Retinal

Scanners have the ability to recognize the unique patterns that exist in each person. By identifying the pattern as an entry, retina scanners can be used to secure devices from unauthorized access.

B. Iris Scanner

Just like retinal scanners, iris scanners can also scan unique patterns inside the eye but in a more complex way. In addition, the use of an iris scanner is also said to be more expensive than a retina scanner.

C. Your Finger print Scanner

Be familiar with this biometric security system. Most smartphones nowadays embed fingerprint scanner biometric technology. Besides being quite safe to use, this biometric technology is also said to be the cheapest compared to other scanners. Fingerprint scanners work by printing fingerprint patterns in three dimensions and then storing them to recognize the same fingerprint pattern.

D. Face Biometric

Everyone has a different facial structure and appearance. Therefore, the biometric security system also relies on the face as an identification of a person's identity. By scanning every structure on the face such as the eyebrows, the distance between the eyes, and the nose, the system can be an alternative to today's security.

E. Voice Recognition

Just like other parts of the body, everyone's voice has a different tone. If you're using the Google Assistant digital assistant, you'll find the trusted voice setting. This system is used only to recognize the voice that was first recorded for authentication. (Prima Fauzi, 2017)

2. Research Methods

3.1. Methods
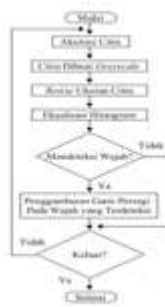Flowchart design from the Viola-Jones
is as shown in Figure 3.1 below:

Figure 2.1 Method flowchart *viola-jones*

Face objects are searched using the Viola-Jones Algorithm which refers to the flow chart above, where the process is:

a. Grayscale images will be scanned per sub-window.
b. Haar-like feature Image classification is based on the value of a feature. It aims to separate the image that is not needed. The haar feature processes the image into several squares, where one box consists of several pixels. Furthermore, the boxes are processed and the resulting difference or difference in the threshold value (threshold) which indicates the dark area and the light area. These values will then be used as the basis for image processing.
c. Ada-Boost functions to look for features that have a high level of differentiation. This is done by evaluating each feature against the largest between facial and non-face considered as the best feature. AdaBoost is a machine learning algorithm that aims to specifically select features that are considered important and conduct training on several classifiers that have been formed. AdaBoost has a series of filters that are efficient enough to classify regions in an image.

d. Cascade Classifier is a cascade classifier. The classification of this algorithm consists of several levels and each level produces a sub-image that is believed to be not a face. Cascade classifier is a classification method whose job is to remove non-faced images using a strong classifier that has been trained by AdaBoost at each classification level.

3.2. Data Source

To be used are several color image files (RGB) with JPG/JPEG/PNG extensions that have been downloaded via *Google Image*. In this study, 10 test images were used for face target detection using the Algorithm *Viola-Jones*.

3. Results and Discussion

The face is one part of the human body that becomes the center of attention when interacting with other people. The definition of face according to the Big Indonesian Dictionary is part of the head; facial expression. According to Wikipedia Indonesia, the face is the front part of the human head which consists of: forehead to chin, hair, eyebrows, nose, cheeks, eyes, mouth, lips, teeth and skin. The face is usually used to express oneself, appearance, and identity. Faces are unique, that is, none of the faces are absolutely identical, even in identical twins. (*Face -Indonesian Wikipedia, Free Encyclopedia*, nd)

The face is a part of the human body that can be recognized using biometric technology. In biometric technology, physical characteristics of humans such as faces can be used as unique information. This unique information can be in the form of characteristics of the facial pattern of each individual. The characteristics of the facial pattern can be measured and analyzed

for the detection process. Therefore, the face can be used as a *feature* or sign to identify someone.

The challenges faced in the problem face detection is caused by the following factors: · Face position. The position of the face in the image can vary because its position can be upright, tilted, turned, or viewed from the side. Components of the face that may or may not be present, such as a mustache, beard, and glasses. Facial expressions. Facial appearance is strongly influenced by a person's facial expressions, for example smiling, laughing, sad, talking, and so on. · Obstructed by other objects. The image of a face can be partially obscured by another object or face, for example in an image containing a group of people. Image capture conditions. The image obtained is strongly influenced by factors such as the light intensity of the room, the direction of the light source, and the characteristics of the sensor and camera lens.

Detection was originally introduced by Haar-Like and then developed by Viola-Jones. Face detection can be viewed as a pattern classification problem where the input is the input image and the output will be determined in the form of a class label from the image. In this case there are two class labels, namely face and non-face. (M Dwisnanto, 2012.

Face target detection methods are widely used in digital image processing. This method is the basis for tracking and recognizing *(recognition)* the person's face. Detection of facial targets can be done in conditions of facial variations, differences in appearance, differences in lighting, and variations in the pose of the face.
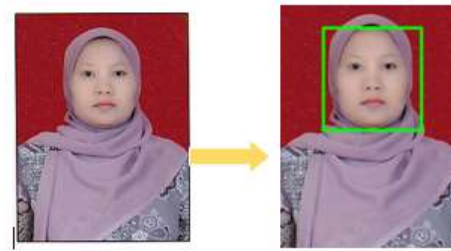

Figure 3.1 Face detection

For the case of taking photos when making SIM, KTP, and Credit Cards, the resulting photo or image generally only contains one face with the same or uniform background, as well as the lighting intensity that has been set previously. In the condition of the image (photo) like this, the face detection process can be done more easily. However, for images with conditions such as: containing more than one face, and containing various backgrounds, different lighting intensities, and various face sizes. An example is the image obtained by terminals, buildings, markets, airports and others. The face detection process can be considered as a pattern classification process from the input image where the output is a label or class category of the image. Where in this case there are 2 class labels namely face and not face.

Detection was originally introduced by Haar-Like and then developed by Viola-Jones. Face detection can be viewed as a pattern classification problem where the input is the input image and the output will be determined in the form of a class label from the image. In this case there are two class labels, namely face and non-face. (M Dwisnanto, 2012)

Face target detection methods are widely used in digital image processing. This method is the basis for tracking and recognizing *(recognition)* the person's face. Detection of facial targets can be done in conditions of facial variations, differences

in appearance, differences in lighting, and variations in the pose of the face.

There are several problems in the face detection process caused by the following factors:

A. Blocked or covered by other objects. In this case the face image is partially blocked by a face or other object. Example: an image containing a group of people.

B. Lighting Conditions. The resulting image is strongly influenced by several factors, namely: the intensity of lighting in the room, the characteristics of the camera lens or sensor.

C. Position of the face. Differences or variations in the position of the face in the image such as: perpendicular, tilted, turned to the side.

D. The face contains components such as: beard, mustache, beard and glasses. Facial expressions. Facial expressions greatly affect the appearance of a person's face. Expressions such as: smiling, angry, sad, talking and others. (et al, IGA Widagda)

Detection with Viola Jones uses a simple Haar-like feature that quickly evaluates new image representations. Viola Jones generates feature sets with integral imagery and boosting algorithms to reduce time complexity. Before being entered into the system, the image must first be searched for its gray value (grayscale). In general, the Viola Jones method has four basic processes, namely:

A. Haar-like feature Image classification is based on the value of a feature. It aims to separate unnecessary images, in this case, the background is not counted.

B. Integral Image Integral image is a data structure and algorithm that sums the values in a subset of the image matrix.

Working Principle of Face Detection with Viola-Jones method Face Detection with Viola Jones uses a simple Haar-like feature that quickly evaluates new image representations. Viola Jones generates feature sets with integral imagery and boosting algorithms to reduce time complexity. Before being entered into the system, the image must first be searched for its gray value (grayscale). In general, the Viola Jones method has four basic processes, namely:

A. Haar-like feature Image classification is based on the value of a feature. It aims to separate unnecessary images, in this case, the background is not counted.

B. Integral Image Integral image is a data structure and algorithm that sums the values in a subset of the image matrix

C. Ada-Boost In practice, none of the features are capable of classifying with a small error [1]. The Ada-Boost algorithm is used to find features that have a high level of differentiation. This is done by evaluating each feature against the largest between facial and non-face considered as the best feature.

Cascade Classifier The characteristic of the Viola Jones method is the existence of a cascade classifier. The classification in this algorithm consists of several levels and each level produces a sub-image that is believed to be not a face. This is done because it is easier to assess non-faced sub images than to assess whether they contain faces. (Tarhini, 2011).

Open CV (Open Source Computer Vision Library), is an open source library developed by Intel that focuses on simplifying programming related to digital images. Open CV already has many features, including: face recognition, face tracking, face detection, Kalman filtering, and various types of AI (Artificial Intelligence) methods. OpenCV has many features that can be utilized, the following are the main features of Open CV, among others:

A. Image and video I/O

With this interface we can read image data from files, or from live video feeds. And can also create image and video files.

B. Computer Vision in general and digital image processing (for low and mid level APIs)

With this interface we can do experiments with various standard computer vision algorithms. It includes line, edge, peak detection, elliptical projection, image pyramid for multi-scale image processing, template matching, and various transforms (Fourier, discrete cosine, distance transform) and more.

C. High level computer vision modules

Open CV also includes "high level" capabilities, such as additional capabilities for face detection, facial recognition, including optical flow

D. Methods for AI and machine learning

Computer vision applications often require machine learning or other AI methods, several methods It is available in the Open CV machine learning package.

E. Image sampling and transformation

Open CV has interfaces for image subregion subtraction, random sampling, rotating, and so on.

F. Methods for creating and analyzing binary images

G. Methods for calculating 3D modeling

This function is very useful for mapping and localization, whether for stereo cameras or single cameras with multiple viewing angles. (Aulia Sidharta, 2017).

Python is a programming language that can execute a number of multi-use instructions directly (interpretively) with object orientation methods. Python is the easiest programming language to understand. Python was created by a Dutch programmer named Guido Van Rossum. Pythonis a programming language that is widely used in the computer world. Python's ability to process images is very limited, for that it is necessary to import libraries from OpenCV. An example of implementing OpenCV in Python is the detection of faces, eyes, mouths, or noses from an image or video. Can also detect all parts of the face. The following is an example of the results of face and eye detection. OpenCV together with Python are used to process images or videos (stack of frames/images) according to their respective purposes, involving a camera to capture images and then processing them on a computer. (Kandir, 2016).

This face detection design is searched using the Viola-Jones Algorithm which refers to Figure 3.1 the flow chart above. The face detection process in this research is implemented using the Python programming language on the sampled

images and photos. The images and photos used are color images (RGB) in JPG/JPEG/PNG format.

The image is the result of a face detection system with OpenCV and Python. The system will search for faces to various image locations. Where initially the input image will be scanned per sub wide, starting from the top left and then repeated iteratively until the bottom right. for every sub-window that is scanned, the Haar feature is applied because there are many Haar features on each sub-window, so feature selection is carried out with AdaBoost. The number of sub-windows in an image is too many, so the Cascade Classifier selects the sub-windows. Sub windows that pass all the Classifier selection stages will be described as faces.

4. Conclusion

Based on the results of this study, it can be concluded as follows:
A. The system can be detected 100% in a frontal state.
B. The system can detect multiple faces in one photo/image.
C. The system can detect human faces.
D. The system can detect an object that resembles a face when the object has the same contour as a human face, such as a statue.

5. Reference

[1] Beham, M. P., & Roomi, S. M. M. (2013). A Review of Face Recognition Methods. International Journal of Pattern Recognition and Artificial Intelligence, Vol. 27(04), 1356005. https://doi.org/10.1142/s0218001413560053.

[2] Boontua, M., Nam-Asa,P., Arwatchananukul, S., & Aunsri, N. (2018). A Study of Features and Classifiers for Multiple Environment Face Recognition System. 3rd.

[3] International Conference on Digital Arts, Media and Technology, ICDAMT 2018, Vol. 1(1).

[4] Lestari,Sir Puji, 2014, Implementasi Pengenalan Wajah Manusia Menggunakan Principal Component Analysis (PCA), Program Studi Teknik Informatika, STMIK Budidarma, Medan.

[5] Nugroho, S., 2004, Sistem Pendeteksi Wajah Manusia pada Citra Digital, Program Studi Ilmu Komputer, FMIPA, Universitas Gadjah Mada, Jogyakarta

[6] Parmar, D. N., & Mehta, B. B. (2014). Face Recognition Methods & Applications. International Journal of Computer Applications in Technology, Vol 4(1), pp. 83–86. Retrieved from http://arxiv.org/abs/1403.0485.

[7] Rahmantha,Nafier, 2014, Penerapan Algoritma Viola-Jones dan Eigenface pada Sistem Pengenalan Wajah, Jurusan Teknik Elektro dan Teknologi Informasi, FT, UGM, Jogyakarta.

[8] Turk,M., and Pentland, A., 2013, Eigenface for Recognition, Journal of Cognitive Neuroscience,Volume : 3.

[9] Tarhini, Ali.2011. Efficient Face Detection Algorithm using Viola Jones method. http://www.codeproject.com/Articles/85113/Efficient-FaceDetection-Algorithm-using-Viola-Jon.

[10] Yanti, Dewi, 2013, Deteksi Wajah Manusia pada Citra menggunakan Dekomposisi Fourier, Program Studi Ilmu Komputer, FMIPA, Universitas Brawijaya, Malang.

[11] Viola, P., and Jones, M., 2004, Robust Real-time Object Detection, International Journal of Computer Vision, Kluwer Academic, Netherlands.

[12] Wang, D., Zheng, G., & Liu, L. (2015). Developmental Network and Its Application to Face Recognition. Proceedings of the 2015 27th Chinese Control and Decision Conference, CCDC 2015, pp. 6310–6315. https://doi.org/10.1109/CCDC.2015.7161951.