Performance Comparation of Real Time Image Processing Face Recognition for Security System

Tri Randi Uetama
Master of Mechanical Engineering
Swiss German University
Tangerang City, Indonesia
randi.gamaplantation@gmail.com

Widi Setiawan
Master of Mechanical Engineering
Swiss German University
Tangerang City, Indonesia
widi_setia@yahoo.com

Edi Sofyan
Master of Mechanical Engineering
Swiss German University
Tangerang City, Indonesia
edisofyan@hotmail.com

Abstract*—This research had been developed a system mainly consists of Arduino microcontoller based hardware and neural network based algorithms. The system has been fully assembled and successfully tested. By using two different methods the point feature detector (PFD) method was used as the first method. An Eigen Feature function was utilized to detect feature point of image. The second method is convolutional neural network (CNN) to recognize human face. Using PFD method, a classification value has been setup <11. The classification value is used as classification category of the program to recognize the subject (face image) correctly. By using PFD method, the response of the system from starting of a face image recognition until opening the locker is 20 second. The CNN method used alexnet to classify the image. At least around 300 training input data are use per person. The face recognition's experiment reached a high recognition's accuracy of 99.99% level and an average response time of 10 seconds. This research presents how the human face can be recognized and used to control the opening of a door lock.

Keywords—eigen feature, feature point detector, convolution neural network, alexnet, classification value.

I. INTRODUCTION

Visual is the most complicated of our sense [1]. Many of the researcher had studied and conduct lots of experiment in order to achieve the most sophisticated technology by utilize visual feedback. Visual feedback method allowed a device or a certain system to be able to controlled by using image processing and make a final decision of movement. A camera is a passive device to capture patterns from optical energy reflected from the scene. A single camera has a limitation in 3-dimensiooonal structures in scanning and lost in resulting in 2-dimentional image. Despite this problem, a human is particularly good to inferring in 3-dimensional nature of scene. In Fig. 1 shows how the 2-dimensional image can inferred as the 3-dimensional object

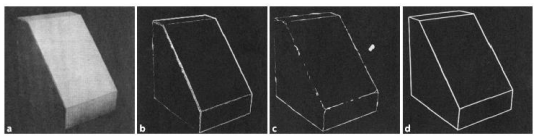

Fig. 1 (a) Original picture; (b) gradient image; (c) connected feature points; (d) reconstructed line drawing [1]

In image processing, MATLAB offers lot of tools that very effective in image signal processing. Image pattern analyses, convolution neural network (CNN), and deep learning is implementable. Additional MATLAB has a function in order to connect microcontroller and camera device.

In this research the image processing method has been implement to control a device by recognizing the human face. To be more applicable the experiment was used solenoid door locker as the actuator.

II. THEORITICAL PERSPECTIVE

A. *Image*

Image is two of dimension of discrete I(m,n) that is output or response of some of sensor at a series of fix positions (m = 1,2, …, M), n= 1,2 …,N) in 2-D cartesian coordinate. An image respectively designated the rows and columns. Image has individual elements or pixels revered by their index or coordinate. An image colour have intensity level or value to every single of pixel location I(m,n).

The size of 2-D pixel stored for each of individual image pixel determine resolution and colour of image. From mathematical view any of 2-D array can be considered as image signal.

Such as measurement data and decision. Image digitization for example image sampling *x*,*y* and gray level quantization. Image represent rectangular array integers. the image size and gray level are usually integer of power of 2. Number of each pixel represent the bright and darkness of the image.

1	8	219	51	69	171	81	41
94	108	20	121	17	214	15	74
233	93	197	83	177	215	183	78
41	84	118	62	210	71	122	38
222	73	197	248	125	226	210	5
35	36	127	5	151	2	197	165
196	180	142	52	173	151	243	164
254	62	172	75	21	196	126	224

Fig 2. Digital image and numerical representation [2]

TABLE I
COMMON IMAGE FORM ATS AND PROPERTIES

Acronym	Name	Properties
GIF	Graphics interchange format	Limited to only 256 colours (8 bit); lossless compression
JPEG	Joint Photographic Experts Group	In most common use today; lossy compression; lossless variants exist
BMP	Bit map picture	Basic image format; limited (generally) lossless compression; lossy variants exist
PNG	Portable network graphics	New lossless compression format; designed to replace GIF
TIF/TIFF	Tagged image (file) format	Highly flexible, detailed and adaptable format; compressed/uncompressed variants exist

B. *Neural Network*

A neural network is very powerful and most common used in image processing. There are many kinds of neural network method was used to recognize an image. Basically neural network components have been inspired by McCulloch and Pitts' in 1943 by their paper "A logical calculus of ideas imminent in nervous activity". The same time Frank Rossenblatt was also developing the computation of human eye.

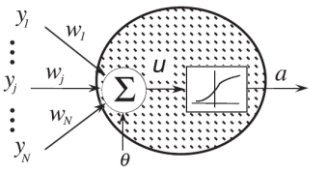

Fig 3. Neuron model of McCulloch and Pitts' [3]

$$u = \sum_{j=1}^{N} w_j y_j + \theta \qquad (1)$$

TABLE II
SUMMARY OF NET FUNCTION

Net function	Formula	Comments
Linear	$u = \sum_{j=1}^{N} w_j y_j + \theta$	Most common equation
Higher order (2nd order formula exhibited)	$u = \sum_{j=1}^{N}\sum_{K=1}^{N} w_{jk}\, y_j y_k + \theta$	u_i is weighted linear combination of higher order polynomial term of input variable. Where N_d is input value and d is the order of polynomial.
Delta ($\Sigma - \Pi$)	$u = \prod_{j=1}^{N} w_j y_j$	Seldom used

a_i is the input of the neuron is related with the input u_i by linear or non linear or its called activation function.

TABLE III
SUMMARY OF ACTIVATION FUNCTION

Activation function	Formula $a = f(u)$	Derivatives $\frac{df(u)}{du}$
Sigmoid	$f(u) = \frac{1}{1+e^{-u/T}}$	$f(u) = f(u)[1-fu]/T$
Hyperbolic tangent	$f(u) = \tanh\left(\frac{u}{T}\right)$	$(1-[f(u)]^2)/T$
Inverse tangent	$f(u) = \frac{2}{\pi} tan^{-1}\left(\frac{u}{T}\right)$	$\frac{2}{\pi T} - \frac{1}{1+(u/T)^2}$
Threshold	$f(u) = \begin{cases} 1 & u > 0; \\ -1 & u < 0. \end{cases}$	$-2(u-m).f(u)/\sigma^2$
Gaussian radial basis	$f(u) = exp[-\|u-m\|^2/\sigma^2]$	$-2(u-m).f(u)/\sigma^2$
Linear	$f(u) = au + b$	a

C. *Convolution Neural Network*

Basically image recognition is classification. Recognizing the image whether recognizing the image of animal or human is the same as classifying. Before CNN the feature extractor has been designed. Therefore it required significant amount of cost. CNN yields the better image classifying. When feature point extractor is deeper. CNN consist a neural network to extract the image and another CNN can classifies feature of the image.

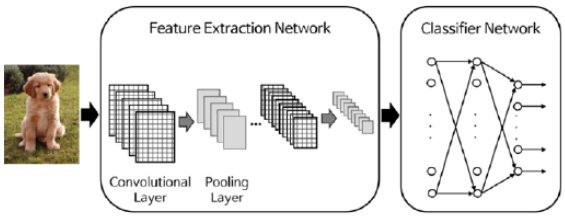

Fig 4. Typical architecture of CNN [4]

The input image enters into extraction network. After extracted the feature signal enter the classification neural network. Then the classification of CNN operates base on the feature of the image and generate the output.

III. PREVIOUS STUDIES

Image recognition research have been investigated in many of studies. A study about face recognition base on color vector binary pattern from multichannel face images has been investigated [5]. The study proposes the novel face descriptor based on color in formation. The study perform face recognition by *local color vector binary pattern* (LCVBPs) feature. The result show the LCBVTs is able to

yield high performance in face recognition. Wang with study "*Representing image base on point image*" has been conducted [6]. The experiment utilize "verge points" to define the curvature image surface. Yeom [7] developed real time 3-d sensing the dynamic biological of microorganism. The study conducted to reconstruct the object by single-exposure on-line (SEOL) digital holography. The study is used pattern recognition by 3-d morphology and 3-d recognition. The result with sphacelaria alga, tribonema aequale alga, and polysiphonia alga are presented. Most research and study in image processing has been performed to recognize specific object for example bacteria, virus, human face, etc. A study to recognize tuberculosis bacteria [8] and vibrio cholera base on colors. Hui [9] with study "*Research on face recognition algorithm based on improved convolution neural network*". had been conducted. the experiment combine *Fisher* criterion to cover the poor property of CNN. The result that combining the Fisher neural network (FNN) and convolution neural network can achieve good performance. In this study utilize *Eigenfeature* point detector and *alexnet* to recognize human face is investigated.

IV. RESEARCH METHODOLOGY

A. Material and Equipment

In this research was used a webcam of computer as the sensor of the image, the input data will calculate by using Eigen Features function to detect point feature to recognize the image pattern. At the end calculating geometry changing of point or utilize alexnet tool in order to learning and classify the image. After that the signal from computer should sent to Arduino port to controlling the relay.

Fig 5. System flow process

The hardware circuit design is shows in Fig. 5 which is consisted kind of electronic devices. The solenoid door locker is required 12 Volt of input voltage, so voltage transformer is needed in order to obtain the voltage nominal value. Rectifier is attached to covert AC to DC and to get more stable voltage, a capacitor with 50 volt and 450 µF was attached.

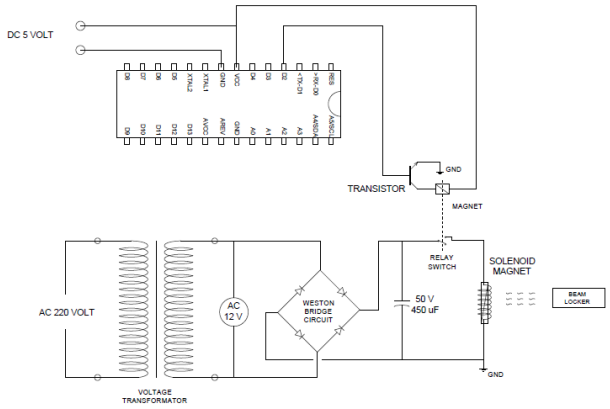

Fig 6. Circuit diagram

B. Software Design

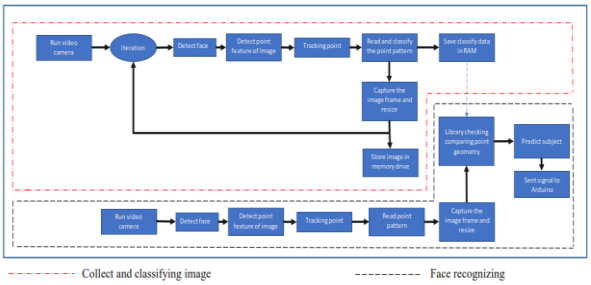

Fig 7. Face recognition base on PFD method block diagram

In the Fig. 7 depict the system work block diagram. The first software marked in dotted red line is to classifying and do looping iteration task to collect the images data. The images data is store into memory drive and classifying data is store in RAM (random access memory) of the computer. The second program is in dotted black line. This is part of recognizing the image from camera in real time. Identifying point feature is used to compare the point feature in the new image and old images. By calling the data classify from data base, the point feature analyzing the geometry point of both of the images in order to predict the subject [10].

V. RESULT AND DISCUSSIONS

A. Performance Test Face Recognition Of PFD Method

TABLE IV
OUTPUT CHANGING VS IMAGE RESOLUTION INPUT VALUE

Image resolution (pixel)	Time (second)	Value of variable classification Level of the image	Output result
28 x 23	0	-	Error
56 x 46	0	-	Error
112 x 92	20	10	Recognized
224 x 184	0	-	Error

TABLE V
DATA RECOGNIZING FACE WITH MORE THAN ONE PERSON

Subject	Time (second)	Value of variable classification Level of the image	Output result
Randi	20	10	Recognized
Lia	21	17	Recognized
Tami	20	27	Recognized

TABLE VI
SUBJECT RECOGNITION EXPERIMENT RESULT BY USING PHOTO OR A PICTURE

Subject	Picture attach on front camera	Time (second)	Value of variable classification Level of the image	Output result
Randi	Randi	20.66	7	Recognized
Randi	Wawan	20.5	-	Un identified person
Randi	Tami	19.89	-	Un identified person

TABLE VII
TIME RESPONSE VS OUTPUT RESULT WHEN ITERATION NUMBER OF LOOP IS CHANGING

Subject	Loop	Time (second)	Value of variable classification Level of the image	Output result
Randi	1	8.8	1	Recognized
Randi	5	9.07	2	Recognized
Randi	10	9.2	1	Recognized
Randi	15	9.44	2	Recognized
Randi	20	9.82	1	Recognized

B. Performance Test Of Convolution Neural Network Method

TABLE VIII
OUTPUT CHANGING VS IMAGE RESOLUTION INPUT VALUE

Image resolution (pixel)	Layers	Learning time until reach 100%	Recognition time (second)	Output result
100 x 100	25	5:38.57	10.05	Recognized
150 x 150	25	5:8.62	10.40	Recognized
200 x 200	25	6:10.72	12.76	Recognized
250 x 250	25	6:21.32	13.86	Recognized
300 x 300	25	6:34.94	15.55	Recognized
350 x 350	25	6:40.56	16.24	Recognized

TABLE IX
DATA RECOGNIZING FACE WITH MORE THAN ONE PERSON

Subject	Layer	Time (second)	Output result
Randi	net.Layer	10	Recognized
Rudy	net.Layer	10.55	Recognized
Usep	net.Layer	9.88	Recognized

TABLE X
SUBJECT RECOGNITION EXPERIMENT RESULT BY USING PHOTO OR A PICTURE

Subject	Picture attach on front camera	Time (second)	Output result
Randi	Randi	10.55	Un identified person
Randi	Rully	11.2	Un identified person
Randi	Usep	10.88	Un identified person

TABLE XI
RESPONSE RESULT WHEN SYSTEM WHEN THE FACE OF HUMAN SET UP WITH DIFFERENT ANGLE

Subject	Angle	Time (second)	Output result
Randi	0°	9.10	Recognized
Randi	30°	9.44	Recognized
Randi	60°	-	Recognized
Randi	90°	-	Recognized

VI. CONCLUSION

The experiment studied here is to analyze the best method to build a system which is robust and easy to create in order to serve another device where security system is required. In this research the face recognition by using feature-point object detection method have been investigated. In training part the *Eigenfeature* function to detect feature point was used. Then by analyzing the point movement per frame to tracking the face. In recognition part the *Eigenfeature* used to detect the new feature point. Estimate the geometric transformation of old points and the new point give the ability to recognize a human face. Using PFD method, a classification value has been setup <11. The classification value is used as classification category of the program to recognize the subject (face image) correctly. By this method sometime it was found error result because the point of the image is not detected completely. The phenomenon led the calculation of geometric point is wrong and make the value of variable classification Level which is contained the calculation give error result. This problem appeared when capturing face in training part is not in centre of the screen frame and led the system making

error in calculation. Geometry point transformation is can be used to recognize facial movement in order to increasing performance of the system. By adding the geometry transformation method the system can differentiate between living human or a picture and run the Arduino to switch the door locker.

The second method to recognize the face is using convolution neural network (CNN) has been investigated. By using *alexnet* tool an image can be predicted. In training part, the image must be contained a human face object to activate the face detector. The image captured and store in the memory or drive directly. Before run the training part the image resized in 227x227 pixel in order to make fit with *alexnet'*s layer. The result of training saved in memory in variable *newnet* where the recognition part should call this variable when running. In recognition part the image capture from camera in real time. The image directly resize in 227x227 pixel to make it fit toward images that had been learned before. After that to run the prediction function to classify the image and compare the result toward the *newnet* data. The experiment result reach high accuracy at 99.99% level in training part and time response average at 10 seconds. This is how the system can recognize the human face. By this method is very robust and accurate because lots of data input can be set make the system can learn more. In the hardware part it found not much error when uploading and running the software but in this research is only used one pieces of Capacitor 50 Volt and 450 µF in order to get DC voltage more stable. The instability of voltage led the solenoid of door locker shaking. To get it more stable, attach another smaller Capacitor 16 volt and 50 Volt and 450 µF will be helpful. Adding another capacitor led the input DC voltage being purer.

ACKNOWLEDGMENT

This study is dedicated with wholeheartedly to my parents and all of my family who have been gave me motivation to achieve more acknowledge and gave me an inspiration and strength. I also thank to Dr. Ir. Widi Setiawan as my advisor who always support and give an advice regarding this thesis and also thank to Edi Sofyan, B.Eng., M.Eng., Ph.D, as my co-advisor who always give me an inspiration to solve many of my problem in this research.

REFERENCES

[1] K. S. Fu, *ROBOTICS: Control, Sensing, Vision, and Intelligence*, New York: McGraw-Hill, 1987.

[2] B. K. P. Hom, *Robot Vision*, MIT-Press, 1987.

[3] M. Ohka, H. Bin, and S. Che, "Object-handling tasks based on active tactile and slippage sensations," *Robot Arms*, 2011, doi: 10.5772/16781.

[4] P. Kim, *MATLAB Deep Learning With Machine Learning, Neural Networks and Artificial Intelligence*, Seoul: Apress, 2017.

[5] A. Kurniawan, *Arduino Programming using MATLAB*. PE Press, 2015.

[6] S. J. Wang, "Representing images using points on image surfaces," *IEEE Transactions on Image Processing*, 14 (8), pp. 1043 - 1056, 2005.

[7] S. Yeom, "Real-time 3-D sensing, visualization and recognition of dynamic biological icroorganisms," Proceedings of the IEEE, 94 (3), pp.550 - 566, April 2006.

[8] S. H. Lee, "Local color vector binary patterns from multichannel face images for face recognition', *IEEE Transactions on Image Processing*, 21 (4), pp. 2347 - 2353, 2012

[9] L. Hui and S. Yu-jie, "Research on face recognition algorithm based on improved convolution neural network," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, 2018, pp. 2802-2805, doi: 10.1109/ICIEA.2018.8398186.

[10] B. Wibowo, E. Sofyan, G. Baskoro, "Prototype design of speed detection mobile application for golfers swing movement using computer vision compared to portable radar and accelerometer systems," 1st Proceedings of The Conference on Management and Engineering in Industry (CMEI 2019)," vol. 1, pp. 49-52, Tangerang, Indonesia, August 2019.