Capital Letter Pattern Recognition in Text to Speech by Way of Perceptron Algorithm

Novan Wijaya [a, 1, *]

[a] Management Informatics Study Prog., AMIK Multi Data Palembang, Jl. Rajawali No.14, Palembang 30113, Indonesia
[1] novan.wijaya@mdp.ac.id*
* corresponding author

ARTICLE INFO

ABSTRACT

Computer vision is a data transformation retrieved or generated from webcam into another form in means of determining decision. All kinds of transformations are carried through to attain specific aims. One of the supporting techniques in implementing computer vision on a system is digital image processing as the objective of digital image processing is to transform digital-formatted picture so that it can be processed in computer. Computer vision and digital image processing can be implemented in a system of capital letter introduction and real-time handwriting reading on a whiteboard supported by artificial neural network mode "perceptron algorithm" used as a learning technique for the system to learn and recognize the letters. The way it works is captured in letter pattern using a webcam and generates a continuous image that is transformed into digital image form and processed using several techniques such as grayscale image, thresholding, and cropping image.

I. Introduction

Computer vision is a data transformation retrieved and obtained from webcam into another form in means of determining a future decision. All kinds of transformation forms are carried through to achieve a few of particular objectives [1]. There are several operations in computer vision starting from object image capturing by camera, object image processing into a more efficient and simple form without omitting representative information of that object, and eventually system analysis to determine the action that will be taken.[2]. One of the applications that is possible to be developed from computer vision is capital letter pattern introduction. The fundamental concept is that of the image is captured by webcam, then the captured image is processed into digital image, furthermore, the analysis of the captured image is deemed to decide whether the image belongs to which letter. Even more, for object image processing on computer vision can employ digital image processing concept [3].

These digital image processing techniques can be used in the recognition of capital letters by way of a webcam in terms of image quality improvements captured by a webcam. Once the computer has obtained a good digital image or the information required by the computer has been obtained, it takes a pattern recognition technique in order for the computer to make a decision to recognize the pattern of the captured image letters. The method that can be used for the pattern recognition process is the method of Artificial Neural Network (ANN).

Artificial neural networks have the ability to learn to solve problems that are rather complicated. This is because the existing knowledge in ANN is not programmed, but through the undergone process of training information. Artificial neural networks are trained using the perceptron algorithm [4]. Perceptron algorithm is an artificial neural network used to classify a pattern of entry into a class or not. Perceptron is also able to be used to classify a pattern belongs to which class, by comparing patterns into each class. Perceptron is a single layer learning algorithm through several process procedures by repeating until it gets the right neural weights [5].

The pattern recognition system and its reading (text to speech) keep on evolving to date, using other methods both from the pattern recognition side, and from the decision-making side. Text to speech is a system capable of converting from a text into a speech or sound. This pattern recognition utilizes several digital image processing techniques in order to obtain accurate image information in accordance with system requirements without compromising the contained important information in the image [6]. From the pattern recognition side using the perceptron algorithm, as well as the maximum approach as a decision-making mode. This paper discusses the use of perceptron algorithm to recognize capital letter pattern recognition in text to speech system.

II. Methods

The introduction of real time letters using the perceptron algorithm is the development of a pattern recognition system that has been created by previous researchers [7]. The pattern recognition system will try to recognize handwriting and will be added to the reading feature of the letter patterns read by the system. Pattern recognition using the webcam as a sensor that is useful for capturing images of letters that will be written on the white board media. The image captured by the webcam is processed by the laptop, and finally the laptop will output the output of the sound pattern of the letters that are read by the system.

The workability principle of the whole system is, first the webcam captures the image of an existing letter on the white board. The image is processed by the means of retrieving the information required by the pattern recognition system letters. The required information is the pixel dimension for the image as well as the binary value contained in each pixel. In order to get the binary value on each pixel then used the technique of digital image processing that is, grayscale image, binary image (threshold), median filter and cropping image [8].

Upon obtaining the desired image information that is the binary value 0 or 1 on each pixel, then those values will be the inputs of the perceptron algorithm as the pattern recognition method. The concept of the perceptron algorithm is to work out the existing patterns in assent with certain rules, up to finally the system generates a special trait for each worked out pattern, in the presence of these special features the system can distinguish existing letter patterns [4].

Image segmentation is a technique for separating the required objects and backgrounds so that the objects in the image are easily analyzed in means of pattern recognition. One of the simplest segmentation techniques is image thresholding [9]. Image mining will separate the image into two areas, i.e. the object area and the background region, the object area can be set in white while the background area is set black, or vice versa [10]. For self image mining using otsu method. The approach adopted by the otsu method is to conduct a discriminant analysis of determining a variable that can distinguish between two or more groups that occur naturally [10]. Discriminant analysis will maximize the variable in order to divide the foreground object (foreground) and background (background). The result of this image mining process is a binary image that has only two grayish degrees i.e. black or white.

III. Results and Discussion

Software design takes substantial role on the establishment of capital letter pattern recognition. Fig. 1 shows the general process of the developed system, while Fig. 2 shows system establishment diagram.

Fig. 3 shows the example of a capital letter that is used as the system input. The letter will followed five process of image segmentation as depicted in Fig. 4. The first stage, digital imaging stage aims at capturing letter picture on the white board by webcam, which eventually is captured
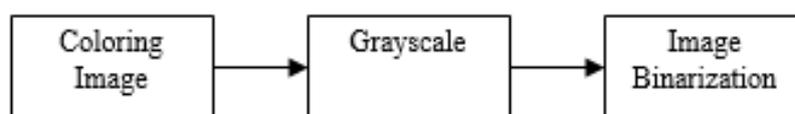
Fig. 1. Thresholding technique image segmentation process
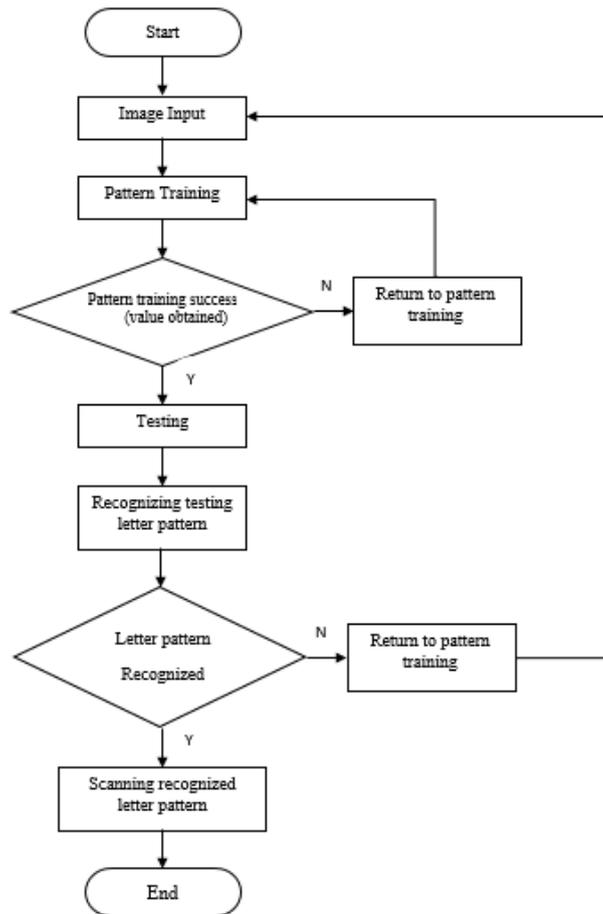
Fig. 2. System establishment flow diagram

and rendered into real time image form. The next step is Image Segmentation.The captured real time image will be processed to obtain necessary information by way of several modes in this segmentation process, i.e. grayscale, thresholding, cropping image, and resize image.

Grayscale image possesses two gradation colors of white up to black in each of its pixel. The result of this stage is depicted in Fig. 5. Fig. 6 shows the result of the following process that is called thresholding. Binary Image is a digital image whose each pixel only shares two values of 0 (black)

Fig. 3. A Letter sample RGB real time image
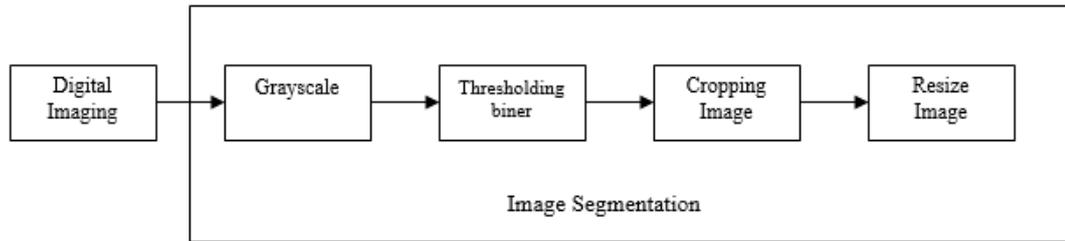
Fig. 4. The process of image segmentation
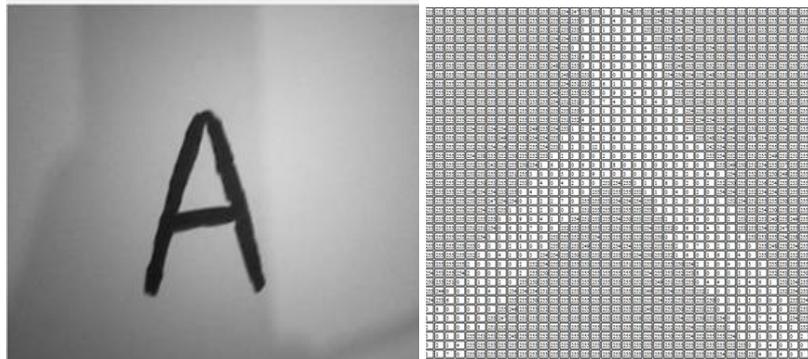
Fig. 5. Grayscale image from real time image ad grayscale image value
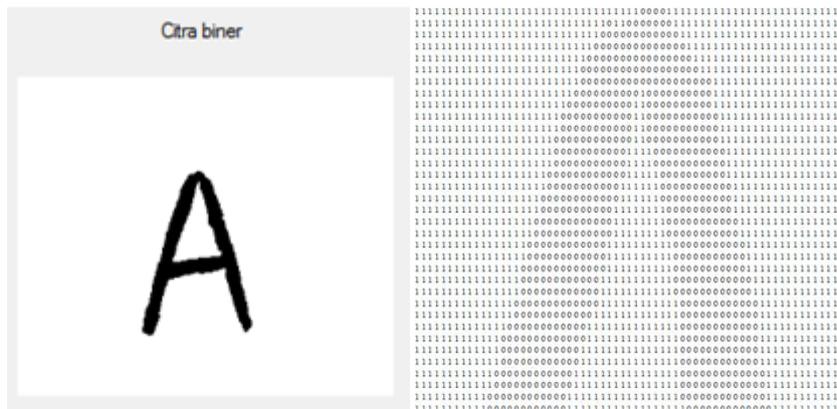
Fig. 6. Threshold image letter A and threshold image value letter A

and 1 (white) or the other way around. The process of digital image processing, binary image can only be obtained out of the conversion result from grayscale to binary image using threshold technique. Afterwards, the image that have undergone several stages would then be cropped to obtain precise information and ease the succeeding stages. Initially, the used image resolution is 640 x 480 pixel. Upon cropping the image, the resolution turns into 20x20 pixel. That amount is singled out to ease the training input on perceptron algorithm. Fig.7 shows the result of cropping image process.

Upon image cropping process, image size would be rendered into 20 x 20 pixel size (Fig.8). Resize is a technique both to reduce size and resolution of an image without removing specific information of that image. The retrieved information of the resized image would then be processed using perceptron algorithm as the employed artificial neural network mode for pattern recognition and learning processes.

Fig. 7. Image upon cropping image process

Fig. 8. Resized image result 20 x 20 pixel in binary form

Fig. 9. Sample data and weighting value data prior to training process

Artificial Neural Network is a mode designed that way to emulate human' way of thinking in learning things or new information. Perceptron Algorithm is used to practice sample data. Sample data of this study is a collection of caputured capital letter image by webcam which then be processed using PC and programs as well as image segmentation techniques. The final result of image segmentation is ninary image, it is an image containing binary value (1 or 0). In view of pixel number on the image in concert with thw forementioned explications i.e. 20 x 20 pixels, thus, the input number of perceptron algorithm would be 400 inputs for each character in regard with the number of binary image pixel.

The first stage to carry out ahead of performing learning stage (training) by way of algorithm perceptron is sample data collection. The collected data are compartmented into a folder as storage and classification media., those data would undertake traing process using algorithm perceptron. Upon the completion of training process, weighting value would be generated for each capital letter and stored in a database and would later be utilized for trial process. Fig. 9 shows the data and weighting value data prior to training process.

Upon completing training process, the trial stage would then be performed. At the trial stage, system workability is reviewed in regard with how far it can distinguish letter patterns that have been worked out and validated.

The letters directly tested using webcam. Over pressing a button on the program display, the system performs several processes starting from capturing the letters on the white board, processing the image into a digital image and processed by way of image segmentation techniques, as well as perceptron algorithm as artificial neural network method. In the end the system tries to recognize existing letter patterns, as well as to read recognizable letters.

After going through a series of processes with the aim of recognizing the patterns of letters trained, the last stage is the testing of letters that have been trained using perceptron algorithm. This testing stage is designed to recognize more than one letter, word, even sentence, but basically the introduction of the tested letter pattern will still be done in the letter.

At the image segmentation process for the testing phase this time, slightly different from the image segmentation in the training process. The difference is when separating one letter with the other letters that exist in an image as when a picture is captured through the camera then the image will be rendered into a digital image by the computer. Therefore, digital images contained several letters in it must be separated first into the letters, so that the test system can recognize the pattern of letters are tested properly.

After each letter is separated then the letters are the input for the test system. Similar to the training process, this testing phase will process the input images using image segmentation techniques. After getting the information input required by the system then, the input will be processed and produce the output as expected.

Decision making mode for this validation system is decision making mode in uncertainty by way of maximax criterion [11]. Uncertainty decision making exhibits decision condition in which the probabilities of potential results are unidentified, in an uncertain circumstance, the decision maker is aware with alternative results in various events, yet the decision maker cannot determine the event' probability nonetheless. Maximax criterion would search for the best (maximum) alternative for every existing option, then make a decision in regards with the maximum value of that outcome, maximum criterion is also entitled as optimistic criterion decision or an alternate with the highest beneficences.

Upon merging the input data as well as the weighting value with maximax criterion, speech (reading) in a form of output would later by generated by computer on the identified letter patterns. At the undertake test (Fig. 10), it results on an information that an image of HELLO WORLD tested letters and the system can distinguish that letter. Whereas, upon multiple tests, as in Fig. 11, system cannot discern one letter. Letter "H" is identified as letter "U", while letter "R" is identified as "K" so that it turns to be "UELLO WOKLD".
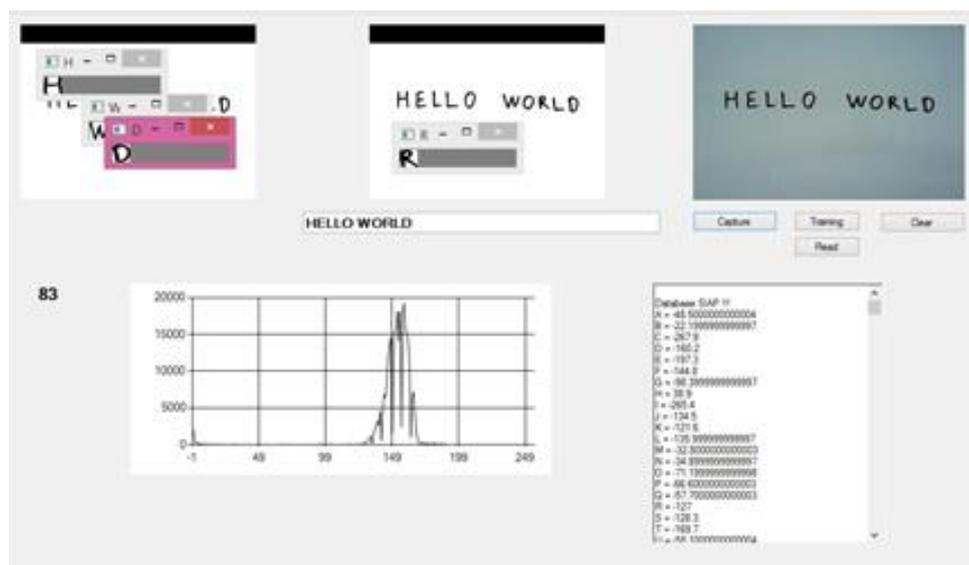
Fig. 10. The interface of the letters that would be tested.
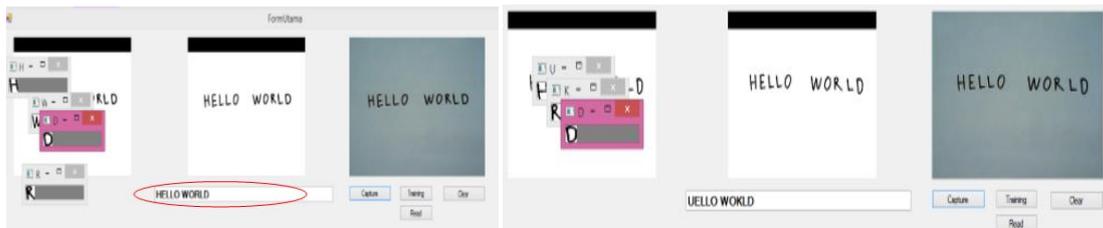
Fig. 11. Text trial " HELLO WORLD"

IV. Conclusion

By way of several image segmentation techniques like grayscale, binarization, as well as image cropping, and supported with perceptron algorithm as capital letter learning mode. Capital letter identification system is designed to be able to run well. From multiple tests, there is one test failing to distinguish letter.

References

[1] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*. USA: O'Reilly Media Inc, 2008.

[2] A. Hamzahan, G. Santosa, and W. Widiarto, "Klasifikasi Objek Dalam Visi Komputer Dengan Analisis Diskriminan," *Makara Teknol.*, vol. 6, no. 1, pp. 24–32, 2002.

[3] G. . Papakostas, E. . Karakasis, and D. . Koulouriotis, "Accurate and speedy computation of image Legendre moments for computer vision applications," *Image Vis. Comput.*, vol. 28, no. 3, pp. 414–423, 2010.

[4] M. R. . Dawson, D. M. Kelly, M. L. Spetch, and B. Dupuis, "Using perceptrons to explore the reorientation task," *Cognition*, vol. 114, no. 2, pp. 207–226, 2010.

[5] I. L. May, "Pengenalan Vokal Bahasa Indonesia Dengan Jaringan Syaraf Tiruan Melalui Transformasi Wavelet Diskret," *Universitas Diponegoro*. 2002.

[6] D. Putra, *Pengolahan Citra Digital*. Yogyakarta: Andi Offset, 2010.

[7] Y.-C. Hu, "Pattern classification by multi-layer perceptron using fuzzy integral-based activation function," *Appl. Soft Comput.*, vol. 10, no. 3, pp. 813–819, 2010.

[8] E. Nugroho, Susilo, and Akhlis, "Pengembangan Program Pengolahan Citra Untuk Radiografi Digital," *J. MIPA*, vol. 1, pp. 46–56, 2012.

[9] Y. Li, D. M. . Tax, and M. Loog, "Scale selection for supervised image segmentation," *Image Vis. Comput.*, vol. 30, no. 12, pp. 991–1003, 2012.

[10] T.-H. Min and R.-H. Park, "Eyelid and eyelash detection method in the normalized iris image using the parabolic Hough model and Otsu's thresholding method," *Pattern Recognit. Lett.*, vol. 30, no. 12, pp. 1138–1143, 2009.

[11] E. D. Handoyo and L. W. Susanto, "Penerapan Jaringan Syaraf Tiruan Metode Propagasi Balik Dalam Pengenalan Tulisan Tangan Huruf Jepang Jenis Hiragana dan Katakana," *J. Inform.*, vol. 7, no. 1, pp. 39–55, 2011.