



# **Development of Text Recognition Prototype with Classification of Neural Networks AND Text-To-Speech in Javanese Scripts Using Incremental Methods**

**<sup>1</sup>Ifan Prihandi; <sup>2</sup>Indra Ranggadara; <sup>3</sup>Boy Yuliandi**

<sup>1,2,3</sup> Faculty of Computer Science, Universitas Mercu Buana

Jl. Raya Meruya Selatan, Kembangan, Jakarta, 11650

<sup>1</sup> [ifan.prihandi@mercubuana.ac.id](mailto:ifan.prihandi@mercubuana.ac.id); <sup>2</sup> [indra.ranggadara@mercubuana.ac.id](mailto:indra.ranggadara@mercubuana.ac.id); <sup>3</sup> [boy.yuliandi@mercubuana.ac.id](mailto:boy.yuliandi@mercubuana.ac.id)

---

*Abstract— Artificial neural networks are information processing systems that have characteristics similar to human neural networks. Learning models need to be done on an artificial neural network before being used to solve problems by examining and correcting any errors that occur during the learning process. Developments in information technology affect particular scientific expertise. Field of image processing does not become one - the only method of solving a problem, but the current image processing combined with artificial intelligence to examine or look for a solution in a variety of applications. Akshara Java is one of the priceless cultural heritages. Form of script and art-making becomes a relic that deserves to be preserved. Not only in Java, but Akshara Java is also used in the Sunda and Bali, although there is little difference in the writing used the same script. The purpose of this research is to create a model of image processing and converted into text into voice so that people can learn Akshara Java and able to preserve the culture of Indonesia and is expected to be a reference for the development of the mobile application development at a later stage.*

*Keywords— Akshara Java, Image Processing, Incremental, Neural Network, Text Recognition*

---

## **I. INTRODUCTION**

The development of information technology has an impact on the diffusion of science or particular skills. The field of image processing is not the only method in solving a problem, but now, image processing is combined with artificial intelligence to research or find a solution in various applications. Today's computer-based learning media has developed very rapidly, many applications that are created and designed to support the learning process become interactive, for example, are applications of digital image processing, speech synthetics or commonly called text-to-speech [1], and apps that use artificial neural network methods. This digital image processing includes character recognition techniques such as alphanumeric characters, handwriting characters, kanji characters, and others. This character recognition technique is generally known as OCR (Optical Character Recognition) technology. While text-to-speech is a process of converting a written text into speech. Artificial neural networks are information processing systems that have characteristics similar to human neural networks.

The learning model needs to be done on an artificial neural network before it is used to solve problems by checking and correcting any errors that occur during the learning process. At a certain level, artificial neural networks can provide the correct response even though the input given is noise or changes by a situation. The advantage of artificial neural networks is the ability to recognize by learning from the pattern of images taught. In the many providers of image pattern recognition, and there is still a lack of pattern recognition in the Javanese script. Where Javanese script is one of the priceless cultural relics. The form of the text and the art of making it become a legacy that deserves to be preserved. The development of information technology has an impact on the diffusion of science or particular skills. The field of image processing is not the only method in solving a problem, but now image processing is combined with artificial intelligence to research or find a solution in various applications.

## II. STUDY LITERATURE

### A. Image Processing

Image processing [2] is a process of image processing and analysis that involves a lot of visual perception. This process has the characteristics of input data and output information in the form of images. The term digital image processing is generally defined as two-dimensional image processing with a computer. In a broader definition, digital image processing also includes all two-dimensional data. Although an image is rich in information, often the vision we have has a decrease in quality intensity, for example containing defects or noise, the color is too contrasting or blurred, of course, such images will be painful to represent so that the information is reduced. For images that are easily disturbed to be served, the image needs to be manipulated into other models of better quality.

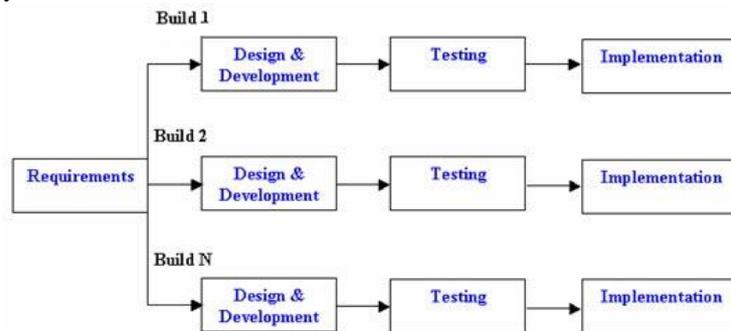


Fig. 1. Incremental Model Cycle

### B. Optical Character Recognition

Image is something that describes an object and usually in two-dimensional form. Model is a representation of the resemblance of an object or object. Digital images are defined as discrete representations of spatial data (layout) and intensity (color) information [3]. OCR can be seen as part of a broader automatic identifier, automatic pattern recognition. In automatic pattern recognition, the pattern recognition system tries to identify whether the received input image matches one of the predetermined images. This system, for example, is used to detect fingerprints, signatures, and even a person's face. Many approaches can be used to develop the making of automatic pattern approaches, including using numerical, statistical, syntactic, neural and production rules (rule-based).

### C. Backpropagation Algorithm

Definition of Backpropagation is a systematic method of artificial neural networks using supervised learning algorithms and is usually used by perceptron with multiple layers of screen to change the weights in the hidden layer. Backpropagation is a controlled type of training which uses a weight adjustment pattern to achieve the minimum error value between the predicted output and the actual output. In this study, the stages used in completing the application are used by adopting an incremental method.

### D. Artificial Neural Network

Based on the architecture [connection pattern], artificial neural networks are divided into two categories, namely Feed-Forward Structure and Feed-Back Structures. In this Final Project the structure used is feed-forward, in this type of system the signal moves from the input then passes through the hidden layer and finally reaches the output unit, has a stable behavior structure.

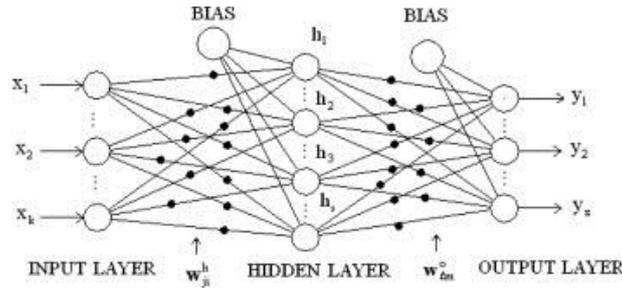


Fig. 2. Feed-Forward Neural Network.

*E. Text-To-Speech*

The transformation from text to sound (speech). This transformation converts speech synthesis, which as much as possible is made to resemble real sounds, adapted to the rules of language pronunciation. TTS (text to speech) is intended to read electronic texts in the form of books, and also to voice texts using integration sound. This system can be used as a communication system. AT & T Bell Laboratories (Lucent Technologies) also has a very long tradition of voice guides (speech synthesis). Coker the first complete was demonstrated in Boston in 1972 and released in 1973. This was based on the articulatory model developed by Cecil Coker [4]. The development of the process of the system of combining synthesis was initiated by Joseph Olive in the mid-1970s [5]. This system is now available for English, French, Spanish, Italian, German, Russian, Romanian, Chinese and Japanese [6]

**III.METHOD**

*A. Sample Selection Method*

The sample selection method used is a method based on the source, namely primary / original / new data that is collected directly with surveys in the field using unique data collection methods and secondary data that have been received by data collection agencies and published to the data user community, and can obtained from the literature or researchers' reports that are used as references by the author. Whereas for taking the sample using Random Sampling procedure.

*B. Sample Selection Method*

The flow in the prototype and algorithm can be described as follows:

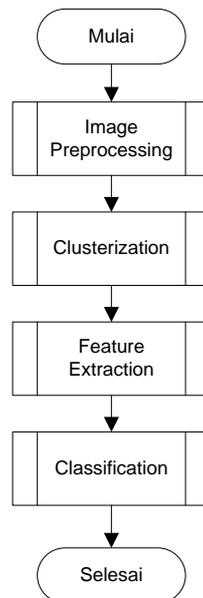


Fig. 3. System Flow Framework

Postprocessing. From the picture above, the system process consists of two stages, namely the preprocessing stage, namely image processing, and postprocessing stages

- 1) Image Preprocessing: this module is a process for removing parts that are not needed in the input image to do the next process.
- 2) Clusterization: This module works to find individual characters in images and calculate the correct sequence of characters in the text.
- 3) Feature Extraction: This module aims to extract unique features of individual characters so that they can be recognized by the classification module.
- 4) Classification: This is the last module on the OCR system that tries to recognize characters using information about character features that have been extracted in the previous stage.

#### IV. RESULT AND DISCUSSION

##### A. Prototype GUI System Testing and Analysis

The process of testing and analysis is carried out to identify whether the system is developed in accordance with the analysis of the system that has been made. This is intended to identify whether proof of concept in the form of a system prototype is appropriate. The first testing process used the Javanese script vowel letters, Swara Script. In the test results it can produce the following:

##### B. Prototype GUI System Testing and Analysis

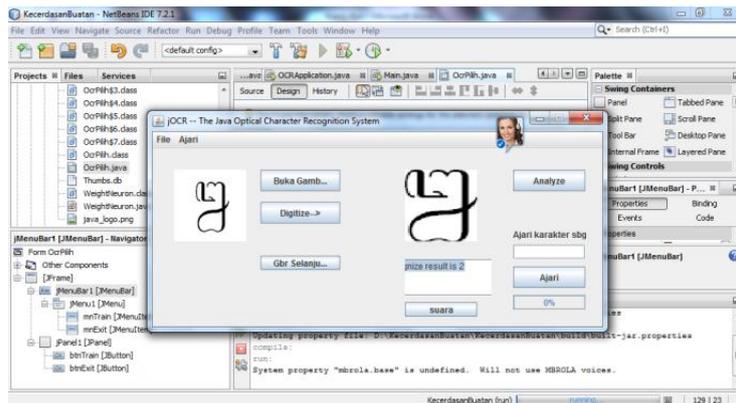


Fig. 4. Test Result on Netbeans

- 1) Test results on Swara script
- 2) Character: Numbers
- 3) Success: 1 character
- 4) Read: 1 character

Tests are carried out only in the process of training data and testing the data used when testing images and system testing is carried out on 16 different images. After the testing process is done, the text is read into the sound that produces:

- 1) Test results of the application of the conversion of letters of the Swara alphabet.
- 2) Total Letters: 5 characters.
- 3) Success Readings: 5 characters.
- 4) Readings Fail: 0 characters

C. Process flow OCR Application

The following is a Picture of Flowchart of Form OCR Application Process Flow, which is explained below:

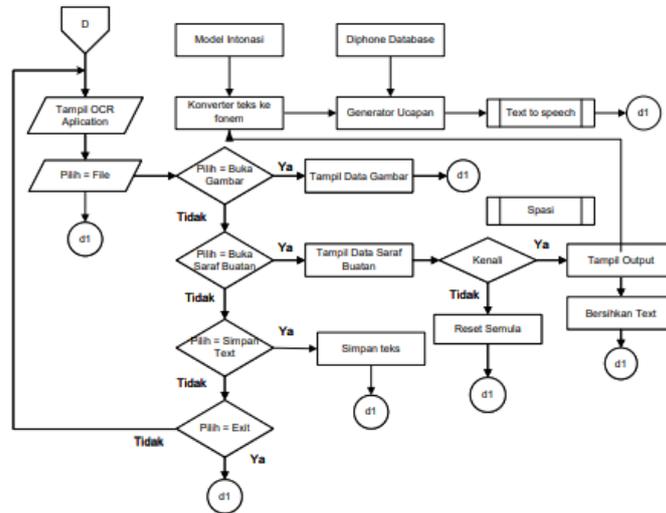


Fig. 5. Flowchart OCR Applications

Form OCR Application is a form used to process images into text (text) and voice (voice) using artificial nerves. below is the display and results of the OCR Application form trial:

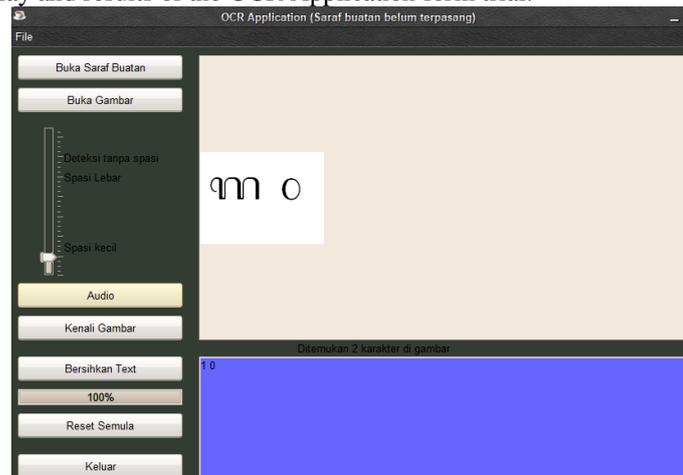


Fig. 6. OCR Testing

testing the data used when testing images and system testing was carried out on 16 different images. After the testing process is done, the text is read into the sound that produces:

- 1) Test results on Swara script
- 2) Character: Numbers
- 3) Success: 2 characters
- 4) Read: 2 characters

D. Interpretation

Based on this research, the training of image data is done repeatedly so that the recognition process can be more accurate. The test results prove that from 5 vowels there are 5 characters that can be read and for conversion into sound, the five characters can be pronounced. This is in accordance with the prototype model proposed even though there are obstacles to reading the letters of Javanese script can only recognize 1 character.

V. CONCLUSION

Based on the discussion carried out in the previous chapter, a number of things can be concluded as follows: (1) Prototype text recognition prototype with neural network classification and text to speech on javascript letters can be applied, this is evidenced by the results of the data obtained are in accordance with the research

objectives. (2) The combination of models used between image processing techniques and artificial intelligence in this case text recognition and text to speech can provide the expected data results (3) Can be developed more than 1 character.

## REFERENCES

- [1] Somov, A., Wildfire safety with wireless sensor networks. EAI Endorsed Transactions on Ambient Systems. pp. 1-11 (2011).
- [2] Motaz, A., Start programming using Object Pascal. Vol. 2, pp. 10-11. Legally Free Computer Books, US (2013).
- [3] G. Acciani, G. Brunetti, et.al, Multiple Neuro Network System to Classify Solder Joints on Integrated Circuits, International Journal of Computational Intelligence Research. ISSN 0973-1873 Vol.2, No.4 [2006], pp. 337-348.
- [4] Fernando de Aguiar Faria, et.al, Machine Vision And Artificial Neuro Networks For Seam Tracking And Weld Inspection, ABCM Symposium Series in Mechatronics - Vol. 4 - pp.768-775.
- [5] Ms. Anuja Bujurge, et.al, ANFIS Based Color Image Segmentation for Extraction of Salient Features: A Design Approach, Int. J. on Recent Trends in Engineering & Technology, Vol. 05, No. 01, Mar 2011.
- [6] A.M. Arymurthy, Suryana, S, Pengantar Pengolahan Citra. Jakarta: PT. Elex Media Komputindo, 1992.
- [7] Deden. M.F. Shiddiq, Yul Y.Nazaruddin, Farida I. Muchtadi, Estimation of Rice Milling Degree using Image Processing, IEEE Journal, 4577-1460, 2011.
- [8] Safitri Juanita, Disja, et.al, Pengembangan Aplikasi Text Recognition Dengan Klasifikasi Neural Network Pada Huruf Hijaiyah Gundul. ISSN : 1693 -9166 Vol.10.
- [9] Arief Maulana Sugiantoro, Aries Kusdaryono, Pengembangan Character Recognition Dengan Klasifikasi Neural Network Dan Text-To-Speech Pada Huruf Vokal Korea, in Prosiding Seminar Nasional Multidisiplin Ilmu Universitas Budi Luhur, ICT 215-224, July 30,2016.