PAPER • OPEN ACCESS

Citizen Id Card Detection using Image Processing and Optical Character Recognition

To cite this article: Wira Satyawan *et al* 2019 *J. Phys.: Conf. Ser.* **1235** 012049

View the article online for updates and enhancements.

Citizen Id Card Detection using Image Processing and Optical Character Recognition

Wira Satyawan, M Octaviano Pratama, Rini Jannati, Gibran Muhammad, Bagus Fajar, Haris Hamzah, Rusnandi Fikri, Kevin Kristian

Artificial Intelligence Department, Premier Optima Sattiga, Jakarta, Indonesia

Email: octav@premieroptima.com

Abstract. Since its emergence in 2011, Indonesian Electronic Id-card has been widely used as authentication or citizen identity. Several issues like deep difficulty in detecting id-card field and also difficulty in character recognition data in id-card should be concerned. In this research, we propose a technique detect electronic Id-card using combination of Image Processing and Optical Character Recognition (OCR). The result, we can obtain 98% accuracy of Id-card detection using our image processing techniques and OCR. This research was embedded in website interface which used by automotive company.

1. Introduction

The development of Information Technology has developed quite rapidly, both in theory and application. A lot of research technology has used to facilitate and accelerate human works. The researches have been implemented to computer and used to accomplishing human works optimally. One example of the development of information technology in business is how to purchase goods. Currently, we don't have to visit the store for purchasing some goods. Purchase of goods can also be done by online. In various businesses, companies need customer data that should be inputted into database for online or offline purchase. Data of customers who buy item by online are usually requested when registering an account, while customers who buy item by offline are usually asked to get their identity. Data of costumer's identity can be obtained from their ID Card. The ID Card that used for this case is citizen ID Card. Previously, customers data inputted manually. That is not efficient process because we need a lot of time to input data one by one. Therefore, we need a system that processes automatically.

Based on that problem, Image Processing technique can be used as an alternative solution of manually input process. This process starts by extracting information in ID Card image. Then, it will be pre-processed to obtain the necessary part of image. Furthermore, Optical Character Recognition (OCR) will be performed in order to recognize text in images. OCR can recognize handwriting and text characters automatically through optical mechanism. OCR is designed to process images consisting of text with little non-text data interference. While the OCR performance depends on the quality of the inputted document [1].

Based on some research above, this study compares the result of character recognition of name and NIK (identity number) in ID Card using two different tesseract models. The first model uses the train data manually that created from five ID Card as data set and training on tesseract 3.05 with the support of software QT-box version 1.08. While the second model uses train data that already contained in tesseract 4.0, which is a data train that contains text data in Indonesian language with different fonts

and using tesseract version 4.0 for OCR which in that version has implemented neural network model that is LSTM.

The contributions of this research are: (1) we propose image processing method for detection citizen id card particularly for Indonesian citizen id card, (2) we try various models of recognition implemented in Tesseract frameworks, (3) we propose website interface that has this system for citizen id card detection and recognition which will be useful for scanning result. The final result of this research is used by one of automotive company in Indonesia which run on website interface platform.

2. Related Works

Recognition process of characters in the image from year to year has growing more. In 2005, Wang et al used Gabor-filters for character recognition with low image quality and for Chinese-readable characters [2]. In 2011, Vikas et al developed document segmentation using histogram analysis [3]. Sreedhar et al in 2012 developed image processing using the Morphological Transformation method and Weber's Law which enhances the contrast of an image [4]. Ryan [5] et al in 2015 conducted research about character recognition on ID Card of Indonesian people using Zhange-suen algorithm divided into 2 algorithms: 3x3 algorithm and pixel-by-pixel algorithm. Valiente [6] using Optical Character Recognition to detect id card combined with cloud technology. Most of previous research using image processing technology combining with Machine Learning to detect citizen id card. The appropriate selection of Image Processing and Machine Learning Techniques can improve accuracy of prediction.

3. Methods

In this research, we start from data collection of Citizen Id Card, then we divide data into training data and testing data. After collecting appropriate data, pre-processing is performed in order to make image that used in forward tasks. Then text area extraction and Segmentation are performed to determine area that should be taken automatically. The last step, Optical Character Recognition (OCR) was used for predicting character in Citizen Id Card.

3.1. Pre-processing

ID card that used in this research has a uniform size 1654×2340 per images. The pre-processing of image is generally divided into 3 parts: grayscale, thresholding, and morphological transformation.

3.1.1   Grayscale

Grayscale is the process of converting an image that previously consisted of 3 RGB layers into a gray image that has 1 layer. Making image to Grayscale is used to obtain optimum Binary image results.

3.1.2   Thresholding

Based on the thresholding formula [7] is defined as:

$$g(x,y)\begin{cases} 1, \text{if } f(x,y) > T \\ 0, \text{otherwise} \end{cases} \tag{1}$$

This thresholding converts the image into binary by selecting the threshold. In this research, we put 100 as thresholding value. For then so that the pixel colour becomes black. And for in other values, so the pixel colour becomes white Therefore the character on the ID Card which originally black will change to white colour while the other colour will change to black colour.

3.1.3   Sobel

Most edge detection methods work on the assumption that the edge occurs where there is a discontinuity in the intensity function or a very steep intensity gradient in the image. Using this assumption, if one takes the derivative of the intensity value across the image and find points where the derivative is maximum, then the edge could be located. The gradient is a vector, whose components measure how rapid pixel value are changing with distance in the x

and y direction. Thus, the components of the gradient may be found using the following approximation [8]:

$$\frac{\partial f(x,y)}{\partial x} = \Delta x = \frac{f(x+dx,y)-f(x,y)}{dx} \tag{2}$$

$$\frac{\partial f(x,y)}{\partial x} = \Delta y = \frac{f(x,y+dy)-f(x,y)}{dy} \tag{3}$$

where and measure distance along the and directions respectively.

3.1.4   *Morphological Transformation*

In this research, eliminate noise in image use morphological transformation technique. Morphological operation is an image processing technique based on the shape of an object. This method applies the element structure to the image input and makes the image output of the same size. The value of each pixel in the image input is based on the pixel ratio with its neighbour in the image input [4]. The morphological operation used in this study consists of four operations: dilation, erosion, opening operation, and closing operation.

- Dilation is a transformation that produces an image that has the same shape as the original image, but has a different size. Structure of element is positioned with its original point and the new pixel value is determined using the equation:

$$g(x,y) \begin{cases} 1, if\ s\ hits\ f \\ 0, otherwise \end{cases} \tag{4}$$

- Erosion is used to reduce the object in the image by reducing the erosion peak and enlarge the width to the minimum area so as to eliminate noise. The structure of element is positioned early on and the new pixel value is determined by [5]:

$$g(x,y) \begin{cases} 1, if\ s\ hits\ f \\ 0, otherwise \end{cases} \tag{5}$$

- Opening on by is obtained from erosion by, then the results are dilated. Generally defined by:

$$A \circ B = (A!\, B) \oplus B \tag{6}$$

- Closing A by B is obtained by dilation A by B, followed by the resulting erosion of the structure by B:

$$A \cdot B = (A \oplus B)!\, B \tag{7}$$

- Tophat is an operation that has some high-pass filtering characteristics. So the opening Tophat operator can detect the wave crest of an image and the closing Tophat operator can detect hollow waves. Tophat opening and closing are defined by [9]

$$OTH_{F,B}(x) = (F - F \circ B)(x) \tag{8}$$

$$CTH_{F,B}(x) = (F \cdot B - F)(x) \tag{9}$$

3.1.5   *Otsu*

An image can be represented by a 2D gray-level intensity function. The value of is the gray-level, ranging from to, where is the number of distinct gray-levels. Let the number of pixels with gray-level be, and be the total number of pixels in a given image, the probability of occurrence of gray-level is defined as [10]:

$$p(i) = \frac{n_i}{n} \tag{10}$$

The average gray-level of the entire image is computed as [10]:

$$\mu_T = \sum_{i=0}^{L-1} ip(i) \tag{11}$$

3.2. *Text Area Extraction*

After pre-processing, the next step to do is text area extraction, to determine the area of the character to be taken. We define the kernel with the size that we want. In this research we determine kernel size 5x5. This kernel is useful for forming a box on everything that contains text.

3.3. Segmentation

Image segmentation is used to determine the part of the text to be retrieved. In this research the part that will be taken is NIK character and name on ID Card. We set the width and height of the kernel box and pixel coordinates. Then it will produce the result of crop ID Card as we want.

3.4. Character Recognition

Character recognition using Optical Character Recognition (OCR) technique with tesseract tools. This research will compare the character recognition model using the data train manually that trained using tesseract 3.05 and data train that already contained in tesseract 4.0 which containing Indonesian language text data with different font.

4. Result

In this research, we made two different models to identify name and ID card character section in ID Card. First model used manual training data and trained using Tesseract 3.05 and QT Box software Version 1.08. Second model used training data in tesseract 4.0 which is Indonesian text with different font.

4.1. Manual Training Data

First step in identifying character was doing scanning ID Card with resolution 1.665 x 2340. And then, image will be done to image processing. Firstly, the result of grayscaling on pre-processing image presented in figure 1.
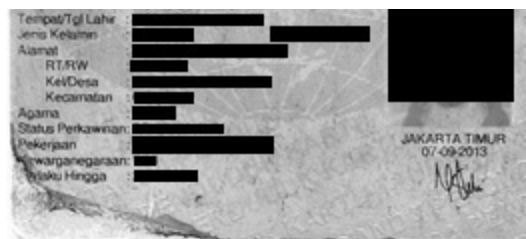

Figure. 1. Greyscale result example

After that, change the image to binary invers with value threshold 90. After changed image to binary invers, this image is transformed using morphological transformation and OTSU as shown in figure 2.
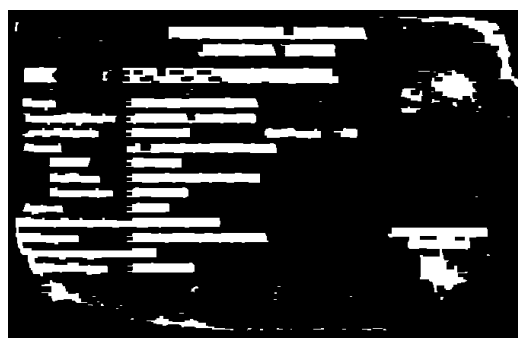

Figure. 2. Morphological transformation and OTSU result example

Next step is extraction text area. In this step, we determine what is kernel we can use for the process. We used the kernel for closing all character in ID card. After that we build the program to cut the area what we want. In this research, we will cut NIK and name area. We select the right kernel in the area we want by setting the kernel height and width depending on the x and y coordinates of the image, so that the selected kernel is the kernel in the NIK column and the name.

After the kernel was selected and was do segmentation the NIK and name on the ID card, the result from segmentation NIK and name are entered the model in training manually using data train tesseract

version 3.05 was assisted with qt-box version 1.08 and we obtained 100% prediction accuration. Next step, the model tested with new data test and the result NIK and name of character recognition use data test is shown in table 1.

Table 1. Result and accuracy character recognition NIK and name data test use data train was created manually use qt-box version 1.08

Ground Truth	Prediction	Accuracy
RANGGA ADIANSIA 3510171910970002	RANGGA ADIANSIA 3510171910970002	100%
KEVIN CHRISTAN AVANTYO 3276022105950009	KEVIN CKRISTAN AVSNTIO 3276022105950009	85%
GIBRAN MUHAMMAD FAJRI 3175082401980006	GIBBAN MUHAMMAD FMAR 3175082401980006	84.21%
HARIES HAMSA 3174101310950005	HARIES HAMSA 3174101310950005	100%
ARNANDI FARHAN 3674030402940001	ARNANOI FABHAN 3674030402940001	84.62%

There are some errors in the character reading results in the model. This is because to the small amount of data train and several letters on the data test is not in the data train used.

4.2. Model Data Train Tesseract 4.0 and OCR Tesseract 4.0

In the second experiment, we used the same pre-processing and segmentation process as the first experiment. The difference from the first experiment is to OCR, we use tesseract version 4.0 which in that version has implemented neural network model that is LSTM.

Training data that used is default data that has been obtained from tesseract that contains text data in Indonesia language with different font. To read the NIK number, we retrained the text data using the font on the NIK number. Based on the model that obtained from the data train, we test using the same data train with the first experiment and we obtained 100% prediction accuration for NIK but we obtained 98.6% prediction accuration for name. After that, do the test using the same test data as the test data in the first experiment, and got the results as shown in table 2.

Table 2. The result and accuracy of NIK and name character recognition on test data using data train made from tesseract 4.0

Ground Truth	Prediction	Accuracy
RANGGA ADIANSIA 3510171910970002	RANGGA ADIANSIA 3510171910970002	100%
KEVIN CHRISTAN AVANTYO 3276022105950009	KEVIN CHRISTAN AVANTYO 3276022105950009	100%
GIBRAN MUHAMMAD FAJRI 3175082401980006	GIBRAN MUHAMMAD FAJRI 3175082401980006	100%
HARIES HAMSA 3174101310950005	HARIES HAMZA 3174101310950005	90,91%
ARNANDI FARHAN 3674030402940001	ARNANDI FARHAN 3674030402940001	100%

The accuracy for NIK character recognition of both methods shows 100%. While the acknowledgment of character recognition of names of both methods does not show accurate accuracy because there are errors of some letter characters. This because of the train data tesseract is not enough so that tesseract could not recognize all characters of letters.

4.3. User Interface

We used Flask to create user interface to make it easier for user to run the program. In first step, we built the simple user interface for users to upload Citizen ID card image with maximum size 1.654 x 2.340. The program will be run the model that has been created and then give a result of ID card and name.

5. Conclusion

Citizen ID card can be detected by using proposed image processing techniques and collaborated with OCR. Image processing techniques in this research consist of preprocessing, text area extraction, and segmentation. OCR proposed for character recognition. This research combines grayscale preprocessing techniques with binary image processing techniques such as Sobel, morphological transformation, and OTSU. Text area extraction uses a kernel that identifies the text area of the NIK and name on the ID card citizen. The experiments with training data made using tesseract 4.0 show that accuracy of detection reach between 90 - 100 % using our propose technique. We also create another model with training data created using qt-box as benchmarks.

6. References

[1] Mithe R, Indalkar S and Divekar N 2013 Optical Character Recognition *Int. J. Recent Technol. Eng.* **2** 72–5
[2] Wang X, Ding X and Liu C 2005 Gabor filters-based feature extraction for character recognition *Pattern Recognit.* **38** 369–79
[3] Dongre V J and Mankar V H 2011 DEVNAGARI DOCUMENT SEGMENTATION USING HISTOGRAM APPROACH *Int. J. Comput. Sci. Eng. Inf. Technol.* **1**
[4] Sreedhar K and Panlal B 2012 ENHANCEMENT OF IMAGES USING MORPHOLOGICAL TRANSFORMATIONS *Int. J. Comput. Sci. Inf. Technol.* **4**
[5] Ryan M and Hanafiah N 2015 An Examination of Character Recognition on ID card using Template Matching Approach *Procedia Computer Science* vol 59 pp 520–9
[6] Valiente R, Sadaike M T, Gutiérrez J C, Soriano D F and Bressan G 2016 A process for text recognition of generic identification documents over cloud computing *IPCV'1International Conf. Image Process. Comput. Vision, Pattern Recognit.* 4
[7] Sitthi A, Nagai M, Dailey M and Ninsawat S 2016 Exploring Land Use and Land Cover of Geotagged Social-Sensing Images Using Naive Bayes Classifier *Sustainability* **8** 921
[8] Vincent O R and Folorunso O 2009 *A Descriptive Algorithm for Sobel Image Edge Detection*
[9] Zeng M, Li J and Peng Z 2006 The design of Top-Hat morphological filter and application to infrared target detection *Infrared Phys. Technol.* **48** 67–76
[10] Ng H-F, Kheng C-W and Lin J-M *A Weighting Scheme for Improving Otsu Method for Threshold Selection*