Food Packaging Search Application From Text Image In Android With Deep Convolutional Neural Network (DCNN) Method

Siti Aisyah*, Fransiska Susilawati Nainggolan, Melva Simanjuntak, Edi Apriyanto Lubis
Faculty of Technology and Computer Science, Universitas prima Indonesia
*sitiaisnah@unprimdn.ac.id

Abstract. Search Application Info on Food Packaging with Image Input based on Android is one application that can make it easier for users to see a description of the ingredients on food packaging. This application uses the Deep Convolutional Neural Network (DCNN) method which is able to show outstanding performance in the field of image recognition, especially in the field of character recognition, because DCNN has a performance capable of extracting high-level features. Detailed information on food packaging is usually written on food packaging with a small font size that makes it difficult for consumers to read it, and on food packaging there is also no selling price of the product. This makes it difficult for consumers who only want to know the information and prices of these food products. A method is needed to get a text version of the food packaging brand image and to become a search keyword on the food information service site. In this report, we will examine digital image processing especially in the field of pattern recognition to analyze the titles contained in food packaging in real time. This research will utilize smartphone cameras that are used to detect objects specifically for food packaging to see the price, composition, brand, net weight and description of the food packaging based on the food packaging detected.

1. Introduction
A product must have information about the product being marketed. A good product must have the information listed both outside and inside the package so that it can provide the users with the same convenience. Information on product packaging becomes a benchmark for the feasibility of its use, especially in food products. So far there are still many consumers who pay less attention to the information printed on the packaging of the products they buy or use. This lack of awareness makes some people suffer losses for the products used. In fact, it is not uncommon for some people to abuse the opportunity to make a profit.

Technology aims to facilitate the fulfillment of human needs [1]. Technology that is currently developing is mobile technology, which is an application that can be used even if the user moves easily from one place to another [2]. In 2013, starting from April to May, a survey was conducted on the use of Android. From the survey results it was found that the OS became the most prominent platform among
its developers, with almost 70% being used by mobile application developers [3]. Neural networks are one of the most widely used fields of science by researchers. Can collaborate neural networks with other fields of science make this field of science popular, like research on the field of chest health [4], skin [5], hepatitis [6]. With a variety of algorithms such as backpropagation [7], Ant Colony Optimization [8]. In addition to collaborating NN has many functions in its use, such as testing, prediction, pattern recognition and so on.

Several studies have been conducted by previous researchers such as A Turnip and D. Soetraprawata conducting research on EEG signals using backpropagation neural networks [9]. Nurul et al used the feature directional element and multi class support vector machine to introduce the introduction of Javanese script writing [10]. One method used for pattern recognition is the Deep Convolutional Neural Network (DCNN). DCNN is one method in the Neural Network. (DCNN) is able to demonstrate outstanding performance in the field of image recognition, especially in the field of character recognition, because DCNN has a performance capable of extracting high-level features. This research focuses on the process. DCNN is studied in various fields because of its fundamentality in image recovery [11]. Some researchers have used DCNN such as Multi-digit number recognition from street view imagery [12-15]. This study focuses on making information retrieval system applications on food packaging using DCNN through the introduction of packaging samples. The stages of writing this study consisted of introduction, Methodology, Result and Discussion, and Conclusion.

2. Methodology
At this stage an analysis of the literature study was conducted to gain an understanding of the Deep Convolutional Neural Network method to be used in solving problems, namely recognizing the text contained in an image.

Character recognition of food packaging images through smartphone cameras in this study consists of several stages. The first stage is the acquisition of food packaging images using a smartphone camera. This stage consists of several steps, namely determining the Region of Interest (ROI) by doing a crop after the image is acquired. Furthermore, the food packaging image will be processed through several stages of pre-processing. This stage begins by changing the image into a grayscaling image followed by a smoothing process, thresholding and erosion.

The results of pre-processing will enter the feature extraction stage and identify using the Deep Convolutional Neural Network. Furthermore, after the preprocessing text is successfully identified, the crawling process is carried out to obtain information about the description of food packaging and then stored in a database of data that is successfully crawled.

The output image from the pre-processing process will be the input for the feature extraction and identification process. In this process, the features will be taken images which will then be identified to find out the characters contained in the image. The process is carried out using Optical Character Recognition technology based on Deep Convolutional Neural Network, namely Google Mobile Vision. The results of this process are A-Z characters, numbers 0-9, and point (.), Comma (,) or space () characters that are usually used to separate the nominal numbers on food packaging.

In this study the resolution used is the resolution of the smartphone used, the higher the quality of the smartphone's camera, the better the resolution for capturing the image on the packaging. From the overall tests that have been carried out on the text characters in the food packaging title using a camera, tests have been carried out on food packaging samples, and test taken on food packaging samples as many as 100 food packs and the best accuracy on text character recognition in food packaging images reaches 93 food packaging that was successfully tested.

Deep Convolutional Neural Network has several stages where each stage will show training accuracy, validation accuracy and cross entropy. In the training process, the first 224x224x3 input image will be processed by the first convolutional layer with a 96 kernel 11x11x3 filter. Then, the results of the first convolutional layer measuring 55x55x48 will be processed by the second convolutional layer with a 256 kernel 5x5x48 filter.
Third, Fourth and Fifth Convolutional layers are connected to one another without pooling intervention or layer normalization. The results of the second convolutional layer will be processed by the third convolutional layer with a 256 kernel 3x3x192 filter. The input image for the fourth convolutional layer has a size of 13x13x192. Will be processed with 256 kernel filter size 3x3x192. After the convolutional layer process, 3 fully connected layers are produced which have 4096 neurons in each layer. The results of the last fully connected layer are 1000-way softmax. Training Process Each Layer uses the Deep Convolutional Neural Network can be seen in the following Table:

| Layer       | Kernel | Size       | Number of Neurons |
|-------------|--------|------------|-------------------|
| Input Image | -      | 224 x 224 x 3 x 1 | 150.528          |
| **Convolutional Layer** |         |            |                   |
| First       | 96     | 55 x 55 x 48 x 2 | 290.400          |
| Second      | 256    | 27 x 27 x 128 x 2 | 186.624          |
| Third       | 384    | 13 x 13 x 192 x 2 | 64.896           |
| Fourth      | 384    | 13 x 13 x 192 x 2 | 64.896           |
| Fifth       | 256    | 13 x 13 x 128 x 2 | 43.264           |
| **Fully-Connected Layer** |         |            |                   |
| -           |        | 2048 x 2   | 4.096             |
| **Fully-Connected Layer** |         |            |                   |
| -           |        | 2048 x 2   | 4.096             |
| **Fully-Connected Layer** |         |            |                   |
| -           |        | 1000 x 1   | 1000              |

3. Theoretical foundation
Artificial neural networks provide calculators that are inspired by the operating structure of the brain and the central nervous system [16]. Digital image processing (Digital Image Processing) is a discipline that studies image processing techniques. The image referred to here is a still image (photo). In order to be processed with digital computers, an image must be presented numerically with discrete values. Representation of continuous functions into discrete values is called image digitization. A digital image can be represented by a two-dimensional matrix f (x, y) consisting of M columns and rows, where the intersection between columns and rows is called pixels (pixel = picture element) or the smallest element of an image [17].

Deep Convolutional Neural Networks (ConvNets) are special cases of artificial neural networks (ANN) which are currently claimed to be the best model for solving object recognition and detection problems. Convolutional Neural Network (CNN) is the development of Multilayer Perceptron (MLP) which is designed to process two-dimensional data. CNN is included in the type of Deep Neural Network because of the high network depth and much applied to image data [18].

4. Results and Discussion
4.1. Dataset
In the initial stage, an analysis of the food packaging will be analyzed, then the food packaging will later become an input to the application requirements. The data that is input is food packaging text. Data in
the form of food packaging images will be the test data tested on the application. The image of food packaging that becomes the test data can be taken by using a smartphone.

4.2. Preprocessing Application
The first stage is the acquisition of food packaging images using a smartphone camera. This stage consists of several steps, namely determining the Region of Interest (ROI) by doing a crop after the image is acquired. Furthermore, the food packaging image will be processed through several stages of pre-processing. This stage begins by changing the image into a grayscaling image followed by a smoothing process, thresholding and erosion.

The results of pre-processing will enter the feature extraction stage and identify using the Deep Convolutional Neural Network. Furthermore, after the preprocessing text is successfully identified, the crawling process is carried out to get information about the description of food packaging and then stored in a database of data that is successfully crawled.

4.3. Post Processing
Text that has been identified is not all stored and the text will then be processed for crawling. For this reason, it is necessary to choose which text will be processed automatically. The selected text is the title text on food packaging. After automatic selection by the system, the system sends the text to the next process, namely the crawling process.

Figure 1. Processing Application
4.4. Crawling Process
To be able to do the crawling process, the system performs crawling begins with searching the string on the website www.lifull-produk.id/ after the search data appears, then to do crawling data, the Depth First Crawling method is used, the crawling method with the search process from root node to node end. Starting from the root node (initial search), go down to the product name node, down to the product data node. If when searching for a product name it is not found, the system will automatically search for another product name according to the name the user is looking for. Once found, the system will continue to the next node, namely product data search.

4.5. Stripping Process
The stripping stage is the stage of cleaning the html code to retrieve the desired data on the website page. The stripping process is carried out to retrieve the title and content of the product data and described in the form of a flowchart. The understanding of flowchart is a part that describes the logic flow of data that will be processed in a program from beginning to end [14].

The flow of the stripping process is that the stripping process begins with getting a link from the user so that it displays an information from the page that has been entered in the link. Stripping system is carried out which is taking some data on a web page by retrieving selected data, starting with taking pictures by storing the image link, then making the process of taking the product name and all product data consisting of price, composition, factory etc.

4.6. Test result
At this stage, the character recognition process in the receipt image begins with the preprocessing stage, then enters the feature extraction stage and introduces using the Deep Convolutional Neural Network (DCNN) implemented through the Java programming language on an Android-based system.

4.7. Examples of test results
The final stage of this research is character recognition in the title of food products with the DCNN method where the text output will display a description of food products in the new window. The following is shown in the picture:
5. Conclusion

Based on the research that has been done, it can be concluded that this study produced an application info on food packaging using Deep Convolutional Neural Network (DCNN) implemented through the Java programming language on an Android-based system. This food packaging info application is only able to recognize food packaging text images found only on food product sites. This accuracy is obtained in conditions with focus and slope when image taking is varied. The failure of food packaging text character recognition is caused by the following things, if there is no food product category to be searched for in the food product site used, there is a pixel disturbance that interferes with the image in the food packaging cover title. The image of the cover title of the food detected has a circular or diagonal shape and has density with other letters.

REFERENCE

[1] Rusman dkk, Pembelajaran Berbasis Teknologi Informasi dan Komunikasi. (Jakarta: Grfindo persada, 2012), hal. 78.

[2] Fahri Rivaldi, 2016. Perancangan Aplikasi Mobile “Kamusku”

[3] Developer Economics Q3 2013 analyst report – http://www.visionmobile.com/DevEcon3Q13 – Retrieved July 2013
[4] D. vally, CH V Sarma, 2015, Diagnosis Chest Diseases Using Neural Network and Genetic Hybrid Algorithm. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 5, Issue 1, 20- 26.

[5] Bakpo, F. S.1 and Kabari, L. G, 2009, Diagnosing Skin Diseases Using an Artificial Neural Network. Artificial Neural Networks - Methodological Advances and Biomedical Applications, 253-270, DOI: 10.1109/ICASTECH.2009.5409725.

[6] Mehdi Neshat, Azra Masoumi, Mina Rajabi and Hassan Jafari, 2014, Diagnosing Hepatitis Disease by Using Fuzzy Hopfield Neural Network, Annual Research & Review in Biology, 2709-2721.

[7] Dewi, Candra dan M. Muslikh, 2013, Perbandingan Akurasi Backpropagation Neural Network dan ANFIS untuk Memprediksi Cuaca. Journal of Scientific Modelling & Computation, Vol. 1, No. 1.

[8] C. Blum and K. Soch, Training feed-forward neural networks with ant colony optimization: An application to pattern classification, Fifth International Conference on Hybrid Intelligent Systems (HIS 2005), Rio de Janeiro, Brazil (2005), pp. 233–238.

[9] A Turnip and D. Soetraprawata, The Performance of EEG-P300 Classification Using Backpropagation Neural Network, Mechatronics Electrical Power and Vehicular Technology, 04 (2013), 81-88.

[10] A. H. Nurul, M. D. Sulistiyo, and R. N. Dayawati, “Pengenalan Aksara Jawa tulisan tangan menggunakan directional element feature dan multi class support vector machine,” in Prosiding Konferensi Nasional Teknologi Informasi dan Aplikasinya, 13 September 2014, Palembang, Indonesia [Online]. Available: http://seminar.ilkom.unsri.ac.id/index.php/kntia/article/view/733/409. [Accessed: 4 February 2017].

[11] Li Xu, et al. Deep Convolutional Neural Network for Image Deconvolution. (2014).

[12] Goodfellow, Ian J, Bulatov, Yaroslav, Ibarz, Julian, Arnoud, Sacha, and Shet, Vinay. Multi-digit number recognition from street view imagery using deep convolutional neural networks. arXiv preprintarXiv:1312.6082, 2013.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25, pages 1106-1114, 2012.

[14] M. D. Zeiler and R. Fergus. Stochastic pooling for regularization of deep convolutional neural networks. CoRR, abs/1301.3557, 2013

[15] Nitish Srivastava, et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research 15 (2014) 1929-1958.

[16] H. R. Maier and G. C. Dandy, “Neural network based modelling of environmental variables: a systematic approach,” Mathematical and Computer Modelling, vol. 33, no. 6-7, pp. 669–682, 2001.

[17] RD. Kusumanto, Alan Novi Tompunu.. Pengolahan Citra Digital Untuk Mendeteksi Obyek Menggunakan Pengolahan Warna Model Normalisasi Rgb, Jurusan Teknik Komputer, Politeknik Negeri Sriwijaya, Palembang 30139, Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2011 (Semantik 2011), ISBN 979-26-0255-0.

[18] I Wayan Suartika E. P, Arya Yudhi Wijaya, dan Rully Soelaiman, Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101, Teknik Informatika, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember (ITS), Jurnal Teknik Its Vol. 5, No. 1, (2016) ISSN: 2337-3539.