Image Processing Identification for Indonesian Cake Cuisine using CNN Classification Technique

Dian Ade Kurnia¹, Andi Setiawan², Dita Rizki Amalia³, Rita Wahyuni Arifin⁴, Didik Setiyadi⁵

¹Departement of Informatics, STMIK IKMI Cirebon, Indonesia
²Departement of Informatics, STMIK IKMI Cirebon, Indonesia
³Departement of Informatics, STMIK IKMI Cirebon, Indonesia
⁴Department of Informatics, Bina Insani University, Indonesia
⁵Department of Informatics, Bina Insani University, Indonesia

Abstract. Indonesia is famous for its traditional food that is popular both domestically and abroad. A number of cakes are among the favorite traditional foods. There are types of cakes that can be processed in Indonesia, such as kue dadar gulung, kastangel, klepon, lapis, lumpur, putri salju, risoles and serabi. The most of types of cakes available, visually the human recognize are easy, however computer vision requires a special technicality in identifying the object of the image to the type of cakes. Therefore, to recognize objects in the form of images from cakes as one of Indonesia's traditional foods, deep learning algorithm techniques can be used, namely Convolutional Neural Network (CNN). In this paper, the CNN Algorithm technique will be applied to 1676 datasets consisting of 80% training data and 20% testing data in which there are images of traditional cakes from Indonesia. The stages are carried out through preprocessing, operational datasets, visualization datasets, modeling techniques, performance evaluations, errors analysis which finally result in the conclusion that performance evaluation reaches 65.00%.

Keywords: CNN, Cakes, Accuracy, images, neural network

1. Introduction

Indonesia is one of the countries with a variety of cultures, from dance to cooking. Cake is an unique culinary specialty. Specific types of cakes that are manufactured to make the taste special, even though the name is the same [1]. For humans, to recognize an object in image is very easy, but in the field of computer vision requires special techniques [2]. In today's age, consumers are fully aware of their own food and diet to care of the patient that are either coming or already there [3][2]. Since people are based on intelligent technology, the provision of an application to track individual diets automatically helps. As either a swollen area of billions of large island nations, In the pacific ocean, Indonesia does have rich water resources that allow A variety of agricultural products and plantations are grown by its people. Traditional cakes are largest island cakes as a form of cultural development. The culture of the ancestors of a nation, that has been transmitted by generation to generation explicitly in the areas or society of Indonesia [2]. Traditional cakes have the characteristics of recipe books created by generation. From generation to generation, use methods, both field tools and basic conventional tools. Traditional cakes are included in conventional dishes, modern cakes. It can be grouped by type. Traditional cake types can be seen in terms of manufacturing, flavors and form [4]. Indonesia is famous
for its traditional food that is popular both domestically and abroad. A number of cakes are among the favorite traditional foods. There are types of cakes that can be processed in Indonesia, such as *kue dadar gulung, kastangel, klepon, lapis, lumpur, putri salju, risoles and serabi* [5]. The most of types of cakes available, visually the human recognize are easy, however computer vision requires a special technicality in identifying the object of the image to the type of cakes. Therefore, to recognize objects in the form of images from cakes as one of Indonesia’s traditional foods, deep learning algorithm techniques can be used, namely Convolutional Neural Network (CNN).

2. **Methodology**

   *a. Object and Dataset*

   The object of this study is a cake-type food that is a type of drawing object consisting of 8 (eight) categories namely *kue dadar gulung, kastangel, klepon, lapis, lumpur, putri salju, risoles and serabi*. The total dataset of 1676 images, taken from digital camera results. The amount of data per category is spread proportionally.

   *b. Convolutional Neural Network (CNN) Algorithm*

   1) **What is the CNN**

   CNN images may be used to recognize and identify objects. CNN is composed of weight, motivation and activation neurons. Convolutional layers are often made up of neurons organized to form a filter, with a length and height [6][7].

   ![Figure 1. Architecture of CNN including input, conf FC layer (Conv=Convolutional FC=Fully Connected)](image)

   Convolution is a significant mathematical analytical process. It is a computational operator that introduces a third function of two functions $f$ and $g$, which displays the domain between $f$ and $g$, which is covered or translated. Its calculation is usually defined by a following formula [5] :

   $$z(t)^{def} = f(t) * g(t) = \sum_{r=-\infty}^{\infty} f(r)g(t-r)$$  \hspace{1cm} (1)

   **Convolutional Layer**

   Neurons shape innovative layers so that a filter of length and height (pixels) is created. The first layer on the extraction layer, for example, is conv. layer with a size of 5x5x3. Length 5 pixels, height 5 pixels, and thickness/number of 3 pieces according to the channel of the image [8][9].

   **Fully Connected**

   The max pooling eliminates systems and numbers of network parameters and increases overfitting facilitations and evaluations. Capes function in synergy with the pattern of spatial blocks. In general, the block shift is the block dimension (E x H). As in the Convolutionary Layer no overlap is found. [10].

   **Activation Function**

   Optimization algorithms are part of the CNN, it is interesting to separate the neural network of proceed. [11]. A nonlinear technique helps a network to solve non-trivial topics is the activation or linear transfer. Each input layer takes a value and performs a math [2]. The activation function in CNN architecture is to evaluate a feature working independently or to construct a feature pattern after convolution or classification. In certain operations, the sigmoids method is also used in analysis. The Tanh, Rectified Linear Unit (ReLU), Leaky ReLU and modulation ReLU were among the most popular activation
functions of ReLU or Rectified Linear Unit. [9]. The input layer Rectified linear unit (ReLU) is used to insert non-linearities in the model via generated output map as shown on the convolution layer. Fig 2 [7].

![Figure 2. Rectified Linier Unit (ReLU) applied in CNN](image)

2) Pre processing

The basic method of image processing is carried out in the pre-processing stage: imaging, redimensioning, cultivation, processing during training. Data of image processing begin for any size image that is then resized 150 x 150 pixels.

3) Training Phase

The studies included image input, pre-processing, identification of artifacts, extraction of features, classifications and cake classification performance.

![Figure 3. The Stage of Research](image)

Several stages of implementation of the method used include load dataset and data visualization, defining data, building models, training and evaluating training results.

c. Load Dataset

The dataset capture is not entirely, will be taken 2 types of cakes and displayed in the form of an image file as many as 10 files from each category, then displayed the image in a size of 150 x 150 pixels.

```
total training kue putri salju images: 174
total training kue klepon images: 200
['18.jpg', '20.jpg', '0.jpg', '27.jpg', '19.jpg', '24.jpg', '2.jpg', '17.jpg', '5.jpg', '1.jpg']
['114.bmp', '29.jpg', '19.jpg', '14.jpg', '1.jpg', '26.jpg', '16.jpg', '9.jpg', '11.jpg', '18.jpg']
```

![Figure 4. Load Dataset Image](image)

d. Define Data

Before modeling, it must first be known the number of datasets used as training data, test data and validation data. Then obtained the result that the amount of training data is 1516 images, validation data is 160 images and test data as much as 160 images. Thus that proportional data training with test data and validation data amounted to 89.45%.
e. **Modeling**

The model proposed in resolving this issue is to use the Convolutional Neural Network (CNN). The reason is that with neural networks can register a large number of image features and can be used for classification. The way this model works is that the image data as input data is inserted into the layer, the layer size at the first convolution called the filter will be arranged 148 x 148 x 16, meaning the input size is 148 x 148 and the number of neurons as many as 16 neurons. Next resize the window (k-size) to one D-dimensional vector, called **pooling**, will eventually be passed one vector resulting from multiple convolution and pooling operations on a multi layer perceptron known as **fully connected** to perform classification.

| Layer (type)      | Output Shape | param # |
|------------------|--------------|---------|
| conv2d (Conv2D)  | (None, 140, 140, 16) | 440     |
| max_pooling2d (MaxPooling2D) | (None, 74, 74, 16) | 0       |
| conv2d_1 (Conv2D)  | (None, 72, 72, 16) | 2320    |
| max_pooling2d_1 (MaxPooling2D) | (None, 36, 36, 16) | 0       |
| conv2d_2 (Conv2D)  | (None, 34, 34, 32) | 4640    |
| max_pooling2d_2 (MaxPooling2D) | (None, 17, 17, 32) | 0       |
| conv2d_3 (Conv2D)  | (None, 15, 15, 32) | 9248    |
| flatten (Flatten)  | (None, 7200) | 0       |
| dropout (Dropout)  | (None, 7200) | 0       |
| dense (Dense)      | (None, 256) | 1843456 |
| dense_1 (Dense)    | (None, 8) | 2056    |

| Total params: | 1,862,168 |
| Trainable params: | 1,862,168 |
| Non-trainable params: | 0 |

f. **Train**

After the construction of the convolution model, further training of the model is carried out. The purpose of training this model is to perform testing and produce validation accuracy. Training is carried out 20 times (epoch), each epoch consists of 24 steps.
3. Result And Discussion

a. Result

The result of this study is accuracy in the accuracy of predicting an image object. Based on the model developed, the accuracy in discriminating the image object in the form of Indonesian cake food amounted to 65%, after being experiment through epoch 20 times and each epoch as many as 24 steps.

Figure 5. Comparison between training accuracy and validation accuracy

Figure 6. Evaluate Result

d. Discussion

The above results show a level of accuracy that is not yet maximal, which is shown with an accuracy result of 65%, this is because cnn algorithms require more detailed and more perfect data. Things to note before convoluted processes are low-resolution images, different of images, different image contours and others.
4. Conclusion

As the final part of this study, it can be done that the classification of Indonesian cakes has been successfully carried out against 8 image categories including **kue dadar gulung, kastangel, klepon, lapis, lumpur, putri salju, risoles and serabi** using the Convolutional Neural Network (CNN) algorithm, so that predictions of accuracy in identifying the object are as large as 65.00%.

The advice and recommendations for further research are on the pre processing of data that should be done in detail, so that it will affect the performance of the model.

Acknowledgment

Thank you to STMIK IKMI Cirebon who have supported the process of conducting research activities that have been conducted by us, ranging from preparation, implementation to the support of funds that have been given.

References

[1] S. P. Erdiana, “Generasi Millenial dan Resep Kuliner Tradisional sebagai Budaya Bangsa Indonesia,” *DeKaVe*, vol. 11, no. 2, pp. 35–42, 2019.

[2] D. J. Attokaren, I. G. Fernandes, A. Srim, Y. V. S. Murthy, and S. G. Koolagudi, “Food classification from images using convolutional neural networks,” *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2017-Decem, pp. 2801–2806, 2017.

[3] E. I. Agustin, R. T. Yuniardi, and A. A. Firdaus, “Voice recognition system for controlling electrical appliances in smart hospital room,” *Telkomnika (Telecommunication Comput. Electron. Control.)*, vol. 17, no. 2, pp. 965–972, 2019.

[4] P. Santosa, *Jajanan Tradisional Indonesia Asli*, no. February. 2019.

[5] M. Xin and Y. Wang, “Research on image classification model based on deep convolution neural network,” *Eurasip J. Image Video Process.*, vol. 2019, no. 1, 2019.

[6] S. Albawi, T. A. M. Mohammed, and S. Alzawi, “Understanding of a Convolutional Neural Network,” *Ieee*, 2017.

[7] F. F. Ting, Y. J. Tan, and K. S. Sim, “Convolutional neural network improvement for breast cancer classification,” *Expert Syst. Appl.*, vol. 120, pp. 103–115, 2019.

[8] X. Zhang, C. Xv, M. Shen, X. He, and W. Du, “Survey of Convolutional Neural Network,” vol. 147, no. Ncce, pp. 93–97, 2018.

[9] R. P. Prasetya and F. A. Bachtiar, “Indonesian food items labeling for tourism information using Convolution Neural Network,” *Proc. - 2017 Int. Conf. Sustain. Inf. Eng. Technol. SIET 2017*, vol. 2018-Janua, pp. 327–331, 2018.

[10] F. Zhu *et al.*, “Towards Unified INT8 Training for Convolutional Neural Network,” pp. 1966–1976, 2020.

[11] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, “Computer vision algorithms and hardware implementations: A survey,” *Integration*, vol. 69, no. July, pp. 309–320, 2019.