Recognition of Indonesian Traditional Cakes using The MobileNet Algorithm

Faruq Aziz
Faculty of Information Technology
Nusa Mandiri University, Indonesia
Email: Faruq.fqs [AT] nusamandiri.ac.id

Abstract— **Indonesia is a country with a variety of cultures, ranging from dance to cuisine and food variations. Cake is one of the unique variations of food include traditional cake. A variety of custom-made cakes will make the taste special, even though the name is the same. Traditional cakes are foods that are part of the ancestral culture that has been passed down from generation to generation explicitly in the region or Indonesian society. Machine learning methods are suitable for consistent and clear object recognition, this requires complex image pre-processing and feature extraction methods. The proposed model of our research is MobileNetv2 which was customized and then we did fine tuning then all of our training datasets do data-augmentation to create new datasets with various patterns so that the train dataset can be more numerous and avoid overfitting and the model can detect cake differences with an accuracy rate of 94% and loss 0.06.**

Keywords-component; CNN, MobileNet Fine-Tuned, Traditional Cake

I. Introduction

Indonesia is a country with a variety of cultures, ranging from dance to cuisine and food variations. Cake is one of the unique variations of food. A variety of custom-made cakes will make the taste special, even though the name is the same [1]. Although for humans, to recognize an object in an image is very easy, but in the future a system in the field of computer vision is needed for object identification and this requires special techniques [2].

The development of the times, humans are also increasingly concerned with the environment and their own bodies, especially assisted by computer vision systems, the availability of applications such as individual diet tracking systems [3], diagnose disease system [4], food quality detection system [5]. As one of the areas with the fourth most population [6] and diverse cultures, including culinary foods many foods including cakes are characteristic of Indonesia.

Traditional cakes are foods that are part of the ancestral culture that has been passed down from generation to generation explicitly in the region or Indonesian society [7]. Traditional cakes have a characteristic in the form of recipe books that are made from generation to generation. From generation to generation, his special methods are both field tools and conventional basic tools. Traditional cakes are included in conventional cuisine, can also include traditional cakes. It can be grouped according to its type. Types of traditional cakes can be seen in terms of making, taste and shape [7].

Indonesia is quite famous for its traditional food which is popular both domestically and abroad. A number of cakes become one of the favorite traditional foods. There are several types of cakes that can be processed in Indonesia, such as Lumpur Cakes, Klepon Cakes, Putri Salju Cakes, Kastengel Cakes, Dadar Gulung Cakes, Risoles Cakes, Lapis Cakes, Serabi Cakes. Most of the existing types of cakes are visually easily recognizable by humans, but computer vision requires special techniques in identifying image objects to types of cakes. Therefore, to recognize objects in the form of images of cakes as one of the traditional Indonesian foods, the Convolutional Neural Network (CNN) – Transfer Learning algorithm technique can be used [3].

Machine learning methods are suitable for consistent and clear object recognition, this requires complex image pre-processing and feature extraction methods [8]. In recent years CNN and Deep Learning algorithms have made tremendous breakthroughs in pattern recognition and classification, object identification, visual localization, and segmentation. Because Deep Learning and CNN-based models can extract features, find hidden structures in unlabeled images, and disorganized data. Deep learning based models do not require human involvement to extract a set of features from the input image and do not perform complex pre-processing on the image. They are inspired by CNN breakthroughs and Deep learning algorithms in object detection, image classification, deep learning research and applications, and CNN in plant disease detection which is currently the center of research [9].

Similar research have been done before with various methods, such as from Tita Karlita, et al, who used MobileNetv2 method with 8 image labels to classify cake types and produced 90% accuracy [10], then there was also from Zaenal Abidin, et al using the method CNN - MobileNetv2 with 1545 images for 8 labels and produces an accuracy of 92.5% [11].

Our research differs from previous research by contributing:

1. The proposed model is MobileNetv2 which was customized and then we did fine tuning.

2. All of our training datasets do data-augmentation to create new datasets with various patterns so that the train dataset can be more numerous and avoid overfitting.

3. The proposed model can detect cake differences with an accuracy rate of 94%.

II. MATERIAL AND METHOD

A. Data Used

The traditional cake image dataset used in this research was taken from the public dataset - Kaggle [12] with a total original dataset of 1664 images where 1392 for training and 272 for validation. Images of Traditional Cake are shown at figure 1.

Figure 1. Example of a figure caption. (figure caption)

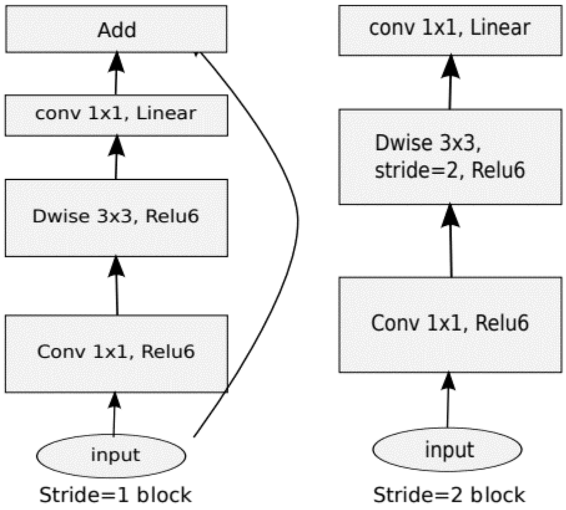
There are 8 labels processed at this research which is Lumpur Cakes, Klepon Cakes, Putri Salju Cakes, Kastengel Cakes, Dadar Gulung Cakes, Risoles Cakes, Lapis Cakes, Serabi Cakes. Which is each label has 174 images dataset.

B. Method

This research proposed the Convolutional Neural Network (CNN) – Transfer Learning is MobileNetv2 method. CNN works by utilizing the convolution process by moving a convolution kernel (filter) of a certain size to an image, the computer gets new representative information from the result of multiplying that part of the image with the filter used [13].

Transfer Learning is a concept so that we can use the existing CNN model for our dataset, generally researchers will use Transfer Learning from the ILSVRC [14] champion. In this study, the method used is MobileNetv2, which has the advantage of being light and fast in computing data and able to be implemented in simpler terms such as the use of mobile devices as the name suggests [15] [16]. The intermediate expansion layer uses lightweight depthwise convolutions to filter features as a source of non-linearity. As a whole, the architecture of MobileNetV2 contains the initial fully convolution layer with 32 filters, followed by 19 residual bottleneck layers [17].

The output of the base model is fed into a pooling layer to get the average value and convert it to a mean or vector. Then the dropout on the fully connected layer is 0.5 to avoid overfitting [10]. There are two activation functions in the hidden layer, namely ReLU and Softmax, where a dropout of 0.5 is given between the two layers. The activation function is used to determine whether a neuron should be active or not based on its threshold number.



Source: [17]

Figure 2. MobileNetv2 Architecture

The output that will be generated in the classification process will have one of eight traditional cake labels, namely: lumpur_cakes, klepon_cakes, putri_salju_cakes, kastengel cakes, dadar_gulung_cakes, risoles_cakes, lapis_cakes, serabi_cakes, along with the accuracy value obtained.

III. RESULT AND DISCUSSION

All experiments implemented using Google Collaboratory use their standard GPU [15]. In this experiment the researcher made a comparison with several other models, namely: VGG16 [18], VGG19 [19], EfficientNet [20].

The VGG16 model is usually implemented for image classification, the ILSVRC runner-up model is a model with 16 layers and uses ReLu as its activation function, but this VGG16 requires high computational power and takes quite a long time [20], with this model we get The results of the validation accuracy are 80.11% with a time of 200 minutes later after being fine-tuned to get the results of 85.54% with a time of 180 minutes. Then we also use a new VGG19 variation model from VGG16, with the same parameters we get 81.20% results with 200 minutes and fine-tuned and produces 85%, 150 minutes with better results than before tuning but not better than VGG16 in accuracy.

Then we also tried to do computations with EfficientNet and got 91% accuracy in 120 minutes, after tuning we got 95.27% with +/- 110 minutes. Meanwhile, the model that we propose is MobileNetv2, which results in an accuracy of 90% without tuning in 90 minutes and then after fine-tuning the result becomes 94.09% with a time of +/- 60 minutes. In terms of accuracy, EfficientNet with Fine-Tuned is superior, but in terms of time and accuracy, MobileNetv2 Fine-Tuned is better, as shown in table below.

TABLE I. COMPARISON CNN – TRANSFE LEARNING

Model Name	Accuracy	Time
VGG16	80.11%	200 min
VGG19	81.20%	200 min
EfficientNet	91%	120 min
MobileNetv2	**90%**	**90 min**

TABLE II. MODEL WITH FINE-TUNED

Model Name	Accuracy	Time
MobileNetv2	94.09%	60 min
VGG16	85.54%	180 min
VGG19	85%	150 min
EfficientNet	95.27%	110 min
MobileNetv2	**94.09%**	**60 min**

IV. CONCLUSION

The CNN method with Fine-Tuning has proven to be successful in extracting the image of Indonesian Traditional Cakes and identifying the different types of cakes. This can be proven by the results of the classification accuracy which is increasing and quite good compared to the previous method. In this study, we use MobileNetV2 which is fine-tuned as the method used. The proposed model obtains a stable accuracy value of 0.94% and a loss of 0.06.

In future research, it is hoped that newer research methods such as Faster R-CNN or YOLO can be used. It is also recommended to add more other types of traditional Indonesian cakes to the dataset.

REFERENCES

[1] S. Wijaya, "Indonesian food culture mapping: A starter contribution to promote Indonesian culinary tourism," *J. Ethn. Foods*, vol. 6, no. 1, pp. 1–10, 2019, doi: 10.1186/s42779-019-0009-3.

[2] D. J. Attokaren, I. G. Fernandes, A. Sriram, 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, doi: 10.1109/TENCON.2017.8228338.

[3] R. A. Rahmat and S. B. Kutty, "Malaysian food recognition using alexnet CNN and transfer learning," *ISCAIE 2021 - IEEE 11th Symp. Comput. Appl. Ind. Electron.*, pp. 59–64, 2021, doi: 10.1109/ISCAIE51753.2021.9431833.

[4] F. Aziz, D. Riana, J. Dwi Mulyanto, D. Nurrahman, and M. Tabrani, "Usability Evaluation of the Website Services Using the WEBUSE Method (A Case Study: covid19.go.id)," *J. Phys. Conf. Ser.*, p. 12103, 2020, doi: 10.1088/1742-6596/1641/1/012103.

[5] K. Kiratiratanapruk, P. Temniranrat, A. Kitvimonrat, W. Sinthupinyo, and S. Patarapuwadol, "Using Deep Learning Techniques to Detect Rice Diseases from Images of Rice Fields," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12144 LNAI, pp. 225–237, Sep. 2020, doi: 10.1007/978-3-030-55789-8_20.

[6] C. M. Annur, "Indonesia Peringkat ke-4 Negara Berpenduduk Terbanyak Dunia | Databoks," *katadata.co.id*, 2020. https://databoks.katadata.co.id/datapublish/2020/12/15/indonesia-peringkat-ke-4-negara-berpenduduk-terbanyak-dunia (accessed Mar. 17, 2022).

[7] A. Wijaya, A. Rahmadi, E. Harmayani, and G. Djarkasi, *The Uniqueness of ASEAN Food*. GSS Djarkasi, 2021.

[8] T. Gressling, *84 Automated machine learning*. 2020.

[9] S. S. Venkatesh, Nagaraju Y, Siddhanth U Hegde, "Fine-tuned MobileNet Classifier for Classification of Strawberry and Cherry Fruit Types," *2021 Int. Conf. Comput. Commun. Informatics*, vol. 1, no. 1, 2021.

[10] T. Karlita and I. Prasetyaningrum, "Indonesian Traditional Cake Classification Using Convolutional Neural Networks," *iCAST-SS 2021*, vol. 647, pp. 924–929, 2022.

[11] Z. Abidin and R. Borman, "Classification of Indonesian Traditional Snacks Based on Image Using Convolutional Neural Network (CNN) Algorithm," *ieeexplore.ieee.org*, vol. 1, no. 1, 2021, Accessed: Mar. 17, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9649707/.

[12] L. Quaranta, F. Calefato, and F. Lanubile, "KGTorrent: A dataset of python jupyter notebooks from kaggle," *Proc. - 2021 IEEE/ACM 18th Int. Conf. Min. Softw. Repos. MSR 2021*, pp. 550–554, 2021, doi: 10.1109/MSR52588.2021.00072.

[13] M. Hussain, J. J. Bird, and D. R. Faria, "A study on CNN transfer learning for image classification," *Adv. Intell. Syst. Comput.*, vol. 840, pp. 191–202, 2019, doi: 10.1007/978-3-319-97982-3_16.

[14] ImageNet.org, "ILSVRC Competition," *ILSVRC*, 2022. https://image-net.org/challenges/LSVRC/ (accessed Mar. 21, 2022).

[15] P. G. J. and N. K. V., "Google Colaboratory : Tool for Deep Learning and Machine Learning Applications," *Indian J. Comput. Sci.*, vol. 6, no. 3–4, pp. 23–26, Aug. 2021, doi: 10.17010/IJCS/2021/V6/I3-4/165408.

[16] Google Inc., "Overfit and Underfit at Deep Learning," *Google Colaboratory*, 2018. https://colab.research.google.com/github/csahat/docs/blob/keras-overfit-trans-id/site/id/tutorials/keras/overfit_and_underfit.ipynb (accessed Jan. 18, 2022).

[17] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.

[18] M. U. Hasan, "VGG16 - Convolutional Network for Classification and Detection," *neurohive.io*, 2018. https://neurohive.io/en/popular-networks/vgg16/ (accessed Jan. 14, 2022).

[19] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Deep

feature based rice leaf disease identification using support vector machine," *Comput. Electron. Agric.*, vol. 175, p. 105527, Aug. 2020, doi: 10.1016/J.COMPAG.2020.105527.

[20] C. Lee, H. J. Kim, and K. W. Oh, "Comparison of faster R-CNN models for object detection," *Int. Conf. Control. Autom. Syst.*, vol. 0, no. Iccas, pp. 107–110, 2016, doi: 10.1109/ICCAS.2016.7832305.