Large Scale Image Classification of Exotic Fruits in Indonesia Using Transfer Learning Method with MobileNet Model

Asyora Dewi Prabandani[1], Novanto Yudistira[1](✉), and Ayu Raisa Khairun Nisa[2]

[1] Informatics Engineering, Faculty of Computer Science, Brawijaya University, Malang, Indonesia
yudistira@ub.ac.id

[2] Aibi Store, Jakarta, Indonesia

Abstract. Exotic fruit is a fruit that is not widely known to the public. In Indonesia, there are many exotic fruits such as rambutan, passion fruit, mangosteen, longan, guava, and many more. Classification of exotic fruit images is needed because of the lack of knowledge from outsiders about exotic fruits in Indonesia. To his end, developing robust artificial intelligence using deep learning is necessary. CNN is the development of the Multilayer Perceptron (MLP) which is designed to process two-dimensional data and in the type of Deep Neural Network because of the high network depth and widely applied to image data. By utilizing the transfer learning method and a little fine-tuning, the efficient model like MobileNet expected to be better than without transfer learning in FruitNet model. Our contribution is applying efficient transfer learning MobileNet for Exotix Fruits in Indonesia which achieves 87% accuracy in average using more than 1000 images. The model performs better than previous model of FruitNet which only reaches 43% accuracy in average.

Keywords: CNN · exotic fruit · image classification · transfer learning

1 Introduction

Image classification of exotic fruits in Indonesia is required because most of outsiders even Indonesians themselves has little knowledge about exotic fruits Convolutional Neural Network (CNN) is development of Multilayer Perceptron (MLP) which is designed to process two-dimensional data. CNN belongs to Deep Neural Network because of its depth network, this method also applied widely in image data program. To get a good model, it takes quite a lot of image data. This will be difficult to produce if the image object is rare. For this reason, a transfer learning method was developed to maximize the performance of the deep learning model. By utilizing the transfer learning method and a little fine-tuning, the result of classification model will be better than without transfer learning method.

On the other hand, CNN is basically a deep learning model that requires a lot of data in the process to recognize image patterns. Training a CNN model from scratch cannot

Y. A. Yusran et al. (Eds.): BIC 2022, AEBMR 235, pp. 675–685, 2023.
https://doi.org/10.2991/978-94-6463-140-1_68

bring out the full potential of CNN itself [1]. Research by Sari et al. (2020) using data augmentation methods in CNN model to increase the variety and quantity of datasets but it was not enough to improve model performance [2]. Thus, the transfer learning method can be a solution to overcome the shortage of training data and improve CNN performance by utilizing the source domain to improve model performance in the target domain. CNN model using transfer learning methods has been carried out in several studies, such as image classification on human facial expressions [3], COVID-19 x-ray images [4] and food [5]. To this end, our research objective is development of deep learning system with MobileNet model using transfer learning to recognize exotic fruits images accurately. To achieve this, step by step process that we propose is:

1. Preparing large dataset of exotic fruits obtained from internet as training and testing data
2. Preparing MobileNet model which has been trained on large ImageNet dataset
3. Fine-tuning transfer learning MobileNet (training) to exotic fruits dataset using standard SGD optimizer and learning rate of 0.001
4. Performing evaluation of trained MobileNet on testing set of exotic fruits dataset.

Trained MobileNet on exotic fruits dataset is possibly applied on mobile software application due to its efficient model does not require high hardware specification like memory and storage. Moreover, this research helps in classifying Indonesian exotic fruits using a more efficient model. Increasing efficiency in disseminating information related to exotic fruit types It can also be developed into automatic system cashier in fruit market such as supermarket or traditional market. Assist in the needs related to data processing.

2 Transfer Learning

Transfer learning is a training model that utilizes pretrained models that train new data to produce model with better performance [6]. This method works by using the knowledge from deep learning models. This model known as problem solver and then this knowledge will be used to solve others new problems. This study uses MobileNetV2 which has been trained using the ImageNet dataset. ImageNet is a dataset with the highest number of images, around 14 million images with 1000 classes.

2.1 MobileNetV2

MobileNetV2 [7] is a CNN architecture that works well on mobile with concern on reduction the required memory. It is based on an inverted residue structure where the residue connections are between the bottle neck layers. The middle expansion layer uses convolution to filter features as non-linear source (Fig. 1). MobileNetV2 utilizes knowledge transfer for adaptation in environment with limited resources in order to solve the problems such as classification, detection and segmentation. MobileNetV2 is a development model of MobileNetV1. In the image classification experiment, MobileNetV2 showed a

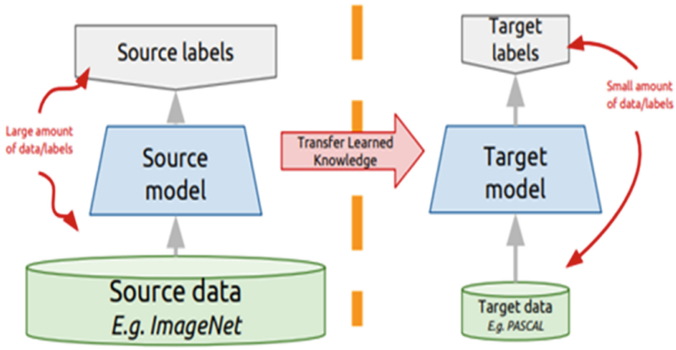


Fig. 1. MobileNetV2

better mean Average Precision (mAP) than MobileNetV1 [8]. The MobileNetV2 model uses Depthwise convolution and inverted block residuals in its architecture.

Depthwise convolution is a specific form of group convolution where the input feature H × W × N divided into N groups. H × W represents the input feature size and N represents the number of input channels with convolution kernel size is K × K. Therefore, the computational cost and the number of depthwise convolution parameters are HWNK2 and K2N. Depthwise convolution is equivalent with collecting the spatial features of each channel separately. It significantly reduces computational costs and number of parameters. However, this convolution method will produce poor information and the ouput data unrelated with input channel.

The reverse block residue in MobileNetV2 is an extension of the residual block in the ResNet model. First, the channel dimensions in the inverted residual block are expanded and then reduced in the opposite manner from the ResNet model. The feature extraction process with high input dimensions can improve the expressive ability of the model. Thus, the inverted residual block can convert the input data into high dimensionality and extract features deeply with convolutions.

2.2 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is a type of neural network used in image data. CNN widely used for image detection and recognition. In general, CNN is no different from the regular neural network. CNN are composed of with their own weights, biases and activation functions. A convolution layer also consists of neurons form a filter that has length and height (pixels). The CNN model is successfully used in various digital image and computer vision operations such as object detection, pattern recognition and image recognition. CNN is a deep learning method that achieved great success on image classification and object recognition. One of the problems that CNN can solve is recognize spatial feature. In addition, an important aspect of CNN is the features become more abstract as the input go deeper into the network.

2.3 Crawling

Crawling is a process in search engine (like Google) can search and scan content on a site such as images, articles, products and many more. A web crawler is a system for the bulk downloading web pages, usually used for web archiving, web data mining and web monitoring [9]. By using this technique, image data is collected based on the 10 classes of fruits.

2.4 FruitNet

FruitNet is a classification model that consists of 12 simple layers and was developed specifically for fruit classification. This model was developed by Siddiqi [10] and showed result that more efficient than MobileNet model, although it has a lower validation score.

3 Method

3.1 Data Collection

Data collection is procedure of collecting information and other facts in the field. Collecting data process is depends on the type of research. In this study, data collection was carried out using image crawling techniques on several search engines such as Google image, Baidu image and Bing image. Crawling is a process where search engines like Google can search and scan content on a site such as images, articles, products and more. A system for downloading the web in bulk which is often used for web archiving, data mining, and web monitoring [9]. A total of 100–200 images can be collected in each class (durian, guava, longan, mangosteen, passion fruit, jackfruit, papaya, rambutan, snake fruit and soursop) with total the images collected are more than 1000 images. The example of each fruit is shown in Table 1.

3.2 Research Instrument

We run our training on Google Collaboratory (Google Colab) which is cloud based editor where python and pytorch based programming are written and run. The specification of Google Colab server is Nvidia K80/T4 GPU, 12 GB/16 GB GPU Memory, 0.82 GHz/1.59 GHz GPU Memory Clock, and 4.1 TFLOPS/8.1 TFLOPS Performance.

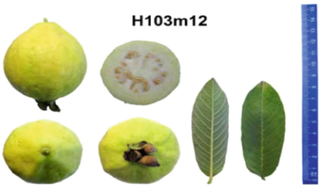
3.3 Pre-processing

Pre-processing data is carried out to prepare for main processing and further analysis. This term can be applied to first processing if there are several steps on process of data prepare for user. According to Alasadi, S. A. & Bhaya, W. S. [21], pre-processing consists of several stages, i.e. integration, cleaning and reduction. First, data cleaning manually; second, transform images to size 224x224; and the last is divide the data into training data, validation data and testing data, as for the division ratio is 8:1:1.

3.4 Architectural Model

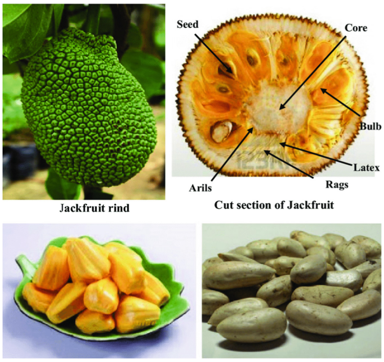
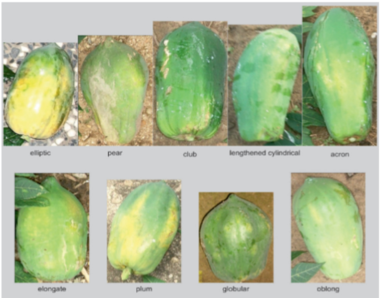
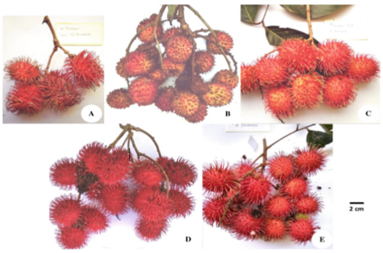
The architectural model built consists of 2 types. The MobileNetV2 model as a model using transfer learning and the FruitNet model by Siddiqi [10] as a model without transfer learning.

Table 1. The example of image fruit

No	Image	Description
1		Durian [11]
2		Guava [12]
3		Longan [13]
4		Mangosteen [14]
5		Passion fruit [15]

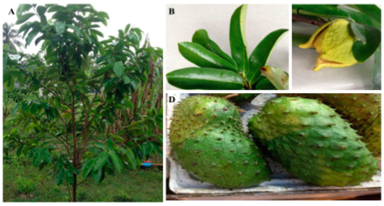
(*continued*)

Table 1. (*continued*)

No	Image	Description
6		Jackfruit [16]
7		Papaya [17]
8		Rambutan [18]
9		Snakefruit [19]

(*continued*)

Table 1. (*continued*)

No	Image	Description
10		Soursop [20]

3.5 Fine-Tuning

Fine Tuning is a technique to retrain the model in the process of learning transfer process. In general, fine-tuning is carried out only on fully connected layers by changing the neurons output according to classes in the dataset. In this study, fine-tuning is performed on all layers. The fine tuning process is conducted with 14 epochs for MobileNetV2 with pre-training and 24 epochs for FruitNet without pre-training depending on the convergence point, Stochastic Gradient Descent (SGD) and learning rate of 0.001. The low batch size aims to save memory usage during the fine-tuning process. The loss function used is cross-entropy which considers neuron output with soft max normalization and distance to label.

4 Result and Discussion

4.1 Crawling

In crawling image process used icrawler library. Icrawler is one of open source libraries available for python. This library only requires keywords to crawl the desired image and max_num (maximum number of images taken) (Table 2).

The results showed that the number of images obtained in each class ranging from 200–300 images. However, from these results, some of images are not relevant with fruit image, for example durian ice, mangosteen peel extract, etc.).

4.2 Training Using Transfer Learning (Fine-Tuning) on MobileNetV2

The model was fine-tuned for 15 epochs. Figure 2 shows the results where the training time is about 28 min.

4.3 Training Without Transfer Learning on FruitNet

Number of models trained is same with model using the transfer learning method, which is 15 epochs. Figure 3 shows the result where the training time is about 43 min. The longer training time with the same epoch range shows that FruitNet has more parameters than MobileNetV2.

Table 2. icrawler library

```
from icrawler.builtin import
BaiduImageCrawler, BingImageCrawler,
GoogleImageCrawler

buah_buahan = ('rambutan', 'durian',
'markisa', 'nangka', 'manggis',
'pepaya', 'kelengkeng', 'salak', 'jambu
air', 'sirsak')

for buah in buah_buahan:
    path = ('F:\Kuliah\Semester
7\PembelajaranMesinLanjut\Image\\'+buah)
google_crawler =
GoogleImageCrawler(storage={'root_dir':
path})
google_crawler.crawl(keyword=buah,
max_num=10)
```

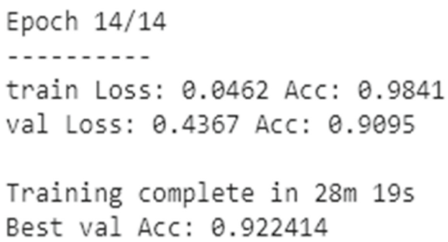


Fig. 2. Training result using transfer learning

4.4 Comparison Training Result Using Transfer Learning on MobileNetV2 and Without Transfer Learning on FruitNet

Comparison between MobilNetV2 and FruitNet are shown by Table 3. The training results show a significant difference between MobileNetV2 with transfer learning and FruitNet without transfer learning. Even though, the parameters on the FruitNet architecture (without transfer learning) are more than MobileNetV2.

```
1 fruitNet = train_model_new(fruitNet, criterion, optimizer, num_epochs=25)

Epoch 22/24
----------
train Loss: 0.0967 Acc: 0.9708
val Loss: 3.9487 Acc: 0.4397

Epoch 23/24
----------
train Loss: 0.0707 Acc: 0.9793
val Loss: 3.6256 Acc: 0.4181

Epoch 24/24
----------
train Loss: 0.0732 Acc: 0.9793
val Loss: 3.7251 Acc: 0.4741
Copying best weight ...

Training complete in 43m 48s
Best val Acc: 0.474138
```

Fig. 3. Training result without using transfer learning

Table 3. Comparison training result on MobileNetV2 and FruitNet model

No	**Class**	**Accuration**	
		MobileNetV2	**FruitNet**
1	Durian	**96.2%**	69.2%
2	Guava	**86.2%**	48.3%
3	Longan	**91.7%**	33.3%
4	Mangosteen	**93.3%**	68.9%
5	Passion fruit	**85.0%**	35.0%
6	Jackfruit	**73.3%**	26.7%
7	Papaya	**68.2%**	50.0%
8	Rambutan	**90.9%**	31.8%
9	Snakefruit	**94.4%**	38.9%
10	Soursop	**90.9%**	27.7%

5 Conclusion

The use of transfer learning is proven to produce a better model with fewer parameters and shorter training time. The lack of available data affects to low accuracy of FruitNet, this can happen because it does not utilize transfer learning method. This result is different with the pre-trained model on MobileNetV2 which has been trained with large amounts of data, so it only needs to do fine-tuning on datasets with more specific domains such as exotic fruits in Indonesia. To conclude, the low accuracy in FruitNet is due to the lack of available data. Unlike the pre-trained model on MobileNetV2 which has been trained with thousands of images beforehand. Hence the classification of Indonesian exotic fruits can be done more efficiently.

References

1. J. Park, R. J. Javier, T. Moon and Y. Kim, "Micro-Doppler Based Classification of Human Aquatic Activities via Transfer Learning of Convolutional Neural Networks," Sensors, no. 16, pp. 1–10, 2016.
2. Y. Li, K. He, D. Xu and D. Luo, "A transfer learning method using speech data as the source domain for micro-Doppler classification tasks," Knowledge-Based Systems, no. 209, 2020.
3. Rahadika, Fadhil Yusuf, et al. "Facial Expression Recognition using Residual Convnet with Image Augmentations." JurnalIlmuKomputer dan Informasi (Journal of Computer Science and Information) 14.2 (2021).
4. Yudistira, Novanto, Agus Wahyu Widodo, and Bayu Rahayudi. "Deteksi Covid-19 pada citra sinar-x dada menggunakan deep learning yang efisien." Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK) 7.6 (2020).
5. Thiodorus, Gustavo, Yuita Arum Sari, and Novanto Yudistira. "Convolutional Neural Network with Transfer Learning for Classification of Food Types in Tray Box Images." 6th International Conference on Sustainable Information Engineering and Technology 2021. 2021.
6. Hao Yang, Minghan Wang, Ning Xie, Ying Qin, and Yao Deng. 2020. Efficient Transfer Learning for Quality Estimation with Bottleneck Adapter Layer. In Proceedings of the 22nd Annual Conference of the European Association for Machine Translation, pages 29–34, Lisboa, Portugal. European Association for Machine Translation.
7. Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
8. Supriadi, Muhammad Fadhlan., Rachmawati, Ema., Arfianto, Anditya. Pembangunan Aplikasi Mobile Pengenalan Objek Untuk Pendidikan Anak Usia Dini. Jurnal Teknologi Informasi dan Ilmu Komputer, [S.l.], v. 8, n. 2, p. 357–364, mar. 2021. ISSN 2528–6579. Tersedia pada: https://jtiik.ub.ac.id/index.php/jtiik/article/view/4363. TanggalAkses: 27 des. 2021 https://doi.org/10.25126/jtiik.2021824363.
9. Olston, Cristopher., dan Najork, Mark. Web Crawling. Foundations and Trends R in Information Retrieval, v. 4, n. 3, p. 175–246. Tersedia pada: https://www.researchgate.net/publication/225844302_Web_Crawling. TanggalAkses: 27 des. 2021.
10. Siddiqi, R. (2020). Comparative Performance of Various Deep Learning based Models in Fruit Image Classification. Proceedings of the 11th International Conference on Advances in Information Technology.
11. Lestari, S., Fitmawati, F. & Wahibah, N.N. (2011) Keanekaragaman durian (Durio zibethinus Murr.) di Pulau Bengkalis berdasarkan karakter morfologi. Buletin Kebun Raya, 14 (2), 2944.
12. LL Valera-Montero, PJ Muñoz-Rodríguez, H Silos-Espino, S FloresBenítez. Genetic diversity of guava (Psidium guajava L.) from Central Mexico revealed by morphological and RAPD markers. Phyton, International Journal of Experimental Botany. 2016; 85:176– 83.
13. Crane, J. H., C.F. Balerdi, S. A. Sargent, and I. Maguire. 2016. Longan growing in the Florida home landscape. FC49. Gainesville: University of Florida Institute of Food and Agricultural Sciences. http://edis.ifas.ufl.edu/mg049

14. Nazre, M.. (2014). New evidence on the origin of mangosteen (Garcinia mangostana L.) based on morphology and ITS sequence. Genetic Resources and Crop Evolution. 61. https://doi.org/10.1007/s10722-014-0097-2.
15. Iif Hanifa Nurrosyidah and Ni Made Mertaniasih and Isnaeni (2020) Inhibitory activity of fermentation filtrate of red passion fruit pulp (Passiflora edulis sims.) against Escherichia coli extended-spectrum beta-lactamase (ESBL) and methicillin resistant Staphylococcus aureus (MRSA). Berkala Penelitian Hayati, 26 (1). pp. 22–25. ISSN 0852–6834
16. Swami, Shrikant & Kalse, Sandeep. (2018). Jackfruit (Artocarpus heterophyllus): Biodiversity, Nutritional Contents, and Health. https://doi.org/10.1007/978-3-319-54528-8_87-1.
17. Aikpokpodion, P.O. (2012). Assessment of genetic diversity in horticultural and morphological traits among papaya (Carica papaya) accessions in Nigeria. Fruits, 67, 173–187.
18. Windarsih, G., & Efendi, M. (2019). Short communication: Morphological characteristics of flower and fruit in several rambutan (Nephelium lappaceum) cultivars in Serang City, Banten, Indonesia. Biodiversitas Journal of Biological Diversity.
19. Sukewijaya, I.M., Rai, I.N., & Mahendra, M.S. (2009). Development of Salak Bali as an organic fruit. Asian Journal of Food and Agro-Industry, 2.
20. Moghadamtousi, S.Z.; Fadaeinasab, M.; Nikzad, S.; Mohan, G.; Ali, H.M.; Kadir, H.A. Annona muricata (Annonaceae): A Review of Its Traditional Uses, Isolated Acetogenins and Biological Activities. Int. J. Mol. Sci. 2015, 16, 15625–15658. https://doi.org/10.3390/ijms160715625
21. Alasadi, Suad A., dan Bhaya, Wesam S. 2017. Review of Data Preprocessing Techniques in Data Mining. Journal of Engineering and Applied Sciences, 12: 4102–4107. Tersedia pada: https://medwelljournals.com/abstract/?doi=jeasci.2017.4102.4107. TanggalAkses: 27 des. 2021 https://doi.org/10.36478/jeasci.2017.4102.4107.