Study on Improving Model Accuracy by Using Non-medical Image Pretraining Set based on ImageNet Weight

Xiangwu Guo[1*]

[1]Software Engineering, South China University of Technology, Panyu District, Guangzhou, China
*Corresponding author: sexiangwu_guo@mail.scut.edu.cn

Abstract
Although modern medical systems have relatively complete skin cancer detection methods, it is still very difficult to detect early skin cancer themselves. Therefore, this paper intends to propose a lightweight model network for patient self-detection of skin cancer. In order to implement a lightweight model, this article chose MoblieNetV2 as part of the model construction. Faced with the inherent difficulty of insufficient in-depth training of lightweight models, this paper tries to use transfer learning to improve model accuracy. Since there is no pre-training set model for skin cancer detection, this paper boldly used ImageNet natural image and training set model to achieve the purpose of improving model accuracy and reducing loss. Through the control experiment set under the same conditions, it is found that the accuracy of the test set using Imagenet can reach 88%, which is very improved compared with the performance without Imagenet. The experimental results clearly show that it is feasible to use natural images and medical image detection. Therefore, this paper can try to apply more natural images to medical image detection to make up for the lack of medical image pre-training set.

Keywords: *MobileNetV2, Transfer Learning, Skin cancer, ImageNet*

1 INTRODUCTION

Skin cancer is one of the most widespread forms of cancer affecting human populations worldwide. Nearly 9, 500 individuals are diagnosed daily in the United States [7]. Nonmelanoma skin cancer is responsible for over one million new cases (excluding basal cell carcinoma) and 64, 000 deaths globally [14]. Nowadays, cancerous skin lesions or melanomas are diagnosed mainly by visual screening of lesions by dermatologists. Based on the most prominent method asymmetry, border, color, diameter and evolution (ABCDE), dermatologists can distinguish between a benign and a malign mole [9]. The diagnosis was then confirmed by dermoscopy, biopsy, and histopathological examination [6].

However, because early skin cancer lesions are so similar to benign skin lesions, it can be very difficult to distinguish them from other skin lesions based solely on visual experience. In addition, further tests to diagnose skin cancer can be time-consuming and costly. Mobile teledermoscopy is a kind of the technology that can solve two of the problems below [8].

Recently, Convolutional Neural Network (CNN) has been widely used in image-related processing task e.g. classification. CNN is great for helping dermatologists diagnose or self-diagnose. CNN can extract skin surface features very effectively, and then realize skin cancer detection through these features. Many people have tried to use VGG, ResNet and other different models to achieve skin detection. There is a paper showed that they used VGG-16 and VGG-19 to classify skin cancer with an accuracy of 65.67% and 68.54%, respectively [11]. It can be seen that the effect of using VGG is not very ideal. So there is a paper would like to employ this VGG model and improve its accuracy based on the current task. However, the model parameters of VGG are very large, which will lead to a long training time [5].

In this paper, a model that allows patients to do skin cancer self-testing at home was also tried to be implemented. In order to enable patients to self-test and help dermatologists to move tests, this paper chose MobileNet, a lightweight model [12], for training. In order to further shorten the training time and improve the accuracy, the method of transfer learning was applied to add the pre-training model on the basis of MoblieNet. But

B. Fox et al. (Eds.): IC-ICAIE 2022, AHCS 9, pp. 804-809, 2023.
https://doi.org/10.2991/978-94-6463-040-4_122

in the absence of a skin cancer-related pretraining model, the Moblienet's own pretraining model for natural objects was used. Finally, it is found that using the natural model can greatly improve the accuracy and help to reduce the training time.

2 METHOD

2.1 Dataset preparation

In this paper, the dataset is taken from the International Skin Image Collaboration (ISIC) Archive [1] which is help reduce melanoma mortality. The dataset is divided in two classes, namely benign and malignant cancer. For the training set, it consists of 1, 440 pictures of benign moles and 1, 197 pictures of malignant classified moles. For the test set, it consists of 360 pictures of benign moles and 300 pictures of malignant classified moles. The pictures have all been resized to low resolution (224×224×3) with RGB format. Figure 1 and Figure 2 indicate the data samples of benign and malignant in the collected dataset.

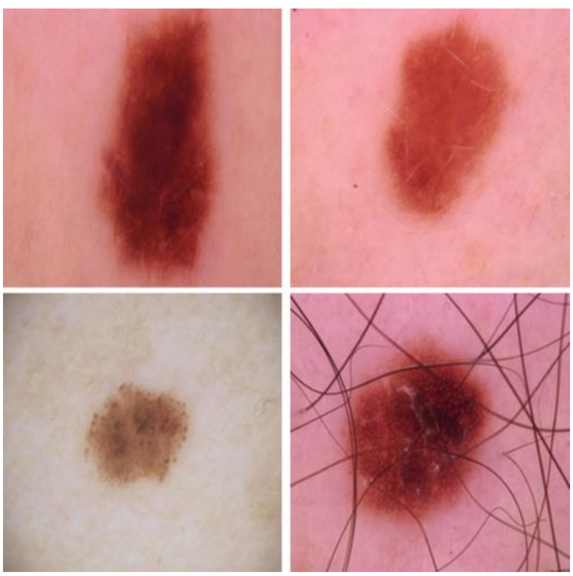

Figure 1: Data sample for the benign cancer.

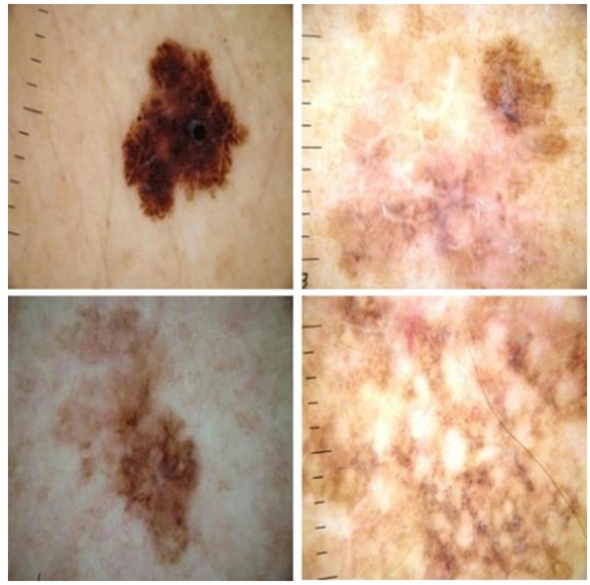

Figure 2: Data sample for the malignant sample.

All Benign images are labelled with "0", and all malignant classified images are labelled with "1". They are shuffled after grouped and then are converted to binary matrices. Finally, the datasets are normalized which effectively reduce the difference of the data set.

2.2 Proposed CNN model

Convolutional neural network (CNN) contains an input layer, an output layer and multiple hidden layers, where hidden layer consists of a convolutional layer, pooling layer, fully connected layer (FC) and various normalization layers [4] It plays an excellent role in image classification. However, limited by its own attributes, CNN still needs a lot of adjustments in the field of cancer detection in order to ensure higher accuracy and less time consumption.

A very important manifestation of transfer learning is Pre-train and Fine-tune which refers to training a network in the source domain, using it directly for data in the target domain, and fine-tuning the target domain data. In many real-world applications, it is more efficient and time saving to use transfer learning instead of recollecting the needed training data and rebuilding the models [13]. To reduce training costs of image classification or detection, this paper tries fine-tuning CNN models pre-trained from natural image dataset to medical image tasks.

So far there is no prominent database of medical images. ImageNet is an image database organized according to the WordNet hierarchy, in which each node of the hierarchy is depicted by hundreds and thousands of images [3]. ImageNet contains more than 20,000 categories of natural images and is the most popular image database available today. It is important to note that ImageNet provides natural images, not medical images. Therefore, the reason for some studies that directly use pretrained model based on ImageNet to predict medical images should be further explained [2] [10].

Objective, there are differences between natural and natural images. This paper wanted to make ImageNet suitable for skin cancer detection by tuning parameters and models. So that the purpose of improving the accuracy and making the model perform better can be achieved.

Table 1: The performance of trained models.

model / performance	MoblienetV2 (without imagenet)	MoblienetV2 (freeze the Moblienet layer)	MoblienetV2 (with imagenet)
Train accuracy	0.7947	0.8070	0.9929
Test accuracy	0.5360	0.7955	0.8826
Training time	About 12 min	About 3min30s	About 12 min

Table 2: The architecture of the employed mobilenetV2.

Layer (type)	Output Shape	Param #
Mobilenetv2_1.00_224 (Functional)	(None, 7, 7, 1280)	2257984
Dropout	(None, 7, 7, 1280)	0
Dense	(None, 7, 7, 256)	327936
GlobalAveragePooling2D	(None, 256)	0
Dropout	(None, 256)	0
Dense	(None, 2)	514

MobileNetV2 is considered in this study due to its excellent performance. MobileNetV2 is a general-purpose computer vision neural network designed for mobile devices. The architecture of it can be found in Figure 3 and Table 2. MobileNetV2 has few model parameters and short training time, which also ensures excellent model accuracy. The attributes of MobileNetV2 are very consistent with the actual working scenes of skin cancer detection in traditional hospitals, and can also be applied to skin cancer detection on mobile.

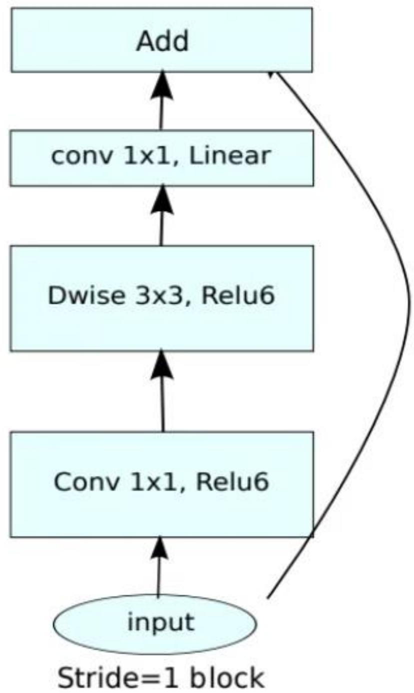

Figure 3: The sample image for the HAM10000 dataset.

MobilenetV2 is based on an inverted residual structure where shortcut connections are located between thin bottleneck layers. The intermediate extension layer uses lightweight deep convolution to filter features as nonlinear sources [12]. This makes the network structure small and accurate, making it particularly suitable for mobile applications. In order to reduce the risk of overfitting, this study choose to add dropout Layer and pooling layer to help training. Finally, the dense layer was used to complete the classification.

As mentioned above, through ImageNet transfer learning, CNN can only use data sets of limited size to learn models with better performance. Another reason for choosing MoblieNet for this paper is that it integrates ImageNet itself which also very effective. The pre-training set is obtained by training a large number of natural images. Using MoblieNet's pre-training set is very helpful for feature extraction of skin cancer detection.

This paper uses binary cross entropy as the loss function. The basic learning rate is 0.00005. When test set accuracy stopped increasing after five consecutive epochs, the learning rate dropped to 0.0000005.

3 RESULTS AND DISCUSSION

3.1 Performance for the Inception V3

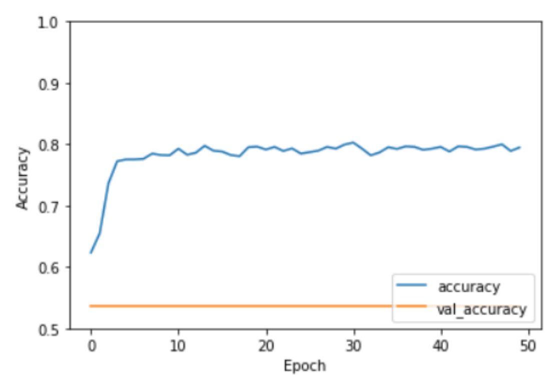

Figure 4: The performance for MoblienetV2(without imagenet).

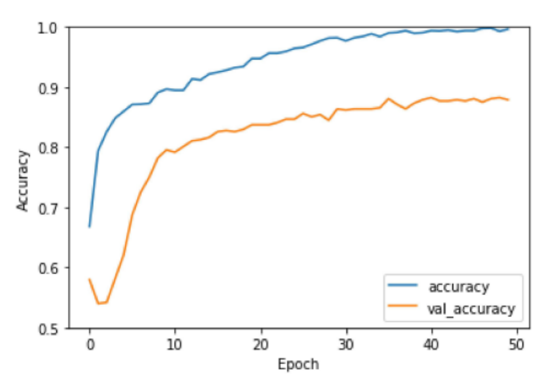

Figure 5: The performance for MoblienetV2(freeze the Moblienet layer).

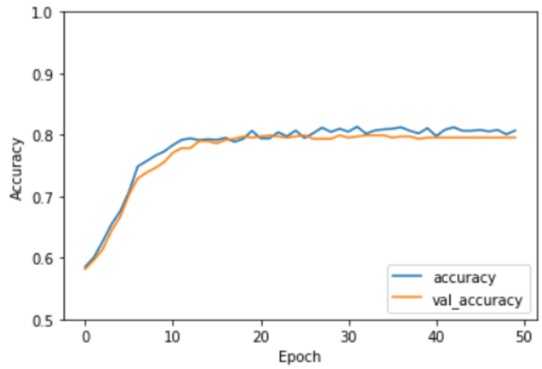

Figure 6: The performance for MoblienetV2(with imagenet).

In the course of the training, this paper used three different training methods to train on the skin cancer data set. To ensure the fairness of comparison, this paper ensures the consistency of the same learning rate, epoch quantity, batch number and other hyperparameters. According to the Table 2, the training duration with moblienetV2 and the pre-training model was about 12 minutes. MoblienetV2 with imagenet can obtain the best training accuracy, which is training accuracy 0.9929 and test accuracy 0.8826 respectively. In comparison, moblienetV2 model without imagenet's pre-training set only obtains the training set accuracy of 0.7947 and test set accuracy of 0.5360 respectively. Figure 4-6 indicate the training curve of three different models.

3.2 Discussion

It can be observed from the Table 1, MoblienetV2 that does not use transfer learning performs poorly. This may be because moblienetV2 itself is limited by a very small number of parameters and the length of training, and because of the high requirements for details in skin cancer detection, MoblienetV2 cannot complete training tasks well under limited conditions.

In addition, MobilenetV2 using pre-training sets has a great improvement in accuracy. The reason may be that the medical image and the natural image in imagenet have similar kernels convolved in the low-level layers. For example, in natural images, kernel will cover the feature extraction of shapes such as triangles and circles and boundaries. The pre-training set contains early feature extraction that can help in skin cancer detection since there are some similar features in this case. Besides, if the pre-training set layer is frozen, it can be seen that the accuracy is only about 0.8. This also demonstrates the incompatibility of natural images in medical detection.

4 CONCLUSIONS

The paper presents a lightweight model for skin cancer detection using MoblienetV2. In order to improve the target of skin cancer detection, through the exploration of transfer learning, this paper adopts ImageNet, a pre-training set of non-medical images, in a new way to help improve the accuracy of model training. Through the comparison experiment under the same conditions, it can be known that the preprocessing model of natural image can also be well applied to skin cancer image, and improve the accuracy of the model. In addition, the gap between natural image imageNet and medical image is also reflected in the comparison test. More complex adjustments to models and parameters are required. In the future, this study will further verify whether natural images can be used in skin cancer detection, such as feature visualization technology for further comparison and research.

REFERENCES

[1] Cancer dataset (2022). www.isic-archive.com.

[2] Chen, Sihong, Kai Ma, and Yefeng Zheng. "Med3d: Transfer learning for 3d medical image analysis." arXiv preprint arXiv:1904.00625 (2019).

[3] Deng, J. (2009). W., dong, r. socher, l. In J., Li,, K., Li,, and, L., Fei, Fei., Imagenet:, A large-scale, hierarchical, image, database., In, IEEE, Conference, on Computer, Vision, and Pattern, Recogni, tion,(CVPR),, pages (pp. 248-255).

[4] Dhillon, A., & Verma, G. K. (2020). Convolutional neural network: a review of models, methodologies and applications to object detection. Progress in Artificial Intelligence, 9(2), 85-112.

[5] Dutta, A., Gupta, A., & Zissermann, A. (2016). VGG image annotator (VIA). URL: http://www. robots. ox. ac. uk/vgg/software/via, 2.

[6] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. nature, 542(7639), 115-118.

[7] Guy Jr, G. P., Thomas, C. C., Thompson, T., Watson, M., Massetti, G. M., & Richardson, L. C. (2015). Vital signs: melanoma incidence and mortality trends and projections—United States, 1982–2030. MMWR. Morbidity and mortality weekly report, 64(21), 591.

[8] Massone, C., Brunasso, A. M., Campbell, T. M., & Soyer, H. P. (2009, September). Mobile teledermoscopy—melanoma diagnosis by one click?. In Seminars in cutaneous medicine and surgery (Vol. 28, No. 3, pp. 203-205). WB Saunders.

[9] Nachbar, F., Stolz, W., Merkle, T., Cognetta, A. B., Vogt, T., Landthaler, M., ... & Plewig, G. (1994). The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin

lesions. Journal of the American Academy of Dermatology, 30(4), 551-559.

[10] Qiu, Y., Chang, C. S., Yan, J. L., Ko, L., & Chang, T. S. (2019, October). Semantic segmentation of intracranial hemorrhages in head CT scans. In 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS) (pp. 112-115). IEEE.

[11] Rezaoana, N., Hossain, M. S., & Andersson, K. (2020, December). Detection and classification of skin cancer by using a parallel CNN model. In 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE) (pp. 380-386). IEEE.

[12] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4510-4520).

[13] S. J. Pan and Q. Yang, (2010) "A Survey on Transfer Learning," in IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, doi: 10.1109/TKDE.2009.191.

[14] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians, 71(3), 209-249.