

A Study of the Classification Method of Bone Marrow Blood Cells Based on MobileVit

Manman Wang^a, Keke Shang^{b*}, Haiming Zhang^c

^{a,b,c}Tiangong University, School of Physical Science and Technology, TianJin,300387,CHN

^aEmail: wangmanman0910@163.com

^bEmail: kkshang@163.com

^cEmail: zhmtjwl@163.com

Abstract

The automatic classification of bone marrow blood cells is of great significance to diagnosis of acute leukemia. We applied the MobileVit network model to morphological classification of bone marrow cells by using transformer based on self-attention mechanism as convolution for global representation and achieve a global representation with reduced calculation parameters. We found that in the bone marrow blood cell classification task, without the use of transfer learning, MobileVit classification achieved 97% accuracy, exceeding the classification accuracy of Resnet and Densenet networks, and reducing FLOPS and params significantly. Compared to other lightweight networks, the MobileVit model demonstrates improved accuracy and comparable FLOPS and params.

Keywords: classification of bone marrow blood cells; morphological diagnosis; Deep learning; MobileVit.

1. Introduction

Morphological classification of bone marrow blood cell is commonly used to diagnose acute leukemia. Clinically, the diagnosis of most hematological diseases still relies on manual microscopic examination by hematologists: manual classification and counting of different blood cells in the bone marrow, combined with other test results and clinical experience for comprehensive diagnosis. [1] This method heavily dependent on the hematologist's experience, is time-consuming and labor-intensive; instruments such as blood cell analyzers and flow cytometry cannot detect morphological abnormalities in blood cells. Therefore, it is of great significance to develop a computer-aided automatic detection system for morphological classification of bone marrow blood cell [2]. The initial work of computer-aided morphological diagnosis was tracked back to 1966, when Prewitt and Mendelsohn used a computer to classify five types of microscopic images of leukocytes in the bone marrow.[3] However, due to the lack of computing power, this work has been suspended for more than 40 years.

* Corresponding author.

In the last two decades, with the development of computer computing power and the development of image processing technology, work in this field has gradually regained the attention of researchers. The early research was mainly based on traditional Machine Learning methods, such as Support Vector Machines (SVM), Random Forest (RF), Fuzzy C-Means (FCM), Naive Bayes Classifiers (NBC), etc.[2]. The disadvantage of these methods is that they require a lot of human supervision and intervention, and need to manually extract features. Later, the rise of Deep Learning further promoted the development of the field.[4] Convolutional Neural Networks, as one of the most important networks in the field of Deep Learning, have been applied to the classification and recognition of bone marrow blood cells.[6] Applications based on Resnet, VGG16, AlexNet have achieved remarkable results[6].

In 2021 the application of VIT (Vision Transformer) based on the self-attention was introduced into visual image processing and soon became popular [9]. In order to combine the advantages of CNNs and VIT models, various VIT models that combine convolutional operations and transformer methods have attracted the interest of many researchers.

However, due to the large number of parameters and scale, it is difficult to run the VIT network on mobile devices. How to build a lightweight, low-latency and high-precision network model is of importance. A lightweight and multipurpose VIT model proposed in 2022, the MobileVIT achieved a Top-1 accuracy of 78.4% with about 6 million parameters on the ImageNet-1k dataset [8].

In our research, the lightweight network MobileVit is combined with the Imbalanced Dataset Sampler module and used in the identification and classification of bone marrow blood cells for the first time, and the final accuracy reaches 97%. Our results show that in the task of bone marrow blood cell classification, the MobileVit can reach 97% classification accuracy without using transfer learning, reaching or even exceeding the classification accuracy of Resnet and Densenet networks, greatly reducing FLOPS and Params. Compared to other lightweight networks, the accuracy is improved and FLOPS and Params are comparable.

2. materials and methods

2.1. MobileVit

The lightweight network model MobileVit used in this paper is a lightweight Vision Transformer that can be used for mobile devices. [8] It combines the advantages of CNN and Vision Transformer, and treats the Vision Transformer based on self-attention as a convolution to extract global feature information, which finally constitutes a low-latency, highly versatile and lightweight network [8].

The Figure 1 shows the network structure of MobileVit, where Conv represents the convolution layer, MV2 represents the MobileNetV2 module, MobileVit Block is the core part of the entire MobileVit, and L represents the number of Transformers in the MobileVit. The input image is fed into four MV2 modules after passing through the 3x3 convolution layers. When (H,W) is (32,32), it enters the MobileVit module, followed by the stacking of MV2 and MobileVit modules, and finally passes through the convolution layer, average pooling layer, and fully connected the layer output.

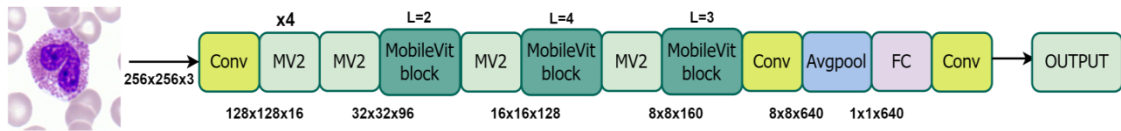


Figure 1: MobileVit structure diagram

2.2. MobileVit-Block

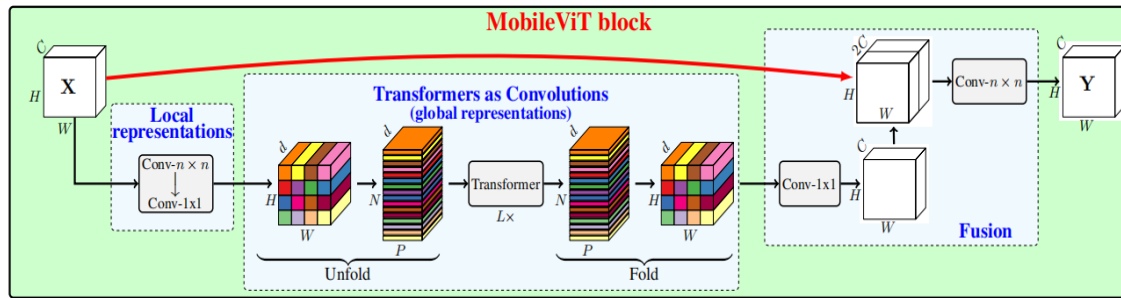


Figure 2: MobilVit-Block

The structure of MobileVit Block is shown in Figure 2. As the core part of the whole network, the overall structure of MobileVit Block can be divided into three parts. The first part is the local information representation module composed of convolution layers for extracting local features; the second part is the global representation module, which uses L Transformers as convolution for extracting global modules; the third part is the feature information fusion and the final output.

The core part of the MobileVit block is the second part (Global Representation). The standard convolution operations are Unfolding, Local Processing, and Fold. In order to have the properties of convolutional network and allow for global representation, the MobileVit Block replaces the Local Processing part with an L Transformer so that the local processing is replaced with a deeper global processing. This allows the global information of the image to be extracted even with small parameters.

2.3. MobileNetV2

Figure 3 demonstrates the structure of MobileNetV2 module. The module first expands the dimensionality of the input data to extract features and then compress the data, which can greatly reduce FLOPS, extract enough features with a small number of parameters and effectively reduce the computation time.[11] This can greatly reduce the amount of calculation, extract enough features even when the number of parameters is small, and effectively reduce the computational time.

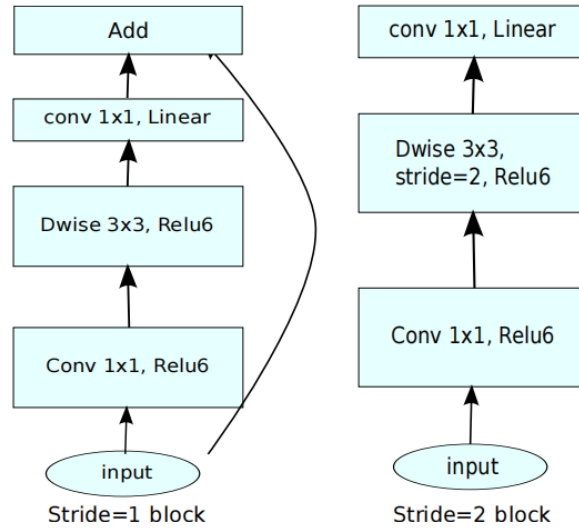


Figure 3: MobileNetV2 module

3. Experiment and analysis

3.1. Selection of datasets

We used the public dataset AML-Cytomorphology_LMU, which contains 15 categories of granulocytes, erythroblast, lymphocytes, etc.[12] Based on FAB criteria for the diagnosis of acute leukemia [13], we selected eight types of bone marrow blood cells that will help with diagnosis of acute leukemia, including primitive granulocytes, promyelocytes, mesophils, late larocytes, rod nucleophils, lobulated nucleocytes, monocytes, and other cells, a total of 6,033 samples. We randomly selected 90% of the samples as the training set and use the remaining 10% for model validation.



Figure 4: Microscopic images of some cells in the dataset

3.2. ImbalancedDatasetSampler module

When training models in deep learning, it is common to encounter situations where the dataset is not balanced. Unbalanced datasets will greatly affect the training effect, resulting in negative effects such as slow convergence in the training process and poor generalization ability. The problem also exists in the public dataset used in this

paper, in which there are three classes of cells with far more samples than others. In this case, it is necessary to ensure that the model we train will not be biased towards the class with a large number of samples.

ImbalancedDatasetSampler[9] can be a good solution to this problem by rebalancing class distributions and automatically estimating sampling weights when sampling in imbalanced datasets, and by combining data augmentation to reduce the risk of overfitting.

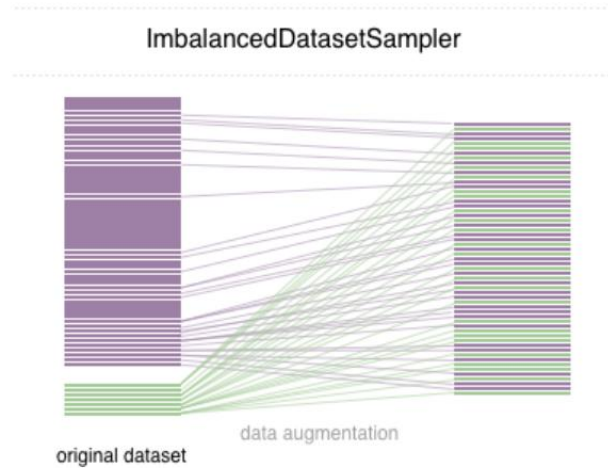


Figure 5: ImbalancedDatasetSampler schematic

3.3. Experimental environment

The operating system is Linux, the programming environment is python 3.7/pytorch 1.6, the GPU is NVIDIA Tesla K80 8G, and the CPU model is Intel(R) Xeon(R) CPU E5-2678 v3.

3.4. Model training

The model in this research adopts the method of de novo training, and uses the MobileVit network to train the bone marrow blood cell classification model. The model hyperparameters are as follows: the batch size is 128, the training epoch is 200, the initial learning rate is 0.001, and the loss function is the cross-entropy loss function. Figure 6 shows the accuracy curve of the validation dataset. Figure 7 shows the change curve of the loss value.

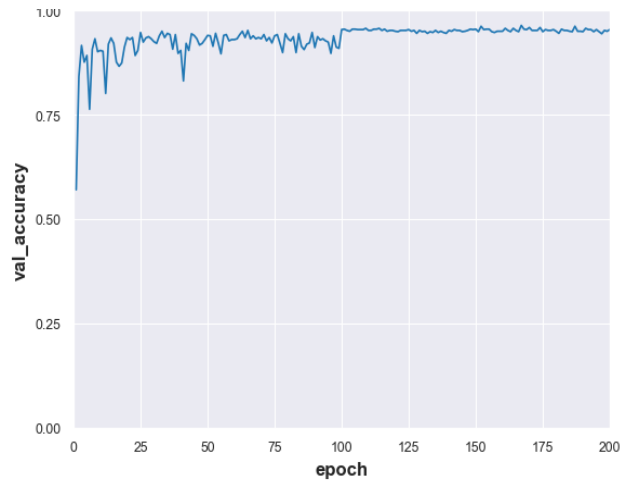


Figure 6: Acc

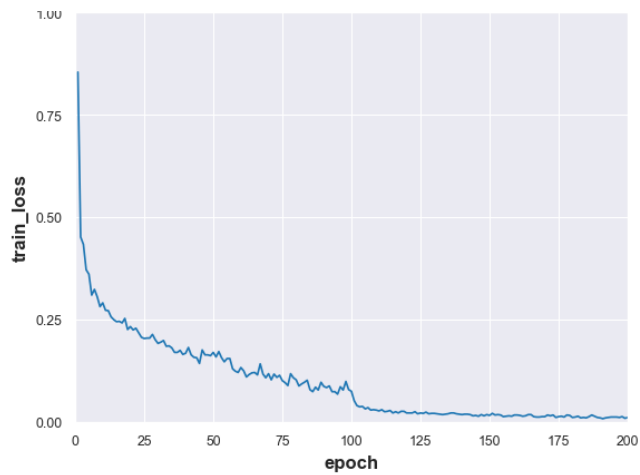


Figure 7: Loss

4. Results

Using the same dataset and setting the same hyperparameters, we compared the test results of MobileVit with some lightweight CNNs and heavy CNNs with high precision. From the three aspects of precision, operation amount and parameter quantity, and the comparison is as follows.

Table 1: Comparison with heavy-weight CNNs

Backbone	Acc	FLOPs	Params
Resnet50	0.950	8.219G	23.524M
Densenet121	0.942	5.725G	6.962M
MobileVit	0.964	1.799G	1.998M

Table 2: Comparison with light-weight CNNs

Backbone	Acc	FLOPs	Params
Mobilenetv2	0.940	495.066M	2.234M
Mobilenetv3	0.937	625.849M	4.304M
MobileVit	0.964	1.799G	1.998M

The advantages of the algorithm in this paper are as follows: improved accuracy, less params, and relatively large FLOPS compared to MobilenetV2.

As shown in the table above, the model network in this paper has higher accuracy than the light-weight CNN and the accuracy is improved by 2.4% compared with the light-weight CNN MobilenetV2. Compared with the heavy-weight CNN, the number of parameters is greatly reduced and the training time is shortened and the accuracy is improved. In summary, the model used in this paper has better performance and can better meet the classification task of bone marrow blood cell images.

5. Conclusion

Classification model with high classification accuracy and lightweight network are of great significance for automated bone marrow blood cell morphology detection systems. Based on the FAB diagnostic criteria [14], eight types of bone marrow blood cell images were selected from the AML-Cytomorphology_LMU public dataset as samples for the identification and diagnosis of acute leukemia [13]. Using these 8 types of bone marrow blood cell images, a dataset was established, and a lightweight classification model based on MobileVit was trained and verified. Our results show that on the same classification task, the MobileVit model performs well in classification accuracy, FLOPS and Params compared with several other neural networks, indicating that the MobileVit model has broad application prospects in this field. In the future work, we can consider to further optimize the model, try to transplant the network model to the mobile device, and implement the bone marrow blood cell morphology analysis system on the mobile device.

References

- [1] Yi Yunlin. "Quality management and clinical significance analysis of bone marrow cell morphology test". *China Contemporary Medicine*, 2011(29):100-101.
- [2] Gu Guiying, Liu Jingxin, Shi Zhangzhen, et al. "Application of computer image recognition in the morphological diagnosis of leukemia". *China Medical Devices*, 2018, 33(6): 6-11.
- [3] Mendelsohn M L, Mayall B H, Prewitt JMS, et al. "Digital transformation and computer analysis of microscopic image In: Barer Red. *advance in Optical and Electron Microscopy*". NewYork: Academic Press, 1968: 245~256
- [4] Fu X, Fu M, Li Q, et al. "Morphogo: an automatic bone marrow cell classification system on digital images analyzed by artificial intelligence". *Acta Cytologica*, 2020, 64(6): 588-596.
- [5] Tian Juanxiu, Liu Guocai, Gu Shanshan, et al. "Research and challenges of deep learning methods for medical image analysis". *Journal of Automation*, 2018, 44(3): 401-424.

- [6] Zhou M, Wu K, Yu L, et al. "Development and Evaluation of a Leukemia Diagnosis System Using Deep Learning in Real Clinical Scenarios". *Frontiers in Pediatrics*, 2021, 9.
- [7] Zhang HTA, Liu JX, Zhao XQ, et al. "Artificial intelligence-assisted diagnostic classification of blood cell microscopic images for acute lymphoblastic leukemia based on VGG16". *China Medical Devices*, 2019, 34(7): 1-4.
- [8] Li Fucui. "Research on morphological diagnosis of leukemia cells based on convolutional neural network combined with migration learning". Jinan University, 2019.
- [9] Dosovitskiy A, Beyer L, Kolesnikov A, et al. "An image is worth 16x16 words: Transformers for image recognition at scale". arXiv preprint arXiv:2010.11929, 2020.
- [10] Mehta S, Rastegari M. "Mobilevit: light-weight, general-purpose, and mobile-friendly vision transformer". arXiv preprint arXiv:2110.02178, 2021.
- [11] Sandler M, Howard A, Zhu M, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks"//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018: 4510-4520.
- [12] Rasha Salim. "A Single-cell Morphological Dataset of Leukocytes"Internet:www.kaggle.com/rashaslim/blood-smear-images-for-aml-diagnosis?select=AML-Cytomorphology_LMU,2020[May.2021].
- [13] Bennett J M, Catovsky D, Daniel M T, et al. "Proposals for the classification of the acute leukaemias French-American-British (FAB) co-operative group". *British journal of haematology*, 1976, 33(4): 451-458.
- [14] MIT."Imbalanced-dataset-sampler"Internet:www.github.com/ufoym/imbalanced-dataset-sampler, 2020[2021].