Development of Neural Network-based Autofocusing Module of Digital Microscopy System

Markosyan Vera,
Bauman Moscow State Technical University
Russia, Moscow
markostanvera@bmstu.ru

Abstract

Autofocus is a key function of any microscopy system. Performance of the entire system depends on the autofocus time consumption. The paper is devoted to development of a neural network-based autofocus module of an automated microscopy system. The following stages of module development are considered: choosing neural network architecture; collecting and processing the database; formation of datasets; training and analysis of neural network models; validation of the best model; embedding the model into the autofocus algorithm. The research of various approaches to architecting neural networks for autofocus purpose is carried out. Selected neural network is tested in terms of accuracy and time efficiency. The obtained results are analyzed, and ways of improvement are considered. The results of the paper allow to increase performance of autofocusing, which can be beneficial for biomedical microsample analysis.

***K*eywords** autofocusing · digital microscopy · convolutional neural network

1 Introduction

Autofocusing issue has been often cited as the culprit for poor image quality in digital pathology [1]. This is not because autofocusing is difficult to do, but rather because of the need to perform accurate autofocusing at high speed and on the fly with the acquisition process.

Here we explore the use of deep convolution neural networks (CNNs) to predict the focal position of the acquired image without axial scanning.

2 Features of the selected convolutional neural networks

2.1 ResNet50

Features of the network ResNet50 due to which it was chosen to complete the task:

- Skip connections to mitigate the vanishing gradient problem and improve training quality.
- Use of 1x1 convolutions to reduce network parameters and computational complexity.
- Use of 3x3 convolutions and max pooling to reduce image dimensionality and extract features.
- Batch normalization for stabilizing feature distributions and accelerating training.
- ReLU activation function for faster training and reduced vanishing gradient effect.
- Data augmentation to improve training quality and prevent overfitting.
- Use of the Adam optimization algorithm for fast and efficient network training.

The ResNet50 network can be chosen for image classification because of its deep architecture with 50 layers, residual connections that allow for better gradient flow and easier training, and state-of-the-art performance on various image classification benchmarks.[2]

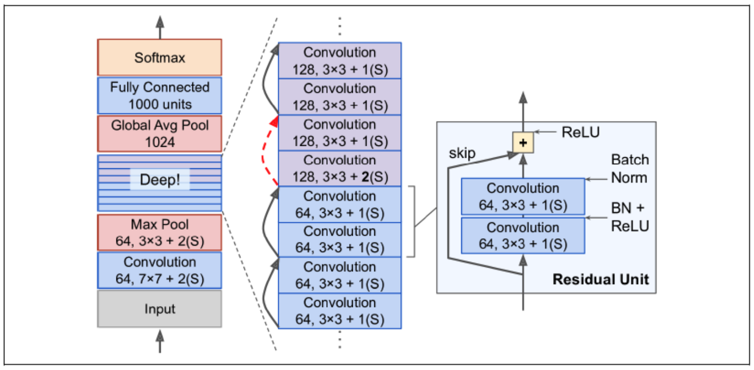

Figure 1: Resnet50 network architecture.

2.2 VGG

Features of the network VGG due to which it was chosen to complete the task:

- Deep architecture with up to 19 layers
- Small 3x3 convolutional filters
- Ability to learn complex features and patterns
- Simple and uniform structure.

VGG network can be chosen for image classification due to its deep architecture [3]

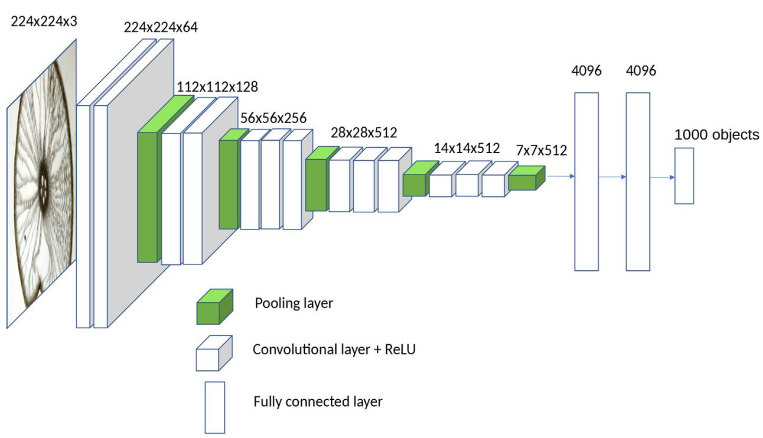

Figure 2: VGG network architecture.

2.3 MobileNetV2

MobileNetV2 builds upon the ideas from MobileNetV1 [1], using depthwise separable convolution as efficient building blocks. However, V2 introduces two new features to the architecture: 1) linear bottlenecks between the layers, and 2) shortcut connections between the bottlenecks1. The basic structure is shown below.

- There are 3 convolution layers in the bottleneck residual block. We know about the last 2 layers which are present in Mobile Net v1. They are depth wise convolution layer followed by a 1 x 1 point to point convolution layer.

- MobileNetv2 still uses the deep separable convolution. But now it has a bottleneck residual block instead of just a deep separable convolution block.
- Another important thing in the bottleneck residual block is a residual connection. It just works the same as in ResNet.
- Each layer has a batch normalization layer and activation function (ReLU6) except the project layer. The projection layer has only batch normalization because as the output from the projection layer is of low dimension, the authors found that introducing nonlinearity there using ReLU6 will decrease the performance.

[4]

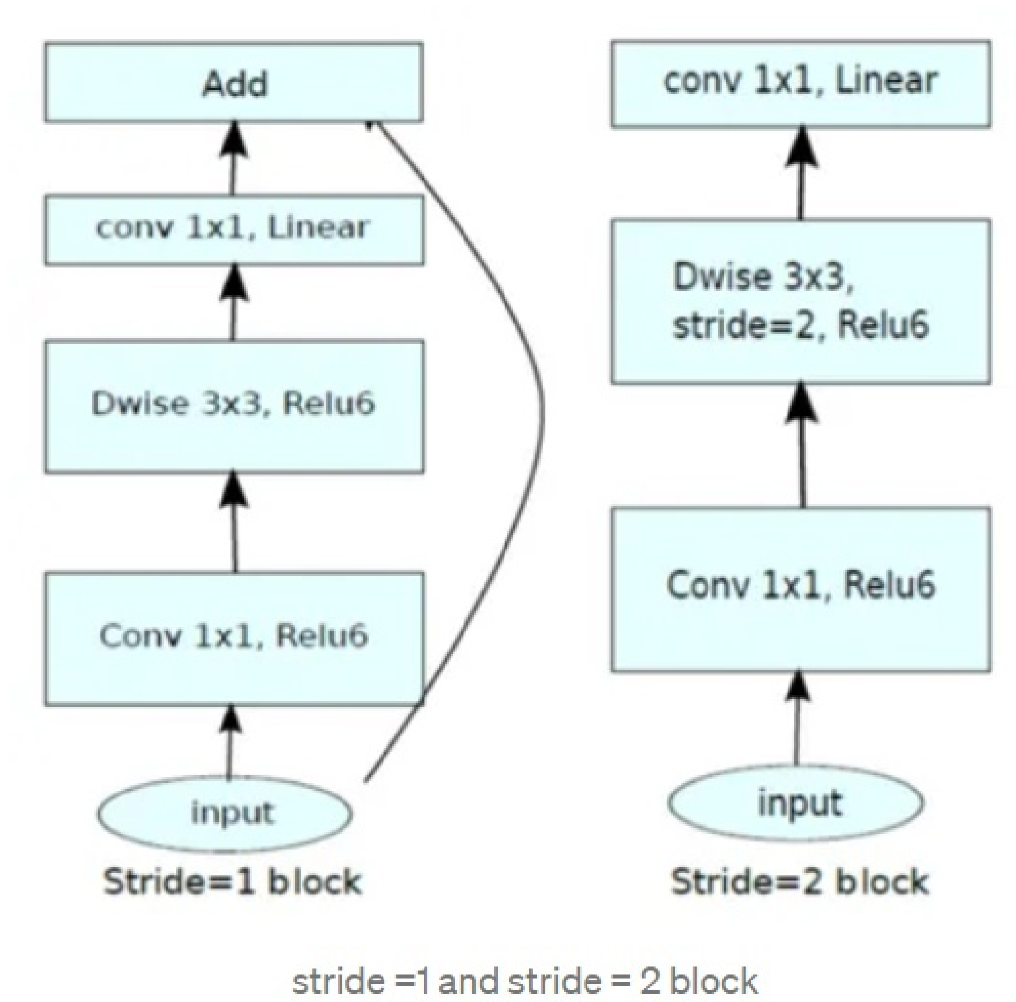

Figure 3: GoogLeNet network architecture.

3 Methods

To conduct the study, a publicly available (from open sources) article [1] dataset was taken. This dataset contains of images of different defocus.

3.1 Classify photos

Previously, all images were divided into 41 classes depending on the defocus.

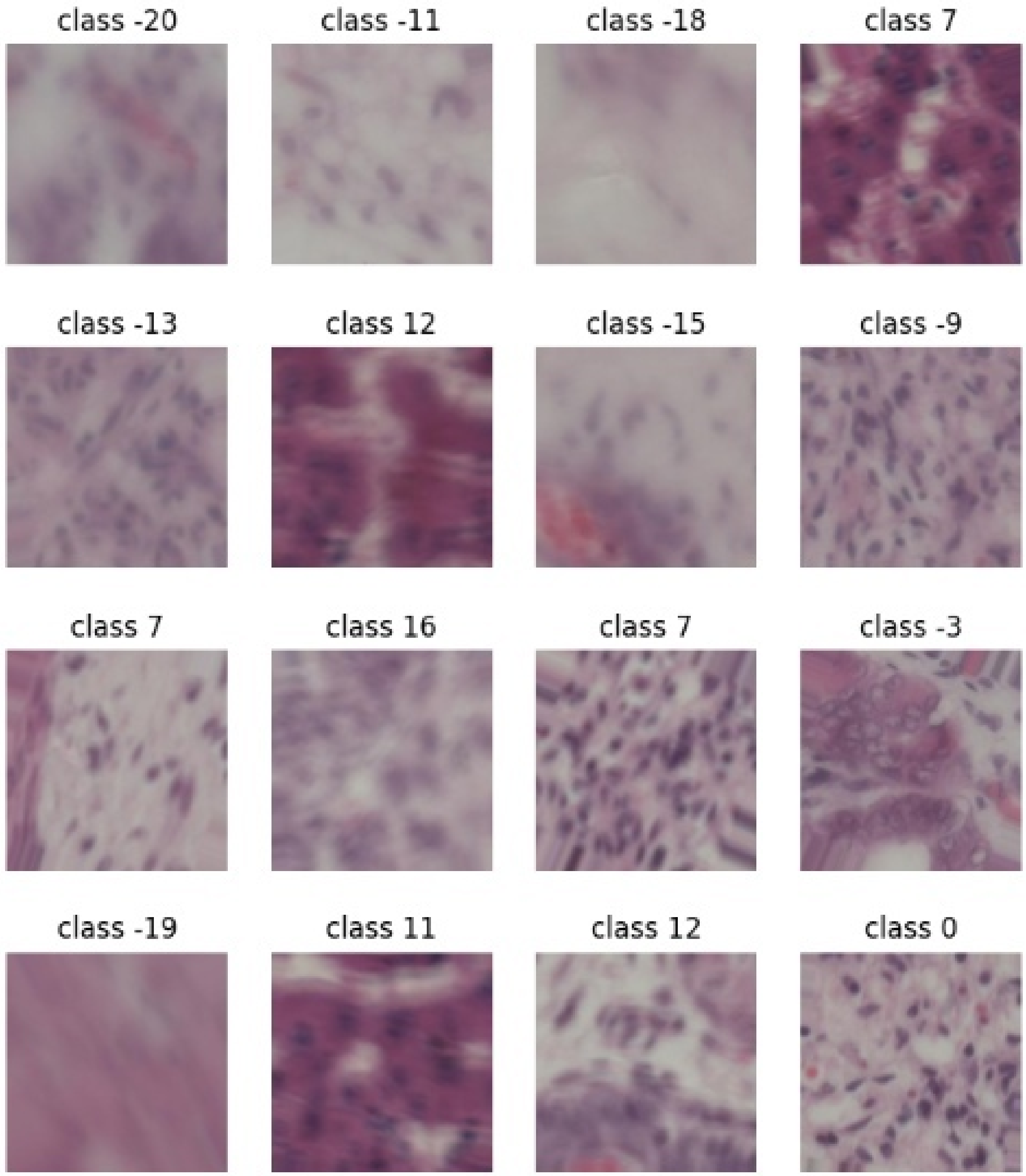

Figure 4: Classes of DataBase.

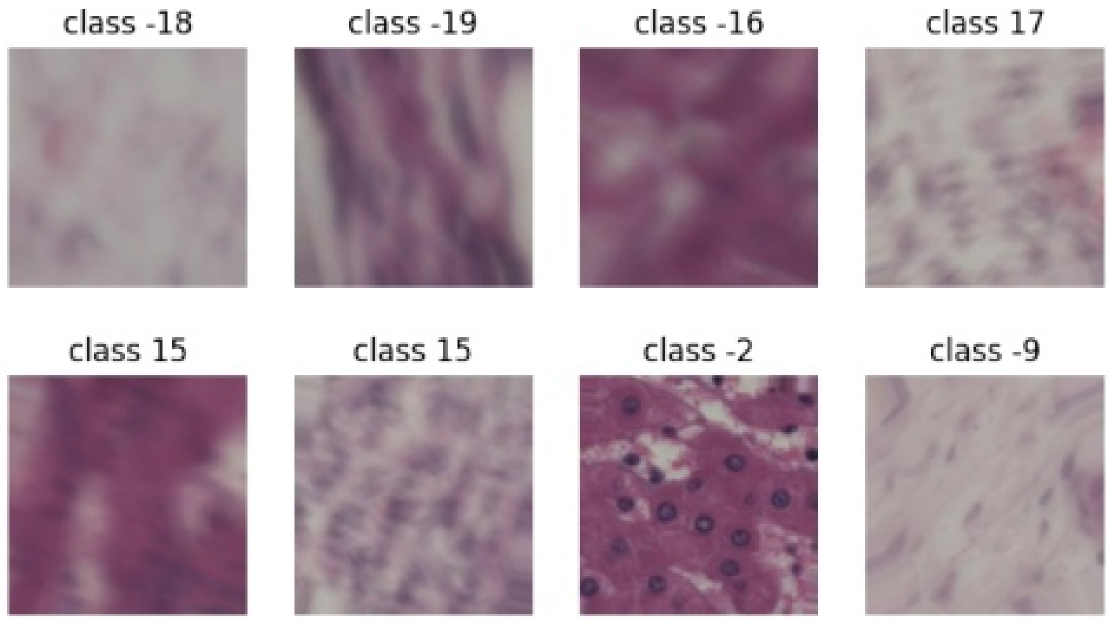

Figure 5: Classes of DataBasee.

3.2 Accuracy obtained using the network MobileNetV2

Like any other deep learning model, the accuracy of MobileNetV2 depends on the specific task it is applied to, as well as the quality of the data it is trained on.

Overall, however, MobileNetV2 is considered one of the most accurate models for image classification. For example, on the ImageNet dataset, which contains more than 1 million images from 1000 different classes, MobileNetV2 shows an accuracy of around 30-35%, which is a very high indicator.

In addition, MobileNetV2 also performs well in other computer vision tasks such as object detection and image segmentation. In general, the use of MobileNetV2 can significantly increase the accuracy of the model and improve its ability to recognize and classify images. MobileNetV2 accuracy was 33.56% at 15 epochs.

4 Graphs

4.1 Graphs of accuracy versus number of epochs

Below are the graphs for each Convolutional Neural Network

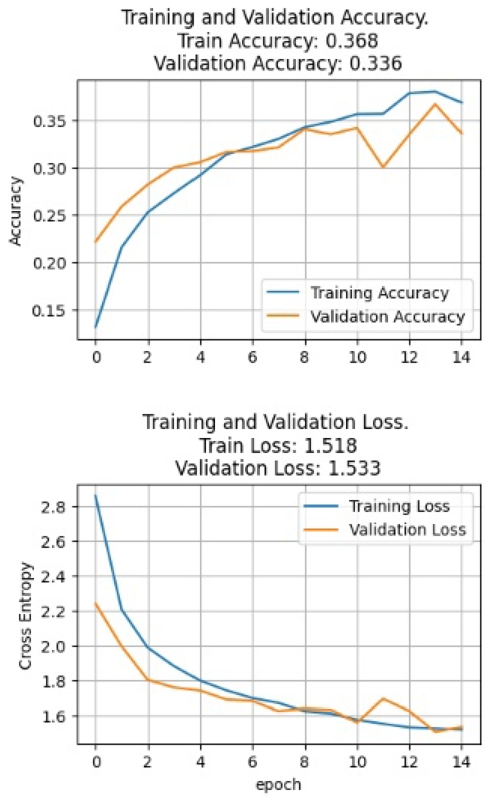

Figure 6: MobileNetV2 accuracy.

Accuracy of each class are shown at the picture bellow. The class 0 has the highest accuracy.

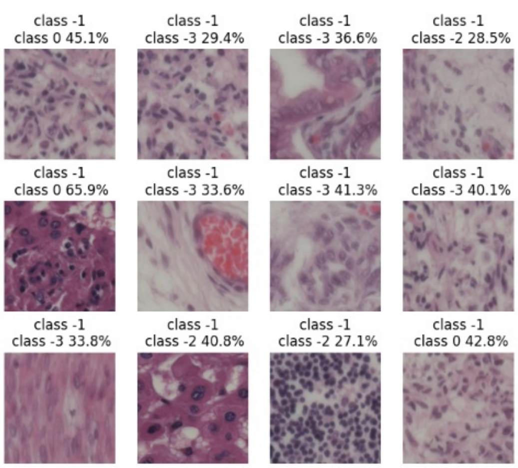

Figure 7: MobileNetV2 accuracy.

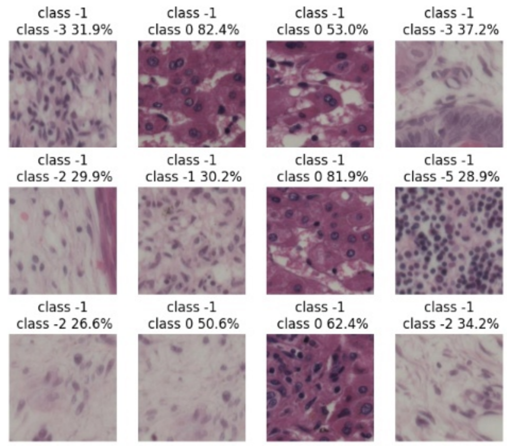

Figure 8: MobileNetV2 accuracy.

5 Comparison of all obtained results

The solution of the project described in this article is presented on the open source platform GitHub in the public repository at the link:

`https://github.com/Vera11113/classification_of_image_focus`

References

[1] Transform- and multi-domain deep learning for single-frame rapid autofocusing in whole slide imaging. *Biomedical Optics Express*, (20/1), 2022.

[2] Zhangshangjie Ge Hongbo Mu Dawei Qi Ruolei Zhang, Yijun Zhu and Haiming Ni. Transfer Learning for Leaf Small Dataset Using Improved ResNet50 Network with Mixed Activation Functions. *Forests*, 2072(13):40–47, 2022.

[3] Andrew Zisserman Karen Simonyan. Very deep convolutional networks for large-scale image recognition. *Visual Geometry Group, Department of Engineering Science, University of Oxford*, 14, 2015.

[4] Arun Mohan. Review On Mobile Net v2. *DataDrivenInvestor*, 2020.