

**Embedded Machine Learning
Summer Term 2024**

Exercise 5

- Return electronically until **Wednesday, June 5, 2024, 09:00**
- Include your names on the top sheet. Hand in only one PDF.
- A maximum of four students are allowed to work jointly on the exercises.
- When you include plots add a short explanation of what you expected and observed.
- Hand in source code if the exercise required programming. You can bundle the source code along with the PDF in a .zip file.
- Programming exercises can only be graded if they run on the cluster in the provided conda environment. Make sure to document additional steps, which you might have taken to run the exercises.

5.1 Activation normalization techniques

In this exercise we are taking a quick look at normalization techniques and the advanced neural network architecture of ResNets.

- Implement the BatchNorm and one other activation normalization technique (Group normalization, layer normalization or instance normalization) for the VGG from the previous exercise. Place the normalization layer in-front of each activation layer. Note: Make sure to use the 2d implementation of the normalization layer. From an implementation standpoint you may be happy to know that the GroupNorm becomes the LayerNorm for $G = 1$.
- Train the VGG for 30 epochs, once without any activation normalization, once with BatchNorm and once with the activation normalization technique of your choice.
- Discuss the results of the previously obtained plot.

(20 points)

5.2 Residual neural network architectures

Residual neural networks have seen great success in the field of machine learning, in particular for image recognition.

In this exercise we will implement a derivative of ResNet, for which you can find more details in the original paper¹.

- Using the template implement a scaled down version of ResNet-18. In particular the ResNet basic block should be implemented as shown in the lecture.
- Train the ResNet for 50 epochs. Make sure to not use any data augmentation to only measure the training time of the network.
- Compare the test accuracy of the ResNet to the VGG11 from the previous exercise and the ConvNet from exercise 3. Note that you can reuse results from the previous exercises here. For all three networks plot the test accuracy over the number of epochs and over the training time.
- For all three networks calculate the number of parameters and MAC operations.
- Discuss how the number of parameters and MAC operations impacts the previous plots and how the different architecture design has impacted these metrics.

(20 points)

¹<https://arxiv.org/pdf/1512.03385.pdf>

5.3 Willingness to present

Please declare whether you are willing to present any of the exercises from this sheet. The declaration should happen on the PDF which you hand in.

The declaration can be made on a per-exercise basis. Each declaration corresponds to 50% of the exercise points. You can only declare your willingness to present exercises for which you also hand in a solution. If no willingness to present is declared you may still be required to present an exercise for which your group has handed in a solution. This may happen if as example nobody else has declared their willingness to present.

- Activation normalization techniques (Section: 5.1)
- Residual neural networks (Section: 5.2)

(20 points)

Total: 60 points