CM2003 Course: Teodor Pstrusinski & Lasse Stahnke

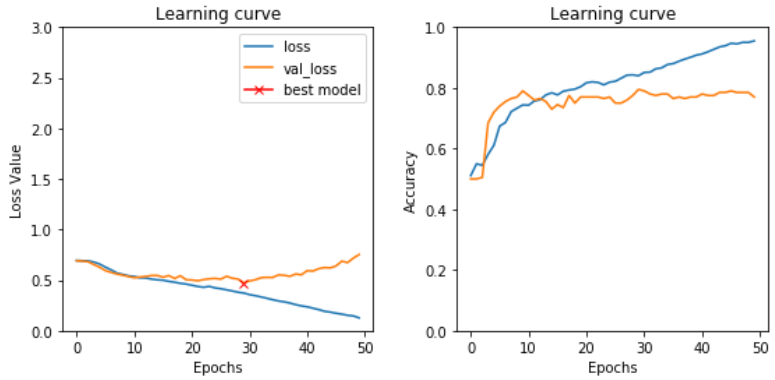
Lab 3

Regularization Techniques

Task1a)

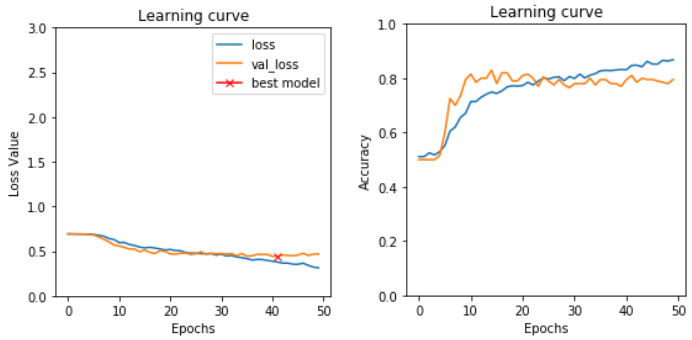
How do you interpret the learning curves?

The learning curves diverge from each other. The validatoin loss is higher than the training loss. Thus, the network seems to overfit the data. Furthermore, the validation accuracy does not substantially increase after 10 epochs.



What is the effect of adding drop out layers?

Adding the dropout reduces the overfitting. This can be seen by the decreased difference between the training and validation loss in the final epochs. However, compared to the model without dropout, the validation accuracy is lower. But overall, the generalization power of the model seems to be higher.



Task 1b)

At which epoch do you observe the same training accuracy as task1a? What is the value of final training accuracy? What is the

effect of the batch normalization layer? Similar to task1a, do this task with and without drop out layers.¶

- Model without dropout: After 10 epochs, the network with Batch Normalization achieved the same training accuracy as the model without Batch Normalization. The final training accuracy is 1 for the model with Batch Normalization.
- Model with dopout: After 17 epochs, the model with dropout and Batch Normalization reaches the same training accuracy as he model with dropout and without Batch Normalization. The final training accuracy of the model with dropout is 0.981.

Therefore, the Batch normalization sped-up training process in both cases, but did not affect the validation accuracy substantially.

Task 1c)

Focus on validation loss & accuracy. Which model resulted in higher validation accuracy? How do you explain the effect of batch normalization?

The model with Batch Normalization resulted in a highrt validation accuracy (0.81 vs. 0.78). The model with Batch Normalization also has slightly lower loss than the model without Batch Normalization. Furthermore, the validation accuracy converges faster when using the Batch Normalization.

Task 1d)

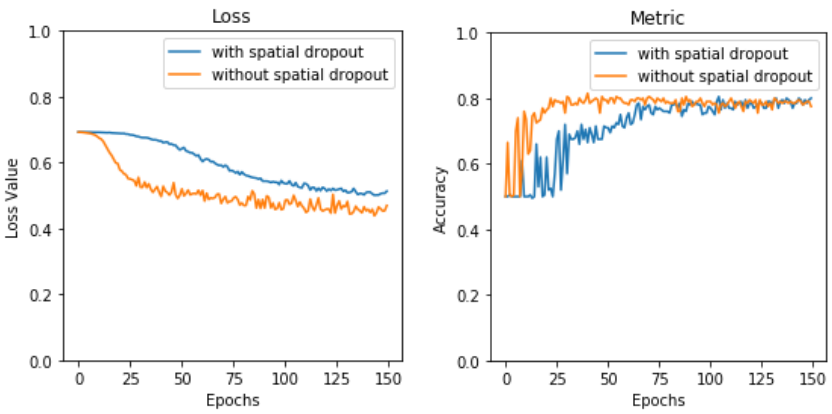
Which model yields more accurate results? Which of them has more generalization power?

The model with batch normalization yiels a slightly better validation accuracy. (0.81 vs. 0.79) Thus, this model is more accurate. However, the difference between the training and valiation loss is higher than in the model without batch normalization. Thus, the model without batch normalization has the higher generalization power.

Task 2a)

Then, run the same model with the same settings but remove all the spatial drop out layers. Which of them converges faster? Why?

The Model with Spatial Dropout converges slower than the model without. Since dropping the channels in the kernel makes the optimization more difficult(because it removes parameters that are helpful for the classification), the model needs more training epochs to achieve the same classification performance.



Task 2b)

In general, discuss how the drop out technique (spatial and normal one) would help the learning procedure.

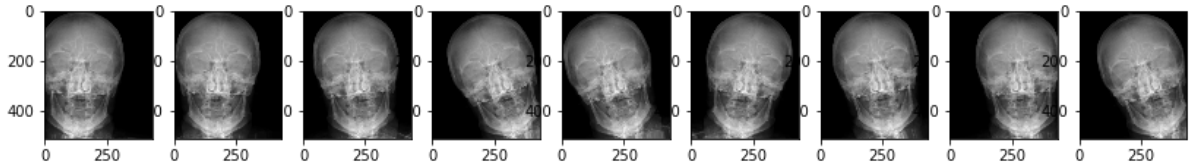
Dropout helps the network learn to not rely only on specific features. Thereby, it regularizes the network and reduces overfitting.

Data Augmentation

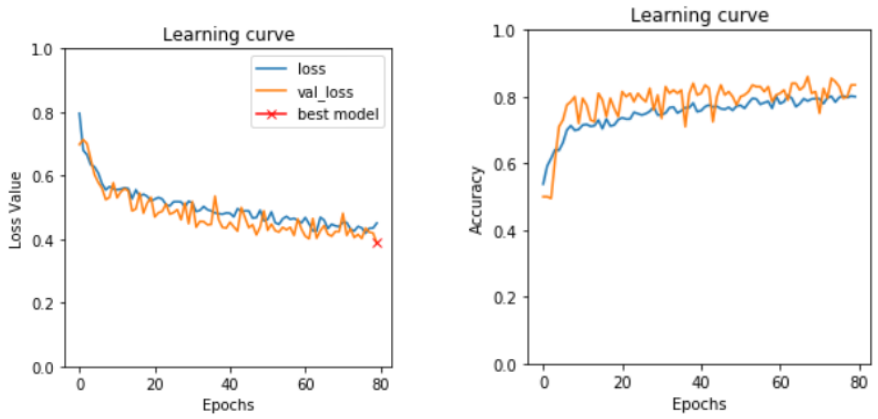
Task 3a

- Rescaling: Rescaling the image resolution with respect to the original resolution
- Rotation: Rotating the image with a specified angle in degrees
- Horizontal flip: Flipping image horizontally
- Vertical flip: Fliiping image vertically
- Intensity rescaling: Rescales intensity values, so that the lower percentile is mapped onto 0% while the upper percentile is mapped to 100%

Task 3b



Task 4



How does the data augmentation impact model training? Why?

When using the data augmentation, overfitting is reduced because the model has more training data.

Transfer Learning

Task 7)

Compare the observed results from transfer learning against the ones you already achieved by training the models from scratch. Describe how transfer learning would be useful for training. How can you make sure that the results are reliable?

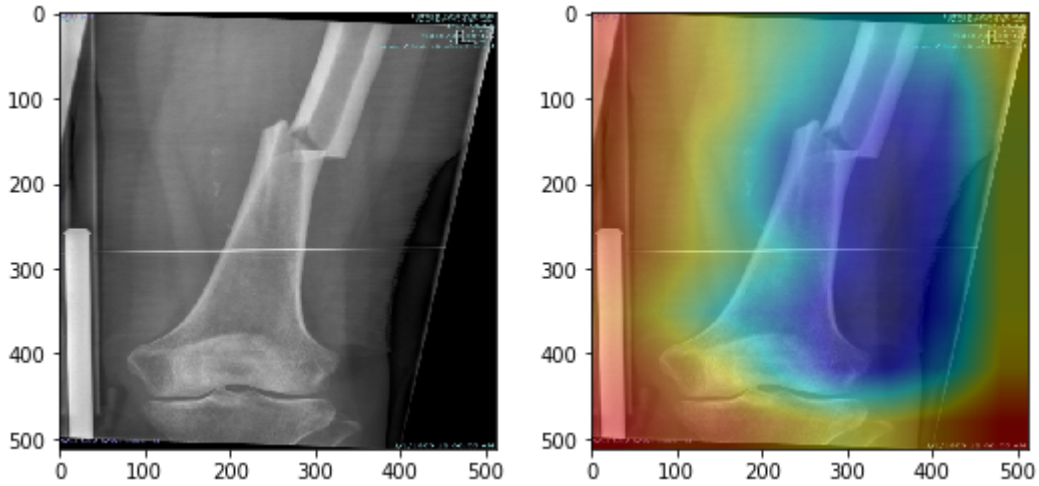
Transfer learning using the pretrained VGG16 model works well on the bone dataset. After transfer learning, a validation accuracy of 1 could be achieved. Training the VGG on the bone dataset from scratch achieved a validation accuracy of 0.9 in comparison.

Transfer learning using the pretrained VGG16 model on the skin dataset didn't work well. The learning curve is very steep in the beginning, but converges quite quickly. Furthermore, it diverged from the training accuracy, indicating overfitting. The transfer learning achieved a validation accuracy of 0.825 compared to ... in the model trained from scratch.

Visualising Activation Maps

Task 8

What can you infer from visualization of the class activation maps?



The network looks ar the corners of the image. Thus, it does not perform the task of classifying broken bones correctly, but seems to have found another connection between images of broken bones that don’t actually include the broken bone itself, but rather properties of the image, such as text on the bottom.