Computer Vision Assignment 3

Alex Dietrich 1637498
D Hema Kumar 1841210

November 2022

1 U-Net Paper

Task: Read the U-Net paper by Ronneberger et al..

Why do the authors state that the U-Net works good for segmentation?

The argumentation of the authors is that the U-Net is able to capture context in its contracting path whilst being able to exercise precise localization in its expanding path. The usual contracting network is supplemented by upsampling operators that replace pooling operators. Further, in order to improve localization in the expanding path, high resolution features from the contracting path are combined with the upsampled output at each step. Succesive convolution layers, consequently, learn to assemble more precise output based on this information. Lastly, they argue that their upsampling part is able to propagate more context to higher resolution layers, as they have a large number of feature channels.

Do you think their argumentation makes sense?

Their argumentation makes sense. Their architecture features a lot of properties that makes it successful. They employ large number of feature channels allowing for better traversing of information among each convolution layer. The contraction path may be pretty standard but their expanding path is the heart of this architecture. The process of appending the learned feature maps from each step of the contraction path to the corresponding expending step is clever, in that it ensures that the features that are learned during contraction are then reused in reconstruction. This seemingly improves localization by multiple magnitudes. This is conceivable, because upsampling is a sparse operation and a good prior from earlier stages helps to better achieve localization.

What are your arguments that the U-Net works well?

A majority of arguments, why we think that the U-Net works well was already stated in the previous question. Though, we think that is not all there is to it. We imagine that the loss calculation in the U-Net may also play a major role in its success. U-Net uses a loss weighting scheme for each pixel such that there is a higher weight at the border of segmented objects. This loss helped the U-Net model segment cells in biomedical images such that even individual cells were easily identified within the binary segmentation map.

2 U-Net Implementation

See code.

3 U-Net Logs

Task: Log train loss, train accuracy, train mean IoU (also called Jaccard index), validation accuracy and and validation mean IoU. To this end you can use the tensorboard graphs and write a short report.

3.1 Training Loss

The model was run for around 35 epochs and after 90 steps(ca. 13 epochs) the training loss was saturated at around 1 and did not improve much until 240steps.

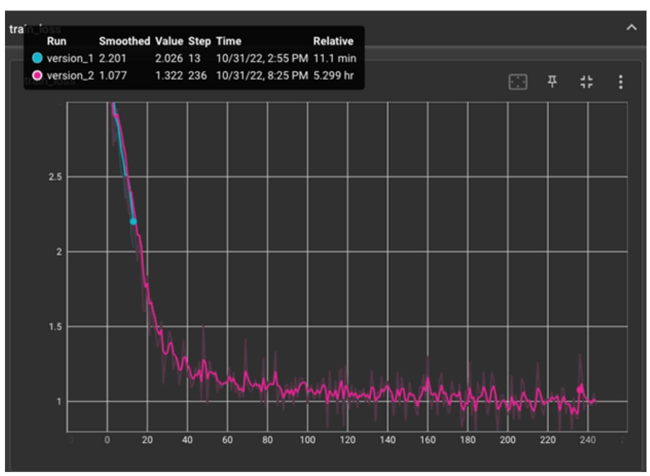

Figure 1: Training loss

3.2 Training Accuracy

The tensorboard graph shows a training accuracy of 75% that was achieved around the 180th step, i.e., at epoch 25 and did not improve much further.

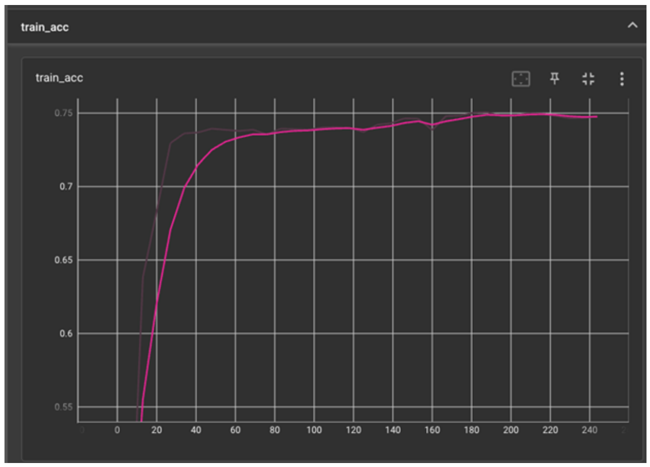

Figure 2: Training accuracy

3.3 Training Mean IoU

The model's training mean IoU reaches a value of about 0.58 after 244 steps. Hereby, the first 40 steps show an enormous rise in mean IoU up until a value of roundabout 0.55 followed by a steady rise until the end of training.

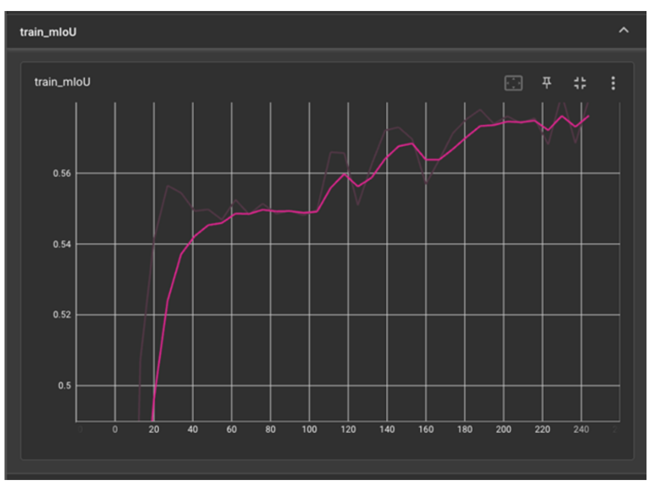

Figure 3: Training mean IoU

3.4 Validation Accuracy

The validation accuracy also reached 74% at time step 90 and did not improve further. At time step 200 it reached 75.24% which represents the peak. Also, the validation accuracy is not very different from the training accuracy meaning the model is probably not overfitting. There is room for improvement of the model through data augmentation techniques.

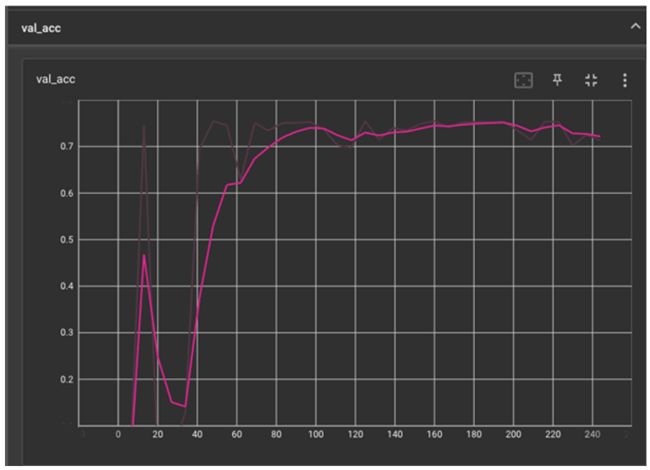

Figure 4: Validation accuracy

3.5 Validation Mean IoU

The validation mean IoU reached shows a fast rise up until the first twenty steps followed by a vast decrease up until step 40. From this point on the validaion mean IoU then grew up until it reached a value of 0.59 at step 180.

4 Augmentations

As augmentations strategies we employed a combination of random horizontal flips, random rotations and random gray scale. The same figures as in the previous exercise are provided for the model where we used the listed augmentations. We achieve similar results to training without augmentations, but one can see a slight tendency of better generation from the model that worked with augmented data when comparing the metrics on the validation set.

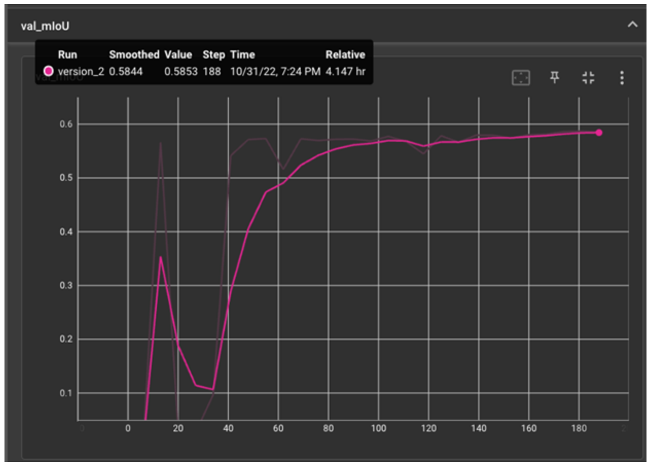

Figure 5: Validation mean IoU

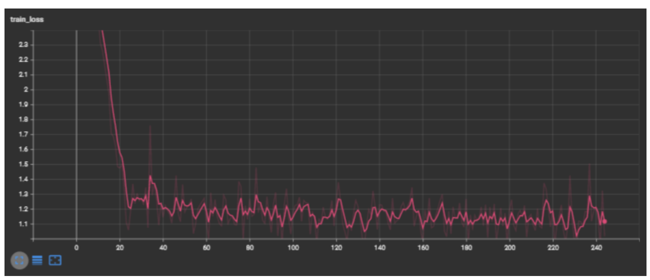

Figure 6: Training loss with augmentations

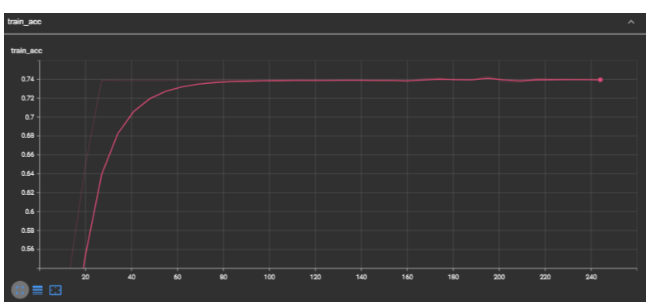

Figure 7: Training accuracy with augmentations

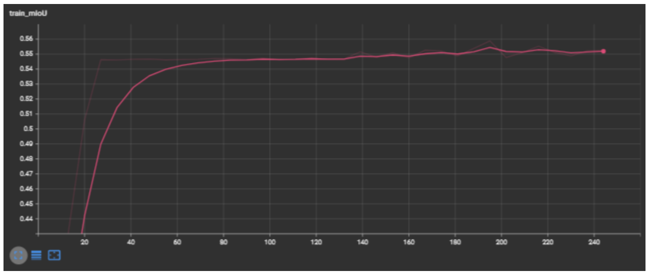

Figure 8: Training mean IoU with augmentations

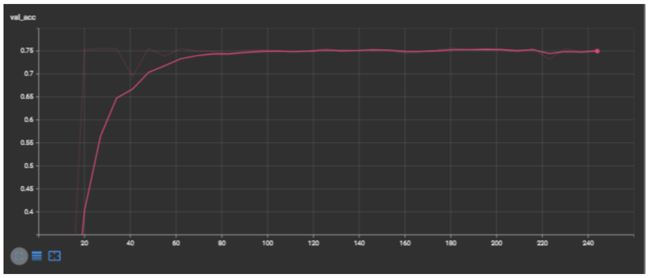

Figure 9: Validation accuracy with augmentations

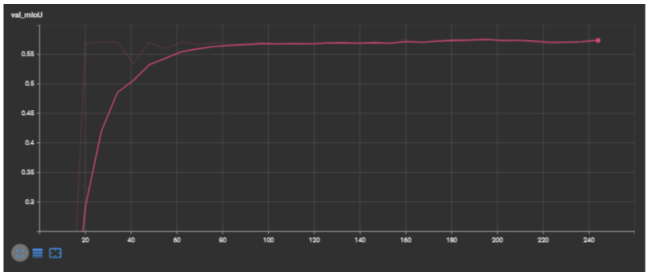

Figure 10: Validation mean IoU with augmentations