Computer Vision Assignment 3

Alex Dietrich 1637498
D Hema Kumar 1841210

November 2022

1   Results

In our implementation of the respective modules and the final vision transformer we followed paper suggestions, various tutorials and open source code implementation. Thus, we adjusted the modules accordingly. For example, we added more parameters like "model_dim" or "ffn_hidden" to most modules to increase customizability or added dropout and activation functions in the fnn module. This gave us the means to adjust model complexity accordingly. In the following, we will present the results for certain parameter constellations that we made during training.

1.1   Transformer(64, 128, 256, 10, 3, ,3, 4)

With this configuration, we were able to achieve a train accuracy of above 99% and a validation accuracy just below 98% as can be seen in figure 1 and 2. Our test accuracy reached a value of 97.87%.
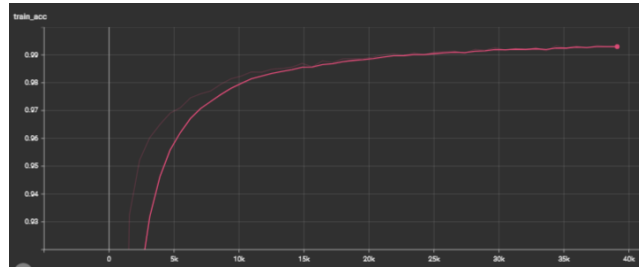

Figure 1: Training accuracy model 1

1.2   Transformer(128, 256, 512, 10, 8, ,8, 4)

This configuration increased the model's complexity on various fronts, such as dimensionality across the board as well as more attention heads and transformer
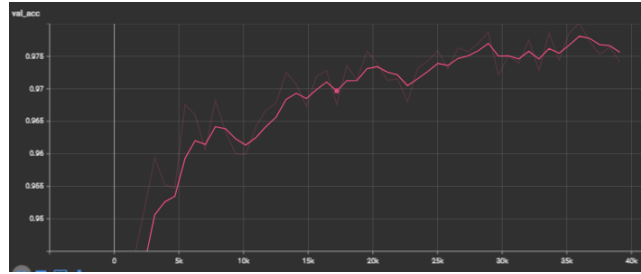
Figure 2: Validation accuracy model 1

blocks. With this we were able to increase the model's performance and achieve a train of above 99% and a validation accuracy of almost 99% as can be seen in figure 3 and 4. Our test accuracy reached a value of 98.45%.
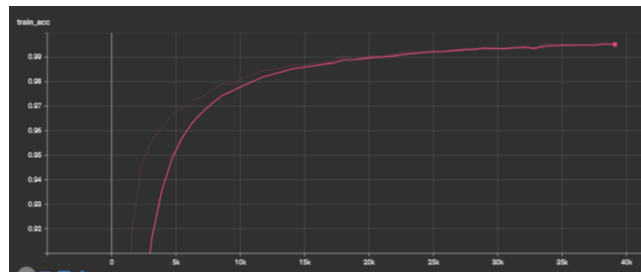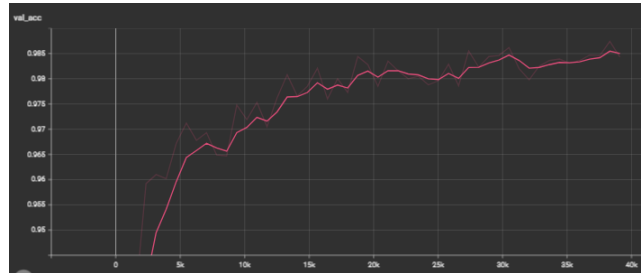

Figure 3: Training accuracy model 2

Figure 4: Validation accuracy model 2