

The autoencoder is trained using Stochastic gradient descent with Nesterov momentum, MSE as the loss function and a learning rate of $5e-4$, which is halved every 25th epoch. To increase the robustness of the autoencoder, we add noise sampled from

$$\mathcal{N}(0, x^2e - 2)$$

to each training sample, where x is the value of the training sample. To help the model converge, we sample from a Glorot normal distribution, like in the case with the stacked hourglass.

After the autoencoder has been trained, the whole network is further trained by following Newell *et al.* [1] as described in Section ??.

1.3 Results

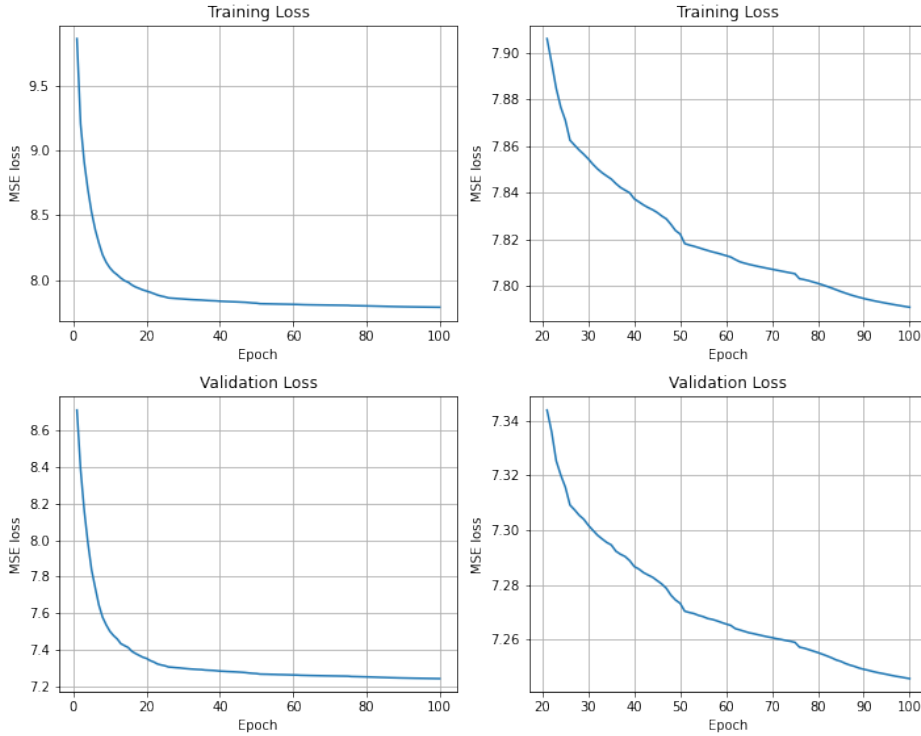


Figure 3: Visualization of the evolution of the training- and validation loss of the autoencoder during training. The left column shows all of the 100 epochs. Right column shows epoch 21 and forward

By training the autoencoder isolated, we get the evolution of the training- and validation loss visualized in Figure ?? . We can clearly see, how the model does not start to overfit, as in the case when we trained the stacked hourglass. We decided to stop the training of the autoencoder, as each update only yielded minor changes to the model. The evolution of training the stacked hourglass with the autoencoder has been visualized in Figure 4. By looking at the figure we can see how the combined stacked hourglass and autoencoder initially performs worse than the original stacked hourglass, however, as the training continues it beats the original stacked hourglass, resulting an validation PCK accuracy of 46.7% - an increase of 7.8% or 3.4 percentage points compared with the original stacked hourglass.

To get an unbiased performance evaluation of the two models for comparison, we evaluate

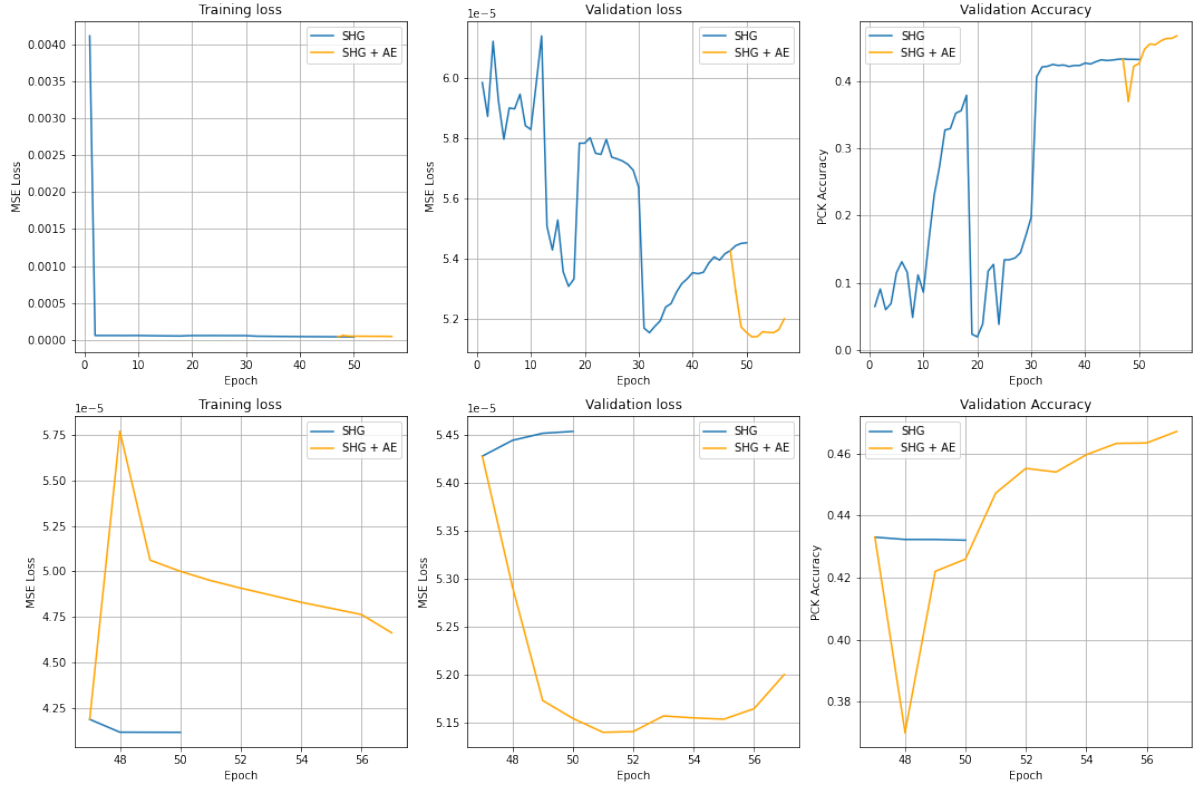


Figure 4: Visualization of the evolution of the training- and validation loss, as well as the PCK validation accuracy of the combination of the stacked hourglass and autoencoder, compared with the evolution of training the original stacked hourglass. The top row is of all of the 57 epochs. The bottom row shows epoch 47 and forward.

our two models on the held-out testing set from section ?? . By doing so we get a PCK score of 44.1% for the standard stacked hourglass and a PCK score of 47.3 for the stacked hourglass with the autoencoder. That is an increase of 7.25% or 3.2 percentage points, by making use of the autoencoder.