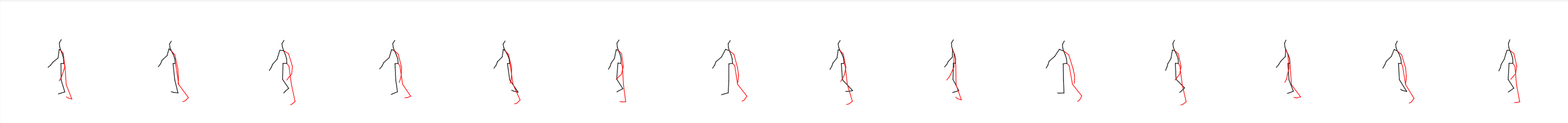
DTU Compute
Department of Applied Mathematics and Computer Science

Generating subject specific human motion using a quaternion based recurrent model

P. E. Møllgaard (s144256), **T. S. Dengsøe** (s144236), **P. M. Jensen** (s144283) & **C. A. Thilker** (s144255)
M.Sc. students at DTU Compute



Abstract

We attempt to generate motion sequences with a quaternion based recurrent neural network which can emulate a specific subject's walking pattern. We were successful in generating data that was unique for each subject, however, proving the generated data was the same as the original data dit not come into fruition.

Method

Motion generation

We use the same network architecture as the long-term pose network from [1], shown in Figure 1.

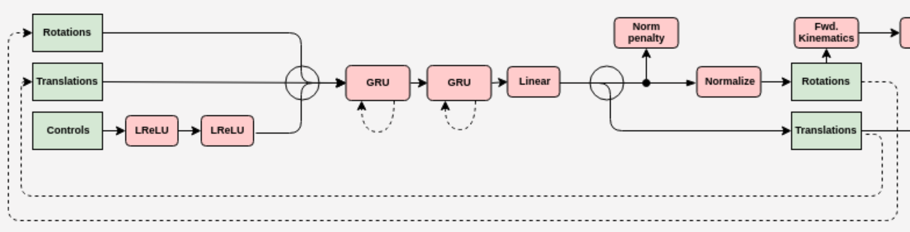


Figure 1: Architecture of long-term generation

Training is done with the same scheme as [1] using the first half of the motion sequence for each subject.

Subject classification

We extend the short-term pose network from [1], shown in Figure 2.

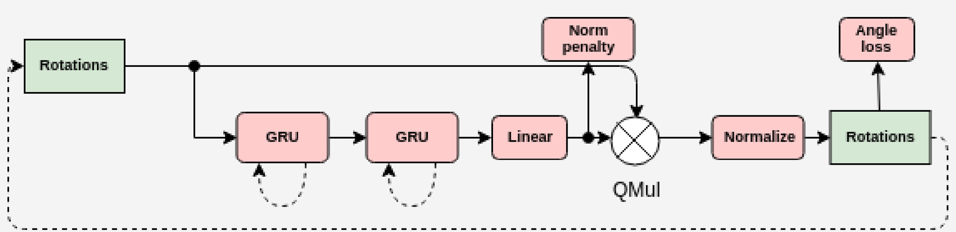


Figure 2: Architecture of short-term generation

The pose network is first trained for short term motion generation. Afterwards, the weights are frozen and two new fully connected layers are connected to the hidden state of the second GRU unit, ending in a softmax unit, see Figure 3. Batch normalization is used after each added linear layer.

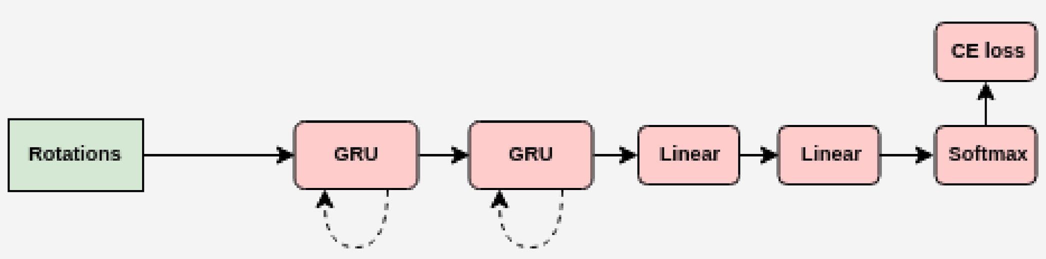


Figure 3: Architecture of classification

When training, we use the first half of the motion sequence for each subject, and validate using the second. Cross entropy is used as loss function.

Results

Motion generation

Figure 4 shows an example of generated motion compared with true data. The two sequences are visually quite similar, which indicates the model is able to generate motion sequences that resembles real motion.

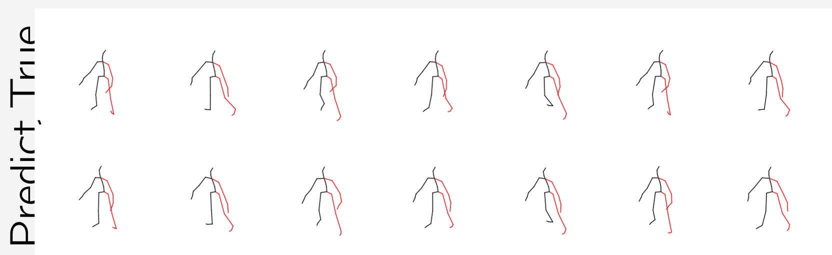


Figure 4: Example frames of generated motion compared to true data.

Figure 5 also shows this similarity, at least for the foot. The hand is more chaotic - and the motion is less similar - but also moves a lot less. Since the loss is based on position, this might explain why the model is less precise here.

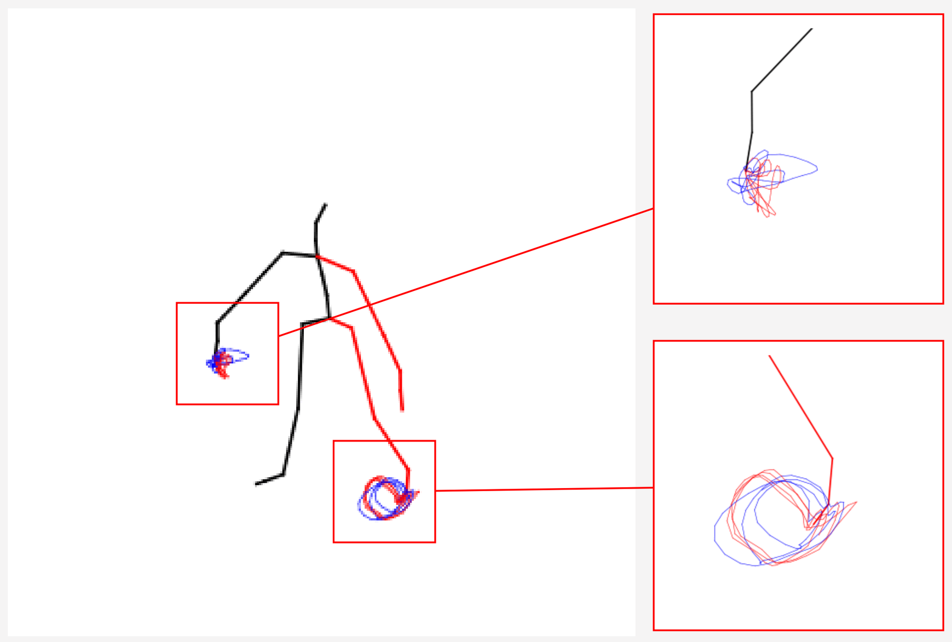


Figure 5: Trajectory of the left foot and right arms for true and predicted

Subject classification

Figure 6 shows that each subject follows a distinct pattern both for real- and generated data, but the combined plot shows a clear separation between real- generated data which indicates dissimilarity.

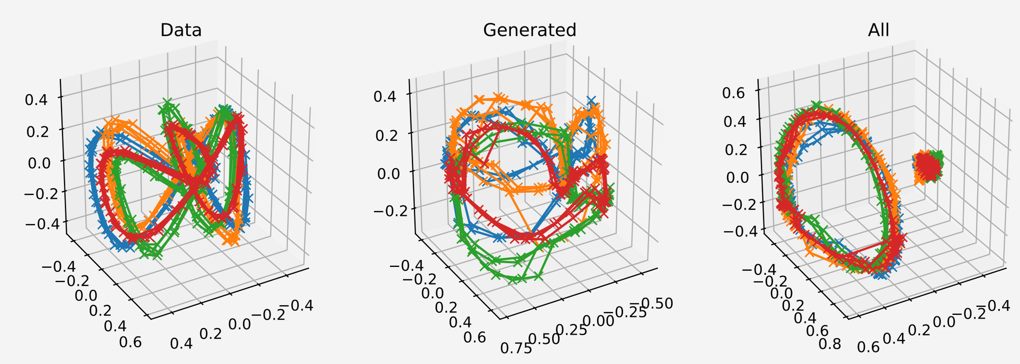


Figure 6: Evolution of hidden state for the second GRU unit shown using a PCA. Each color corresponds to a different subject.

From Figure 7 and Table 1 it is clear that the classifiers work best for the kind of data they were trained on. The classifier trained on both data and generated motion performs best when validated against generated data.

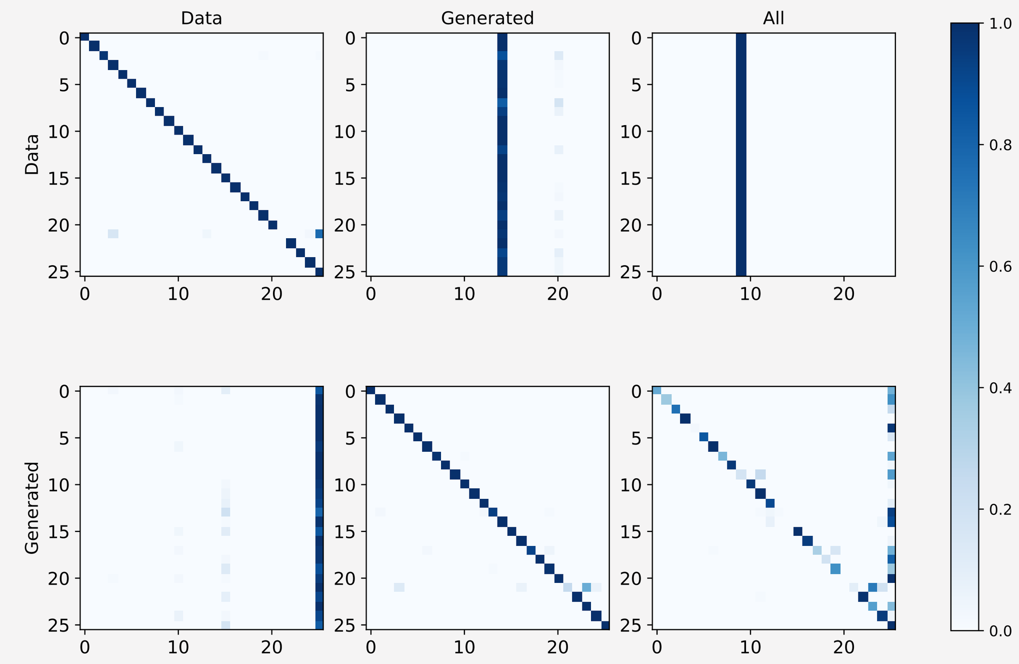


Figure 7: Confusion matrices for the classifier networks. Top: results for networks evaluated against real data. Bottom: results for networks evaluated against generated data. Left to right: network trained on real data, generated data, and both real and generated data.

Table 1: Performance measures for classifier networks.

Validated on	Data			Generated		
Trained on	Data	Gen.	All	Data	Gen.	All
AUC	1.00	0.50	0.92	0.47	1.00	0.99
F1	0.96	0.04	0.04	0.04	0.97	0.61
Validation accuracy	0.96	0.04	0.04	0.04	0.97	0.61

As walking is a periodic motion we quantify the difference in motion using the sum of the log-spectral difference wrt. the x, y, and z component for each joint. Figure 8 shows an example spectrogram on the left. Qualitatively there is some similarity, but the peaks are not completely aligned. Figure 8 right shows that generated sequences are similar to each other, but there is no clear correspondence between a generated sequence and the subject it was generated from.

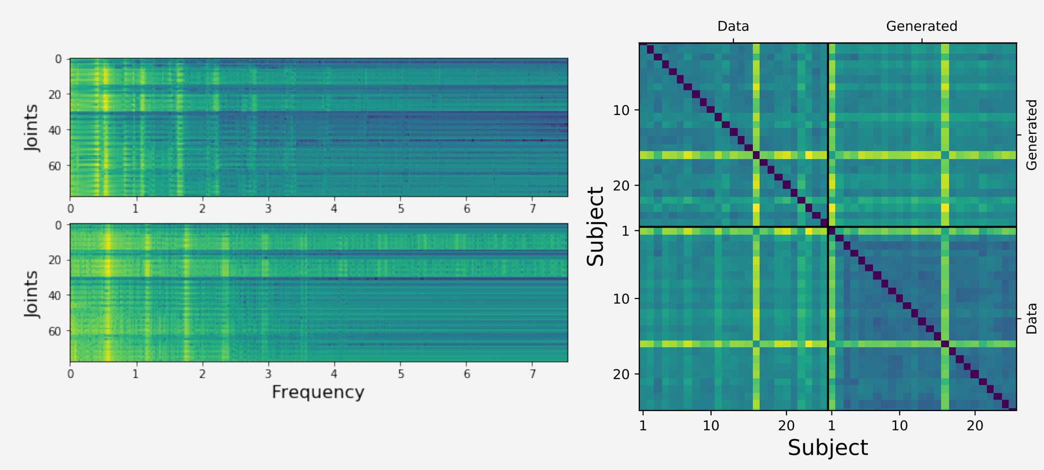


Figure 8: Left: Spectrograms of the x,y,z positions for all the joints of subject 4. Top left: Raw data. Bottom left: Generated data. Right: Distance matrix between subjects for true- and generated data.

Future work

For future work it could be of interest to:

- Use subject specific skeletons
- Investigate better similarity measures for motion
- Investigate necessity of extra control inputs
- Investigate other network architectures such as GANs and VAEs.

References

[1] Dario Pavllo, David Grangier, and Michael Auli, "Quaternet: A quaternion-based recurrent model for human motion," *CoRR*, vol. abs/1805.06485, 2018.