1 Discussion

- PCK: Scaler ikke med input størrelse
- Glorot Initialisering - kan være dårligt, se EML forelæsning om optimization.
- Acc. var stadig stigende - måske skulle jeg lade den træne længere
- Måske skulle jeg have gjort brug af gap statistics til at finde optimals k
- Curse of dimensionality ved clustering
- Se kap. 14.6 "Learning Manifolds with Autoencoders" i deep learning book. Til sidst er der en diskussion om brugen af Autoencoder til manifold learning
- Euclidisk afstand er måske ikke det rigtige?
- Man kan ikke gøre brug af den rigtige centroid for hvert cluster (grundet skip-connections). De centroids vi har fået kan eventuelt ligge langt væk fra de rigtige centroids, resulterende i misvisende centroids.
- Latent space af AE sættes til 50, idet vi ved shape analysis har set, at de resterende dimensioner er støj. Dette bygger sig dog på fulde skeleter og ikke alle skeleter, som modellen ellers trænes på

1.1 Summary of Obtained Results

In Section **??** we succefully implemented and trained a stacked hourglass, consisting of a single hourglass. We did this by following the configuration details described in Newell *et al.* [1] and Olsen [2]. The developed model has a validation PCK accuracy of 0.433 and a test PCK accuracy of 0.441.

In Section **??** we gained an understanding of how the devleoped model works by exploring the different components of the model. We could verify, that the skip-connections of the model were used for recreating details that are lost during the encoder-phase of the model, as argued by Newell *et al-* [1] and Olsen [2]. We then used PCA to gain an understanding of the structure of the latent space of the model. By doing so we came to the conclusion, that the model had learned the differences between people standing up and people sitting down, as well as possibly discovering some redundancy in the model, as principal componet 50 and above seemed to act as noise. Lastly, we used clustering to gain an understanding of how the model works. By doing so we learned, that the model knows the difference between fully-annotated people and not-fully annotated people, as well as knows the difference between stationary people and moving people. Here we also identified some possible reasons for inaccuracies of the model, as these classifications were not always correct.

In Section **??** we used our knowledge of the model to improve the performance of the model. This was done by developing and training an autoencoder, which was placed in the model. By doing so, both the validation and test PCK accuracy increase to 0.467 and 0.473, respectively.

1.2 Comparison of Models

1.3 Hvorfor er mine resultater dårligere/bedre end Newell/Camilla?

- Forskelle imellem min(e) modeller og deres
- Skal også komme ind på hvorfor mine valg var bedre end de valg de valgte

1.3.1 Forskelle

- Batch normalization - ved ikke placering
- Autoencoder
- Noget andet data
- Gør også brug af $v = 1$
- Anden spredning ved gaussian filter ved $v = 1$
- Camillas model er kun trænet på $v = 2$, hvilket nok gør den bedre når $v = 2$
- Forskelligt antal stacks
- PCK med fixed normalization konstant straffer folk tættere på kammeraet mere end folk længere væk

1.4 Future Work

If we were to work further with this project, it would be ideal to explore the effects of stacking multiple modified hourglasses end-to-end. By doing so we would not only hope that the performance of the model to increase further, but we would also hope we could obtain the same accuracy as Newell *et al.* experiences [1], however with fewer stacks. For instance, we could hope that by stacking 2 modified hourglasses, we would achieve the same results as Newell *et al.* achieves with 4 standard hourglasses.