

## **Project Memo (11/01/2018)**

Our project is on track according to the proposal. We have completed the Computer Vision part (self-supervised task for classifying frame misalignment between a pair of frames within a video). We implemented the embedding network to generate embeddings for frames, and a classification network to classify the misalignment between two such images (i.e., embeddings) on the Moving MNIST data set, and another data set generated by us where we have greater control over the parameters of the environment. For both, we are getting classification accuracies over 90%. Furthermore, we also trained a decoder on top of the embeddings to reconstruct the image, which are generally also very good reconstructions (and the corresponding MSE losses are very low).

Our next steps are to generate t-SNE visualizations of the learnt embedding space to see if the distances make sense, and then to move on to the Reinforcement Learning part of the project, where the task is to initialize the embeddings with the learnt ones and train an agent to play games (for now, we would be considering the game of Pong from Atari) by watching a very small set of expert demonstration videos.