

Toward image inbetweening using Latent Model



Paulino CRISTOVAO, Yusuke TANIMURA
Hidemoto NAKADA and Hideki ASOH
University of Tsukuba

National Institute of Advanced Industrial Science and Technology



1

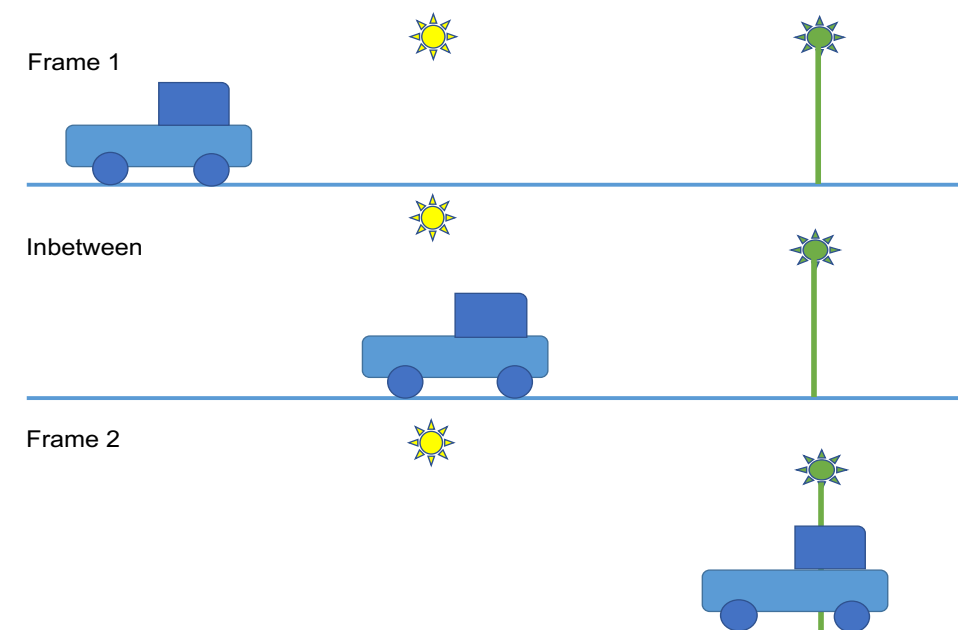
Motivation

- Generate image inbetween based on Generative models using variational autoencoders.

Restrictions with previous approaches

- Cannot capture what is not present in the picture
- Often produce blurry image

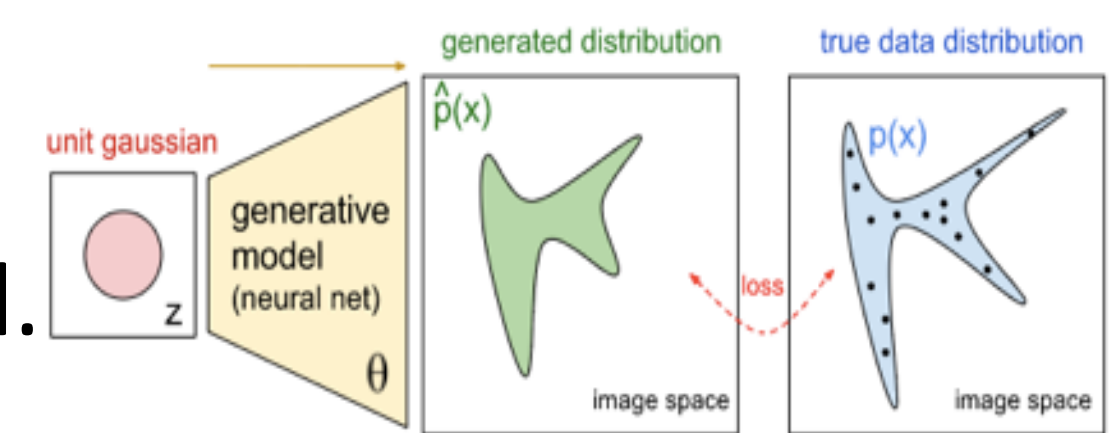
Long Term Goal – Image inBetween



2

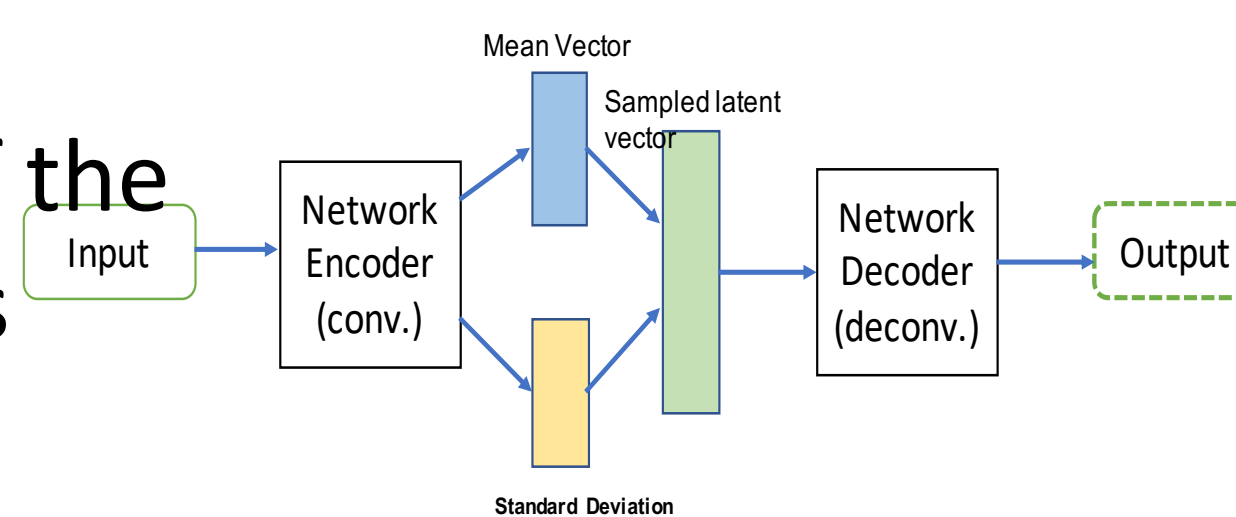
Generative Models

- Learns rich and hierarchical probabilistic model.



Variational Autoencoder (VAE)

- Learns a latent representation of the hidden structures of its input data.

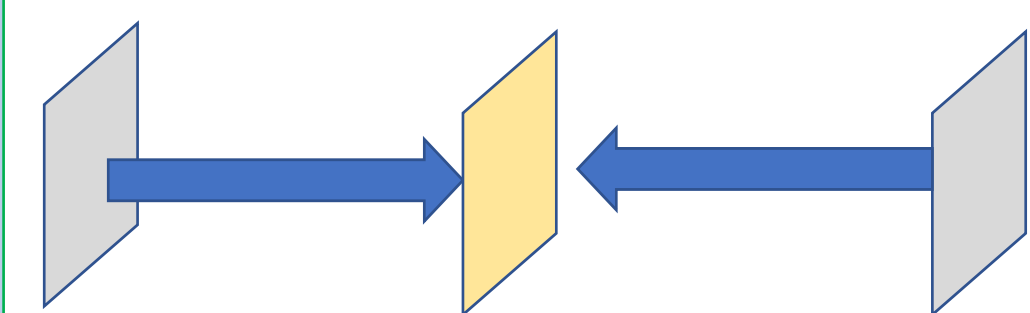


3

Proposed Method - Latent-variable based inbetweening

Existing methods: 'pixel-based'

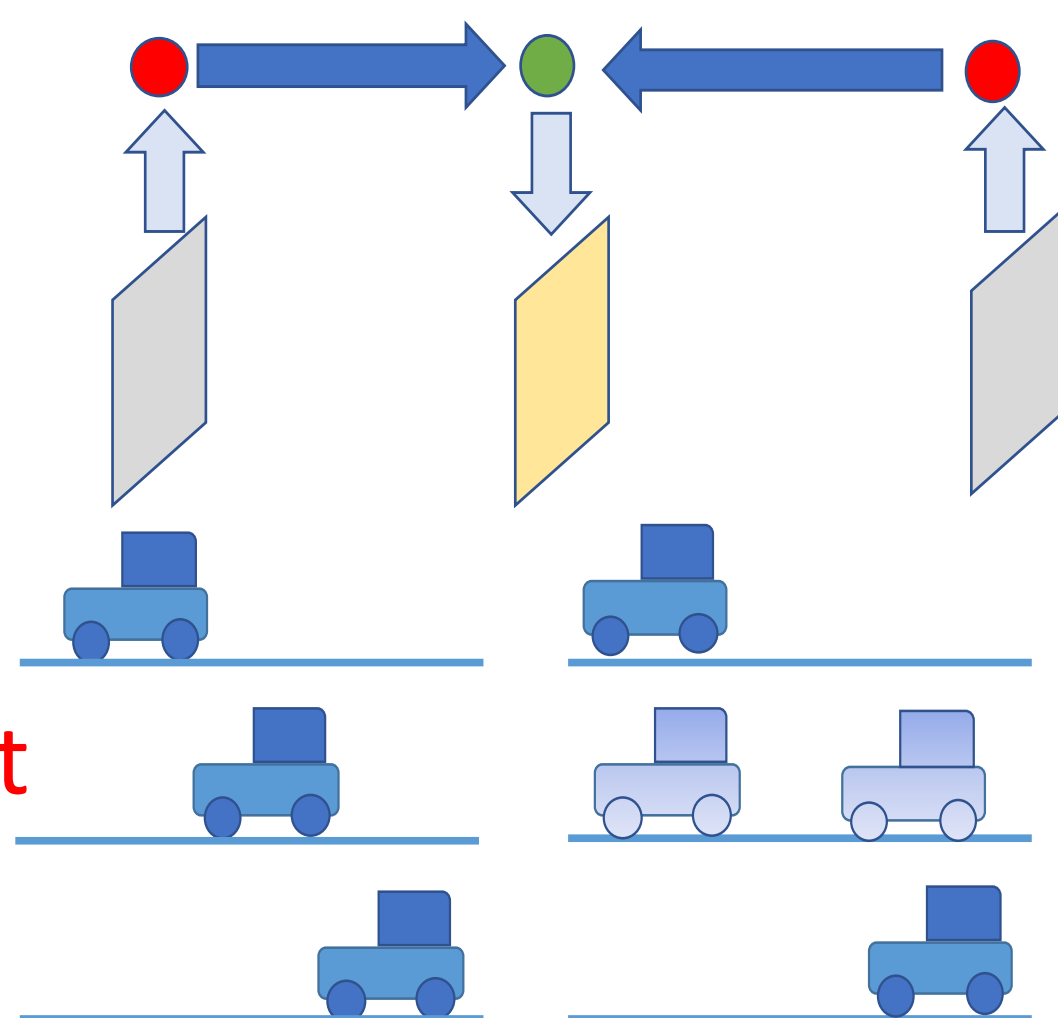
- Directly generate inbetween frame, from other frames



Need to control the latent space structure

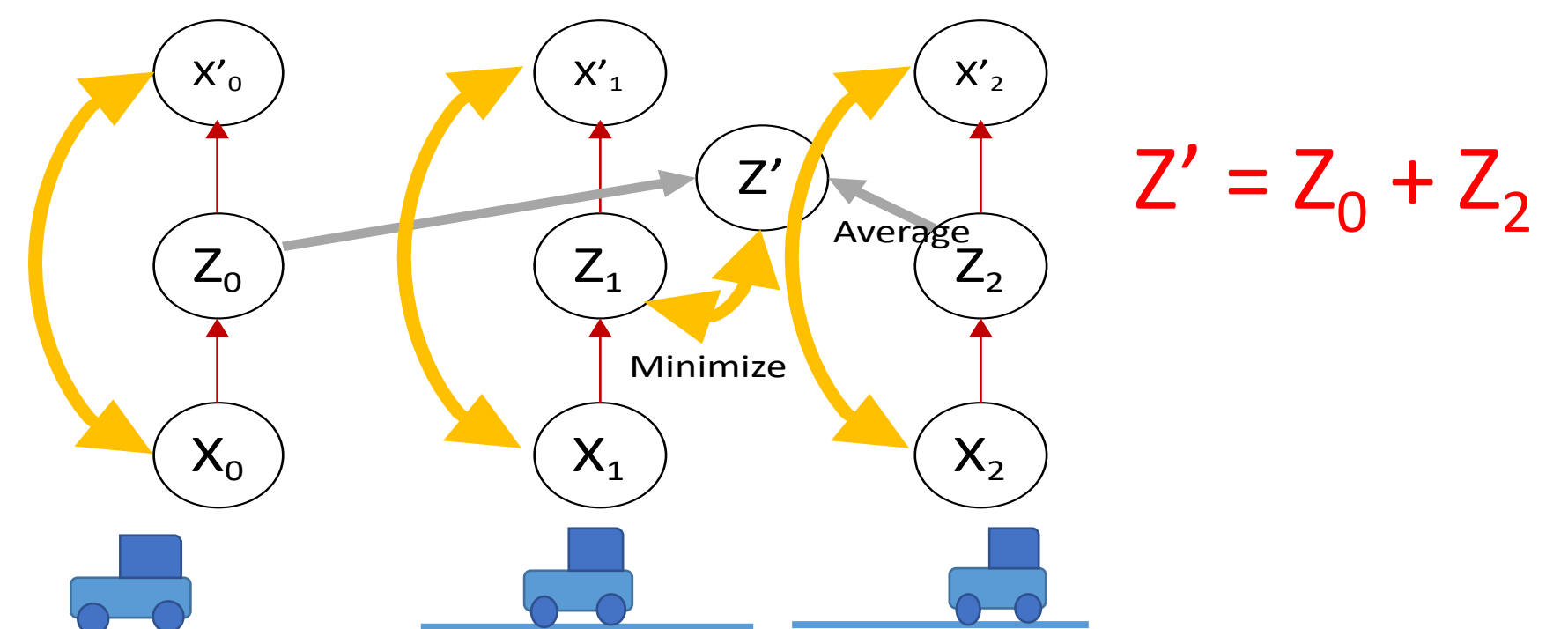
'Latent-variable' based method

- Interpolate in the latent space and generate inbetween frame



4

How to have 'such' latent space?



Proposed – Loss Function

$$l_{(x_0, x_1, x_2)} = l_{VAE}(x_0) + l_{VAE}(x_1) + l_{VAE}(x_2) + \alpha \left(D_{KL} \left(q(x_1) \parallel \frac{q(x_0) + q(x_2)}{2} \right) \right)$$

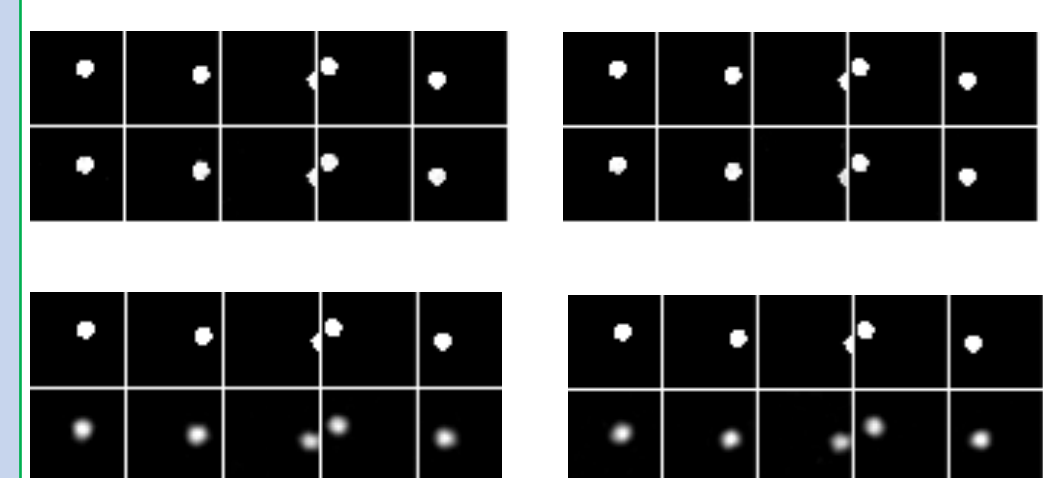
Goal Minimize: difference of (Z_1 and Z')

5

Evaluation

- Image reconstruction
- Image inbetween

Image Reconstruction



Dataset

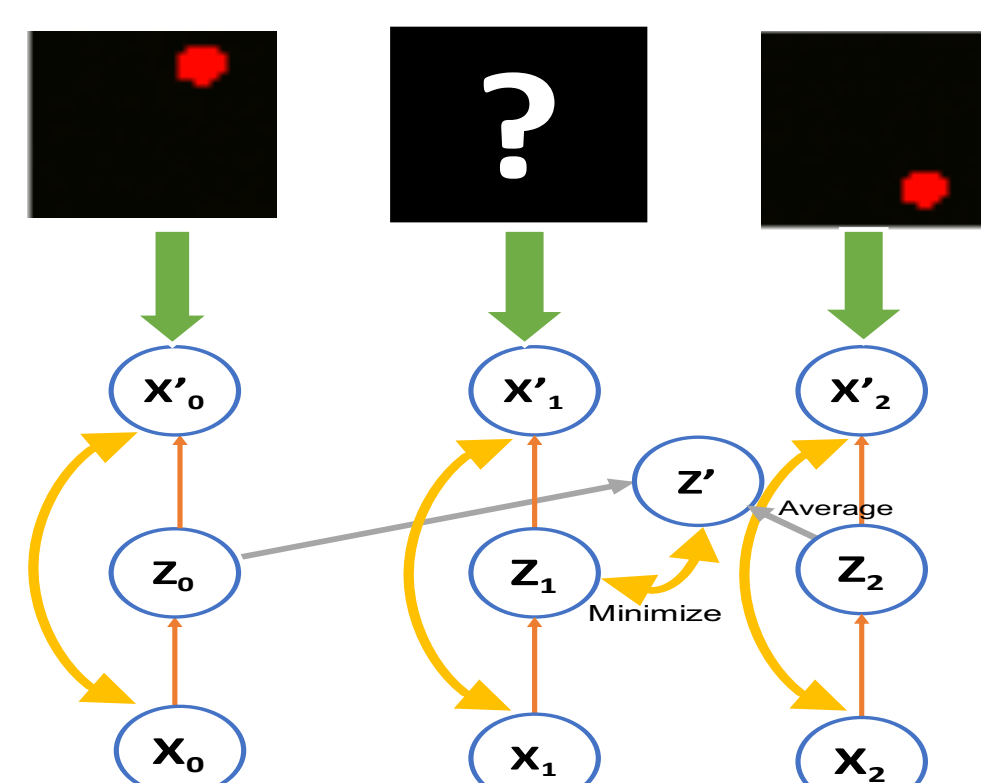
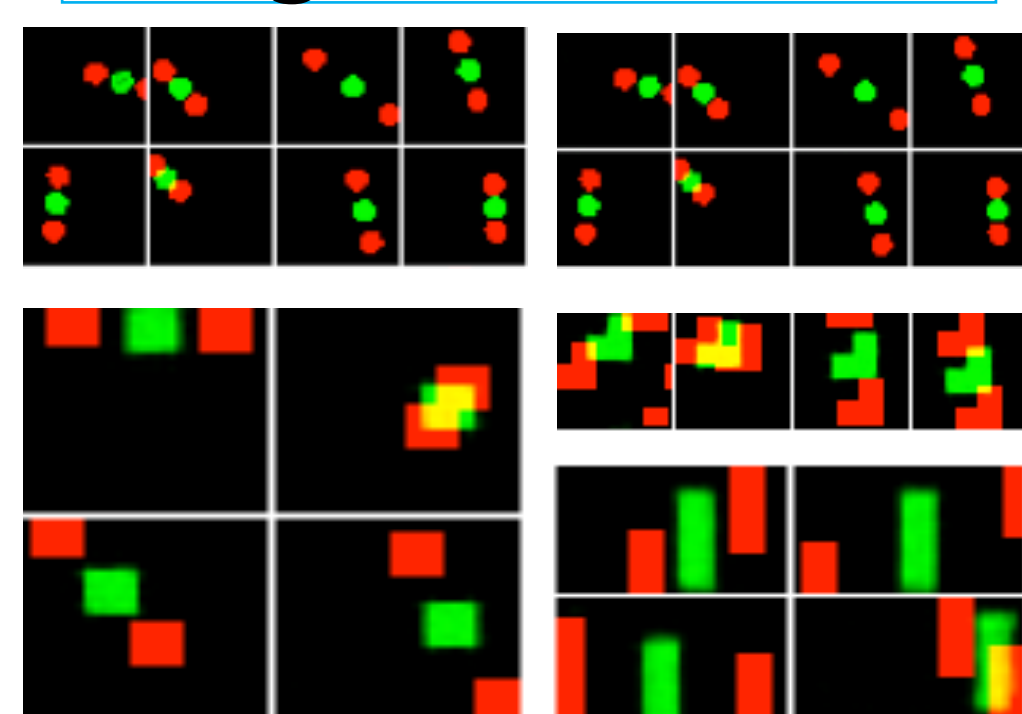


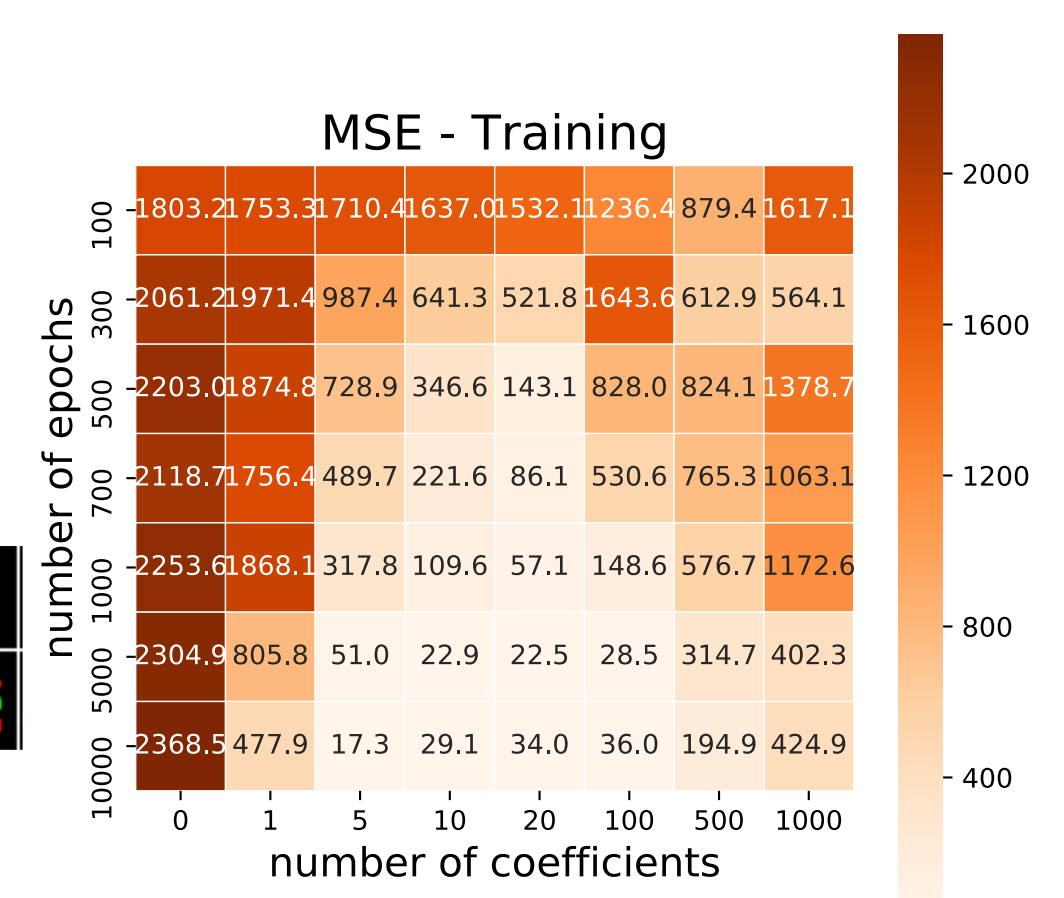
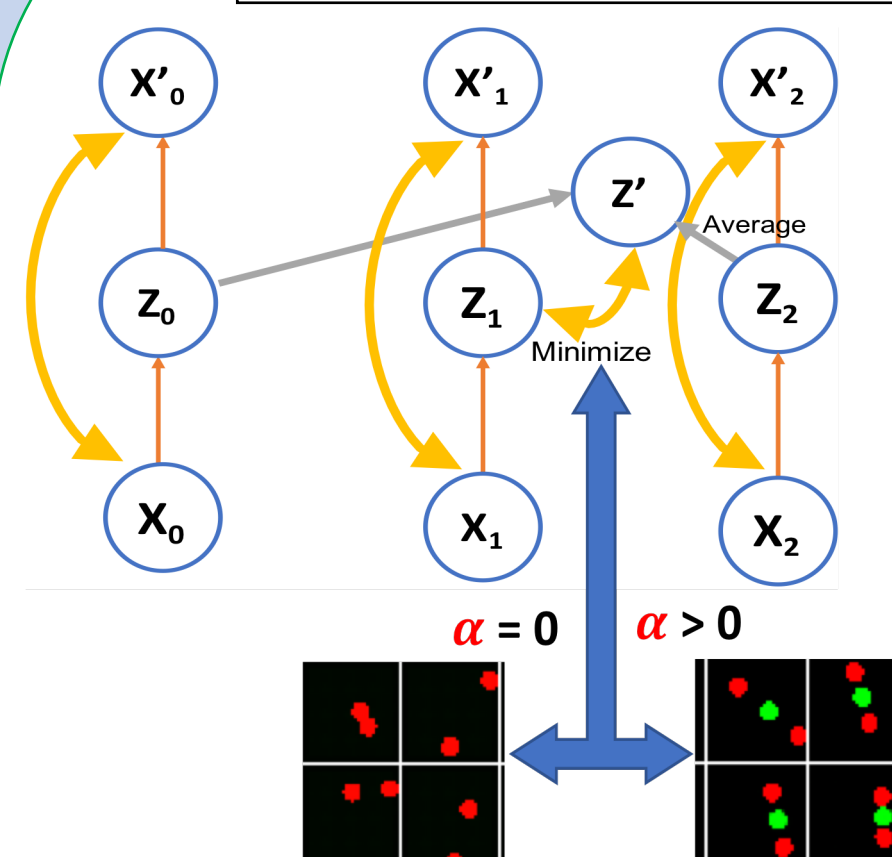
Image Inbetween



6

Contribution

Evaluation - MSE



- Drawback: Network loss increases

Finding:

- With zero coefficient we have no reconstruction
- Increasing the coefficient network loss increase

7

Conclusion

- Model predicts the spatial location of the object. With the strong interpolation term we still can reconstruct a fair image. For future, we will work with complex objects and video.

Acknowledgement

This paper is based on results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).
This work was supported by JSPS KAKENHI Grant Number JP16K00116