Image Frame Reconstruction Using Event Data

Rishabh Samra*

Dr. Kaushik Mitra

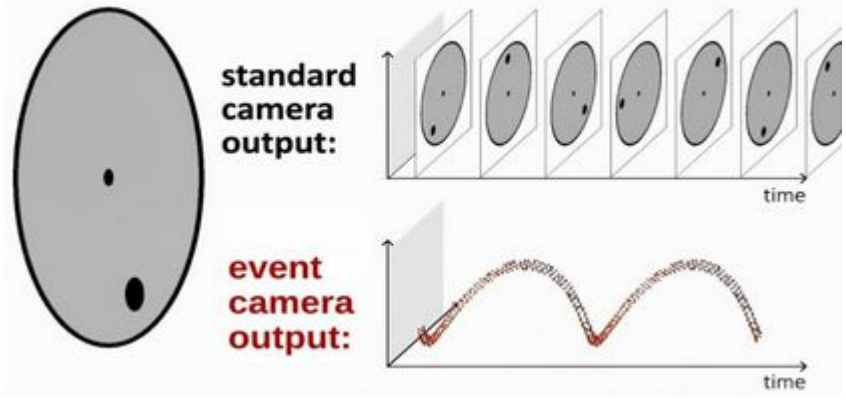
ACKNOWLEDGEMENT

STANDARD CAMERA:

- Pixels: Globally triggered
- Output: Full image frames at fixed frame rate

EVENT CAMERA (LIKE DVS):

- Smart pixels: Independent and Asynchronous
- Output: Sequence of events (local brightness changes)



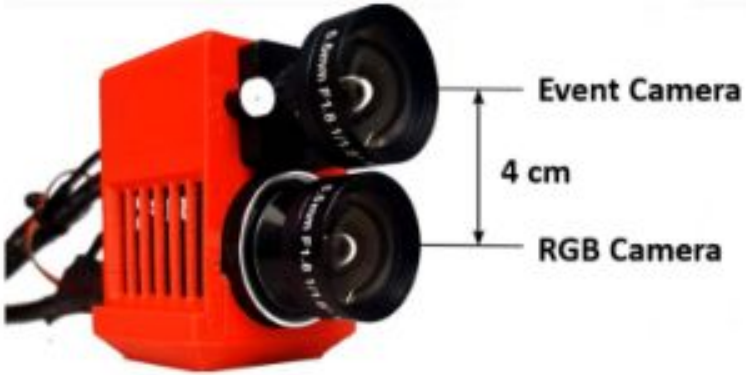
Properties of an Event Camera

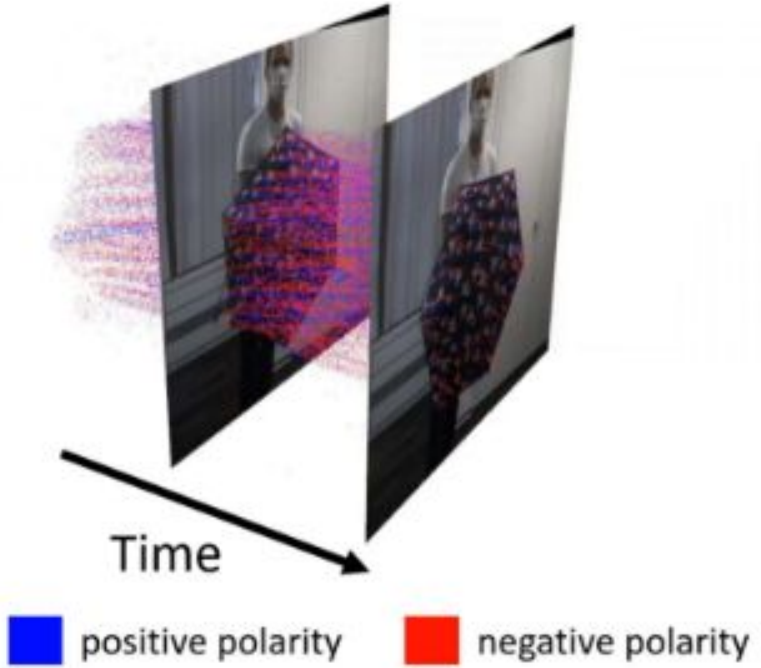
- Novel sensor that measures motion in the scene.
- Low latency(~1 μs)
- High Dynamic Range(140 dB instead of 60 dB)
- Ultra-low Power(mean: 1mW vs 1 W)

Traditional Vision Algorithms could not be used because :

- Asynchronous Pixels
- No intensity information
  (only binary intensity changes)

EXPERIMENTS



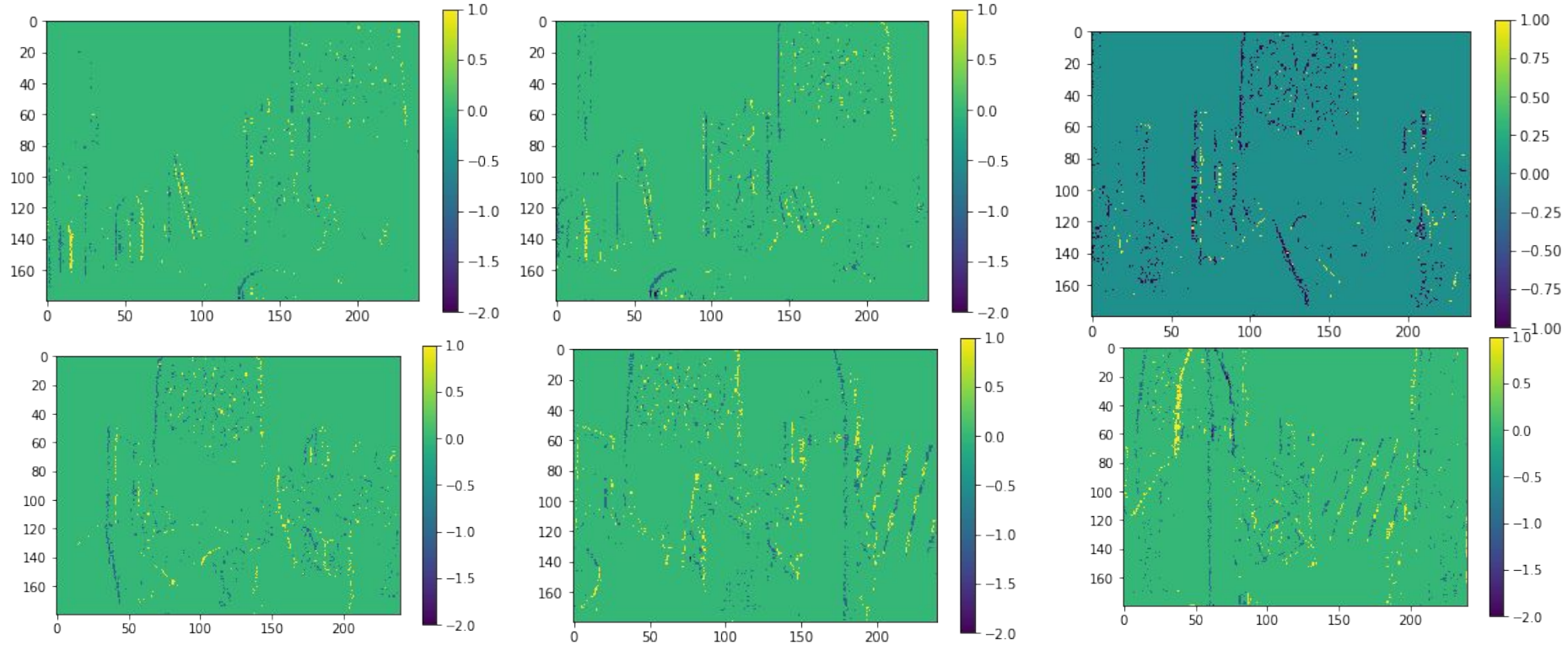


1. SLIDER_DEPTH DATASET

Ground Truth Images

Reference: *http://rpg.ifi.uzh.ch/davis_data.html*

EVENT FRAMES

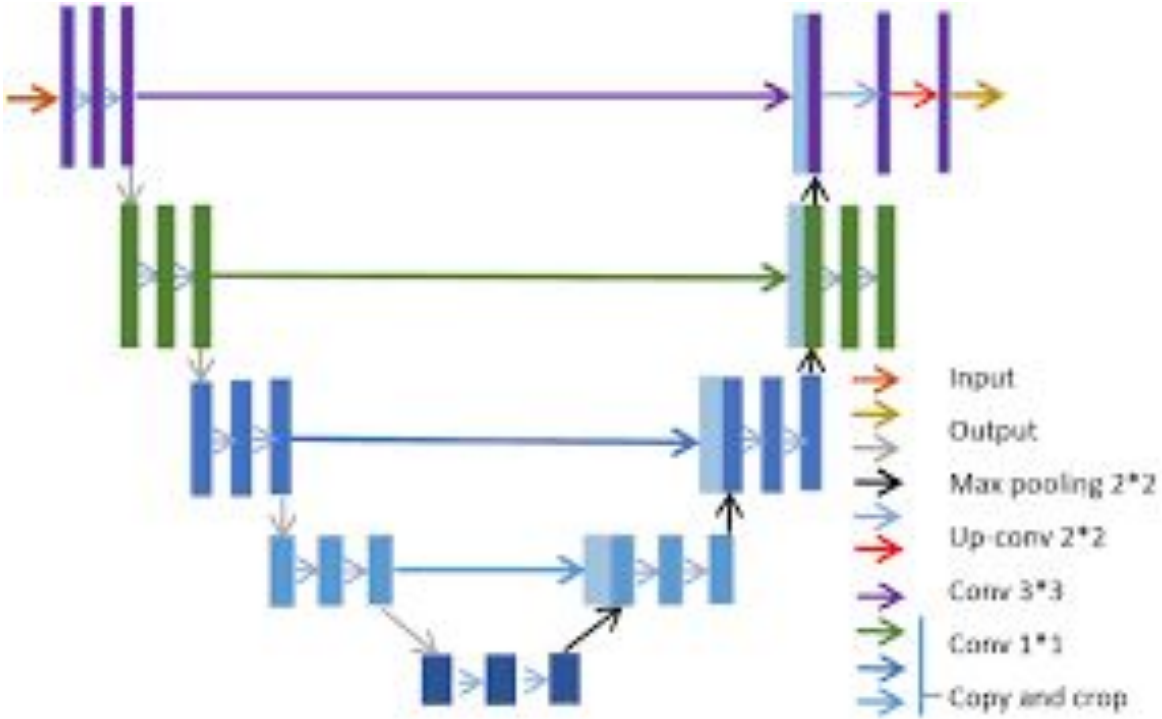


(Sequence from Left to Right) : 0 , 81 , 330 , 458 , 635 , 766

SLIDER_DEPTH DATASET

TRAINING WITH A U-NET ARCHITECTURE

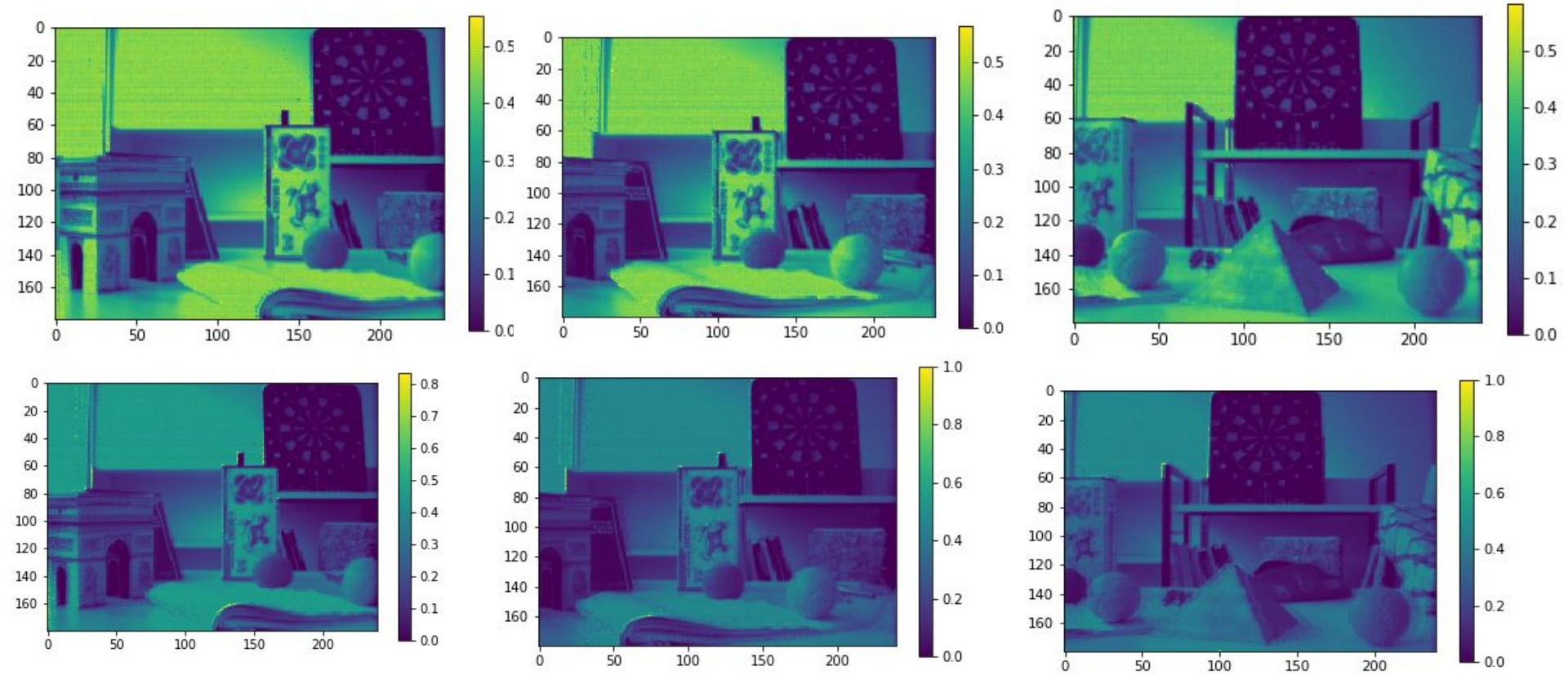
- U-Net Architecture was used to train a model where data is the images generated using forward formula, $I_n = I_0 \exp\{\gamma(E_1+E_2+...+E_n)\}$ and labels were generated using backward formula $I_n = I_m \exp\{-\gamma(E_{n+1} +...+ E_m)\}$. (m>n); $\gamma = 0.2$



Adam Optimizer
(Beta =0.9,0.999)

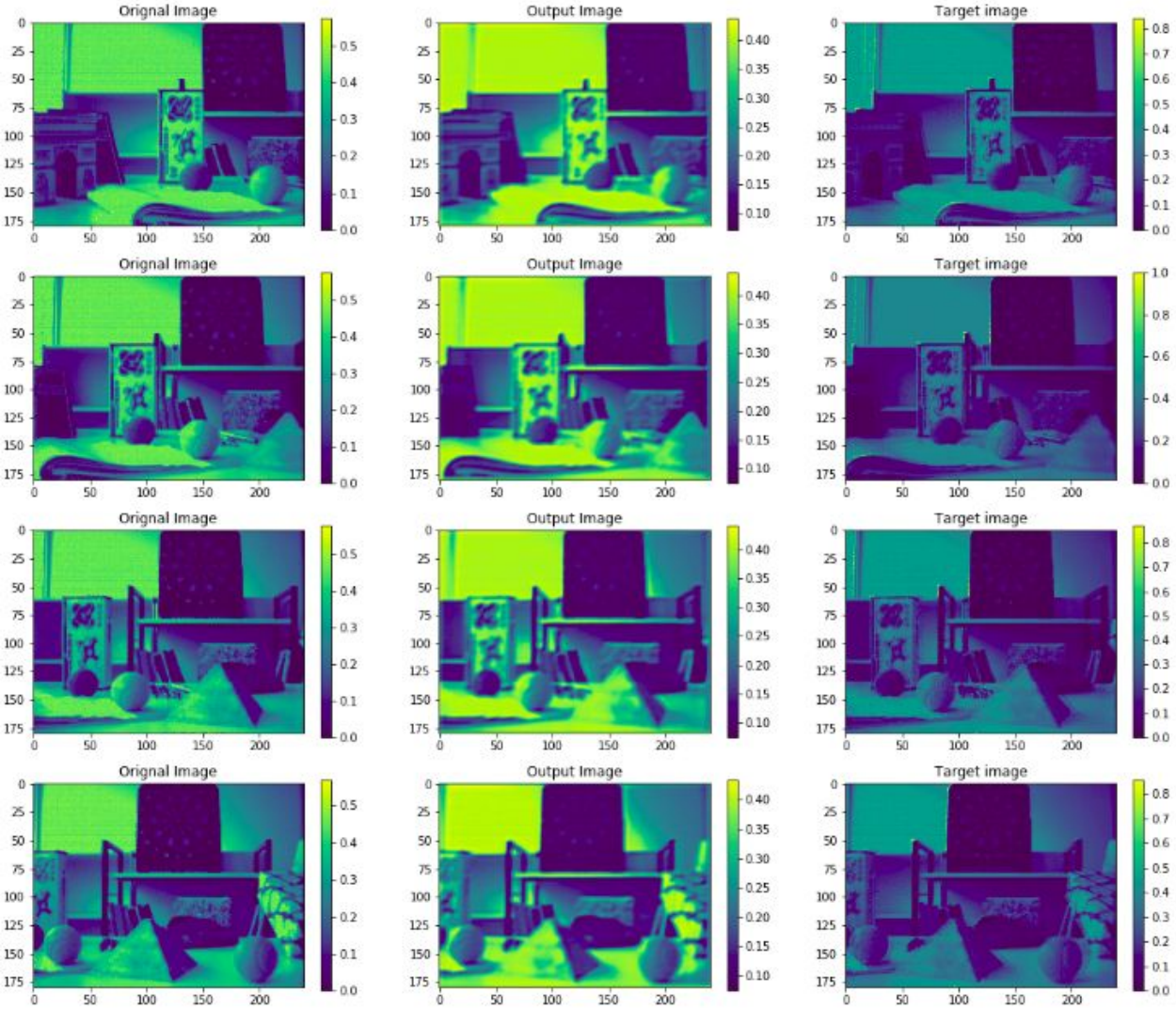
Learning Rate:0.0001
Exponential LR with $\gamma = 0.9$

Intermediate Image Frames Generated Using Forward & Backward Formula

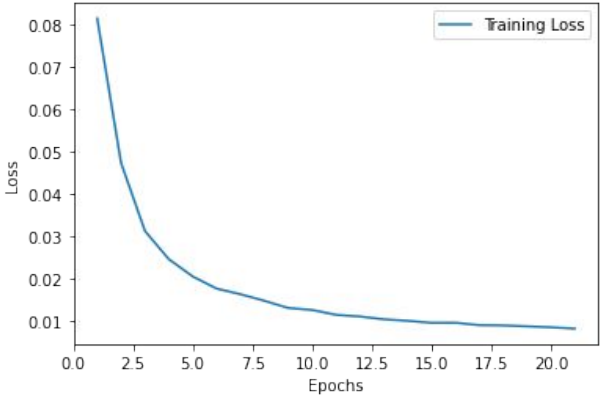


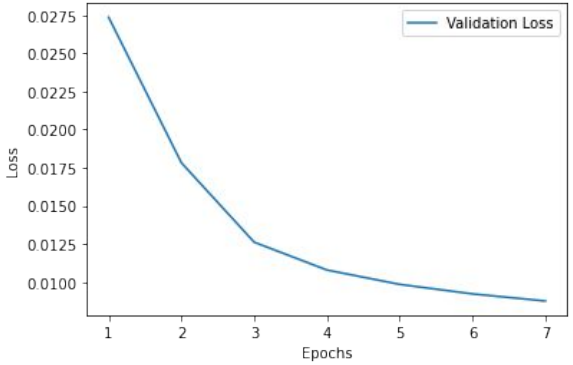
UP: Data **Down**: Label

0 , 81 , 330


Orignal Image
Output Image
Target image
Orignal Image
Output Image
Target image
Orignal Image
Output Image
Target image
Orignal Image
Output Image
Target image

RESULTS

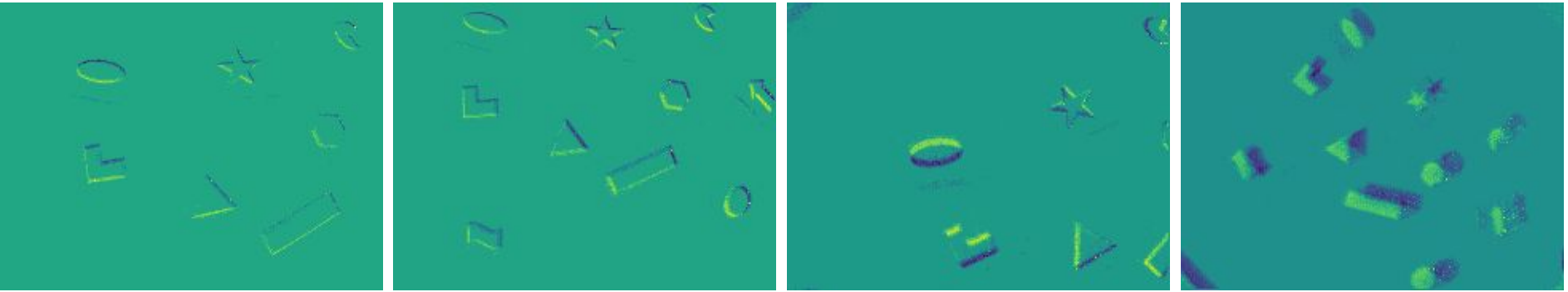

Training Loss
Loss
Epochs


Validation Loss
Loss
Epochs

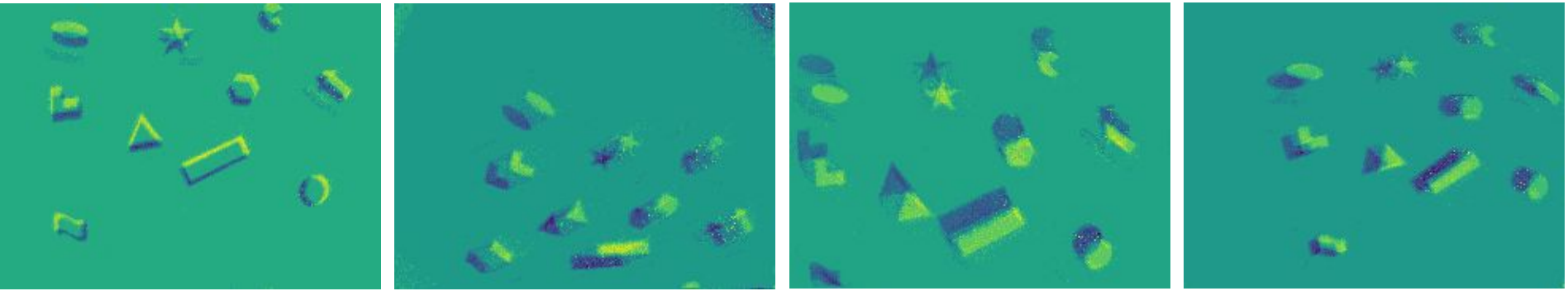
2. SHAPES_6DOF Dataset

EVENT FRAMES

4, 206, 409, 650

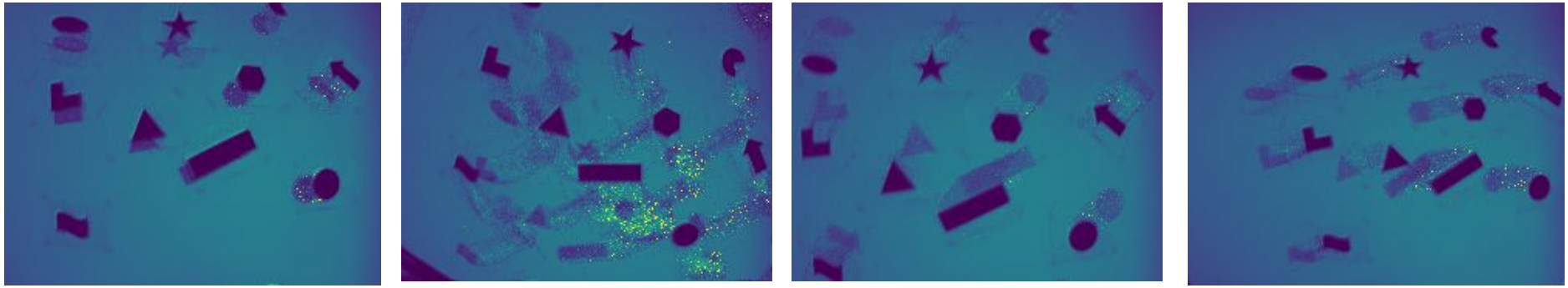


802, 967, 1100, 1321

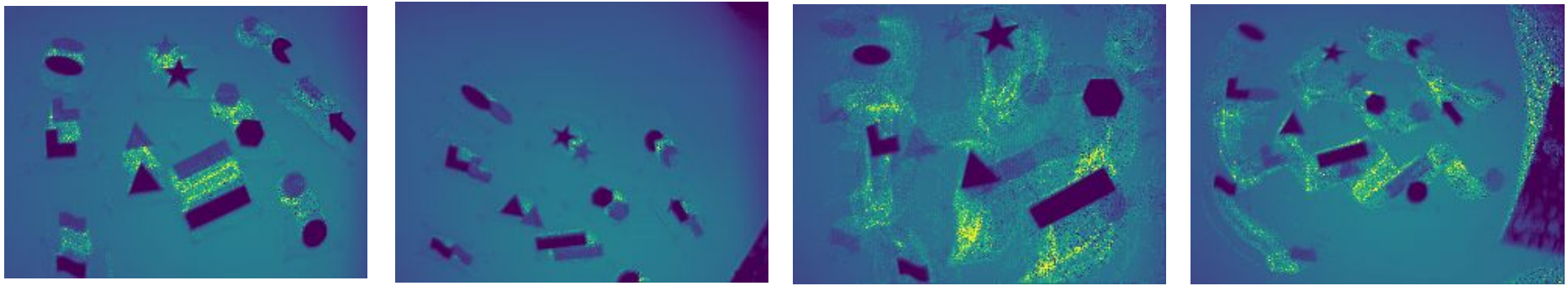


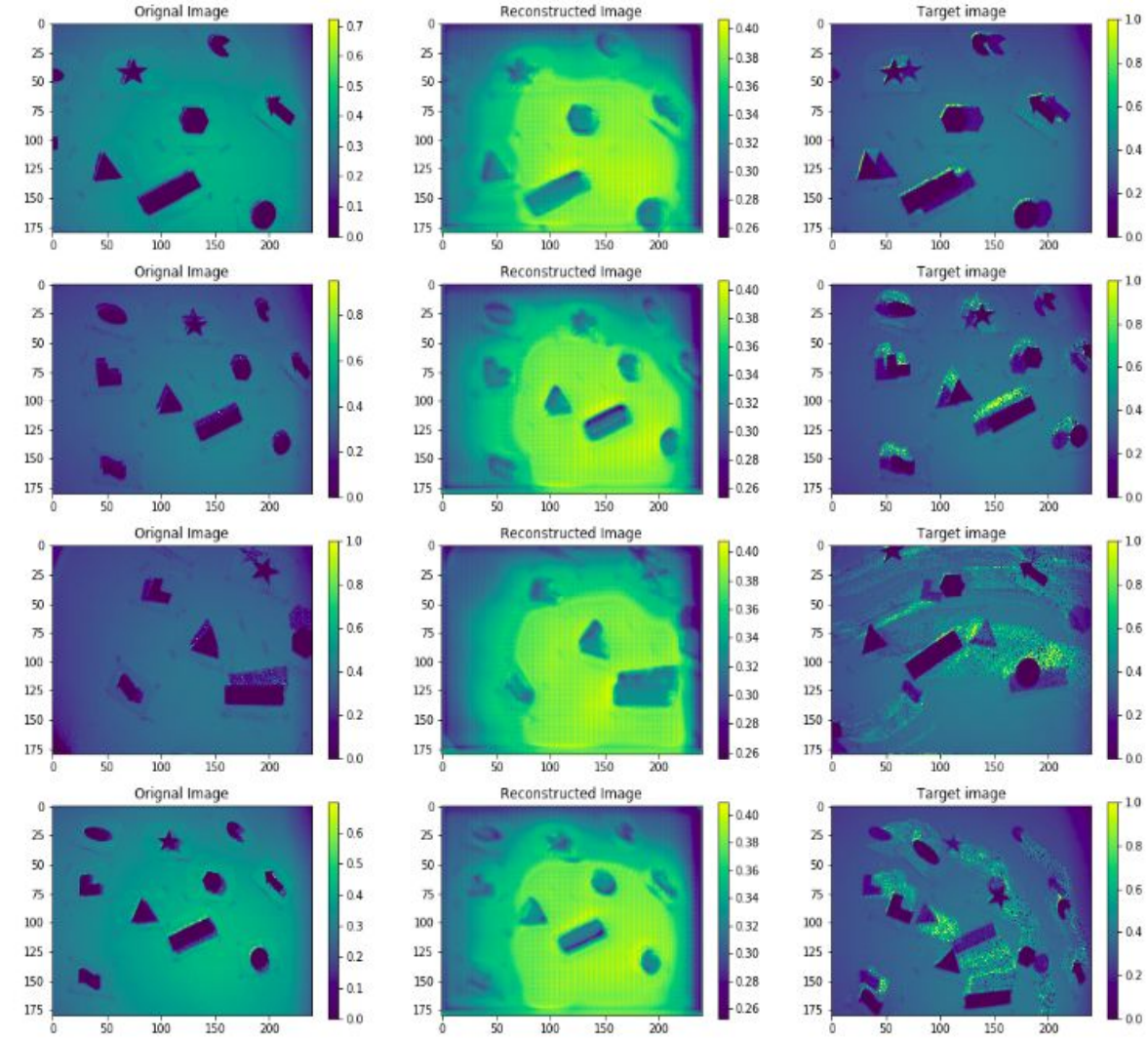
Data

802, 967, 1100, 1321

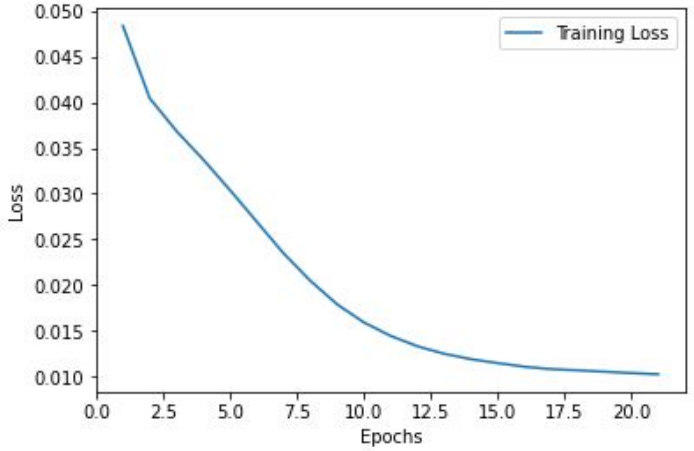


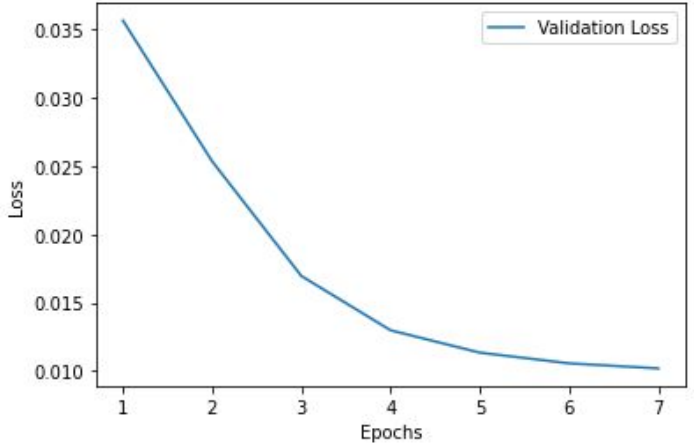
Labels




Orignal Image
Reconstructed Image
Target image
Orignal Image
Reconstructed Image
Target image
Orignal Image
Reconstructed Image
Target image
Orignal Image
Reconstructed Image
Target image

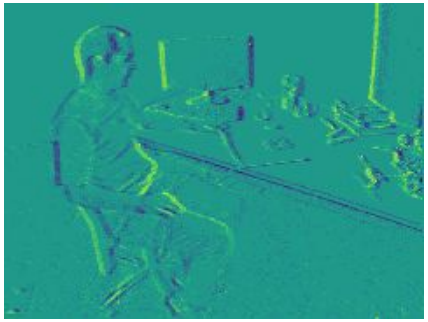
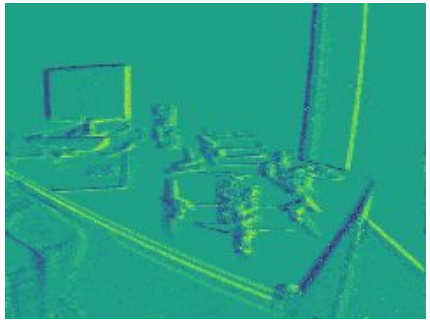
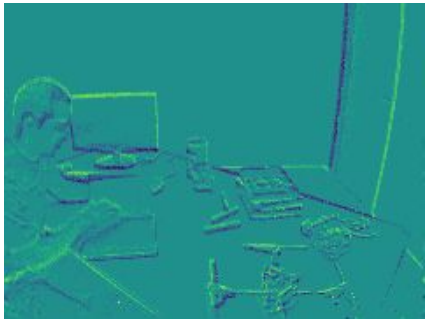
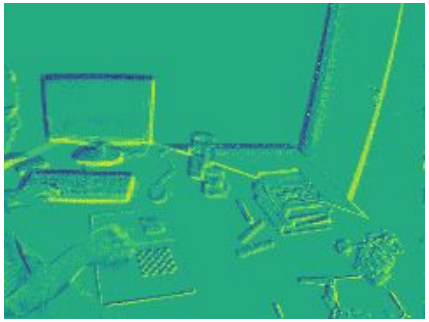
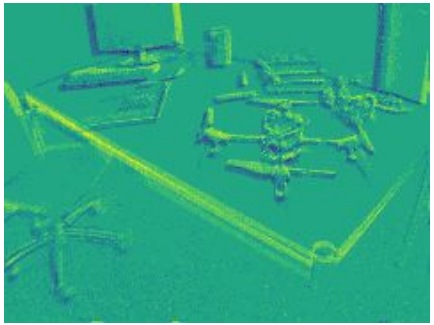
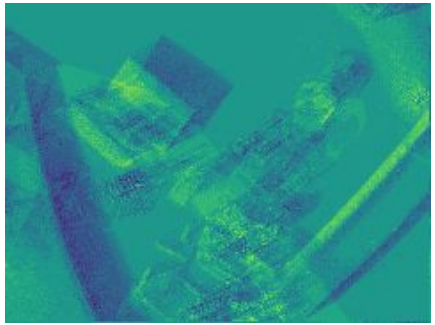
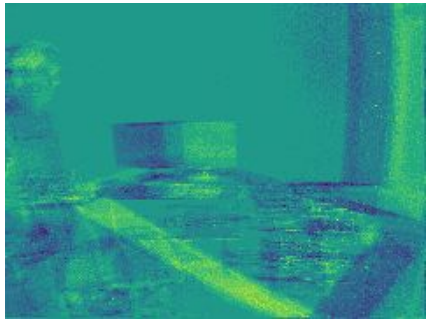
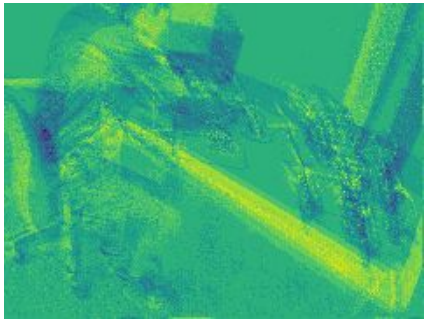
Results


Training Loss
Loss
Epochs


Validation Loss
Loss
Epochs

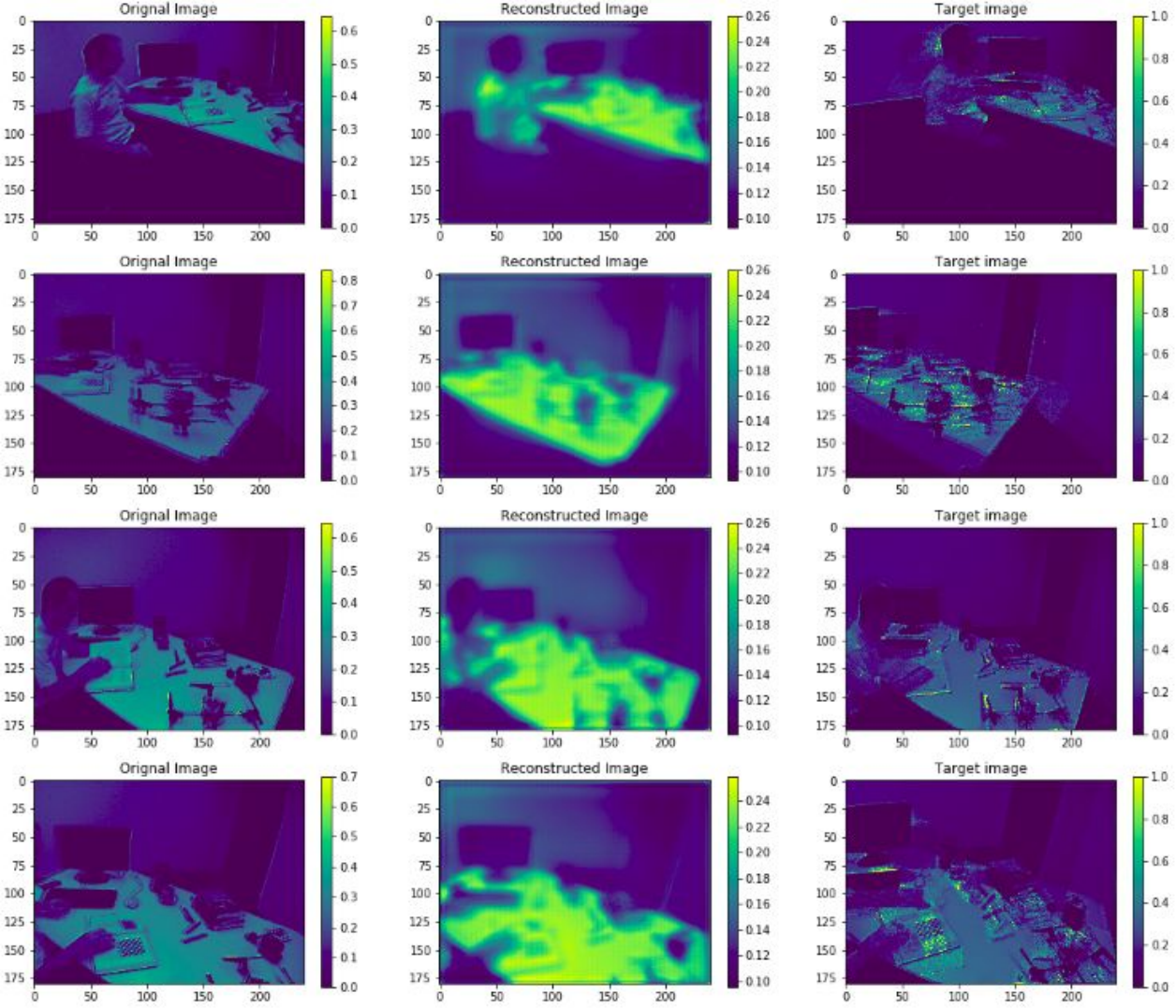
3. Dynamic_6DOF Dataset

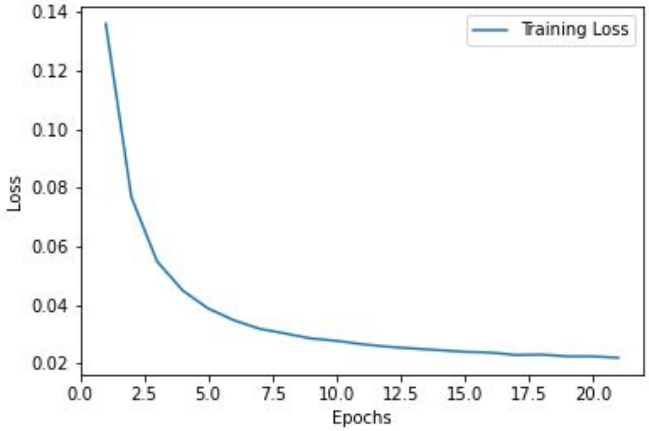
Event Frames

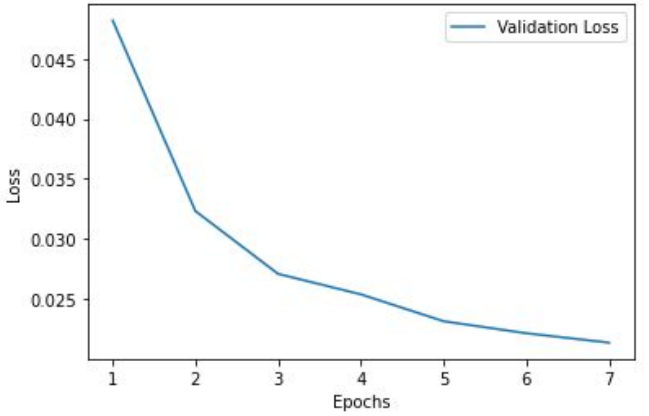









(Sequence from left to right): **50 , 200 , 350 , 500 , 650 , 800 , 950 , 1250**

Results


Orignal Image
Reconstructed Image
Target image
Orignal Image
Reconstructed Image
Target image
Orignal Image
Reconstructed Image
Target image
Orignal Image
Reconstructed Image
Target image

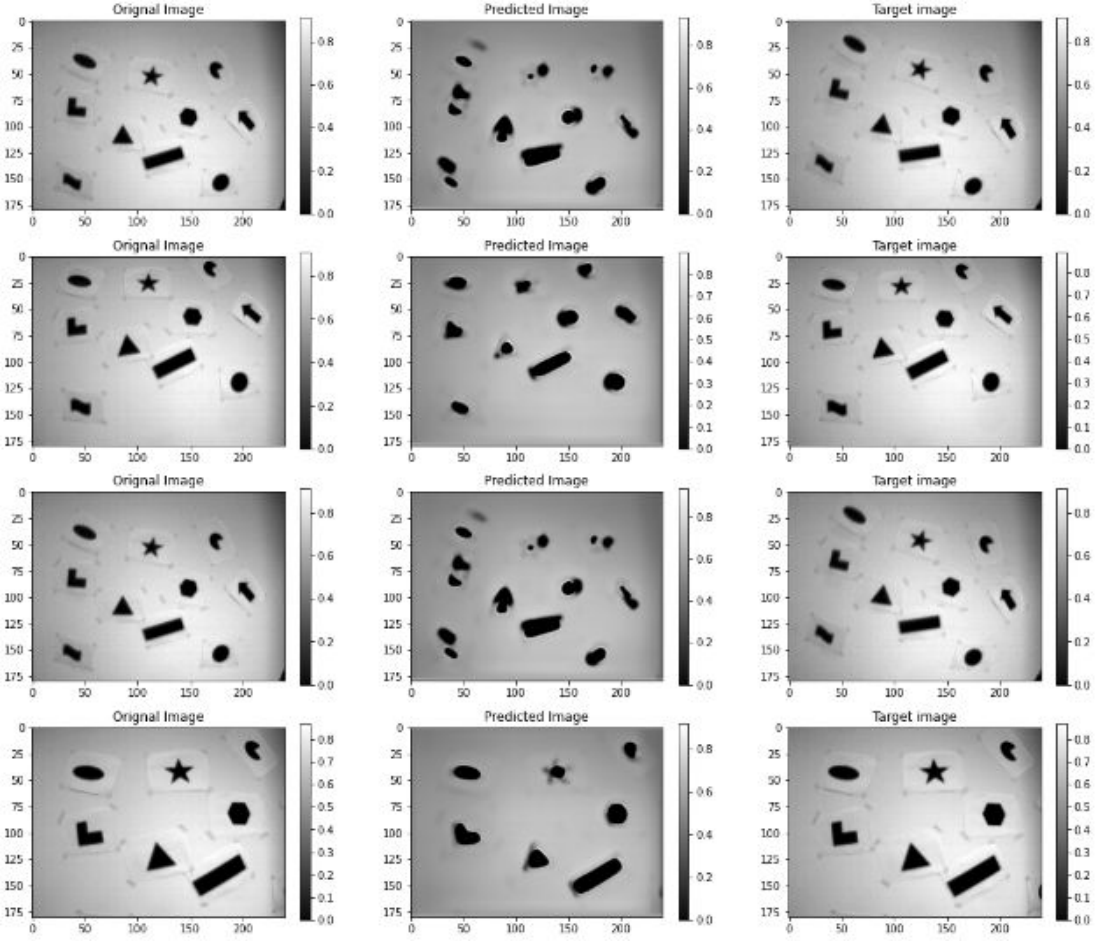

Training Loss
Loss
Epochs

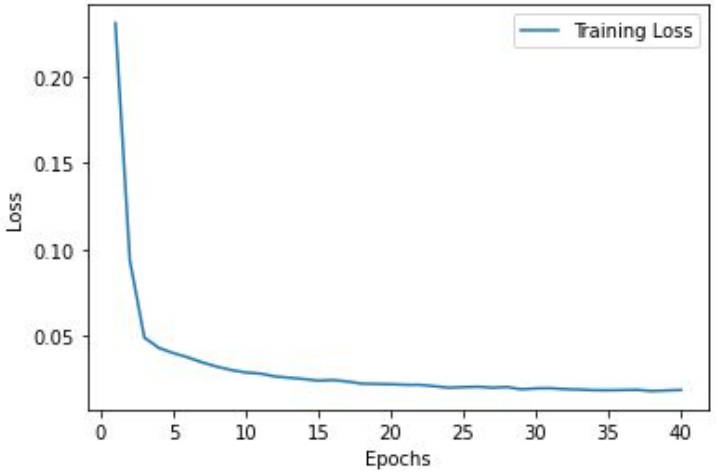

Validation Loss
Loss
Epochs

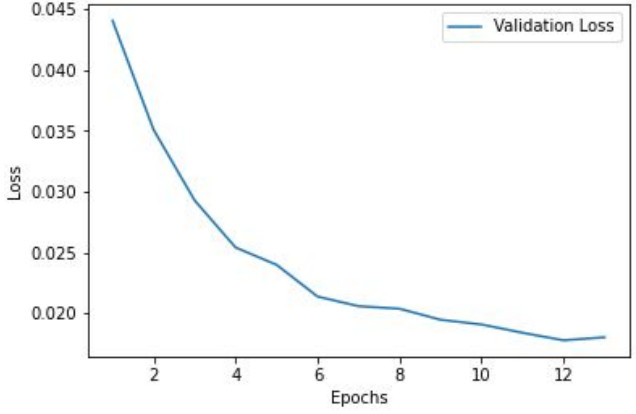
	MSE	
	Previous Methods	**Ours**
Slider_Depth	0.05	0.01
Shape_6dof	0.02	0.01
Dynamic_6dof	0.05	0.021

A FEW MORE EXPERIMENTS

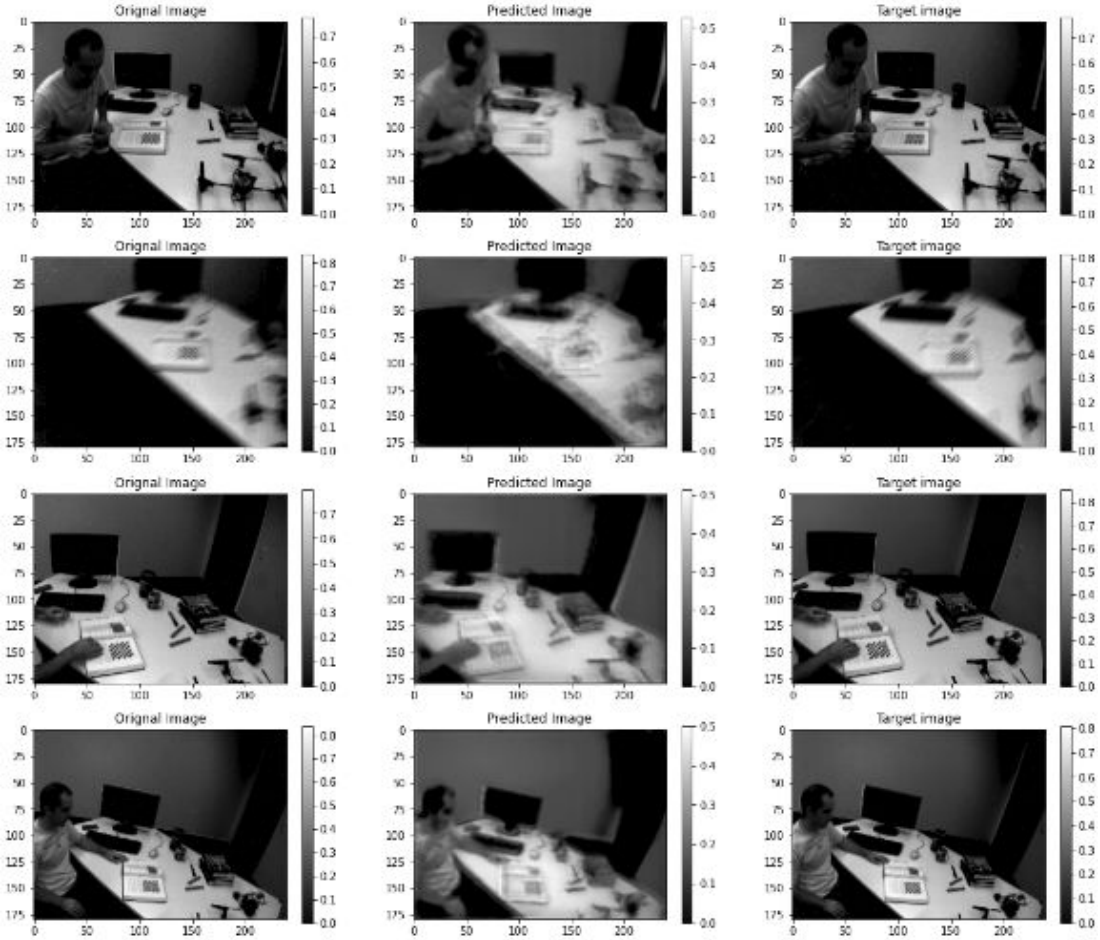
Input: (I0,E0) **Label:** I1

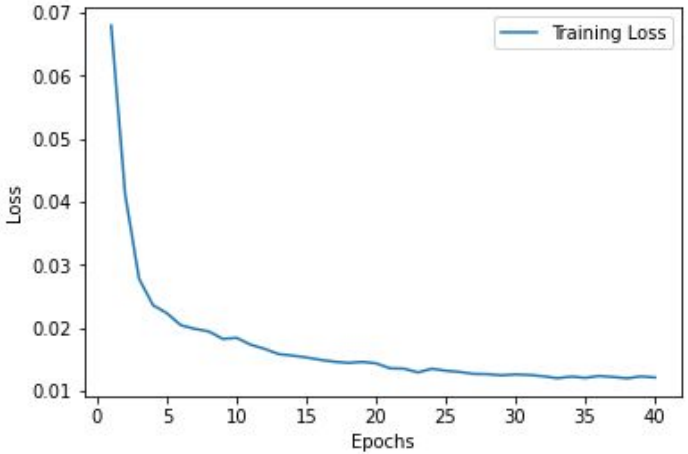


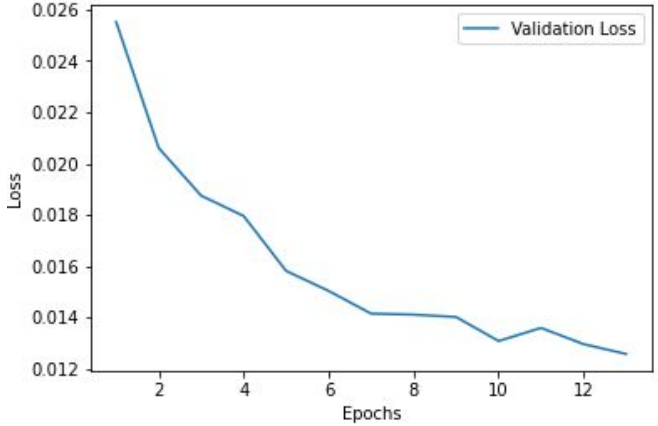




Input: (I0,E0) Label: I1 (Normalized)

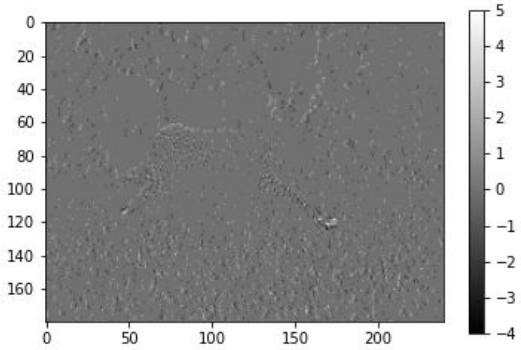


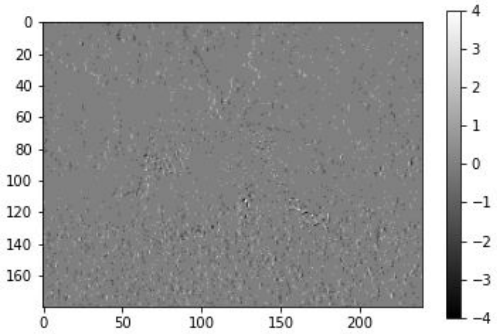


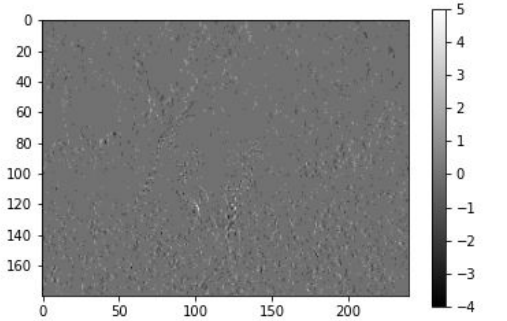


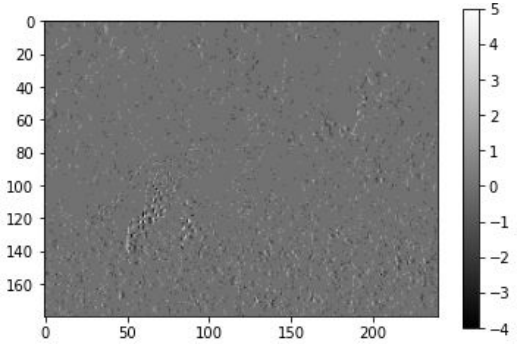
Cheetah_Running

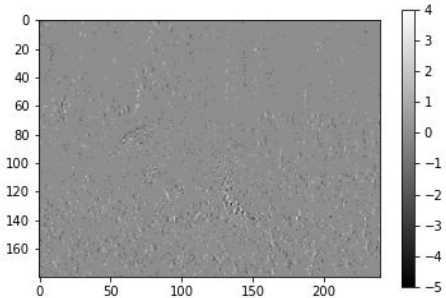
Event Frames

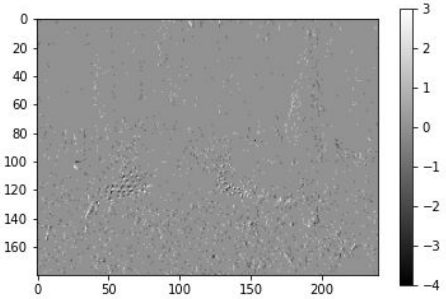








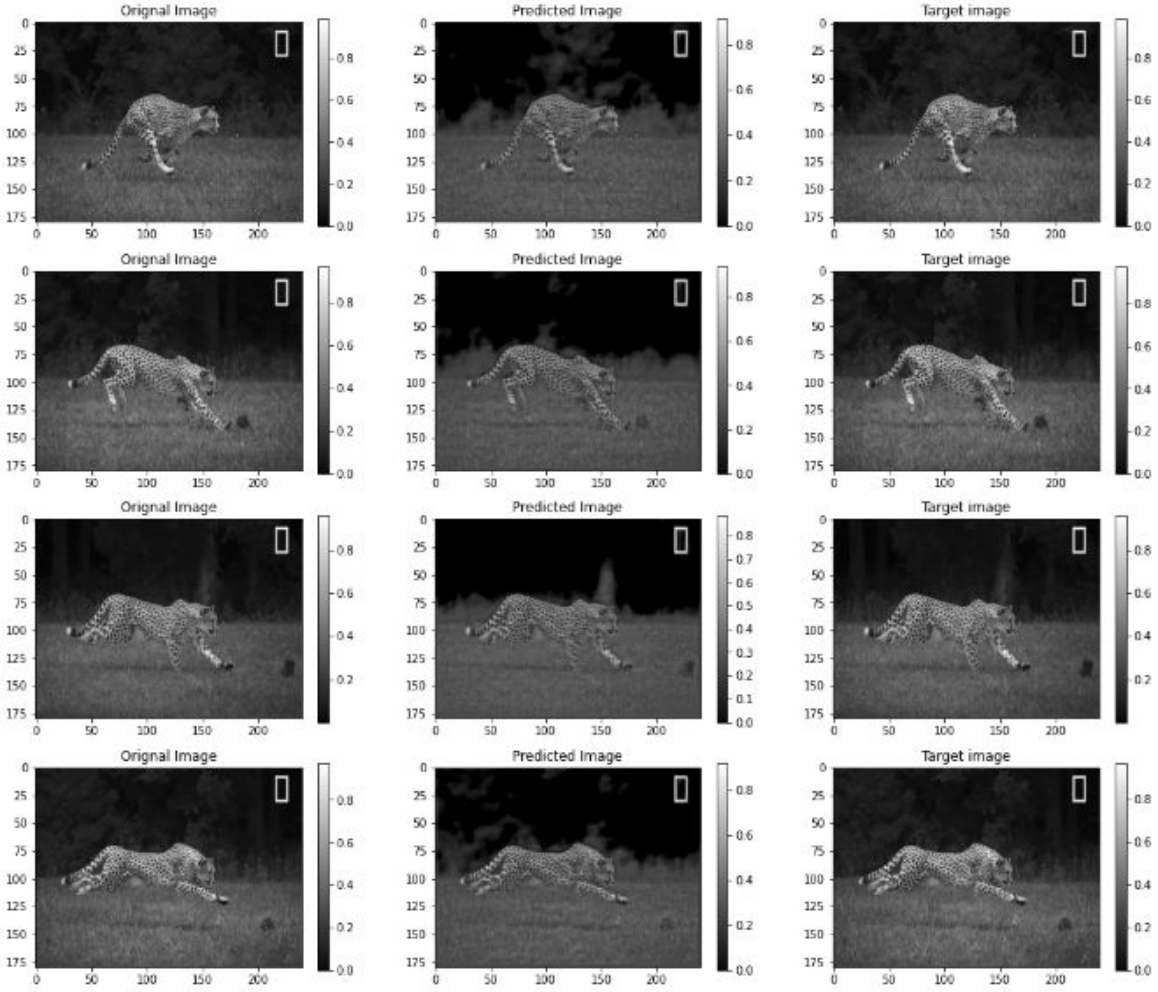


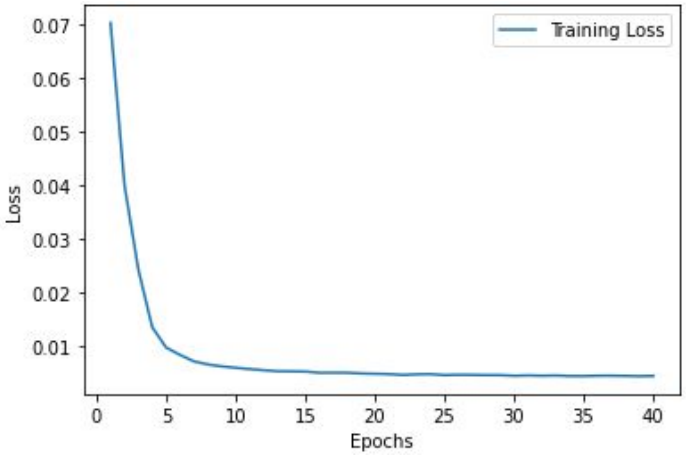


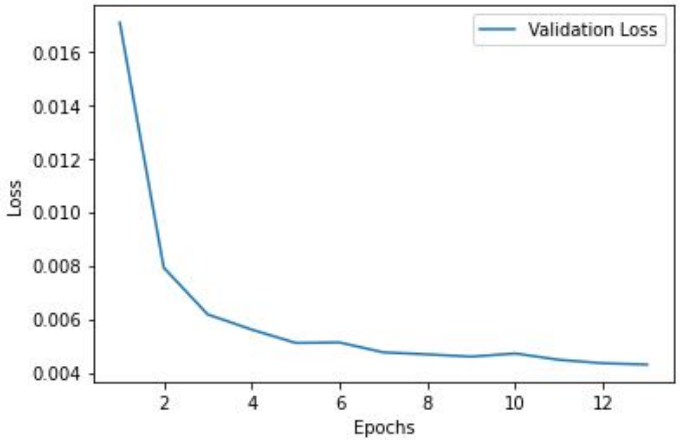
2 , 75 , 158 , 334 , 587 , 825

Source: [ESIM Simulaor: events from video]

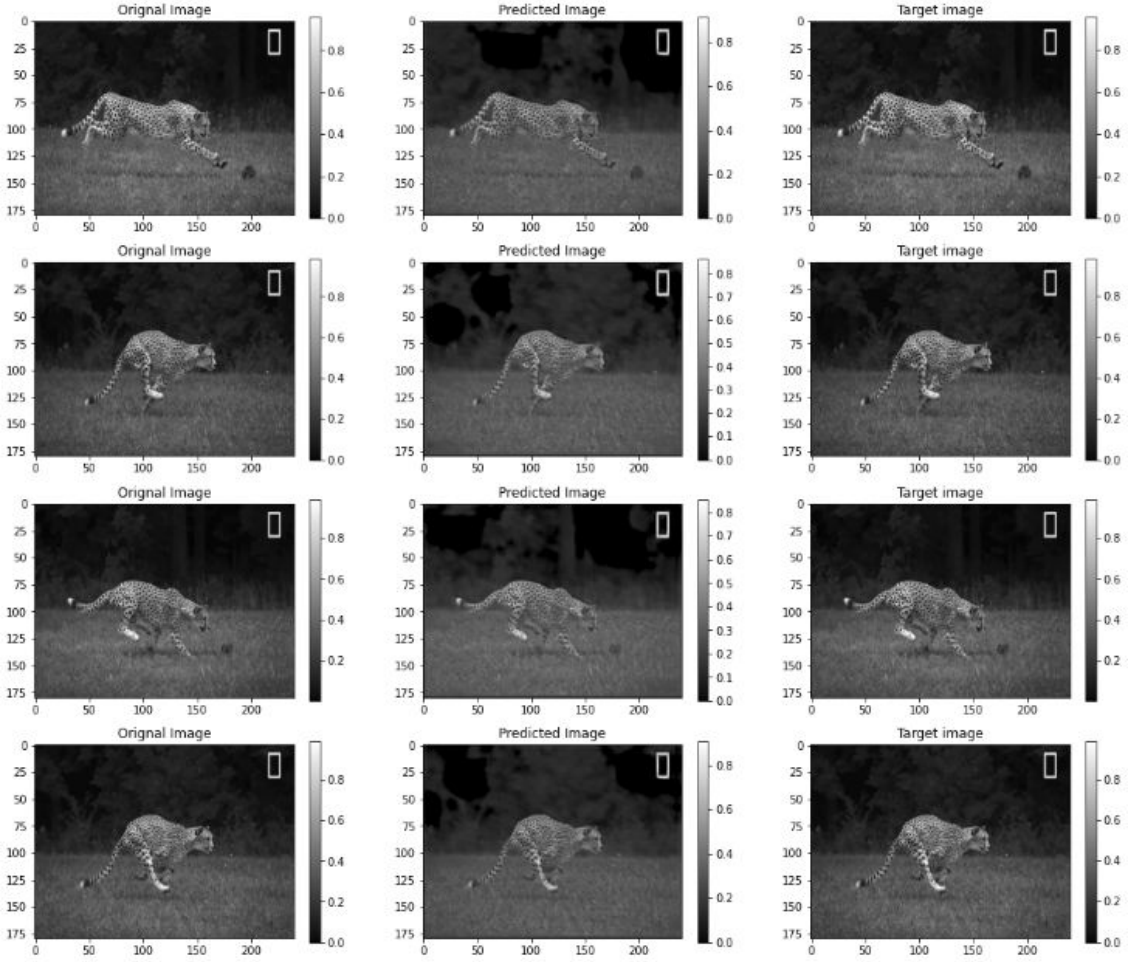
Input: *(I0,E0)* ***Label:*** *I1* (L2 Loss)

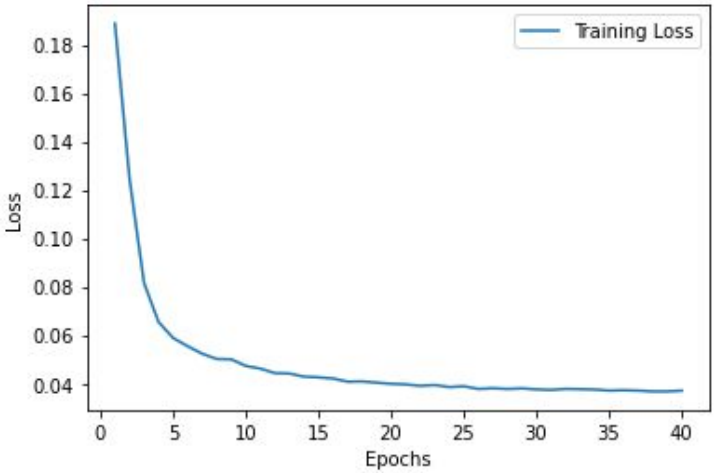


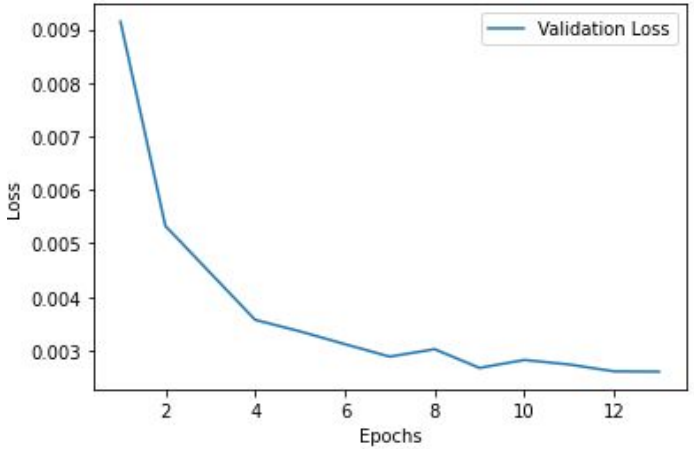




Input: *(I0,E0)* **Label:** *I1* (L1 Loss)







Input:(I0,E0….E9) **Label:** I10