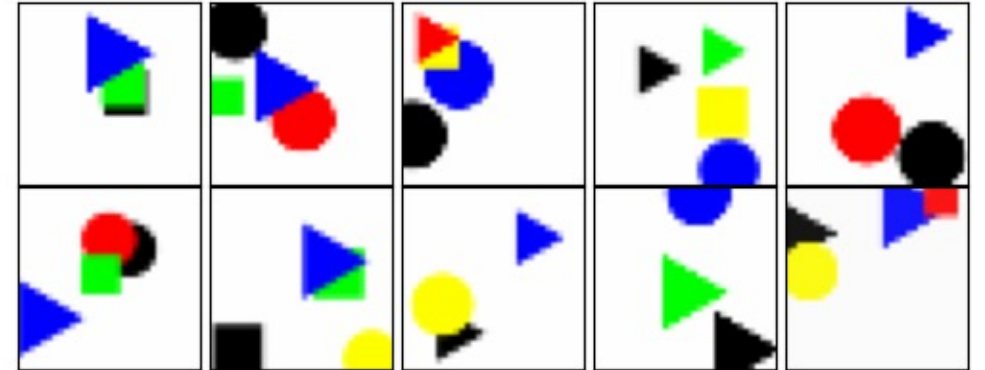# CGAN Project

Project Presentation
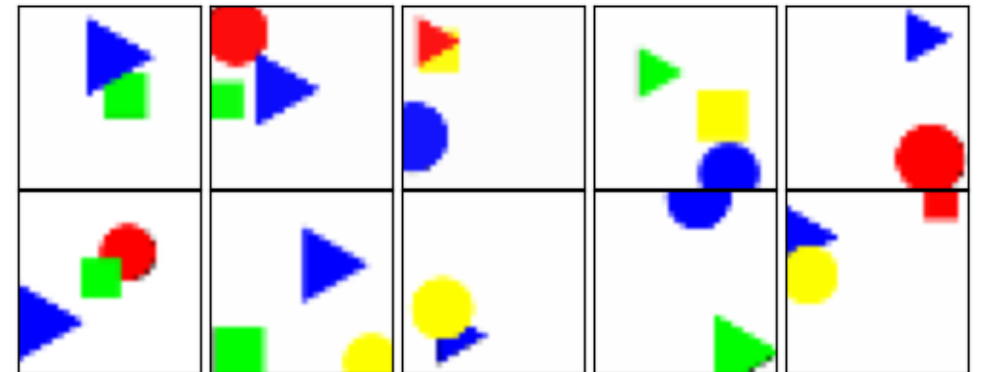
Presented by: *Mabin Varghese*

Task And Objective

- Dataset: Flying objects -> set of frames

- Extract FF and LF of every sets (Total 300 sets)

- Build a Generative adversarial network


Training input/Real images

Last Frames

Model Architecture

1. Generator Model
   i. UNET generator model -> Typical for Image segmentation
   ii. Encoder and decoder
   iii. Skip connections -> preserve spatial features and capture low- and high-level feature

2. Noise – Random noise with a normal distribution

Conv+BatchNormalization+ReLU
Pooling operation
Upsampling Layer
Skip-Connection
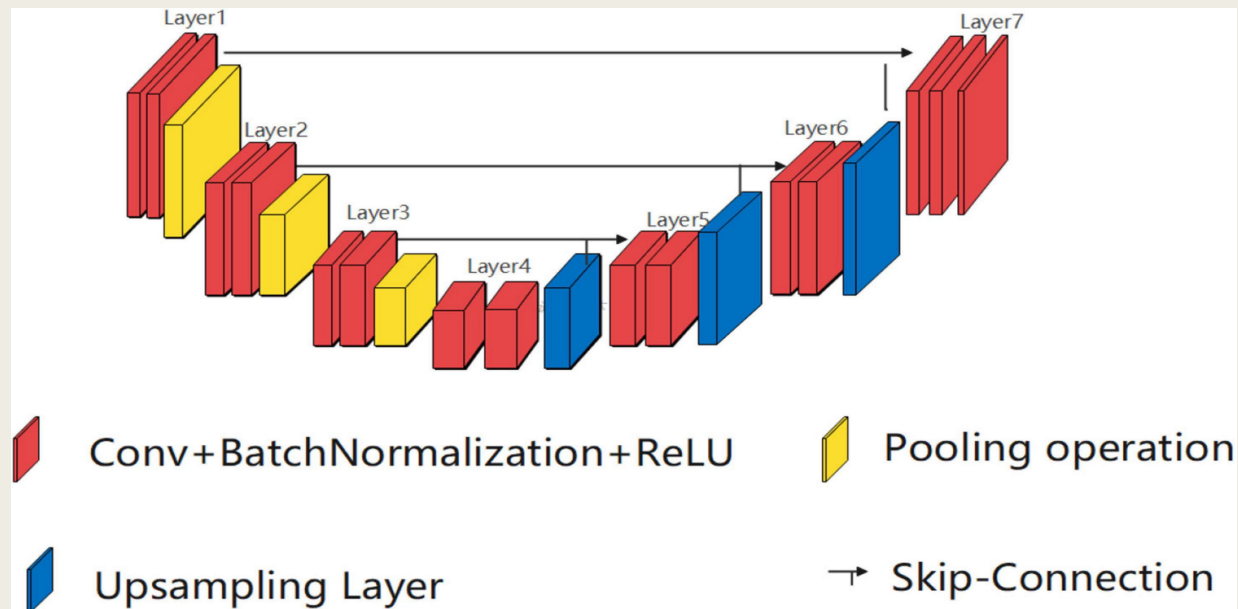
https://www.frontiersin.org/files/Articles/841297/fnagi-14-841297-HTML-r2/image_m/fnagi-14-841297-g001.jpg

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 32, 32, 32]	5,792
Conv2d-2	[-1, 64, 32, 32]	18,496
Conv2d-3	[-1, 128, 16, 16]	137,344
BatchNorm2d-4	[-1, 128, 16, 16]	256
ReLU-5	[-1, 128, 16, 16]	0
Conv2d-6	[-1, 256, 8, 8]	524,544
BatchNorm2d-7	[-1, 256, 8, 8]	512
ReLU-8	[-1, 256, 8, 8]	0
Conv2d-9	[-1, 512, 4, 4]	2,097,664
BatchNorm2d-10	[-1, 512, 4, 4]	1,024
ReLU-11	[-1, 512, 4, 4]	0
Conv2d-12	[-1, 512, 2, 2]	4,194,816
BatchNorm2d-13	[-1, 512, 2, 2]	1,024
ReLU-14	[-1, 512, 2, 2]	0
Conv2d-15	[-1, 512, 1, 1]	4,194,816
BatchNorm2d-16	[-1, 512, 1, 1]	1,024
ReLU-17	[-1, 512, 1, 1]	0
ConvTranspose2d-18	[-1, 512, 2, 2]	4,194,816
BatchNorm2d-19	[-1, 512, 2, 2]	1,024
ReLU-20	[-1, 512, 2, 2]	0
Dropout2d-21	[-1, 512, 2, 2]	0
ConvTranspose2d-22	[-1, 256, 4, 4]	4,194,560
BatchNorm2d-23	[-1, 256, 4, 4]	512
ReLU-24	[-1, 256, 4, 4]	0
Dropout2d-25	[-1, 256, 4, 4]	0
ConvTranspose2d-26	[-1, 128, 8, 8]	1,572,992
BatchNorm2d-27	[-1, 128, 8, 8]	256
ReLU-28	[-1, 128, 8, 8]	0
Dropout2d-29	[-1, 128, 8, 8]	0
ConvTranspose2d-30	[-1, 64, 16, 16]	393,280
BatchNorm2d-31	[-1, 64, 16, 16]	128
ReLU-32	[-1, 64, 16, 16]	0
Dropout2d-33	[-1, 64, 16, 16]	0
ConvTranspose2d-34	[-1, 32, 32, 32]	98,336
BatchNorm2d-35	[-1, 32, 32, 32]	64
ReLU-36	[-1, 32, 32, 32]	0
Dropout2d-37	[-1, 32, 32, 32]	0
ConvTranspose2d-38	[-1, 3, 32, 32]	867
BatchNorm2d-39	[-1, 3, 32, 32]	6
Tanh-40	[-1, 3, 32, 32]	

Discriminator Model

i. Patch GAN discriminator model -> maps a NxN array to the patch from the real images.

ii. Series of convolutional layers, batch normalization, leaky ReLU activations, dropout layers, and a final classification layer with a sigmoid activation function.

iii. Input the condition in discriminator also
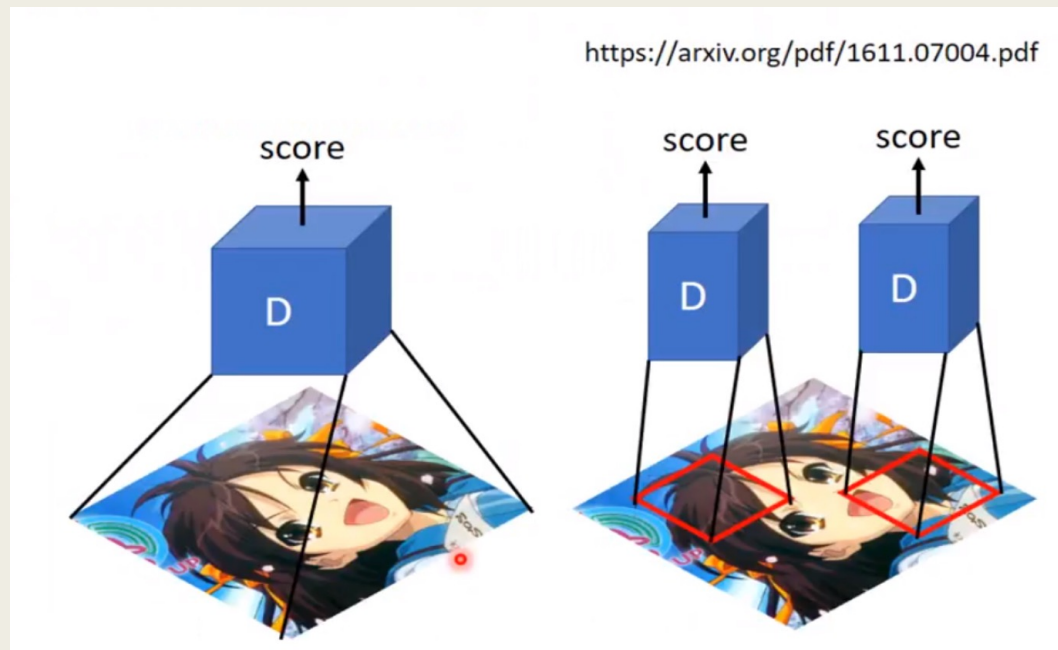

```
discriminator = Disc()
discriminator = discriminator.to(device)
summary(discriminator, input_size=[(6,32,32)], device=device.type)
```

```
----------------------------------------------------------------
        Layer (type)           Output Shape         Param #
================================================================
            Conv2d-1       [-1, 64, 16, 16]           6,208
       BatchNorm2d-2       [-1, 64, 16, 16]             128
         LeakyReLU-3       [-1, 64, 16, 16]               0
            Conv2d-4        [-1, 128, 8, 8]         131,200
       BatchNorm2d-5        [-1, 128, 8, 8]             256
         LeakyReLU-6        [-1, 128, 8, 8]               0
         Dropout2d-7        [-1, 128, 8, 8]               0
            Conv2d-8        [-1, 256, 4, 4]         524,544
       BatchNorm2d-9        [-1, 256, 4, 4]             512
        LeakyReLU-10        [-1, 256, 4, 4]               0
         ZeroPad2d-11       [-1, 256, 6, 6]               0
        Dropout2d-12        [-1, 256, 6, 6]               0
           Conv2d-13        [-1, 512, 3, 3]       2,097,664
      BatchNorm2d-14        [-1, 512, 3, 3]           1,024
        LeakyReLU-15        [-1, 512, 3, 3]               0
        ZeroPad2d-16        [-1, 512, 5, 5]               0
        Dropout2d-17        [-1, 512, 5, 5]               0
           Conv2d-18          [-1, 1, 2, 2]           8,193
      BatchNorm2d-19          [-1, 1, 2, 2]               2
        AvgPool2d-20          [-1, 1, 1, 1]               0
         Sigmoid-21          [-1, 1, 1, 1]               0
================================================================
Total params: 2,769,731
Trainable params: 2,769,731
Non-trainable params: 0
----------------------------------------------------------------
```

Dataset & Transformation

Original Dataset

- Training – 300 pairs
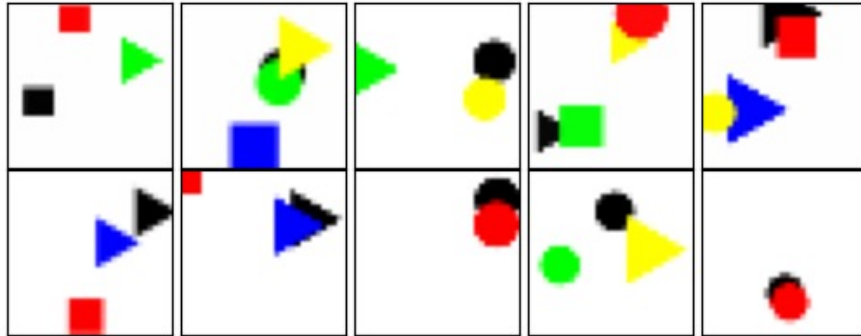- Validation – 60 pairs
- Testing - 60 pairs

Augmentation -> to increase pairs
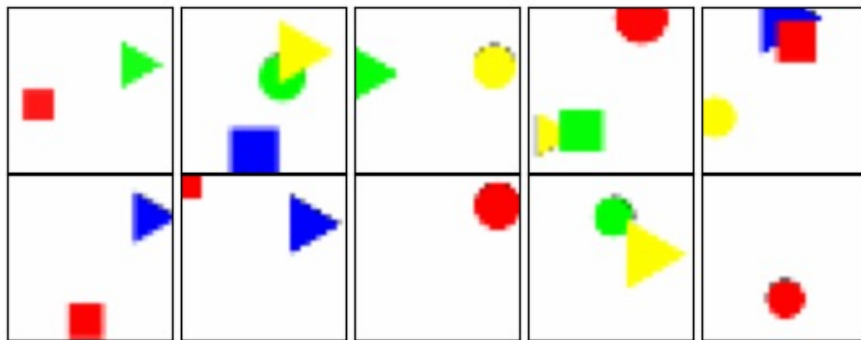
- Color jitter
- Gaussian blur

New Dataset

- Training – 882
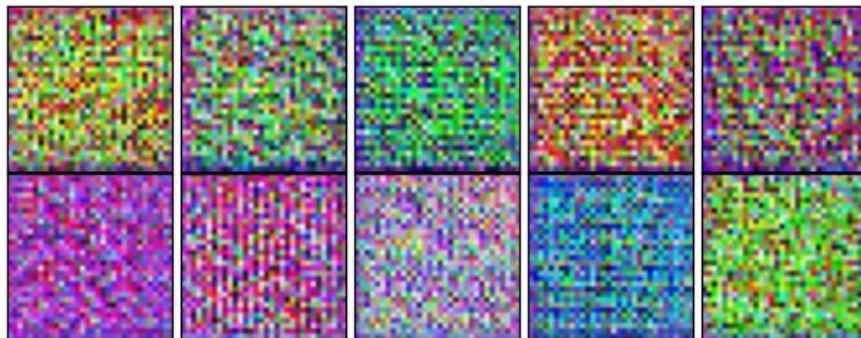- Validation– 252
- Testing - 126

Training input/Real images

Last Frames

Generated images/Fake images
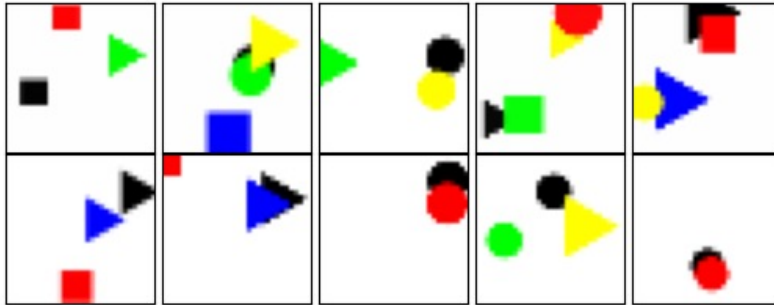
First Epoch Before Training

Conditional GAN Training

1. Initialized Xavier weights
2. Applied learning rate scheduling based on validation losses
3. Simultaneous Gen() and Disc() Training

Evaluation while training
   1. Losses
   2. L1 Loss – Assessing similarity b/w generated image and real image
   3. SSIM scores
   4. Visual inspection

Training input/Real images
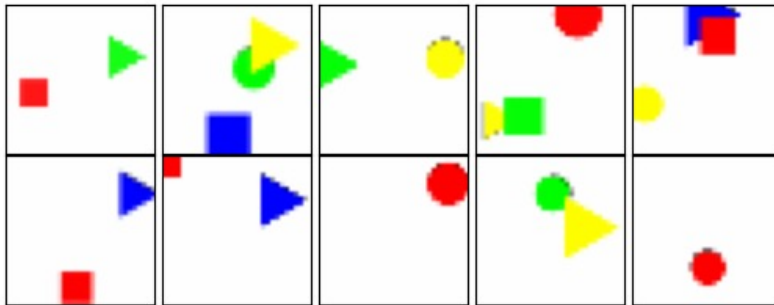
Last Frames

Generated images/Fake images
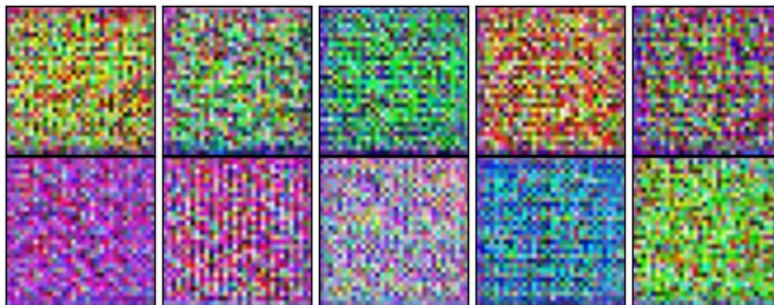
Visualization on the 1st Epoch

SSIM Score : -0.083

After 10 Epoch

Training input/Real images
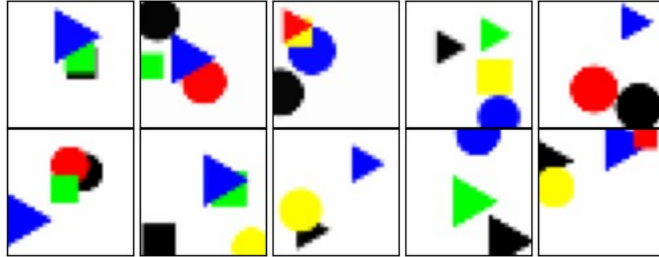
Last Frames
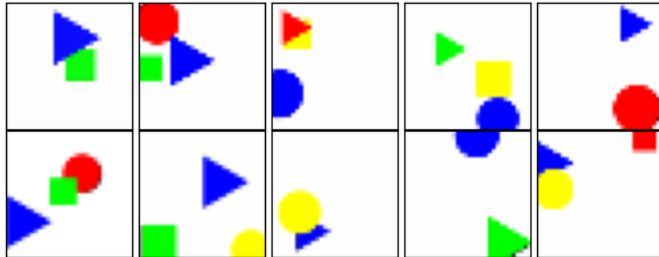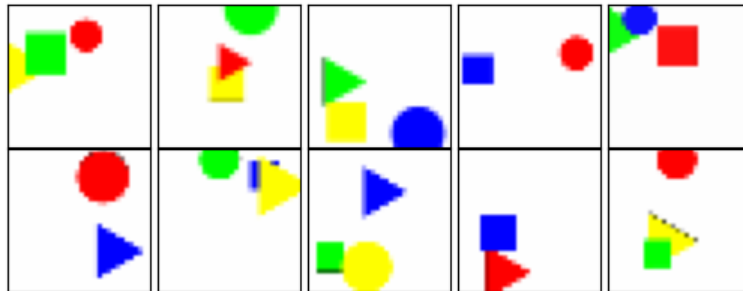

Generated images/Fake images
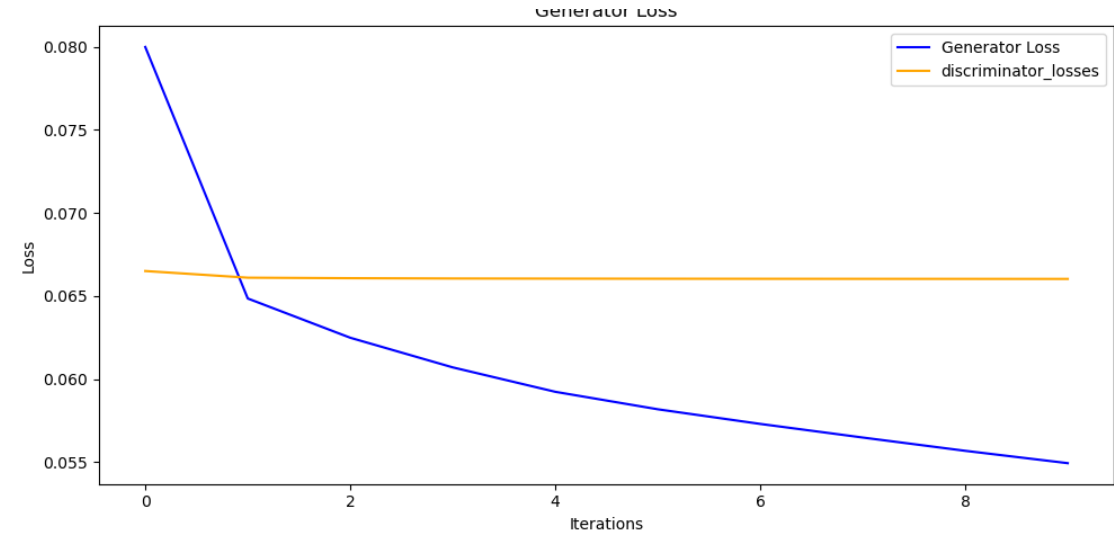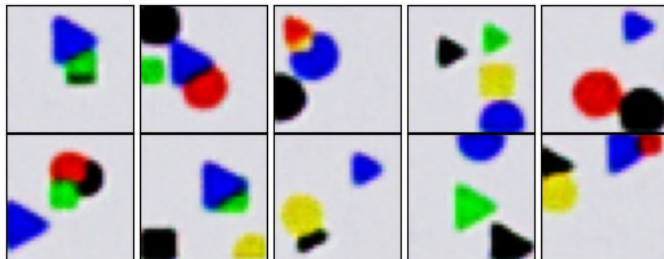
Generator Loss

Epoch 6/500, Val Disc Loss: 6.6192686557769775, Val Gen Loss: 2.437118724698112
Final Average L1 Loss: Test-: 299.7445445212107 Val-: 302.6081597994244 Train-: 301.9561856782355
Epoch 7/500, Discriminator Loss: 6.603388356513717, Generator Loss: 5.729565869112944
Epoch 7/500, Val Disc Loss: 6.618022209122068, Val Gen Loss: 2.384256725273435
Final Average L1 Loss: Test-: 295.1511125715952 Val-: 297.9273597399394 Train-: 297.38670709181804
Epoch 8/500, Discriminator Loss: 6.603120730307096, Generator Loss: 5.6475960073016935
Epoch 8/500, Val Disc Loss: 6.640234069218712, Val Gen Loss: 2.287672980437203
Final Average L1 Loss: Test-: 286.8669184427413 Val-: 289.99678199253384 Train-: 289.29934480022683
Epoch 9/500, Discriminator Loss: 6.6028856771603195, Generator Loss: 5.567617116331243
Epoch 9/500, Val Disc Loss: 6.641473514693123, Val Gen Loss: 2.2058519106062633
Final Average L1 Loss: Test-: 281.4507105993846 Val-: 284.54173178899856 Train-: 283.7893042434641
Epoch 10/500, Discriminator Loss: 6.60265444898281, Generator Loss: 5.4932345203261255
Epoch 10/500, Val Disc Loss: 6.638700905300323, Val Gen Loss: 2.1360248563781616
Final Average L1 Loss: Test-: 275.43047590861244 Val-: 278.2971281853933 Train-: 277.7097190048149
Testing DT - Last 10 Average L1 Loss Trend: Decreasing
Average SSIM Score for Training: 38.12792512018833
Average SSIM Score for Validation: 38.02091590049586
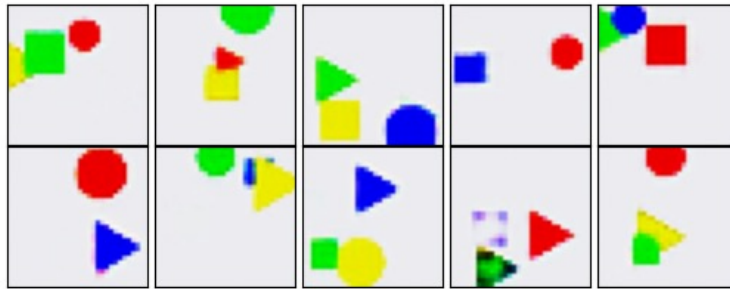Average SSIM Score for testing: 38.626941468241036
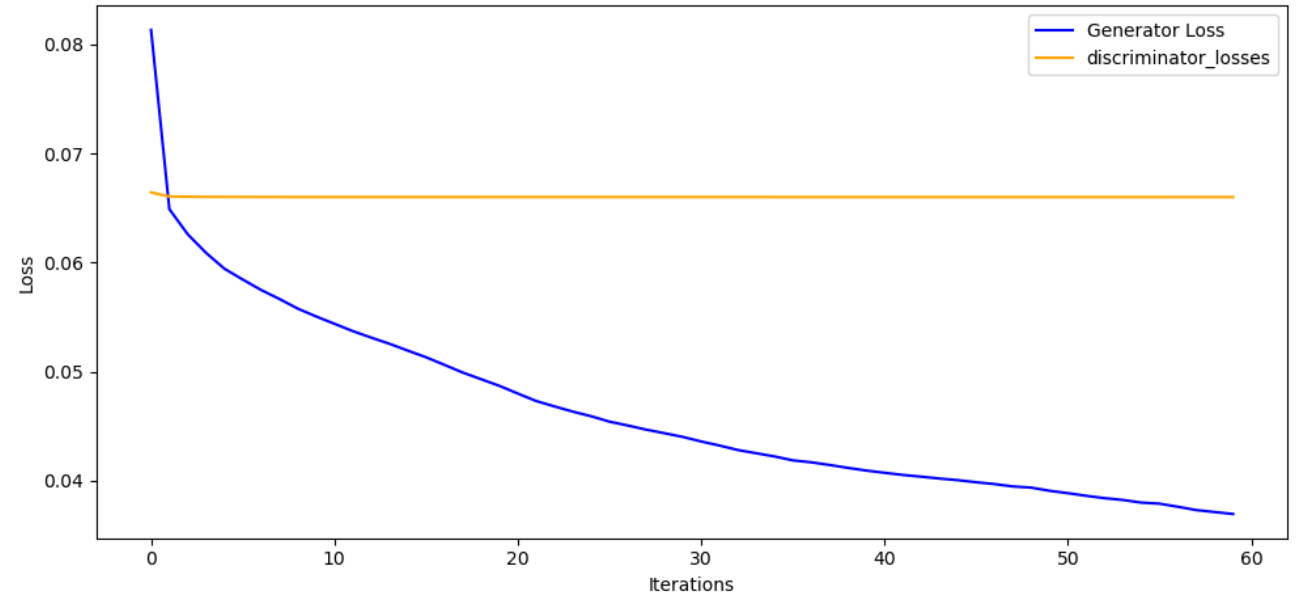
After 70 Epoch

Training input/Real images
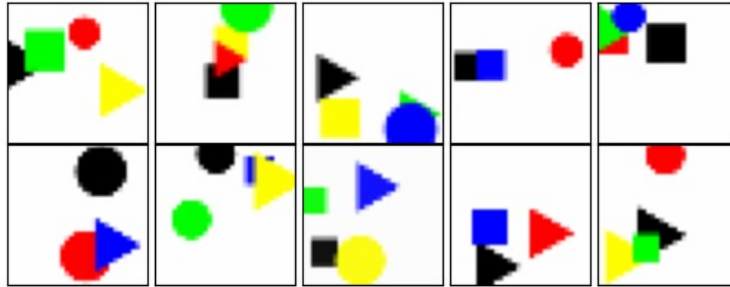
Last Frames

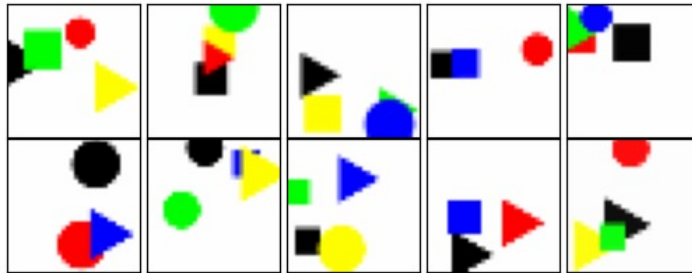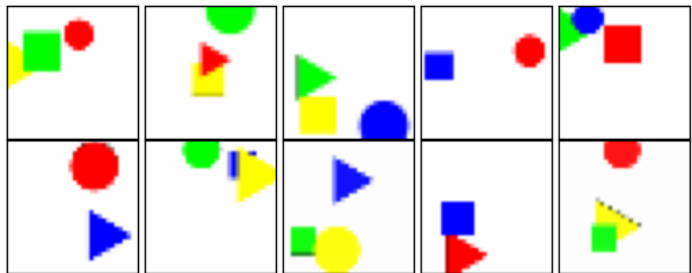**Generated images/Fake images**

Generator Loss

Final Average L1 Loss: Test-: 115.05315701166788 Val-: 115.71146666057525 Train-: 110.62344776935318
Epoch 68/500, Discriminator Loss: 6.601439960419185, Generator Loss: 3.6098242481819898
Epoch 68/500, Val Disc Loss: 6.601313844559685, Val Gen Loss: 0.3593471448218066
Final Average L1 Loss: Test-: 111.88131357942308 Val-: 112.47358329239347 Train-: 107.2018522591818
Epoch 69/500, Discriminator Loss: 6.60143817633458, Generator Loss: 3.602707237343129
Epoch 69/500, Val Disc Loss: 6.601471087289235, Val Gen Loss: 0.3460363337090091
Final Average L1 Loss: Test-: 108.53355484349387 Val-: 109.2031237388414 Train-: 103.68656983721554
Epoch 70/500, Discriminator Loss: 6.601438000629278, Generator Loss: 3.593542151440298
Epoch 70/500, Val Disc Loss: 6.600648022833325, Val Gen Loss: 0.3390749206855183
Final Average L1 Loss: Test-: 107.6504233337584 Val-: 108.5324189256108 Train-: 103.3987434340172
Testing DT - Last 10 Average L1 Loss Trend: Not Strictly Decreasing
Average SSIM Score for Training: 83.09536146627163
Average SSIM Score for Validation: 80.83135344953132
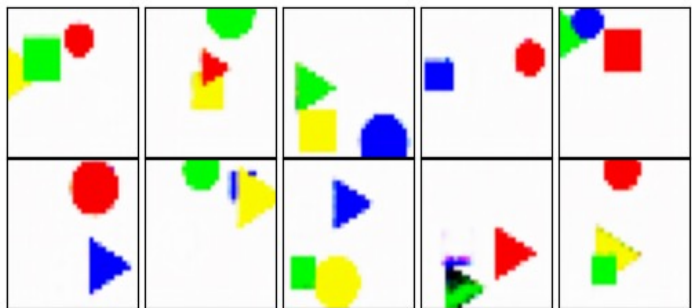Average SSIM Score for testing: 81.05532626800729

After 450 Epoch

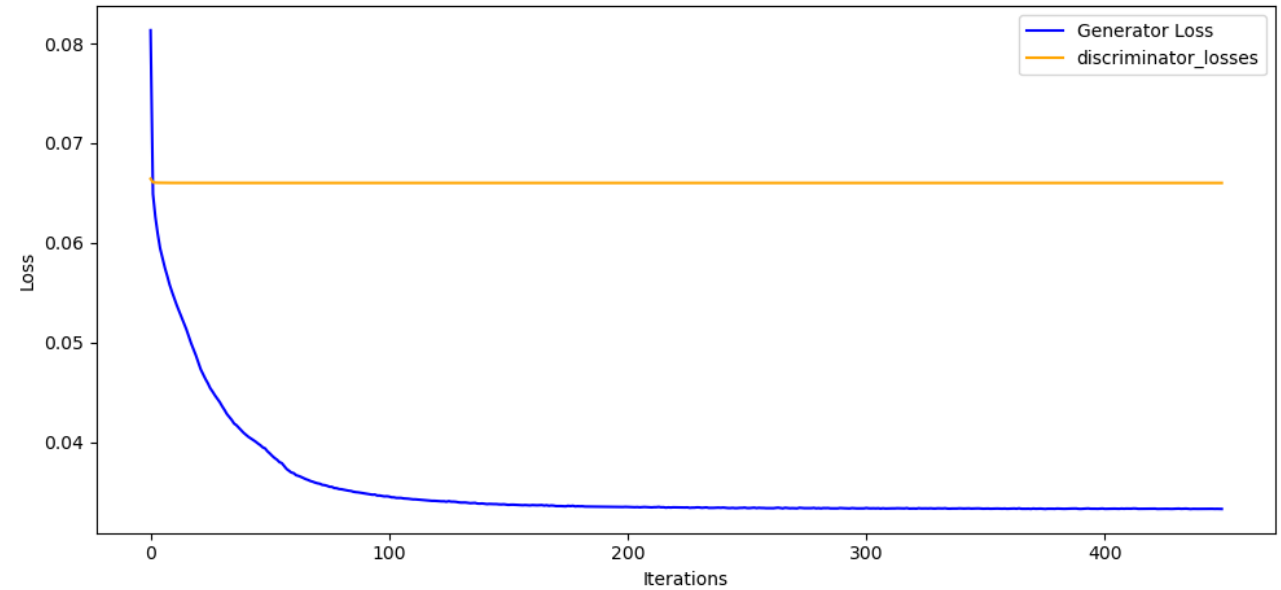Training input/Real images

Last Frames
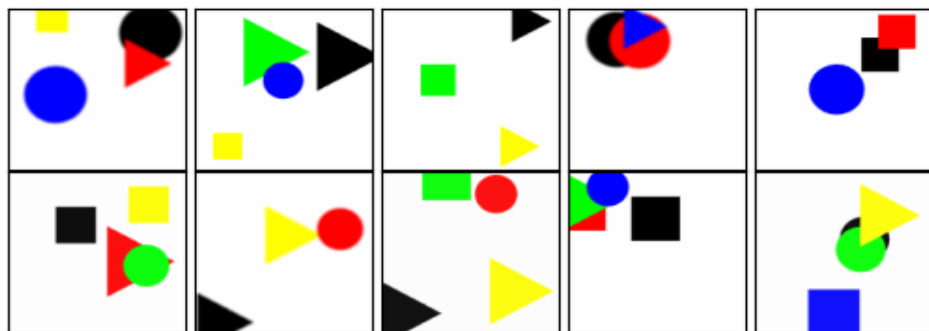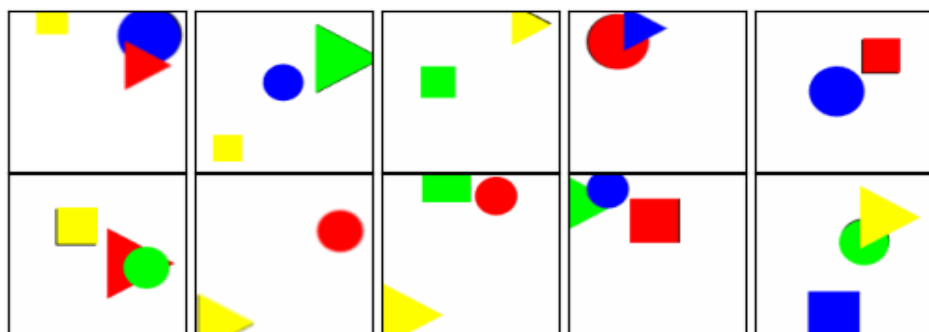

Generated images/Fake images

Generator Loss

Epoch 448/500, Discriminator Loss: 6.601401940494979, Generator Loss: 3.3305991398774575
Epoch 448/500, Val Disc Loss: 6.601404954516699, Val Gen Loss: 0.12546464087559828
Final Average L1 Loss: Test-: 16.603139894349233 Val-: 17.3287567431255 Train-: 11.252043261921324
Epoch 449/500, Discriminator Loss: 6.601401899947601, Generator Loss: 3.3294837558620913
Epoch 449/500, Val Disc Loss: 6.601388634197296, Val Gen Loss: 0.12607862775598372
Final Average L1 Loss: Test-: 16.82409765346656 Val-: 17.555720793704193 Train-: 11.554057783787213
Epoch 450/500, Discriminator Loss: 6.601401940494979, Generator Loss: 3.3293133960559524
Epoch 450/500, Val Disc Loss: 6.601415219761077, Val Gen Loss: 0.12735636028948996
Final Average L1 Loss: Test-: 16.99660283823808 Val-: 17.588565214758827 Train-: 11.501546374726999
Testing DT - Last 10 Average L1 Loss Trend: Not Strictly Decreasing
Average SSIM Score for Training: 94.72890512918323
Average SSIM Score for Validation: 91.70754764283377
Average SSIM Score for testing: 91.94685407911362
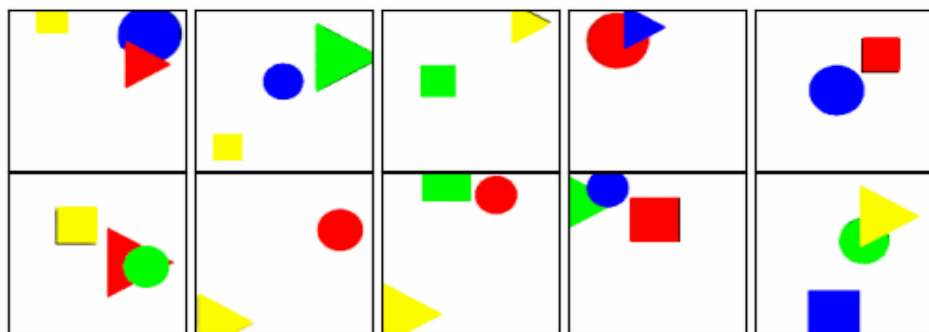
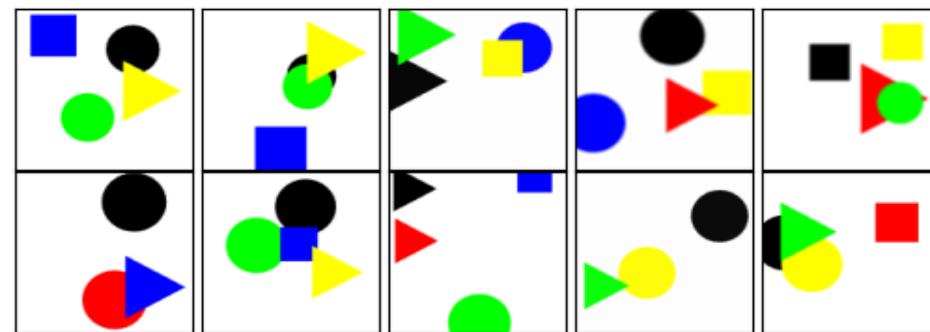# Evaluating Test Dataset
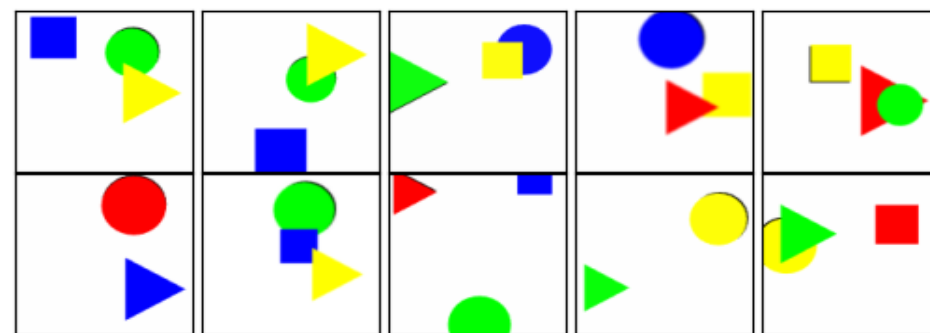

Training input/Real images

First Frame

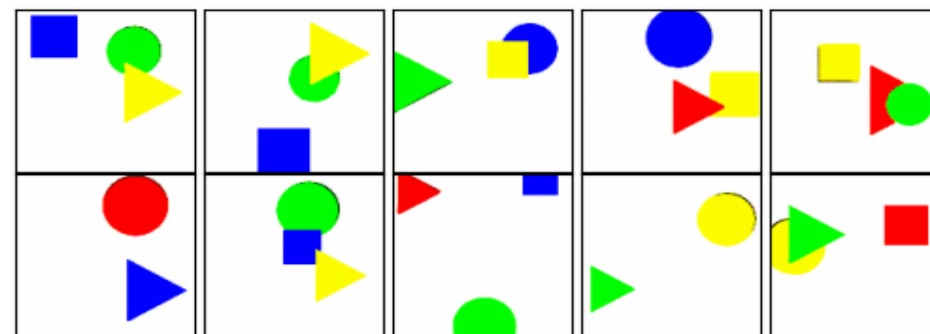Last Frames
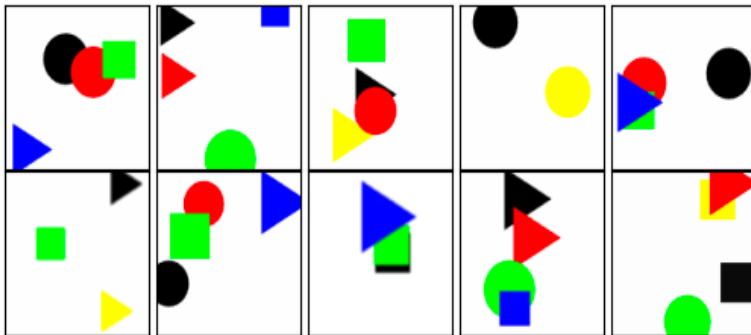
Last Frame

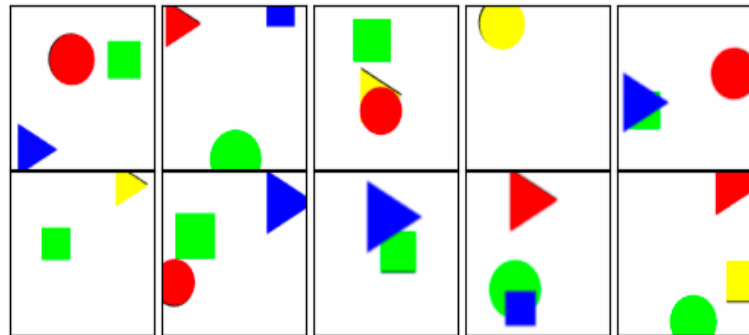Generated images/Fake images

Generated Last Frame

Evaluating Test Dataset

```
Epoch 98/100, Discriminator Loss: 6.601402467610884, Generator Loss: 3.3399948862945146
Epoch 98/100, Val Disc Loss: 6.6013446875980915, Val Gen Loss: 0.07010654413274356
Final Average L1 Loss: Test-: 6.1906495737651035 Val-: 6.892448808583948 Train-: 3.394786059400257
Epoch 99/100, Discriminator Loss: 6.601402737926734, Generator Loss: 3.340074588922687
Epoch 99/100, Val Disc Loss: 6.601355851642669, Val Gen Loss: 0.06552046896623713
Final Average L1 Loss: Test-: 6.170129063465293 Val-: 6.630496521081243 Train-: 3.791787400799476
Epoch 00106: reducing learning rate of group 0 to 4.4372e-04.
Discriminator Learning Rate changed: 0.000499236317115515 -> 0.0004437212386522697
Epoch 100/100, Discriminator Loss: 6.601402481126676, Generator Loss: 3.3405757545073285
Epoch 100/100, Val Disc Loss: 6.601365028865754, Val Gen Loss: 0.06155784108809062
Final Average L1 Loss: Test-: 6.159258015926868 Val-: 6.259872433951212 Train-: 3.6009003159055766
Testing DT - Last 10 Average L1 Loss Trend: Not Strictly Decreasing
Average SSIM Score for Training: 99.11197097605452
Average SSIM Score for Validation: 98.32972945494417
Average SSIM Score for testing: 98.29212238850143
```


Training input/Real images   Last Frames   Generated images/Fake images