

Department of Computer Science

Advanced Machine Learning (UE15CS415)

# Problem Statement:

Using CNNs to identify an object ( partial or complete) from an image and reconstruct it.

|  |  |
| --- | --- |
| **Name** | **USN** |
| Harshita Singh | 01FB15ECS123 |
| Likhita Navali | 01FB15ECS156 |
| Varnika Srivastav | 01FB15ECS336 |

# Different approaches

#vgg

## Reconstruction part

**1st approach**:

Trained a GAN on MNIST data. Generated partial images to test the model. This image was reconstructed using the above model. The images in MNIST are loaded as one-hot format. The images are on gray scale. The model consists of a generator and a discriminator. The generator network has 2 hidden layers each of 128 neurons which are fully connected and has an output layer with 784 neurons. The discriminator network has 2 hidden layers each of 784 neurons which are fully connected and has an output layer with one neuron and sigmoid as the activation unit. The loss function used is sigmoid cross entropy logits. Adam optimizer is used to train generator and discriminator. Fake images are generated by the generator and discriminator tries to differentiate between the real and fake image.

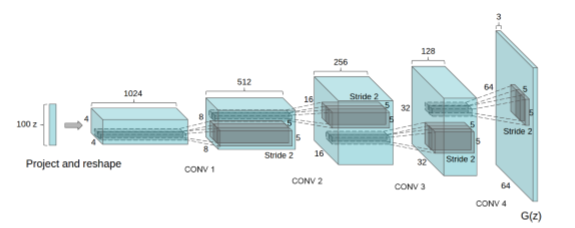
**2nd  approach**:

An attempt to convert the chair dataset into MNIST data format is made so that it could be loaded to train the tensor-flow GAN model and not use pretrained GAN model. First the chair images dataset were gray scaled. The MNIST dataset is already labeled but the issue with the chair images dataset was that it had to be manually labeled. We were able to overcome this issue we tried creating a matrix depicting the intensity of each pixel in the images, but we were unable to format it according to MNIST data.

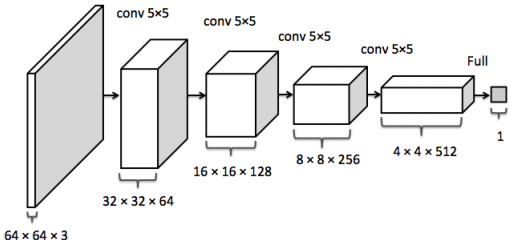
**3rd approach:**

A DCGAN model was used to reconstruct a face image since we could not find sufficient data for chairs/furniture. The CelebA dataset was used to try and train the model. However, because of the lack of hardware support we could not train the model.

The generator model architecture is as shown below



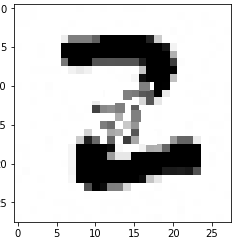
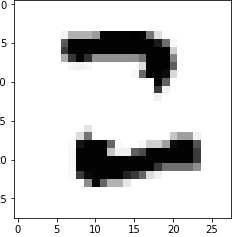
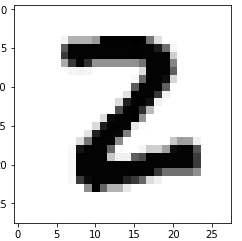
The discriminators model architecture is as shown below



An idea to use variational auto-encoders to simplify the problems faced while training our GAN was presented. However, it requires dataset of image with partial chair and image with complete chair to be fed to the model at the same time as X and Y respectively to train the model. Hence we could not implement the approach because of lack of dataset.

# Tests and Results

1. The results of reconstructing the MNIST images are shown in the following images



Actual Image Partial Image Reconstructed Image

2.The results of conversion of colored image to gray scaled image are shown below





Actual Image Gray scaled Image