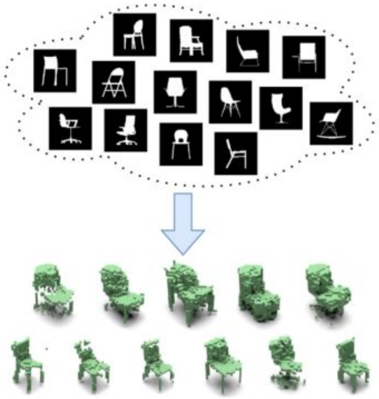
3D Shape Induction from 2D Views

Kiran Dapkar(170050020)
Rishav Arjun(170100051)
Saurabh Parekh(170100016)
Sumit(170050111)



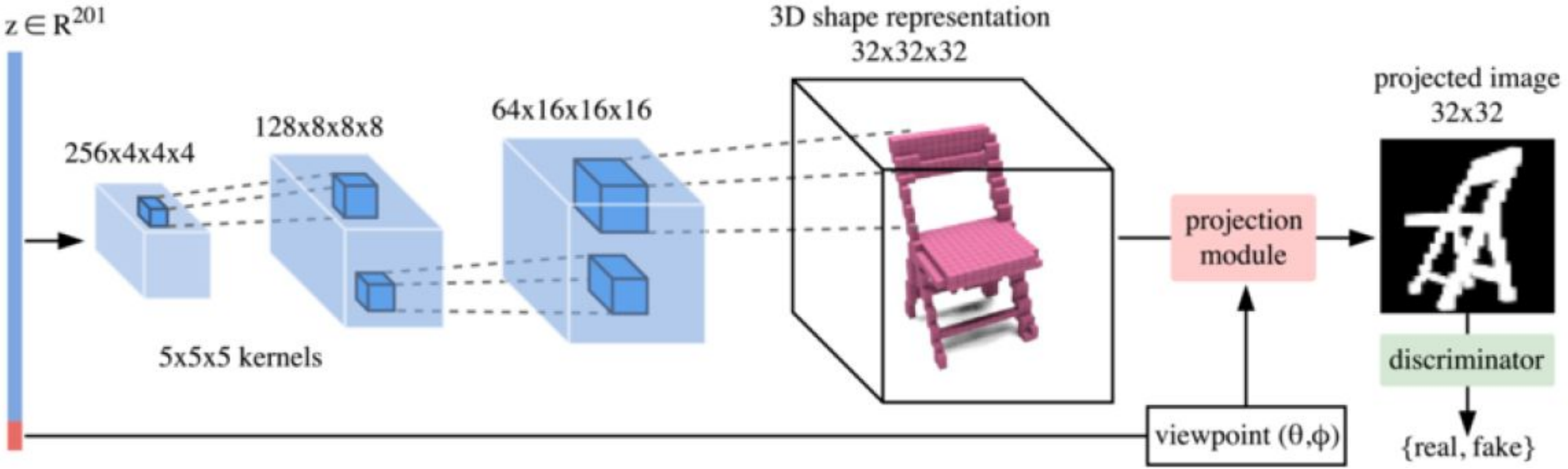
Task Description

We are trying to learn a generative model of 3D shapes given a collection of images of an unknown set of objects taken from an unknown set of views.

Challenge

- We are assuming that shapes are rendered as binary images bounded by silhou-ettes
- Thus, no shading cues available
- Moreover, which instance was used to generate each image, the viewpoint from which the image was taken, or even the number of underlying instances are not provided.

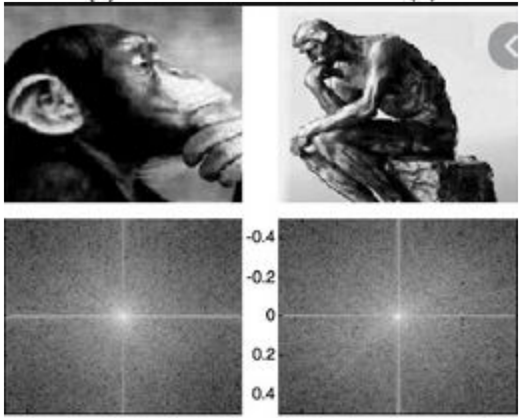
Available method or related work to solve the problem

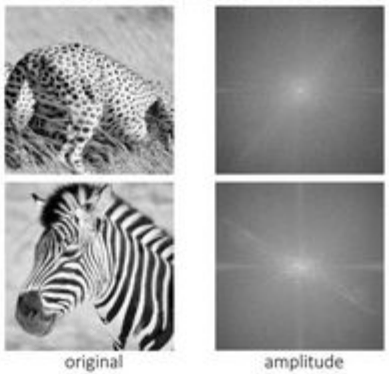


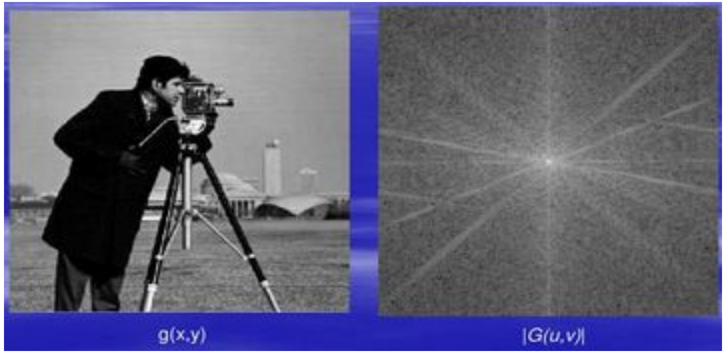
Outline of method(or extension to earlier method) we are proposing

Fourier transform of a natural image(without patterns) generally looks like a star. We have taken samples from a star-like distribution and taken its inverse fourier transform. We are hoping to get better results as, now, the generator will be generating images using z which is like something “natural”.Some examples of fourier transform of natural images:-

(a) (b)







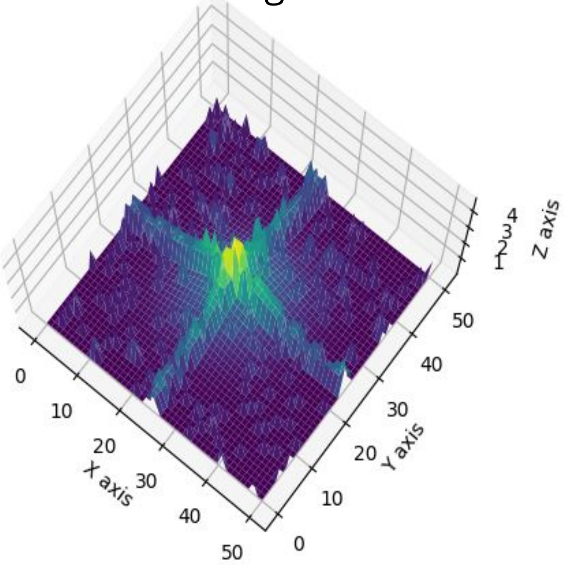
Images taken from results of googling “fourier transform of natural images”

Method contd.

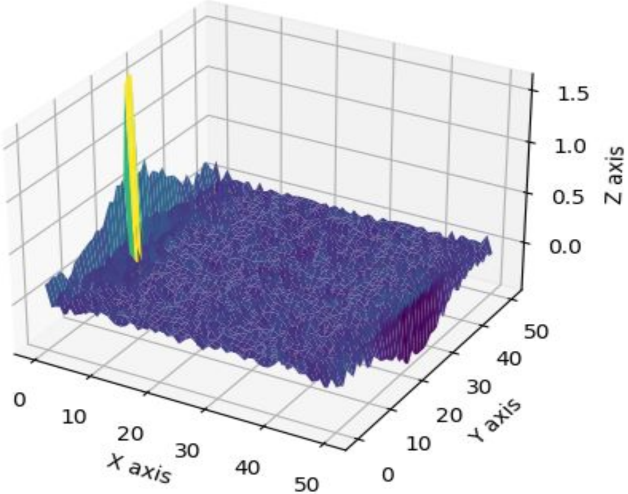
fz is a combination of gaussian and 3 (horizontal strip, vertical strip and one spread across the whole region) uniform distributions

Instead of uniform random noise, we are using this z

fz = Fourier transform of a “natural” image

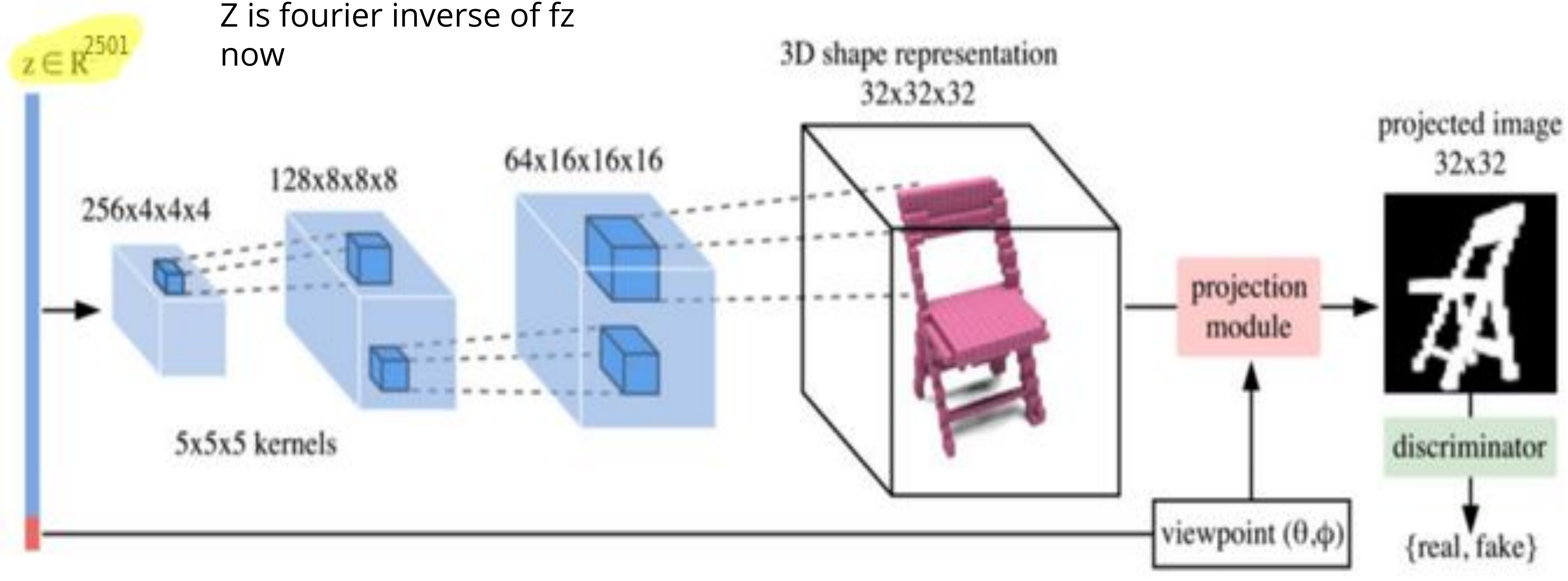


z = inverse fourier transform of fz



Method contd.

Z is fourier inverse of fz
now



Loss

- The generator G aims to transform samples drawn from a simple distribution P that appear to have been sampled from the original dataset.
- The goal of the discriminator D is to distinguish synthetic samples (created by the generator) from real samples (drawn from a data distribution D).
- The discriminator predict whether a given 2D image is from real dataset or obtained by projecting the 3D voxel, generated by the Generator.

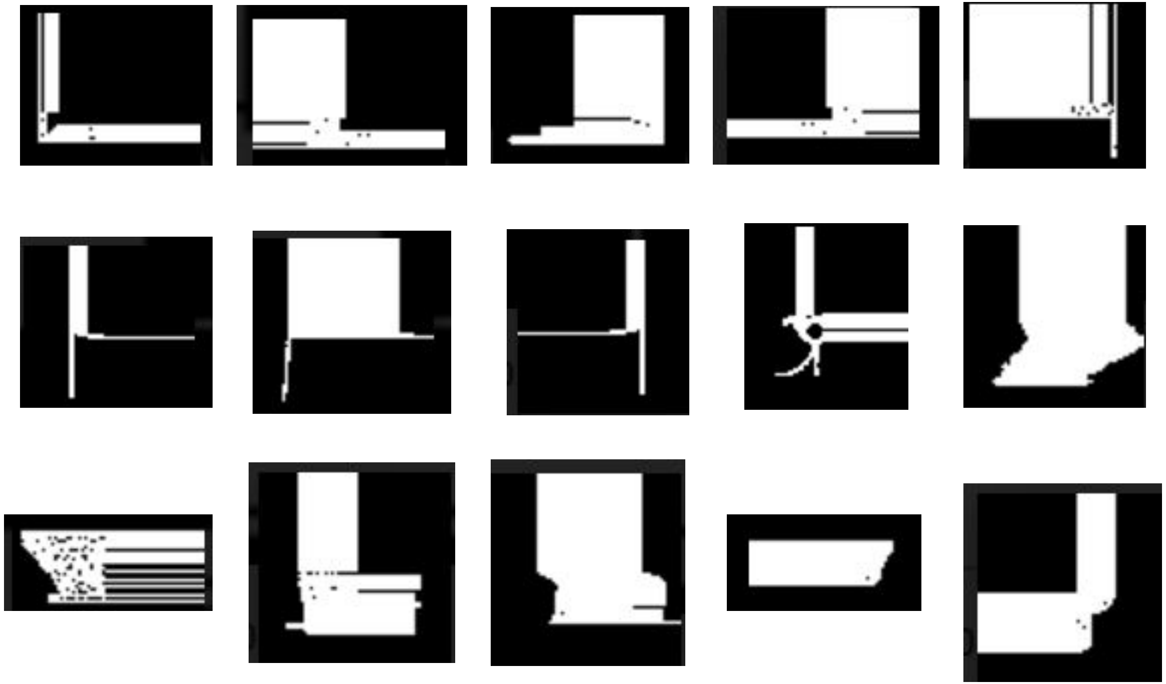
$$\min_G \max_D \mathbb{E}_{x \sim \mathcal{D}}[\log (D(x))] + \mathbb{E}_{z \sim \mathcal{P}}[\log (1 - D(G(z)))].$$

Experiment Details and Results

Results of paper

We run the model given in paper on one class “Aeroplane” because of lack of computation resource and poor quality of 2D images for other classes.

2D images for ModelNet10 classes



Bed

Chairs

Toilet

Conclusion

- This paper was unique in the sense that it generates 3D images by training the Generative model using 2D images.
- We are yet to test our hypothesis of using better latent variables.

References

- https://arxiv.org/pdf/1612.05872.pdf
- https://github.com/matheusgadelha/PrGAN
- https://towardsdatascience.com/gan-by-example-using-keras-on-tensorflow-backend-1a6d515a60d0
- https://3dusher.com/3d-viewer/obj
- https://www.kaggle.com/balraj98/modelnet10-princeton-3d-object-dataset