Classifying 3D objects as a whole
JunYoung Gwak, Christopher Choy (not in class)
This project is part of a research at CVGL, Stanford

ABSTRACT
In this project, we approached classification of a 3D objects by using the whole voxel representation as an input to the neural network. There has been a lot of works that tries to classify a 2D image to correct class but there was a little effort to classify 3D model as a whole. This is partially due to lack of training data and difficulty of building intuitive hand-engineered features on 3D space. We have access to enough data with advent of ShapeNet and random data augmentation. With enough data, we could approach this problem without engineering features using 3D convolutional neural network.

NETWORK
We convoluted on 3D filters to use a whole voxel as an input to the network. Our network is configured as following:

(3dconv + ReLu + pool) x 4
+ (FC + Relu + Dropout) x 2 + SoftMax

Input: (32, 1, 32, 32)
Conv1 filter: (32, 1, 3, 3, 3), Pool1 filter: (2,2,2)
Conv2 filter: (64, 32, 3, 3, 3), Pool2 filter: (2,2,2)
Conv3 filter: (128, 64, 3, 3, 3), Pool3 filter: (2, 2, 2)
Conv4 filter: (256, 128, 3, 3, 3), Pool4 filter: (2, 2, 2)
FC5 output: 1024, FC6 output: 1024

RESULT

57 classes classification result:
Test error: 0.2046

POSE ESTIMATION
By adding another SoftMax layer classifying model pose into 8 bins of \([-\pi, \pi]\) and defining cost as \(\text{total} \cdot \text{cost} + \lambda \cdot \text{pose} \cdot \text{cost}\), we enforced the network to learn model pose along with its class. We observed that pose errors come from bins nearby or 180° apart.
Test error: 0.1264

CONCLUSION
• We demonstrated that classification of 3D objects as a whole can be achieved with low error using 3D convolutional neural network.
• We analyzed that the network has learned simple gradients on layers closer to input and class-specific data on layers closer to output.
• As an extension of this work, we additionally enforced the network to learn the pose of the model by adding another softmax with weighted cost.