Reaction Report

O-CNN: Octree-based Convolutional Neural Networks for 3D Shape Analysis
Rushikesh Dudhat

What I like about this paper: The authors of this paper have proposed a new method called as Octree Based CNN (O-CNN) for 3D shape analysis (classification, segmentation and retrieval). The key idea is to represent the 3D shapes with Octrees and then perform CNN operations only on the sparse octants containing the surfaces of 3D shapes. Intuitively and so does computationally as proved by the authors, this method is a nice way to perform convolutions and pooling only where required, saving the computational time and memory significantly. I liked the concept of shuffled keys which is very efficient in retaining the 3D location of each octant. The corresponding x, y and z coordinates are kept track right from the very starting node with minimum memory consumption. Compared to the full voxel based solution the memory and the computation cost using OCNN is less than by the factor of n (voxel size). Moreover, the operation is further optimized by using threads to compute k+1 th information (hash tables) in kth iteration. The authors have used LeNet, to clearly demonstrate the advantages of using the Octree based representations. Unpooling for segmentation is also done very efficiently using switch variables as the octants are stored in continuous manner. The continuous storage of the octants also aids in efficient convolution, Batch Normalization operations.

What I don't like about this paper: The concept of the average normal vectors which is mentioned several times throughout the paper is not very clear. The fact that that it represents the orientation of the 3D sampled surface inside the octant, however, it is not clear weather just the magnitudes are considered or the projections along the x, y and z directions are taken. In the octree data preparation, the truncated bounding cube is used but this may occlude some of the points in complex shapes. Additionally, only the 3D surface is considered while making the Octrees, this might lead to information loss as the volume inside the surface is not captured.

Future Directions: Instead of the face centers of the truncated bounding cube, the vertices of the dodecahedron can be used to create the octrees. This might help in extracting intricate features and have a slight effect on the accuracy. It might be good to see the inclusion of the complete volume in the octree representation and apply OCNN. I believe that this might improve accuracy for higher resolution without serious performance overhead. It would also be interesting to see the performance of the OCNN using deep neural nets. Also, the geometry of the shape can be considered while creating the octree. Maybe this can be done by using the tangent to the 3D shape surfaces and the nonchanging tangent surfaces can be added to the same octant. This might reduce the computational costs significantly.