

# Enhanced convONet: Leveraging Remaining 3D Coordinates

Youjin Song

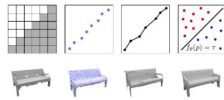
Electrical Engineering, Stanford University

## Motivation

- **Implicit neural representations** have gained popularity for learning-based 3D reconstruction
- Precise and more memory-efficient way than conventional methods such as voxel, point clouds, and mesh representation.
- The project aims to achieve **fine-grained implicit 3D reconstruction** from **noisy point clouds** or low-resolution voxel representations.

## Related Work

- Occupancy Network(Onet)[1] introduced a new representation based on learning a **continuous 3D mapping**, fully-connected network architecture



- Convolutional Occupancy Network(ConvONet)[2] suggests using **convolutional layer** for integrating local information.

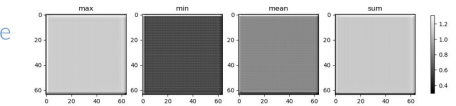
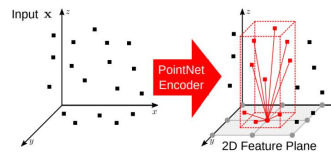
## References

- [1] L. Mescheder, M et.al, "Occupancy networks: Learning 3d reconstruction in function space," CVPR, 2019.  
[2] S. Peng, et.al, "Convolutional occupancy networks," iECCV 2020

## Main idea

- Previous model(ConvONet) : **discards** non-projected coordinates during orthographic projection onto canonical planes for **2D plane encoder**
- This method integrates the non-projected 3<sup>rd</sup> dimensions for feature extraction

ex) projecting 3d points to the xy plane

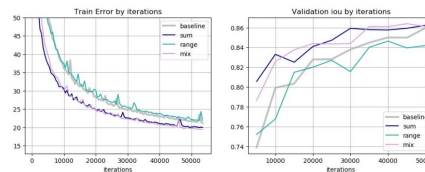


Multiply weights obtained from simple statistics from the non-projected coordinate :

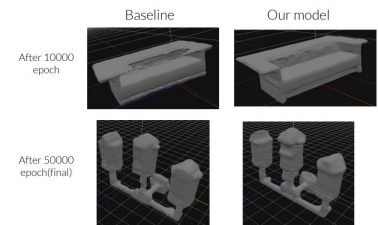
- **sum** (total number of points corresponding to the projection)
- **range** (difference between the maximum and minimum coordinate values)
- **mean** (average coordinate values)

## Experimental Results

### Quantitative Analysis



### Qualitative Analysis



- Faster Convergence during training
- Better IOU
- **Range** of the 3<sup>rd</sup> coordinate does not contribute to the performance