3D SEMANTIC SEGMENTATION IN DENSE URBAN ENVIRONMENTS

Arnau Ruana Ramos

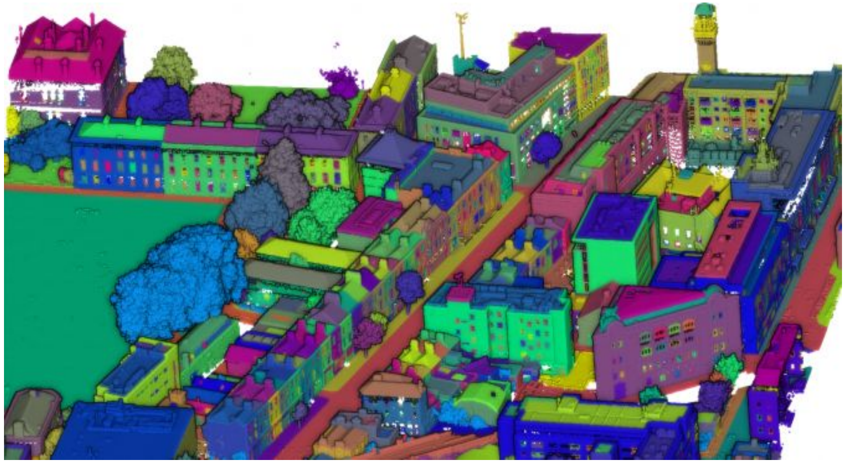
Rafel Palomo Avellaneda

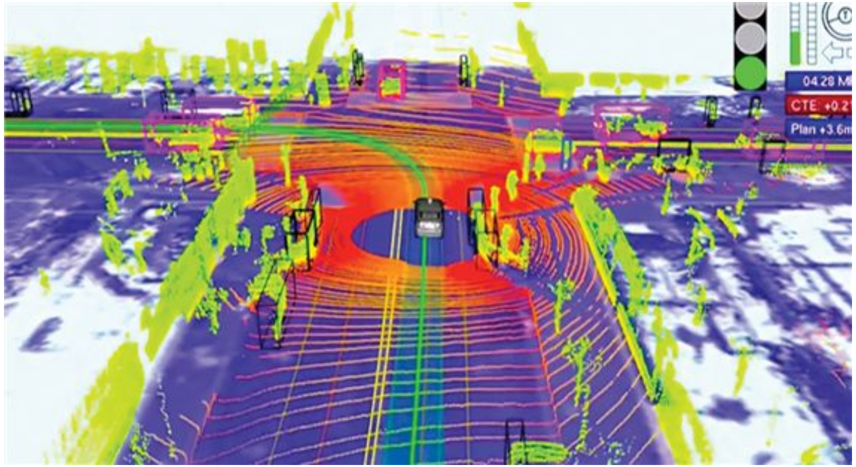
Soraya Gonzalez Sanchez

Advisor: Mariona Carós Roca

Motivation of the project

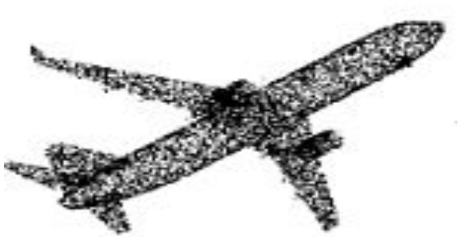
- Complex and highly important task in various fields, such as computer vision, robotics and virtual reality.
- Better spatial understanding.
- Lidar sensors brings real distance data. On the other hand, with cameras you could only get estimates.
- Independence from ambient lighting.
- Robustness against adverse environmental conditions.





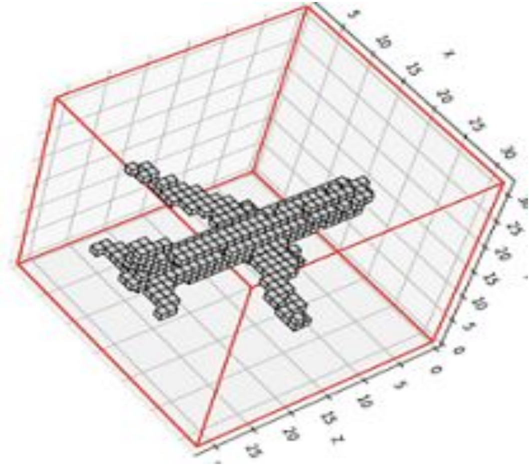
3D data representation

Raw



unordered points scattered

Volumetric

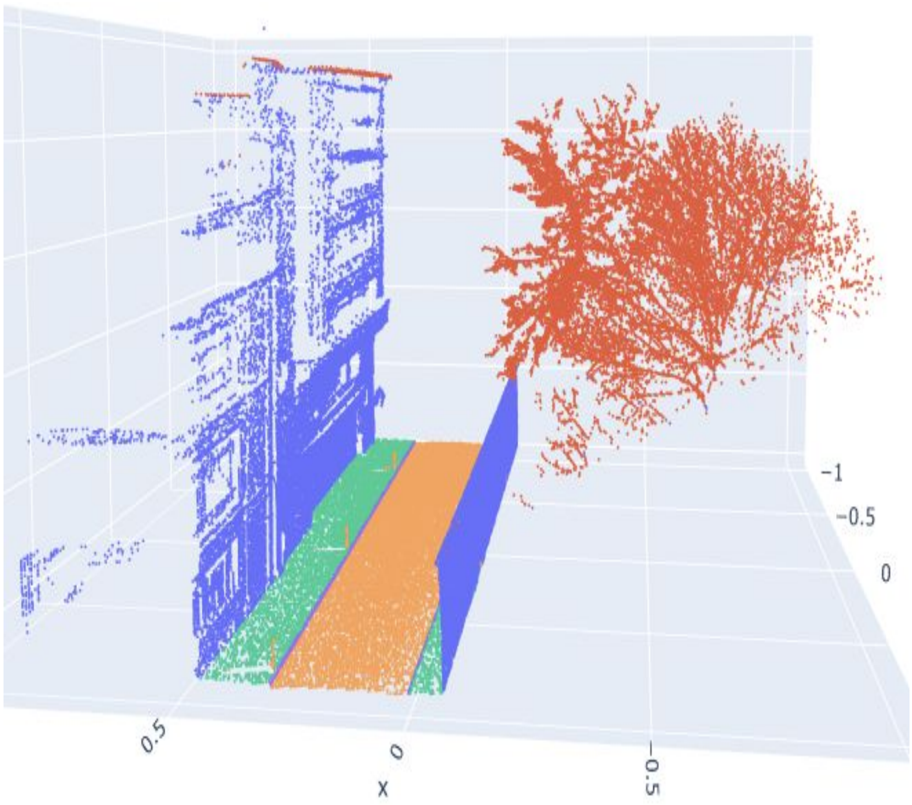


structured points in voxels

Our proposal

PointCloud

- Extends torch.Tensor
- DataFrame behaviour
- Get separate info (coords, feats, labels)
- Random sampling
- Shuffle of points
- Rotation (axis ‘z’)
- Plotting 3D



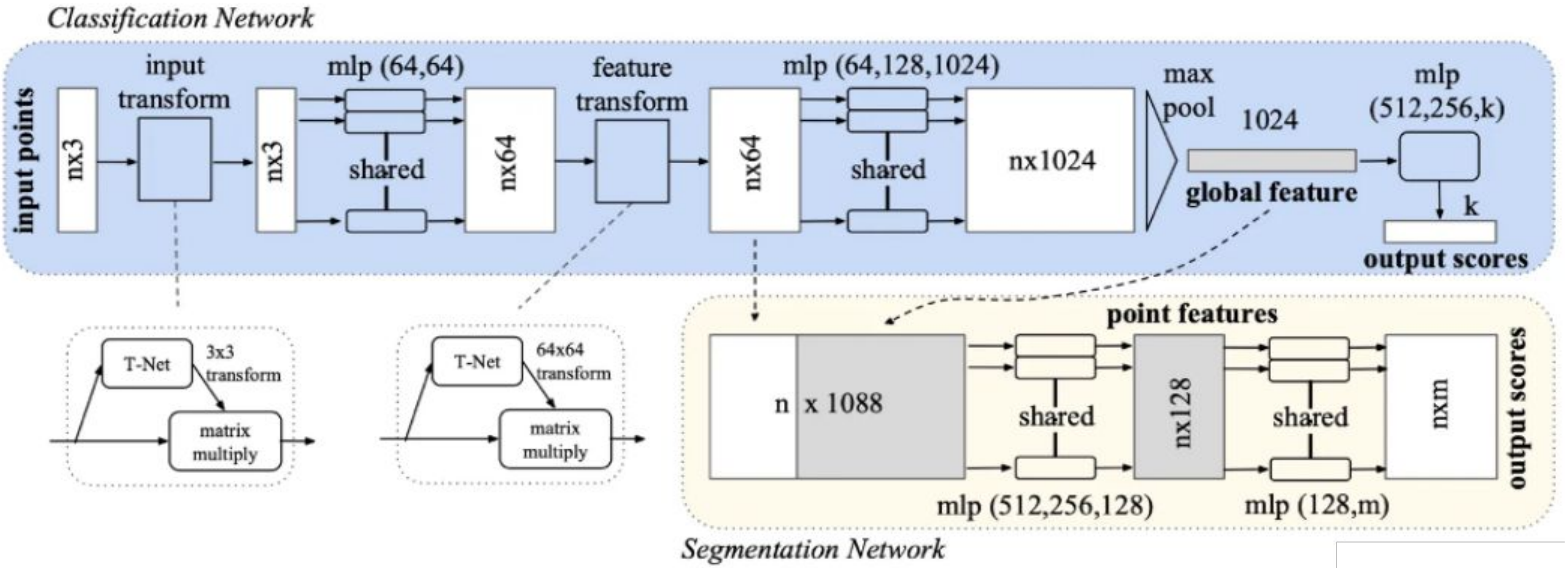
PointNet

- INVARIANT TO PERMUTATIONS OF POINTS ✅
- CAPTURE LOCAL STRUCTURES FROM NEARBY POINTS ✅
- INVARIANCE UNDER TRANSFORMATIONS ✅
- EUCLIDEAN SPACE ✅
- SEMANTIC SPACE ❌
- RELATIONSHIPS BETWEEN A POINT AND ITS NEIGHBORS ❌
- CAPTURE GEOMETRY OF THE OBJECTS ❌

Dynamic Graph CNN

- ✅ INVARIANT TO PERMUTATIONS OF POINTS
- ✅ CAPTURE LOCAL STRUCTURES FROM NEARBY POINTS
- ✅ INVARIANCE UNDER TRANSFORMATIONS
- ✅ EUCLIDEAN SPACE
- ✅ SEMANTIC SPACE
- ✅ RELATIONSHIPS BETWEEN A POINT AND ITS NEIGHBORS
- ✅ CAPTURE GEOMETRY OF THE OBJECTS

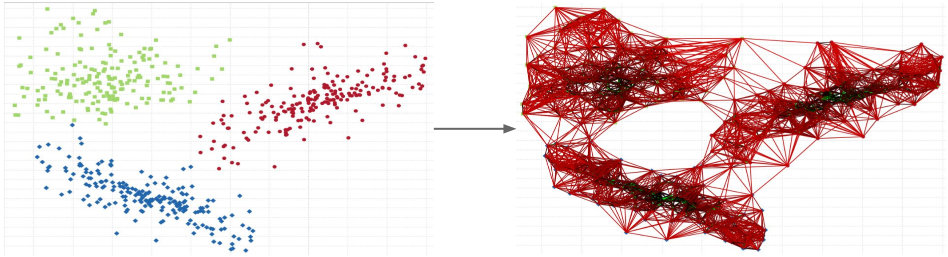
Point Net

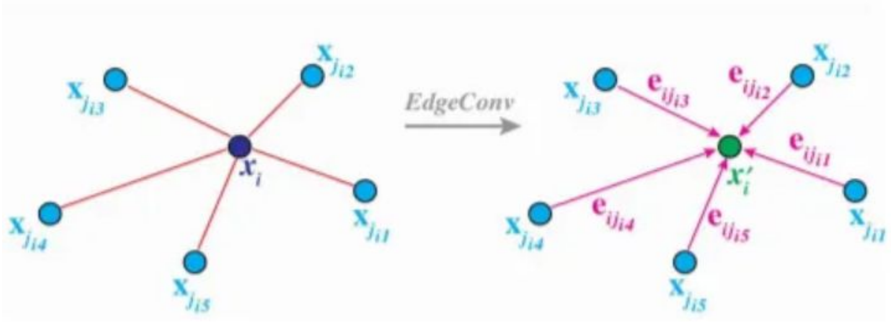


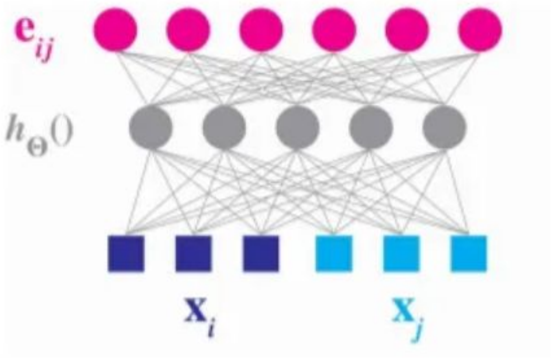
PointNet architecture

Dynamic Graph CNN

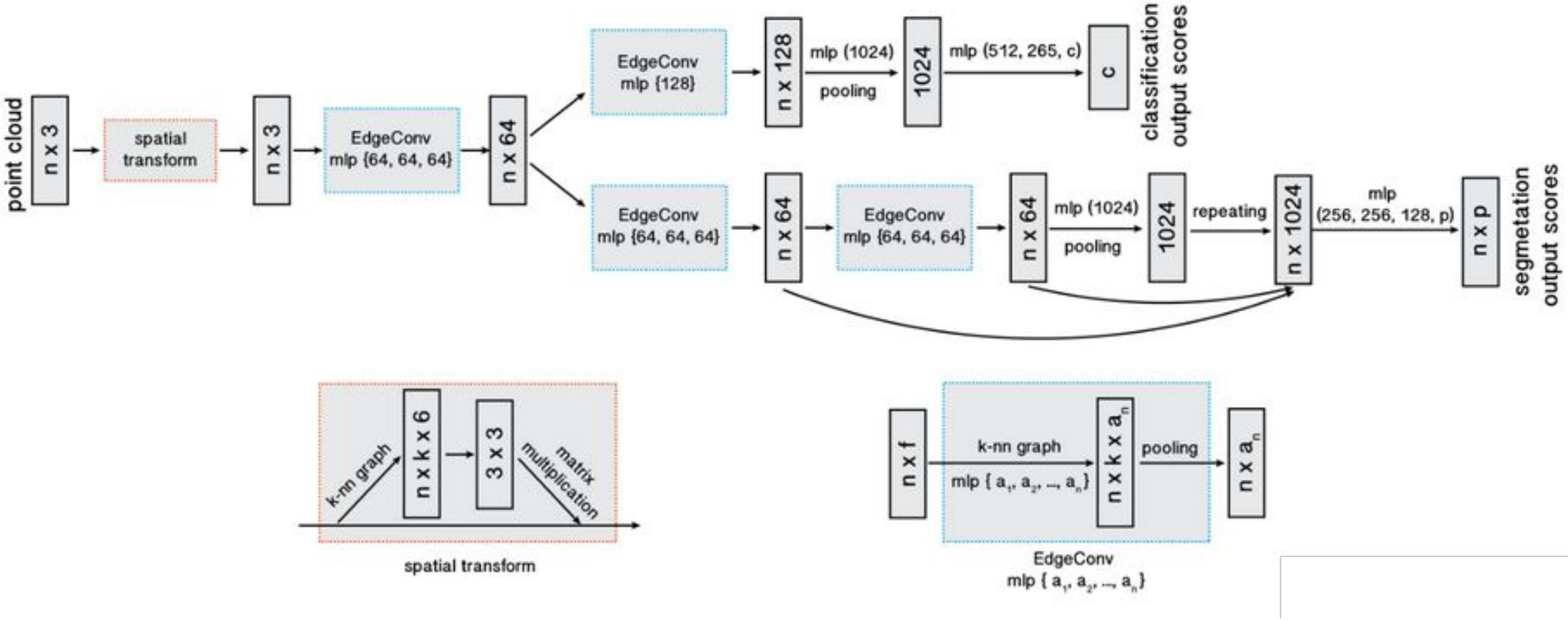
- KNN graph
- Edge Conv
- Embeddings
- Grouping points both in Euclidean space and in semantic space.







Dynamic Graph CNN



Computational Requirements

- Google colab
  - Concat trainings using state_dicts
- Google Cloud
  - Only 8 CPUs were used
- Local PC
  - GPU 3070, 16 GB RAM

Milestones

STATE OF THE ART

Dataset selection

Model selection

DATASET PREPARATION

Data understanding

Data preprocessing

MODEL TRAINING

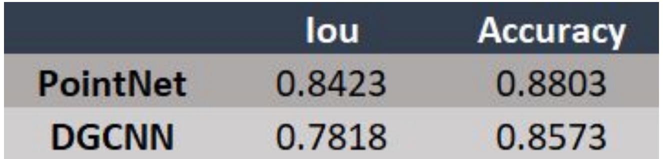
PointNet implementation
PointNet evaluation

DGCNN implementation
DGCNN evaluation

PROJECT EXTENSION

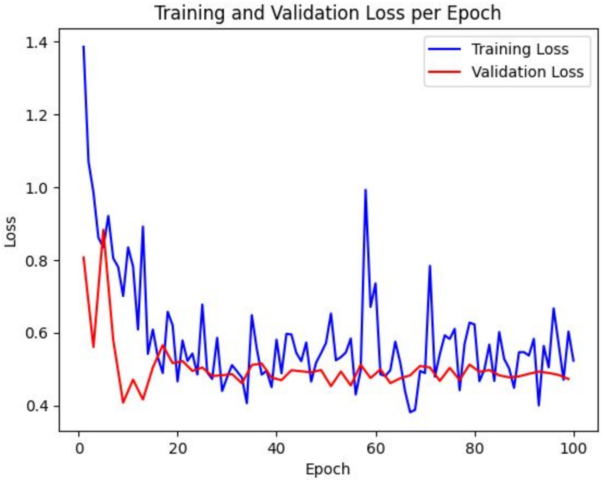
Instance Segmentation

Results

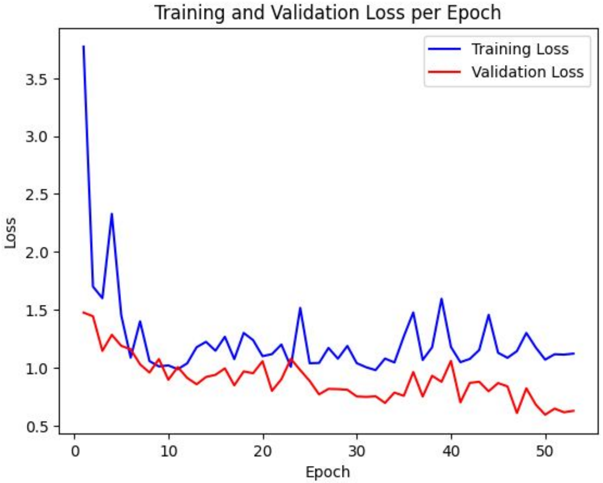


	Iou	Accuracy
PointNet	0.8423	0.8803
DGCNN	0.7818	0.8573

PointNet

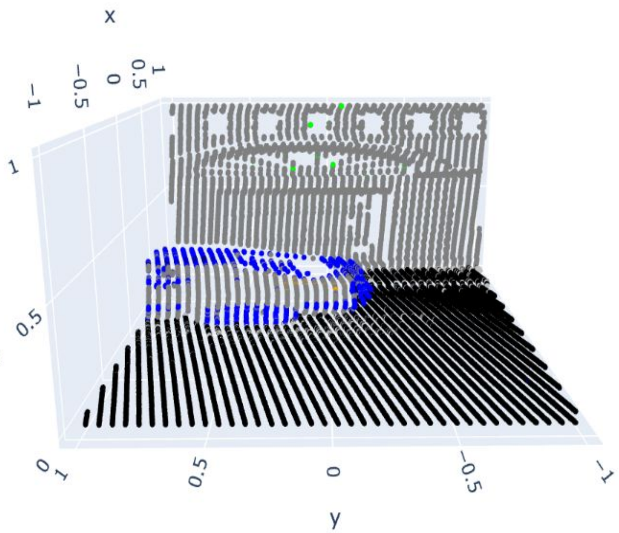


DGCNN

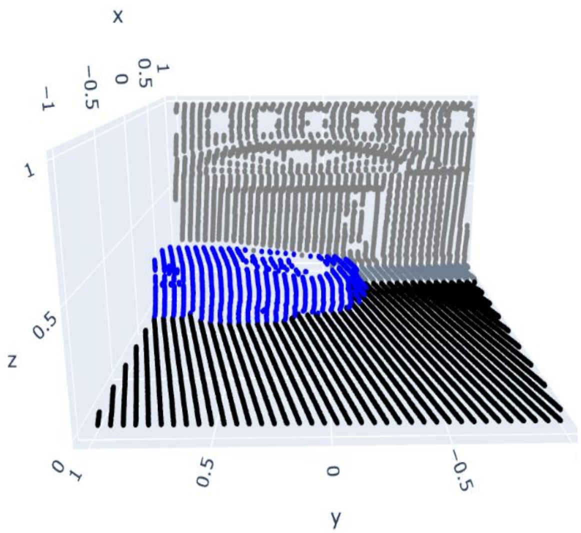


Graphical results

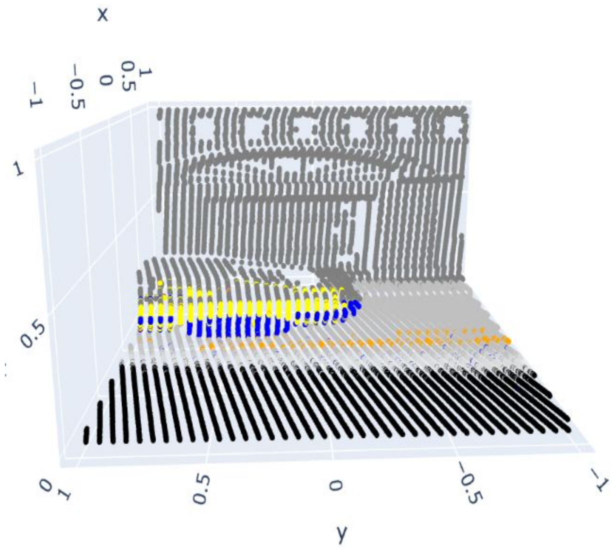
DGCNN



Ground truth

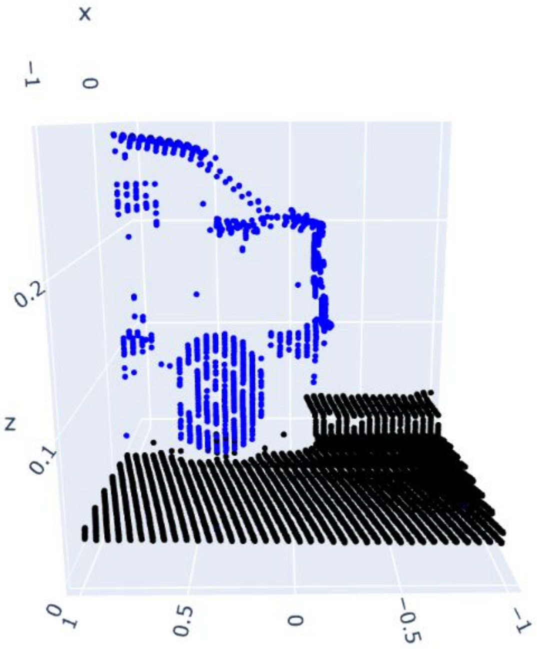


PointNet

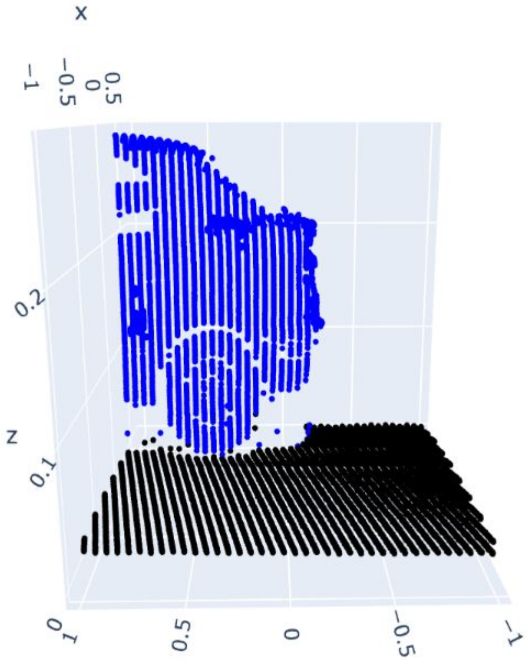


Graphical results

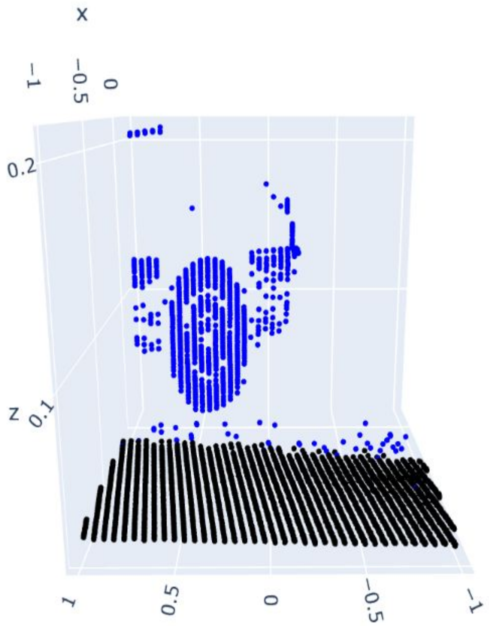
DGCNN



Ground truth



PointNet



Conclusions

Challenges Faced

- **Dataset Selection:** limited availability, binary data format, and incomplete annotations.
- **Training:** computational requirements extremely high for training stage.
- **Data**: Learn to treat data structure not seen in class. Architectures, batches, transformations, plotting…
- **Class imbalance:** highly unbalanced dataset

What have we learned?

- Importance of quality data
- Importance of using the right model
- Data preprocessing and normalization
- Hyperparameter optimization
- Evaluation and performance metrics
- How PointCloud architectures works
- Use of Google Cloud
- Use efficiently the available hardware

Thanks for your attention