

Point Cloud Occupancy with Dynamic Planes

Computer Vision Course Project

Master's Degree in Artificial Intelligence and Robotics

Eugenio Bugli (1934824)

Academic Year 2024/2025



SAPIENZA
UNIVERSITÀ DI ROMA



Table of Contents

1 Introduction

► Introduction

► Dataset

► Architecture

► Reconstruction

► Results

► Improvements



What is the Addressed Problem

1 Introduction

- We would like to learn the **Occupancy** values of points inside a bounding box
- We learn the features and the dynamic planes
- During Inference we reconstruct meshes with Multiresolution IsoSurface Extraction



Table of Contents

2 Dataset

- ▶ Introduction
- ▶ **Dataset**
- ▶ Architecture
- ▶ Reconstruction
- ▶ Results
- ▶ Improvements



FAUST Dataset

2 Dataset

This Dataset is composed by high-resolution **human scans** of 10 different bodies in 30 different poses.

- Each samples inside the training set is composed by the scan and the registration, while we have only the scan for the test set.
- The test set is composed by 200 scans, while the training has 100 scans.
- About 80 % fo the initial training set has been used for training, while the other 20 % has been used for validation.



Data Processing

2 Dataset

Starting from the meshes given by the Dataset, some pre-processing is needed:

- We need to randomly sample **3000** points from the mesh surface and add a perturbation (Gaussian noise) to them.
- We need to randomly sample **2048** points from the bounding box that contains the original mesh.

These noisy clouds are then used to learn the features and the geometry of the object through the Encoder, while the other points are used in the decoding part of the network.



Data Augmentation

2 Dataset



Label Generation

2 Dataset

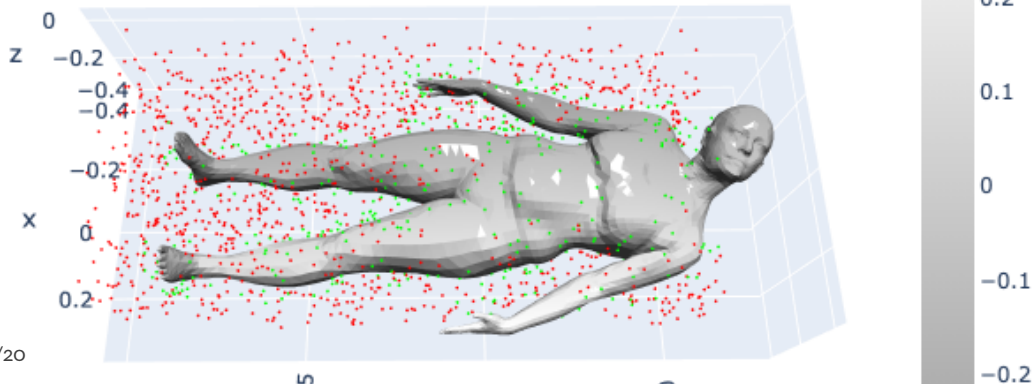




Table of Contents

3 Architecture

► Introduction

► Dataset

► **Architecture**

► Reconstruction

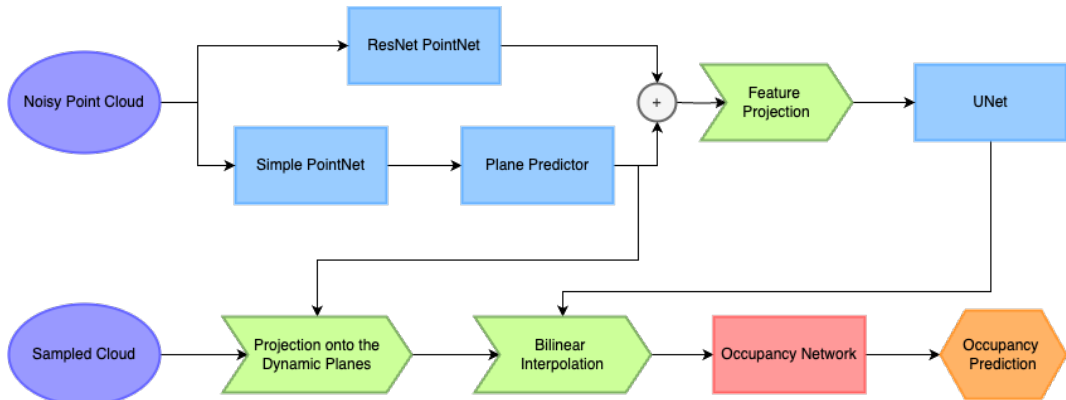
► Results

► Improvements



Architecture design

3 Architecture



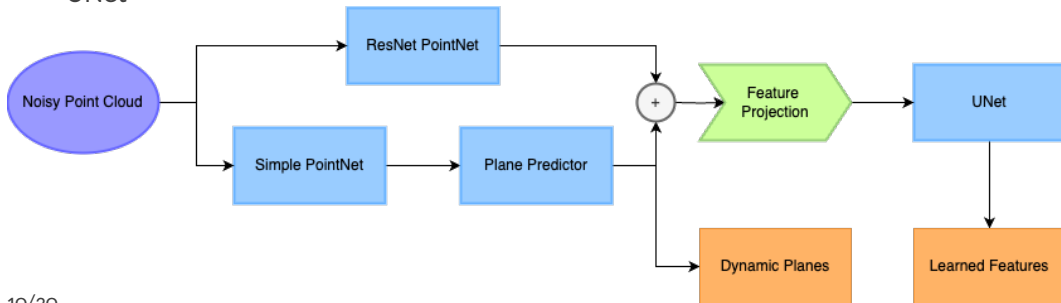


Encoder

3 Architecture

The Encoding part takes in input the Noisy Cloud and it's composed by the following steps:

- ResNet PointNet
- Simple PointNet + Plane Predictor
- Feature summation + projection
- UNet





Decoder

3 Architecture

The Decoding part is composed by:

- Feature Projection and Bilinear Interpolation
- Occupancy Network

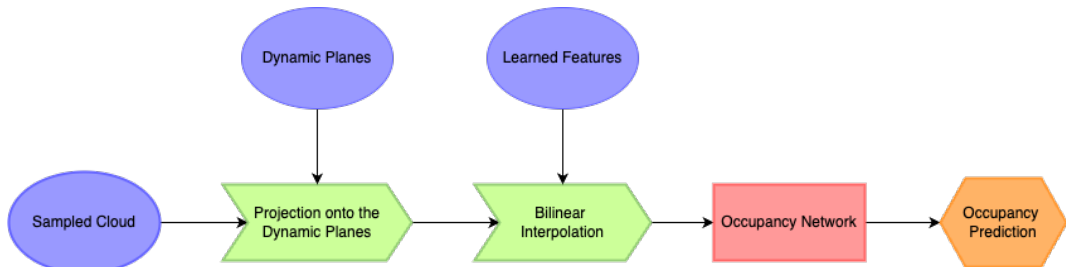




Table of Contents

4 Reconstruction

- ▶ Introduction
- ▶ Dataset
- ▶ Architecture
- ▶ **Reconstruction**
- ▶ Results
- ▶ Improvements



Multiresolution IsoSurface Extraction (MISE)

4 Reconstruction

- Create a grid over all the bounding box
- Evaluate the occupancy of each corner of the voxels
- Define the Active voxels as the one composed with at least one occupied corner and one not
- Subdivide each Active voxels into 8 subvoxels ($2 \times 2 \times 2$ grid) and evaluate the new points occupancy
- Repeat until the desired resolution is obtained

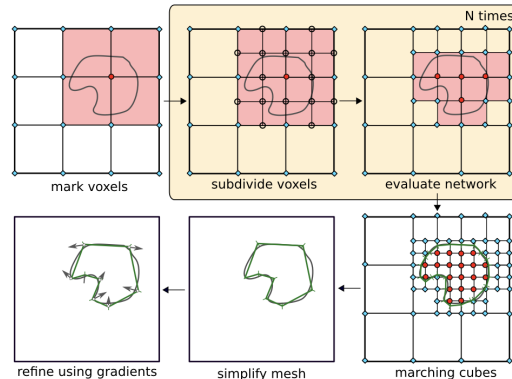




Table of Contents

5 Results

- ▶ Introduction
- ▶ Dataset
- ▶ Architecture
- ▶ Reconstruction
- ▶ **Results**
- ▶ Improvements



Metrics

5 Results

In order to evaluate the performance of our model, the following metrics have been used:

- Chamfer Distance :

$$CD(A, B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} \|a - b\|_2^2 + \frac{1}{|B|} \sum_{b \in B} \min_{a \in A} \|b - a\|_2^2$$

- IOU : $IoU(A', B') = \frac{|A' \cap B'|}{|A' \cup B'|}$
- F-Score:

Add each formula



Loss Performances

5 Results

Insert here plots



Metrics Performances

5 Results

Insert here just a table with metrics, gpu usage various types of sampling



Table of Contents

6 Improvements

- ▶ Introduction
- ▶ Dataset
- ▶ Architecture
- ▶ Reconstruction
- ▶ Results
- ▶ Improvements



Possible Changes and Future Improvements

6 Improvements



Point Cloud Occupancy with Dynamic Planes

Thank you for listening