# POINTNET++: DEEP HIERARCHICAL FEATURE LEARNING ON POINT SETS IN A METRIC SPACE

Charles R. Qi, Li Yi, Hao Su, Leonidas J. Guibas

Available at `https://arxiv.org/abs/1706.02413`

# INTRODUCTION

- **Context:** Analyzing geometric point sets, such as those from 3D scans, is crucial for applications like autonomous vehicles.
- **Challenges:** Point clouds must be invariant to permutations, and their local structures vary due to factors like scanning density.
- **Prior Work:** PointNet directly processes point sets but lacks the ability to capture local geometric structures effectively.
- **Our Approach:** Introducing PointNet++, a hierarchical network that processes points in a metric space to capture fine to coarse geometric structures.

# INTRODUCTION (CONT'D)

- **Design Goals:** PointNet++ uses overlapping local regions and a recursive application of PointNet to abstract point features hierarchically.
- **Unique Aspects:**
  - Overlapping partitions defined by a farthest point sampling (FPS) algorithm to ensure coverage.
  - Multi-scale neighborhood processing to handle non-uniform densities effectively.
- **Advantages:** Adapts to data and metric, ensuring efficient and effective feature extraction compared to traditional CNN approaches.

# PROBLEM STATEMENT

- **Metric Space Definition:** $\mathcal{X} = (M, d)$ where $M \subseteq \mathbb{R}^n$ and $d$ is the distance metric derived from Euclidean space.
- **Challenge:** Non-uniform density of $M$ in ambient space, impacting the effectiveness of local pattern learning.
- **Goal:** Develop function $f$ that can process $\mathcal{X}$ to produce semantic outputs, such as classification or segmentation.
- **Relevance:** Understanding and processing the varying densities and scales of point sets to extract meaningful semantic information.

# REVIEW OF POINTNET

- PointNet processes unordered point sets directly, approximating any continuous set function.
- Uses a multi-layer perceptron (MLP) to encode spatial information of each point.
- Aggregates features using a max pooling operation, ensuring permutation invariance.

# HIERARCHICAL POINT SET FEATURE LEARNING

- Builds a hierarchical grouping of points to abstract local regions progressively.
- Integrates multiple levels of set abstraction:
  - Sampling layer selects centroids via farthest point sampling (FPS).
  - Grouping layer forms local regions around these centroids.
  - PointNet layer encodes local regions into feature vectors.

## ROBUST FEATURE LEARNING UNDER NON-UNIFORM SAMPLING

- PointNet++ uses multi-scale grouping (MSG) and multi-resolution grouping (MRG) to handle varying densities.
- Adapts to local density changes, improving robustness and detail capture:
  - MSG captures multiple scales simultaneously.
  - MRG combines features across different resolutions.

## FEATURE PROPAGATION FOR SET SEGMENTATION

- Hierarchical feature propagation strategy interpolates features from sampled to original points.
- Inverse distance weighting ensures smooth feature transition across points.
- Combines interpolated and directly linked features for enhanced context:

$$f^{(j)}(x) = \frac{\sum_{i=1}^{k} w_i(x) f_i^{(j)}}{\sum_{i=1}^{k} w_i(x)} \quad \text{where} \quad w_i(x) = \frac{1}{d(x, x_i)^p}, \ j = 1, \ldots, C \tag{1}$$

# OVERVIEW OF DATASETS

- **MNIST:** 60k training and 10k testing samples of handwritten digits.
- **ModelNet40:** 9,843 training and 2,468 testing samples of 3D CAD models.
- **SHREC15:** 1,200 shapes from 50 categories, using five-fold cross-validation.
- **ScanNet:** 1,513 scanned and reconstructed indoor scenes with 1,201 for training and 312 for testing.

# CLASSIFICATION EXPERIMENTS ON MNIST AND MODELNET40

- Evaluated on MNIST (2D) and ModelNet40 (3D).
- **MNIST:** Images converted to 2D point clouds.
- **ModelNet40:** 3D point clouds sampled from mesh surfaces.
- Networks trained with 512 points for MNIST and 1024 points for ModelNet40.
- Additional features (normals) and more points used for higher performance.

# RESULTS OF 2D AND 3D POINT SET CLASSIFICATION

# EVALUATING ROBUSTNESS TO POINT DENSITY VARIATION

- Tested robustness by randomly dropping points during test.
- Demonstrated strong performance despite reduced point density.
- Multi-scale and resolution strategies maintain high accuracy.

## SEMANTIC SCENE LABELING ON SCANNET DATASET

- Goal: Predict semantic labels for points in indoor scans.
- Compared with baselines on the ScanNet dataset.
- Hierarchical feature learning captures multi-scale geometric features.

# VISUALIZATION OF FEATURE LEARNING ON 3D POINT CLOUDS

- Visualization of learned 3D point cloud patterns from ModelNet40.
- Illustrates the network's ability to recognize various geometric features.