

# 卒業論文概要書

## Summary of Bachelor's Thesis

Date of submission: 01/31/2023 (MM/DD/YYYY)

学科名 Department	情報通信	氏名 Name	小野蒼平	指導 教員 Advisor	渡辺 裕 ㊞
研究指導名 Research guidance	オーディオビジュアル 情報処理研究	学籍番号 Student ID number	1W192080-6		
研究題目 Title	局所構造を考慮した点群分類法における点群サンプリングについて On Point Cloud Sampling in Point Cloud Classification Methods Considering Local Structure				

### 1. まえがき

近年、三次元形状をデジタル記述する手法として点群が注目されている。点群は三次元座標を持つ点の集合として形状を表現する手法であり、他の表現手法と比較すると計測のしやすさにおいて優れている[1]。本研究では、点群のクラス分類 (Classification) を扱う。点群には、順不変性、移動不変性、局所性の3性質があり、これらを考慮したネットワークを構成する必要がある。PointNet [2]は、そのうちの順不変性、移動不変性を考慮したネットワークであり、分類およびセグメンテーションを行うことが可能である。さらに、局所性を満足できていないという PointNet の欠点を克服したものが PointNet++ [3]である。サンプリングとグルーピングを行い PointNet によって局所的な特徴を得るといった工程を階層的に繰り返すことが PointNet++の特徴であり、性能の改善に成功している。このネットワークはサンプリングに FPS: Farthest Point Sampling [4]を使用している。FPS は点群の輪郭を明瞭に残してサンプリングすることができる手法であるが、計算に時間がかかるという欠点がある。そこで、本研究ではサンプリング層の一部及び全体を別のサンプリング手法に変更し比較することで、より実用的な性能と所要時間のバランスを考察する。

### 2. 従来手法

#### 2.1 PointNet

PointNet は点群データを入力して分類およびセグメンテーションを行うネットワークである。最大要素を出力する Max Pooling 関数によって順不変性を、T-Net というネットワークで入力点群から得たアフィン行列をアフィン変換として作用させることで移動不変性を獲得している。

#### 2.2 PointNet++

PointNet++は PointNet と同じく点群データを入力して分類およびセグメンテーションを行うネットワークである。サンプリング層でダウンサンプリングを行って選択された点を、グルーピング層でグループ化し、PointNet によって局所性を得る。サンプリング層とグルーピング層は二つずつ存在する。サンプリング手法は FPS である。FPS は、ランダムに最初の点を選択し、その点から最も距離が離れた点を二つ目の点として選択する。次に、その点から最も距離が離れた点を三つ目の点とし、これを繰り返す。また、グルーピング手法は rNN [5]である。

### 3. 比較条件

PointNet++のサンプリング層において用いられるサンプリング手法を別のサンプリング手法に変更してクラス分類実験を行い、分類性能及び計測時間を比較する。以下、二つあるサンプリング層のうち一つ目を第1サンプリング層、二つ目を第2サンプリング層とする。第1サンプリング層で1,024個の点を512個に、第2サンプリング層で512個の点を128個にサンプリングする。この2つのサンプリング層のうちどちらか一方または両方を別のサンプリング手法に置き換えて実験を行う。使用する代替サンプリング手法は、全ての点を無作為に選択するランダムサンプリング (RS: Random Sampling) と、無作為に選んだ一つ目の点から等間隔に点を選択する系統抽出法 (SRS: Systematic Random Sampling) の2種類である。PointNet++の手法を FPS-FPS とし、RS-RS, SRS-SRS, FPS-RS, RS-FPS, FPS-SRS, SRS-FPS の計7条件で比較する。

#### 4. 評価実験と結果

クラス分類実験を 7 条件の下で実行する。学習とテストには ModelNet10 [6]を用いる。このデータセットは 4,899 種類の点群データがあり、トレーニング用とテスト用に分かれている。これを用いて 7 条件それぞれにおいて学習を行う。Epoch 数は 200 に設定する。ModelNet10 には 10 種類のクラスがあり、予測と実際のクラスが一致するかどうかを調べ、Instance Accuracy,

$$Acc_i = \frac{\text{total hits}}{\text{total predicts}} \quad (1)$$

Class Accuracy ( $C = 10$ )

$$Acc_c = \frac{1}{C} \sum_{class} Acc_{class} \quad (2)$$

の 2 種類の評価指標を用いる。ただし、本概要書には Instance Accuracy のみ結果として述べる。クラス毎の正解率  $Acc_{class}$  も記録する。また、プログラムの開始から終了まで全て含んだ計算時間を記録する。計算環境における GPU は GeForce GTX 1080 Ti を使用する。

各条件における分類性能と計算時間の関係を図 1 に示す。

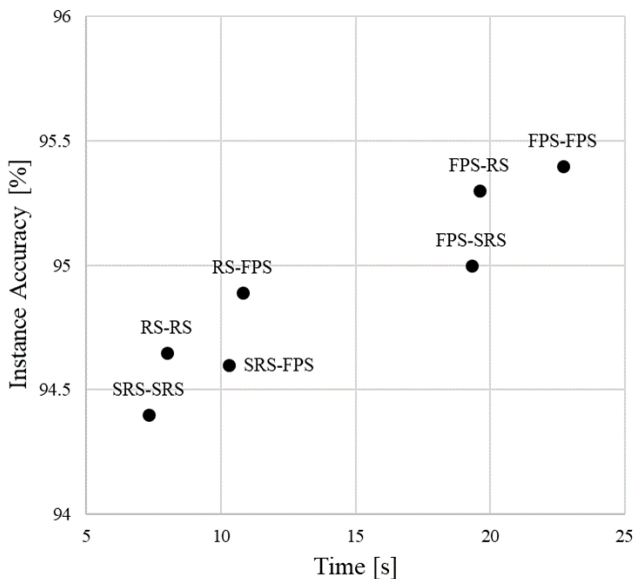


図 1 Instance Accuracy と計算時間の関係

この図より、FPS-FPS の分類性能が最も高く、第 2 サンプル層を RS に変更した FPS-RS がそれに近い正解率を示していることが分かる。また、RS は SRS より計算時間が長いですが分類性能が高

く、そのバランスは RS の方が優れているということも分かる。

また、クラス毎に見ると、FPS-FPS の正解率が最大とならないクラスが存在した。

#### 5. 結論

本研究では、点群分類手法の研究基盤となっている PointNet++において、サンプリング層の一部及び全てを別のサンプリング手法に変更し比較することで、より実用的な性能と所要時間のバランスを考察した。

実験により各条件での分類性能、計算時間およびクラス毎の性能を調査した結果、以下のような特性があることが分かった。まず、局所性を得るためのサンプリング手法としてランダムサンプリングは系統抽出法よりも適している。次に、ランダムサンプリングを使用することで計算時間を短くすることが可能で、初めに FPS を使用してからランダムサンプリングを行えば分類性能の低下も防ぐことができる。また、FPS-FPS の正解率が低いクラスが存在したことの根拠を解明できれば FPS の欠点の把握と改善につながるので、これについてのさらなる考察も必要である。

#### 参考文献

- [1] 千葉直也, “深層学習を用いた三次元点群処理入門,” 第 230 回 CVIM 研究発表会チュートリアル, May, 2022
- [2] C.R. Qi, H. Su, K. Mo, and L.J. Guibas, “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation”, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.652-660, Jul. 2017.
- [3] C.R. Qi, L. Yi, H. Su, and L.J. Guibas, “PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space”, International Conference on Neural Information Processing Systems (NIPS'17), pp.5105-5114, Dec. 2017.
- [4] Y. Eldar, M. Lindenbaum, M. Porat, and Y.Y. Zeevi. “The FarthestPoint Strategy for Progressive Image Sampling” IEEE Transactions on Image Processing, vol.6, no.9, pp.1305-1315, Sep.1997.
- [5] Jon L. Bentley, “A Survey of Techniques for Fixed Radius Near Neighbor Searching,” Stanford Linear Accelerator Center, Technical Report, SLAC-186; STAN-CS-75-513, Aug. 1975.
- [6] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, “3D ShapeNets: A Deep Representation for Volumetric Shapes,” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1912-1920, Jun. 2015.

2022 年度 卒業論文

局所構造を考慮した点群分類法における点群サンプリングについて

On Point Cloud Sampling in Point Cloud Classification Methods  
Considering Local Structure

提出日 2023 年 1 月 31 日

指導教員 渡辺 裕 教授

早稲田大学基幹理工学部 情報通信学科

1W192080-6

小野 蒼平

## 目次

第1章	序論.....	1
1.1	研究背景.....	1
1.2	関連研究と問題点, および研究目的.....	1
1.3	本論文の構成.....	2
第2章	関連研究.....	3
2.1	まえがき.....	3
2.2	PointNet.....	3
2.2.1	Max Pooling.....	3
2.2.2	T-net.....	4
2.3	PointNet++.....	4
2.3.1	サンプリング層.....	4
2.3.2	グルーピング層.....	5
2.4	サンプリング手法について.....	5
2.4.1	FPS: Farthest Point Sampling.....	5
2.4.2	ランダムサンプリング (RS: Random Sampling).....	5
2.4.3	系統抽出法 (SRS: Systematic Random Sampling).....	5
2.5	むすび.....	5
第3章	比較条件.....	6
3.1	まえがき.....	6
3.2	比較条件.....	6
3.3	むすび.....	7
第4章	実験と結果.....	8
4.1	まえがき.....	8
4.2	実験.....	8
4.2.1	学習方法.....	8
4.2.2	評価方法.....	8
4.3	実験結果.....	10
4.4	考察.....	12
4.4.1	RS と SRS の比較.....	12
4.4.2	RS と FPS の比較.....	13

4.4.3	クラス毎の分類性能 .....	13
4.5	むすび .....	13
第5章	結論 .....	14
5.1	結論 .....	14
5.2	今後の課題 .....	14
謝辞	.....	15
参考文献	.....	16
図一覧	.....	17
表一覧	.....	18
研究業績	.....	19

# 第1章 序論

---

## 1.1 研究背景

産業のデジタル化が進み、近年は様々な現場でデジタル技術の需要が拡大している。特に、現実の物体や現象を計算機上で再現する技術は需要が非常に高く、あらゆる分野の発展のために必要不可欠である。これを実現するためには三次元形状をデジタル記述する手法が必要であり、工夫の末様々な手法が考えられてきた [1]。例として、ボクセル、メッシュ、多視点画像、深度画像、そして点群が挙げられる。

点群とは、三次元座標を持つ点の集合として三次元形状を表現する手法である。データによっては、これに加え色や反射強度などの付加情報が与えられる。前述の手法と比較すると、点群は計測のしやすさにおいて優れている。点群データは写真測量やレーザーキャナーを用いて取得することができ、1回の計測で広範囲のデータを得られるため短時間で行うことが可能である。点群のタスクの中でも最も基本的なものの一つとして、クラス分類が挙げられる。点群の分類は顔認証システムなどに利用されるほか、物体検出やセグメンテーションなどの他の多くのタスクにも繋がるため重要である。

## 1.2 関連研究と問題点、および研究目的

点群のクラス分類とは、与えられた点群データがどのカテゴリに属するかを推定する技術である。二次元の画像でも同様のタスクがあるが、三次元点群には考慮しなくてはならない性質があり、扱いの難しさが課題となっていた。三次元点群の代表的な性質は順不変性、移動不変性、局所性の3つである [2]。順不変性とは、点に順序がなく、入れ替えても情報が変わらないという性質である。このため、順序を入れ替えて入力しても出力が変わらないネットワーク構造が必要である。移動不変性とは、平行移動や回転移動を作用させても情報が変わらないという性質である。同様にこちらにも対応したネットワーク構造が必要である。局所性とは、空間的距離が近い点同士が互いに関連性を持つという性質である。上記のように点群は順不同であるため、近傍点群の特徴量を得るために工夫が必要になる。PointNet [2]は、そのうちの順不変性、移動不変性を考慮したネットワークであり、分類およびセグメンテーションを行うことが可能である。さらに、局所性を満足できていないという欠点を克服したものが PointNet++ [3]である。サンプリングとグルーピングを行い PointNet によって局所的な特徴を得るといった工程を階層的に繰り返すことが PointNet++の特徴であり、性能の向上に成功している。このネットワークのサンプリング層では、FPS: Farthest Point Sampling [4]を使用している。FPS は点群の輪郭を明瞭に残してサンプリングすることができる手法であるが、計算に

時間がかかるという欠点がある。

そこで、本研究では PointNet++ のサンプリング層の一部及び全てを別のサンプリング手法に変更し比較することで、より実用的な性能と所要時間のバランスを考察する。実験により各条件での分類性能、計算時間およびクラス毎の性能を調査し、その傾向について模索する。

### 1.3 本論文の構成

以下に本論文の構成を示す。

第1章 研究の背景および目的について述べる。

第2章 関連研究および本研究で用いるサンプリング手法について述べる。

第3章 本研究の比較条件について述べる。

第4章 評価実験と、その結果について述べる。

第5章 結論と今後の課題について述べる。

## 第 2 章 関連研究

### 2.1 まえがき

本章では、従来の点群分類手法である PointNet および PointNet++ について述べる。ここでは、それらの関連研究のうち基本的に分類についてのみ述べる。また、本研究における比較実験の主題であるサンプリング手法について述べる。

### 2.2 PointNet

PointNet は点群データを入力して分類やセグメンテーションを行うネットワークである。従来の研究では点群の性質を考慮するためにボクセルや画像などに変換して入力する必要があるものが大半であったが、PointNet は点群データをそのまま入力して使用するネットワークであり、学習が容易であるという特徴がある。

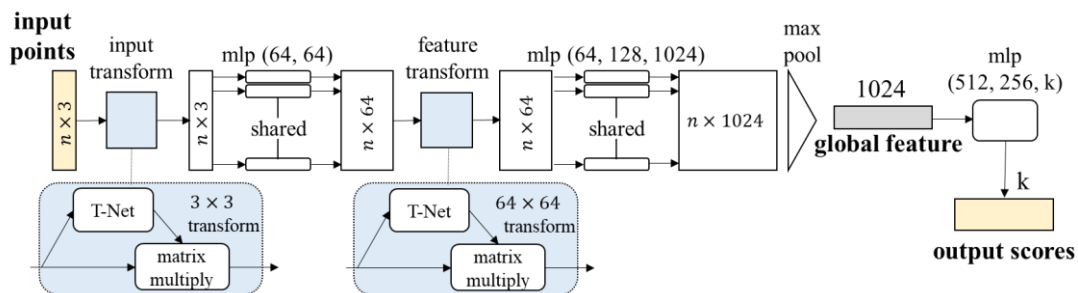


図 1 PointNet のアーキテクチャ (Classification)

図 1 のように、入力情報は点群の座標データである。本ネットワークでは単純化のために色などの付加情報は使用しない。出力は、分類クラスごとの正解率である。

以下、特に重要な構成要素である Max Pooling 関数と T-net について説明を行う。

#### 2.2.1 Max Pooling

Max Pooling は入力要素のうち最大の要素を出力する関数である。非常にシンプルな関数であるが、入力要素の順序に関わらず同じ結果が出力されるという特性が順不変性の獲得に適している。PointNet はネットワークの最終層で Max Pooling を用いることで順不変性を獲得している。



## 2.2.2 T-net

T-net は三次元点群を入力すると出力としてアフィン行列を得るニューラルネットワークである。出力されたアフィン行列をアフィン変換として入力点群に対して作用させることで、近似的に移動不変性を獲得している。

## 2.3 PointNet++

PointNet++は、PointNet の著者らによって提案された新たなネットワークである。クラス分類およびセグメンテーションを行うことが可能。PointNet では取り入れることのできなかった局所構造の特徴を取り入れた。

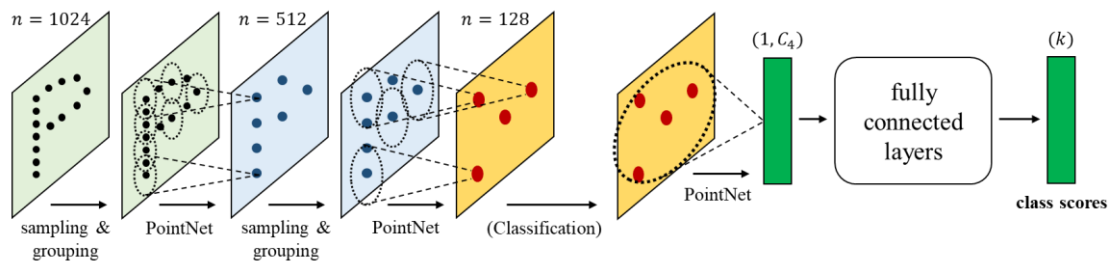


図2 PointNet++ のアーキテクチャ (Classification)

PointNet++は、PointNet と同様に三次元点群を入力情報とし、クラスごとの正解率を出力情報として得る。PointNet を階層的に適用し、サンプリングとグルーピングを繰り返すことで局所性を獲得した。

以下にサンプリング層、グルーピング層についての説明を行う。

### 2.3.1 サンプリング層

三次元点群の重心の情報を入力すると、指定の個数にダウンサンプリングされて出力される。ダウンサンプリングとは、情報量を失わないようにデータ数を減らす手法であり、計算量の短縮のためにデータの入力時に行われるのが一般的である。今回はその用途ではなく、点のグループを作るための代表点を選ぶために用いられる。サンプリングで使用される手法はFPS: Farthest Point Sampling である。この手法については後述する。サンプリングは2回行われ、クラス分類のプログラムでは1回目に1024個の点を512個に、2回目に512個の点を128個にサンプリングしている。なお、本研究において実験に使用したベースのプログラムでは、点群データの入力時のサンプリングにもFPSが使用されている。

### 2.3.2 グルーピング層

三次元点群とその重心の情報を入力すると、点集合のグループが出力される。グルーピングには rNN: radius-Nearest Neighbor [5]が使用されており、サンプリングされた点を中心とした半径  $r$  の球の内部にある点の集合がグループ化される。

## 2.4 サンプルング手法について

### 2.4.1 FPS: Farthest Point Sampling

最初の点をランダムに選択し、その点から最も距離が離れた点を二つ目の点として選択する。次に、その点から最も距離が離れた点を三つ目の点とし、これを繰り返す。第 1 章でも述べたように、FPS は計算量こそ多いが点群の輪郭を明瞭に残してサンプリングすることができる。

### 2.4.2 ランダムサンプリング (RS: Random Sampling)

点群の中から所定の数の点を重複しないよう無作為に選択する。点群の形状を残せるように考案された手法と比較するとまばらな箇所が出てきやすいが、計算量は非常に少ない。

### 2.4.3 系統抽出法 (SRS: Systematic Random Sampling)

母数と選択数から抽出間隔を算出し、点の並びをその抽出間隔で選択数と同数に分割する。その最初の区間の中から無作為に一つ目の点を選択し、その点から抽出間隔ごとに点を選択する。無作為に選択するという工程が 1 回のみであるため、ランダムサンプリングよりも計算量が少ない。点の並びに周期性がある場合は偏りが生じることがある。また、点が抽出間隔ごとに特徴の区別された層となっていればランダムサンプリングより精度が上がる場合もある。

## 2.5 むすび

本章では、従来の点群分類手法である PointNet および PointNet++のクラス分類について述べた。また、サンプリングの手法について述べた。

## 第3章 比較条件

---

### 3.1 まえがき

本章では，第4章で行うクラス分類実験の比較条件について述べる．

### 3.2 比較条件

以下，二つあるサンプリング層のうち一つ目を第1サンプリング層，二つ目を第2サンプリング層とする．前章でも述べたように，第1サンプリング層で1,024個の点を512個に，第2サンプリング層で512個の点を128個にサンプリングする．PointNet++ではどちらの層もFPSが用いられる．この二つのサンプリング層のうちどちらか一方または両方をランダムサンプリングまたは系統抽出法に置き換えたプログラムを用意し，元のプログラムと合わせて7条件で比較する．ランダムサンプリングと系統抽出法の組み合わせについては実験しない．

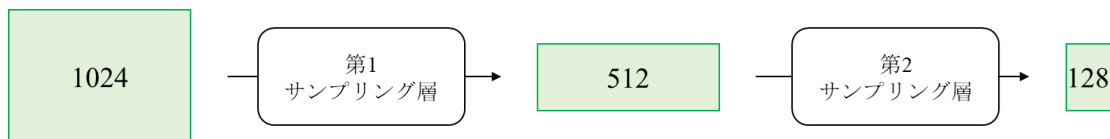


図3 PointNet++の二つのサンプリング層

また，前章で述べたデータ入力時のサンプリングに用いられるFPSについては変更しない．

以下に，各条件についての説明を述べる．

- FPS-FPS (Original)

PointNet++で使用される条件である．

- RS-RS

両方のサンプリング層をランダムサンプリングに置き換えた条件．

- FRS-FRS

両方のサンプリング層を系統抽出法に置き換えた条件．RS-RSとの比較に特に注目する．

- **FPS-RS**

第2 サンプル層のみをランダムサンプリングに置き換えた条件.

- **RS-FPS**

第1 サンプル層のみをランダムサンプリングに置き換えた条件. FPS-RS との比較に特に注目する.

- **FPS-SRS**

第2 サンプル層のみを系統抽出法に置き換えた条件. FPS-RS との比較に特に注目する.

- **SRS-FPS**

第1 サンプル層のみを系統抽出法に置き換えた条件. RS-FPS および FPS-SRS との比較に注目する.

### 3.3 むすび

本章では, PointNet++のモデル構造のうちのサンプリング層を FPS から別の手法に入れ替えて行う実験の7つの比較条件について述べた.

## 第4章 実験と結果

---

### 4.1 まえがき

本章では、第3章で述べた比較条件で行うクラス分類実験の方法と評価基準、実験結果および考察について述べる。

### 4.2 実験

#### 4.2.1 学習方法

データセットには ModelNet10 [6]を使用する。ModelNet10 は、4,899 種類の点群データがあり、トレーニング用とテスト用に分かれている。学習にはトレーニング用の 3,991 種類を使用する。1 種類につき 10,000 個の点があり、これを FPS でサンプリングし 1,024 個の点を使用する。また、ModelNet10 は三次元座標の他に色の情報を持つが、ここでは座標のみ使用する。第3章で述べた7種類の条件それぞれにおいて学習を行う。Epoch 数は 200 に設定する。

#### 4.2.2 評価方法

テストデータセットには ModelNet10 のテスト用の 908 種類を使用する。ModelNet10 は bathtub, bed, chair, desk, dresser, monitor, night\_stand, sofa, table, toilet の 10 種類のクラスがあり、予測と実際のクラスが一致するかどうかを調べ、Accuracy を算出する。Accuracy は、Instance Accuracy,

$$Acc_i = \frac{total\ hits}{total\ predicts} \quad (1)$$

Class Accuracy ( $C = 10$ )

$$Acc_c = \frac{1}{C} \sum_{class} Acc_{class} \quad (2)$$

の2種類の評価指標を用いる。クラス毎の正解率  $Acc_{class}$  も記録する。

Instance Accuracy は、ミクロ平均法またはマイクロ平均法と呼ばれる計算方法で、各クラスのサンプル数に偏りがある場合、サンプル数の大きいクラスの影響が大きくなる。対して、Class Accuracy は、マクロ平均法と呼ばれる計算方法で、クラスのサンプル数に過度な影響を受けず平等である。サンプル数に応じて重み付けをしたい場合は前者を、逆にその影響を抑えたい場合は後者を参照することが望ましい。

各クラスのカテゴリとテストデータの個数を表 1 に、視覚化したときの大まかな形状を図 4 に示す。

表 1 ModelNet10 のクラス一覧

Class	Category	Number of test data
0	bathtub	50
1	bed	100
2	chair	100
3	desk	86
4	dresser	86
5	monitor	100
6	night_stand	86
7	sofa	100
8	table	100
9	toilet	100

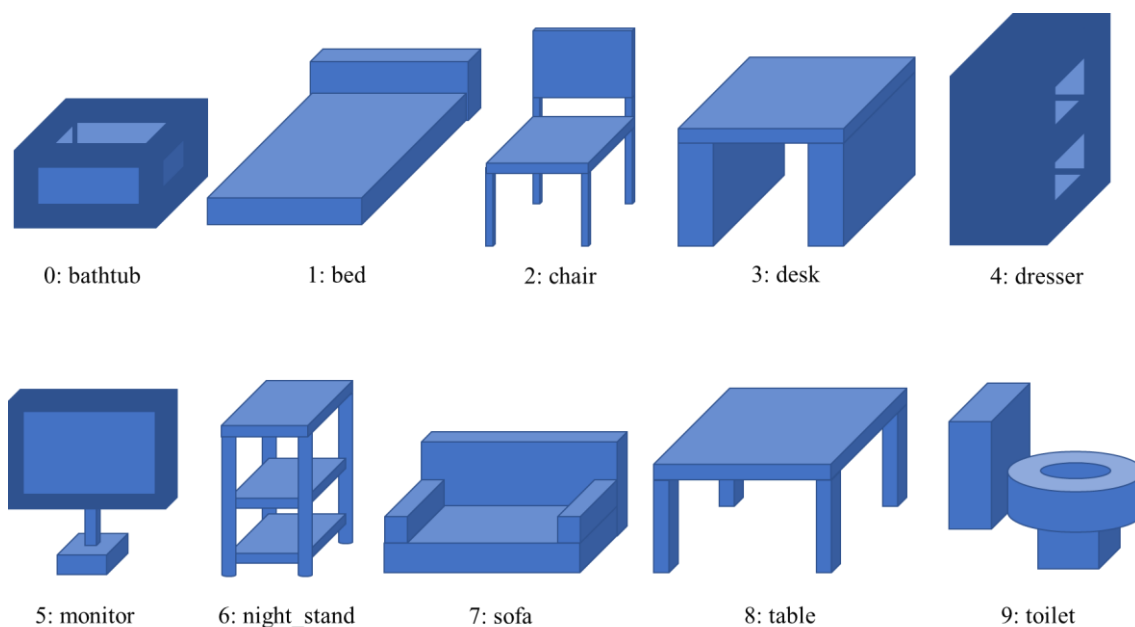


図 4 ModelNet10 のカテゴリ

また、テストの所要時間も測定する。プログラムの開始から終了まで全て含んだ時間を記録する。計算環境における GPU は GeForce GTX 1080 Ti を使用する。

本研究では、この分類性能と時間の二つの観点から分類方法の性能を評価する。

### 4.3 実験結果

各条件での実験結果について述べる。数回テストを行ったところ数値の散らばりが大きいことがわかった。そこで、10回テストを行いその平均や分散を記録した。分散の小ささも、分類性能や計算時間より優先度は低いが一つの性能指標として考慮する。

各条件における分類性能の平均を表2に示す。FPS-FPSの性能が最も高かったが、他条件でも大きな性能の低下はなかった。Instance AccuracyとClass Accuracyでは全条件でInstance Accuracyの数値が上回った。また、Instance AccuracyとClass Accuracyのどちらを参照しても条件間の順位は変動しなかった。

表2 各条件における分類性能の平均

Method	Instance Accuracy [%]	Class Accuracy [%]
FPS-FPS	<b>95.4</b>	<b>95.3</b>
RS-RS	94.65	94.6
SRS-SRS	94.4	94.3
FPS-RS	95.3	95.1
RS-FPS	94.89	94.88
FPS-SRS	95.0	95.0
SRS-FPS	94.60	94.5

次に、各条件における分類性能の分散を表3に示す。RS-FPSの条件で分散が最小になった。こちらは、Instance AccuracyとClass Accuracyのどちらを参照するかで部分的に順位変動が見られた。

表3 各条件における分類性能の分散

Method	(Instance Accuracy) [-]	(Class Accuracy) [-]
FPS-FPS	$3.26 \times 10^{-6}$	$3.32 \times 10^{-6}$
RS-RS	$2.71 \times 10^{-6}$	$2.79 \times 10^{-6}$
SRS-SRS	$4.65 \times 10^{-6}$	$4.63 \times 10^{-6}$
FPS-RS	$4.61 \times 10^{-6}$	$5.32 \times 10^{-6}$
RS-FPS	<b><math>1.50 \times 10^{-6}</math></b>	<b><math>1.40 \times 10^{-6}</math></b>
FPS-SRS	$8.99 \times 10^{-6}$	$8.64 \times 10^{-6}$
SRS-FPS	$3.40 \times 10^{-6}$	$3.74 \times 10^{-6}$

次に、各条件における計算時間の平均を表4に示す。FPSを別の手法に置き換えた条

件では計算時間が短縮された。

表 4 各条件における計算時間の平均

Method	Time [s]
FPS-FPS	22.7
RS-RS	8.0
SRS-SRS	<b>7.3</b>
FPS-RS	19.6
RS-FPS	10.8
FPS-SRS	19.3
SRS-FPS	10.3

各条件の Instance Accuracy と計算時間の比較を図 5 に示す。RS-RS は SRS-FPS より上位の性能である。また、RS は SRS より高性能であることが分かる。

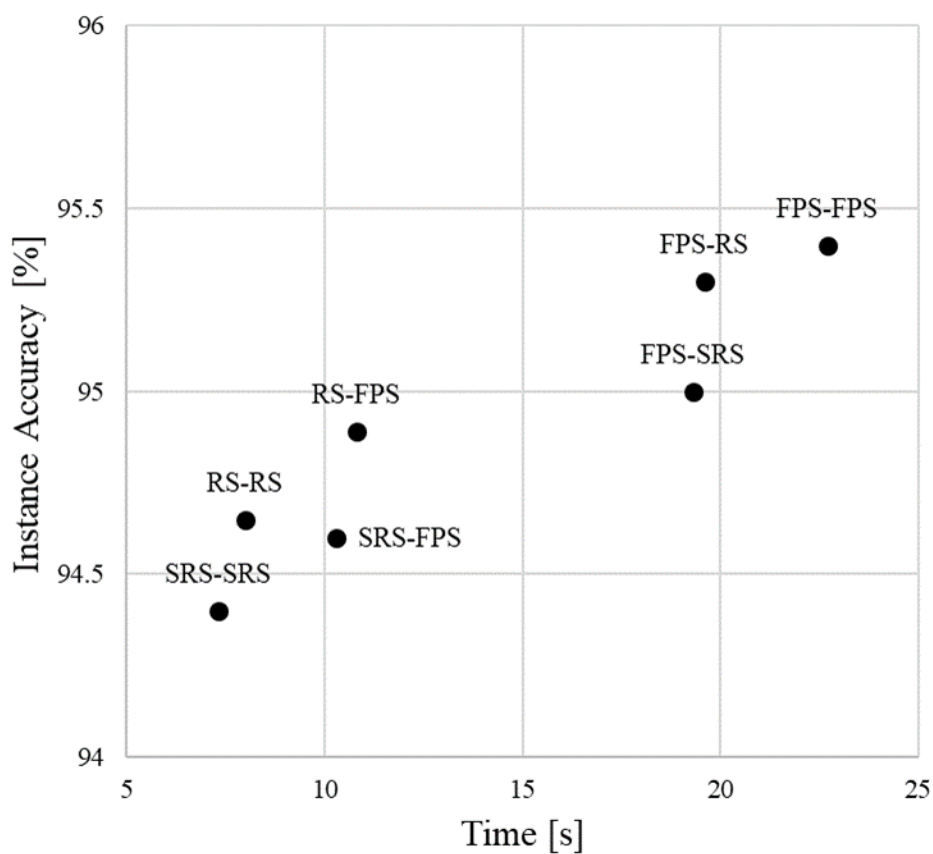


図 5 Instance Accuracy と計算時間の関係

各条件におけるクラス毎の分類性能を 10 回評価した結果を表 5,6 に示す。正解率で



他の条件が FPS-FPS を上回るクラスが存在した。

表 5 各条件におけるクラス毎の分類性能の平均 (クラス 0~5)

Method	0 [%]	1 [%]	2 [%]	3 [%]	4 [%]
FPS-FPS	<b>100</b>	<b>100</b>	<b>100</b>	90.2	90.6
RS-RS	<b>100</b>	97.4	<b>100</b>	90.9	88.8
SRS-SRS	<b>100</b>	99.5	<b>100</b>	90.3	90.0
FPS-RS	99.2	99.3	<b>100</b>	93.8	<b>92.2</b>
RS-FPS	<b>100</b>	<b>100</b>	<b>100</b>	<b>95.2</b>	90.0
FPS-SRS	99.8	<b>100</b>	<b>100</b>	92.2	92.0
SRS-FPS	<b>100</b>	<b>100</b>	<b>100</b>	94.9	91.5

表 6 各条件におけるクラス毎の分類性能の平均 (クラス 5~9)

Method	5 [%]	6 [%]	7 [%]	8 [%]	9 [%]
FPS-FPS	<b>99.0</b>	79.5	99.0	<b>94.8</b>	99.4
RS-RS	<b>99.0</b>	79.4	98.4	93.7	98.0
SRS-SRS	<b>99.0</b>	72.9	99.0	93.5	99.0
FPS-RS	98.0	77.2	99.9	93.8	98.0
RS-FPS	<b>99.0</b>	80.3	<b>100</b>	85.2	99.0
FPS-SRS	98.2	<b>80.6</b>	99.2	89.0	98.7
SRS-FPS	<b>99.0</b>	74.3	98.6	87.2	<b>100</b>

また、分類の誤りの傾向を調べたところ、すべての条件に共通してクラス 3: desk とクラス 8: table, クラス 4: dresser とクラス 6: night\_stand を混同する例が多く見られた。

## 4.4 考察

### 4.4.1 RS と SRS の比較

RS は SRS より計算時間の短さでわずかに下回るものの分類性能では上回った。図 4 から分かるように、この 2 項目のバランスでは RS の方が優れている。系統抽出法は点の並びに周期性が無ければランダムサンプリングと差がないものであるが、このような結果になったということは入力点群の並びに何らかの周期性があったということになる。点群には順不変性の性質があるため、入力の順番によって偏り方が左右される SRS は点群のサンプリングに使用するべきではないといえる。

#### 4.4.2 RS と FPS の比較

FPS-RS と RS-FPS を比較すると、分類性能は前者が上回り、計算時間の短さは後者が上回った。特に、FPS-RS は FPS-FPS と比較しても正解率で 0.1 ポイントしか下回らずに計算時間を 3 秒短縮しているため実用的な条件となる。FPS は形状の輪郭を取りやすいサンプリング方法であるため、後から行うよりも初めに行うのが効果的である。ただし、第 2 サンプリングを FPS から変更することで分散は大きくなる場合がある。

#### 4.4.3 クラス毎の分類性能

クラス 0: bathtub, クラス 1: bed, クラス 2: chair, クラス 5: monitor, クラス 7: sofa, クラス 9: toilet はどの条件でも非常に高い正解率であった。クラス 3: desk, クラス 4: dresser, クラス 8: table は少し劣るが高い正解率を出し、クラス 6: night\_stand が最も低くなった。また、クラス 3 は FPS-FPS の正解率が最も低く、クラス 4, 6 も FPS-FPS は最上位ではなかった。

### 4.5 むすび

本章では、比較実験に用いたデータセット、実験方法、性能について述べた。基本的に計算時間を短縮する条件を採用すると、Instance Accuracy が低下するという比例関係にあることが分かった。ただし、ほぼ同一の計算時間であっても、Instance Accuracy が多少改善される場合がある。このような例は、RS-FPS が SRS-FPS とほぼ同じ計算時間であっても、0.3%程度分類精度が改善される点に見られる。したがって、FPS 探索の条件設定によっては、分類性能の改善が可能であると考えられる。

## 第5章 結論

---

### 5.1 結論

本研究では、点群分類手法の研究基盤となっている PointNet++において、サンプリング層の一部及び全てを別のサンプリング手法に変更し比較することで、より実用的な性能と所要時間のバランスを考察した。

実験により各条件での分類性能、計算時間およびクラス毎の性能を調査した結果、以下のような特性があることが分かった。局所性を得るためのサンプリングの手法としてランダムサンプリングと系統抽出法を比較すると、ランダムサンプリングの方が適している結果になった。ランダムサンプリングも取り入れ方によって効果の現れ方が変わり、初めに FPS を行った後に使用すると FPS とランダムサンプリング双方の利点を生かすことができると分かった。また、カテゴリによってはランダムサンプリングを使用することで FPS を使用した場合の分類性能を上回るというケースもあった。

### 5.2 今後の課題

今回の結果を生かして、第 1 サンプリング層の初めの何割かで FPS を使って残りをランダムにサンプリングした場合の分類性能と計算時間を検証する予定である。

また、カテゴリによって分類性能に違いが出た根拠を解明できれば分類性能と計算時間の双方を向上させる手法を提案できる可能性があるため、さらなる原因究明が必要である。

## 謝辞

本論文の執筆に際して、丁寧なご指導をしてくださり、快適な研究環境を与えてくださった渡辺教授に感謝いたします。

毎週のゼミ活動において、研究の方向性に関して丁寧なアドバイスを下さった早稲田大学情報通信研究センター招聘研究員の原潤一博士に感謝いたします。

研究室における温かい環境を提供してくださり、本論文の執筆においても丁寧なアドバイスを下さった渡辺研究室の皆様感謝いたします。

最後に、私をここまで育ててくださり、常に心を支えてくださり、生活を支えてくださっている家族に感謝いたします。

## 参考文献

- [1] 千葉直也, “深層学習を用いた三次元点群処理入門,” 第 230 回 CVIM 研究発表会 チュートリアル, May, 2022.
- [2] C.R. Qi, H. Su, K. Mo, and L.J. Guibas, “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation” , 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.652-660, Jul. 2017.
- [3] C.R. Qi, L. Yi, H. Su, and L.J. Guibas, “PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space” , International Conference on Neural Information Processing Systems (NIPS'17), pp.5105-5114, Dec. 2017.
- [4] Y. Eldar, M. Lindenbaum, M. Porat, and Y.Y. Zeevi. “The Farthest Point Strategy for Progressive Image Sampling” IEEE Transactions on Image Processing, vol.6, no.9, pp.1305-1315, Sep.1997.
- [5] Jon L. Bentley, “A Survey of Techniques for Fixed Radius Near Neighbor Searching,” Stanford Linear Accelerator Center, Technical Report, SLAC-186; STAN-CS-75-513, Aug. 1975.
- [6] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, “3D ShapeNets: A Deep Representation for Volumetric Shapes,” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1912-1920, Jun. 2015.

## 図一覧

図 1	PointNet のアーキテクチャ (Classification).....	3
図 2	PointNet++ のアーキテクチャ (Classification).....	4
図 3	PointNet++の二つのサンプリング層 .....	6
図 4	ModelNet10 のカテゴリ .....	9
図 5	Instance Accuracy と計算時間の関係.....	11

## 表一覧

表 1 ModelNet10 のクラス一覧.....	9
表 2 各条件における分類性能の平均.....	10
表 3 各条件における分類性能の分散.....	10
表 4 各条件における計算時間の平均.....	11
表 5 各条件におけるクラス毎の分類性能の平均 (クラス 0~5).....	12
表 6 各条件におけるクラス毎の分類性能の平均 (クラス 5~9).....	12

## 研究業績