卒 業 論 文 概 要 書

Summary of Bachelor's Thesis

Date of submission: 01/30/2024 (MM/DD/YYYY)

学科名 Department	情報通信	氏名 Name	范洪睿	指 導 教 員 Advisor	渡辺 裕 印
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	1W202288-1		
研究題目 Title	残差 MLP と注意機構を導入した点群分類手法に関する研究 Research on Point Cloud Classification Method with Residual MLP and Attention Mechanism Integration				

1. まえがき

点群データは物体や環境の 3D 構造を表現するデータ形式であり，現実世界での分析や認識において重要な役割を果たしている．自動運転やロボットアームなどの応用領域では，点群処理のリアルタイム性と高い精度が要求される．

従来の点群分類手法として PointNet++[1]があるが，複雑な局所特徴の解析を行うため，膨大な計算が必要となる．そのため，処理時間の大幅な増加を招くという問題がある．そこで，本研究では PointNet[2]をもとに，残差 MLP と注意機構を導入することで，既存の高速で効率的な点群データ処理の利点を継承しつつ，更なる高精度を達成する点群分類手法を提案する．

2. 関連研究

PointNet は，3 次元点群データに対して，深層学習を適用したエンドツーエンドの点群分類とセグメンテーションの手法である． 最大要素を出力する MaxPooling 操作により，順序のない不規則な点群データを効果的に処理できるようになる．

残差モジュール[3]は入力データと残差を直接学習することで，特徴を捉える能力を向上し，学習効果を向上させる重要な手法となっている．

注意機構は，ニューラルネットワークにおいて，重要な特徴を強調し，モデルの表現力を向上させるための手法である．本研究で使われる CBAM[4]は注意機構の一種で，チャンネルアテンションと空間アテンションにより，チャンネルおよび空間ごとの情報を同時に考慮することができる．

3. 提案手法

提案手法では，PointNet のモデル構造に残差 MLP と注意機構を導入する．具体的には，Res MLP + CBAM Block の導入により，提案手法を実現する．提案手法のモデル構造を図 1 に示す．

Res MLP + CBAM Block は残差 MLP フレームワークと CBAM を直列に結合した構造となっている．こより優れた性能を発揮し，表現力の向上が見込まれる．Res MLP + CBAM Block で採用した残差モジュールは，勾配を効果的に伝達し，ショートカット接続により，ネットワークの学習効率と全体的な性能が向上する．Res MLP + CBAM Block で採用した CBAM は，チャネル方向の Channel Attention と，空間方向の Spatial Attention により，入力データの重要なチャンネルと空間情報に効果的に焦点を当て，入力の抽象的な表現を向上させることで，モデルの表現力をさらに向上させる．この構造は実験と分析に基づいて選択され，より優れた性能を発揮し，表現力の向上が見込まれる．

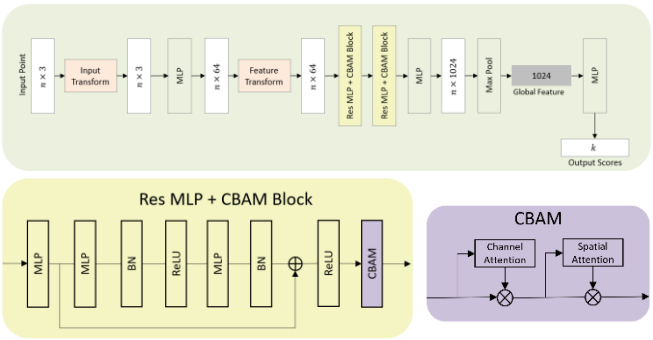

図 1 提案手法のモデル構造

4. 実験

4.1 有効性検証実験

本実験では，ModelNet40 データセット[5]を使用する．ModelNet40 データセットは，点群の分類タスクに使用されるデータセットであり，40 の異なるカテゴリを含んでいる．提案手法，バックボーンの PointNet，および局所特徴を分析する代表的な点群分類手法である PointNet++の 3 種類の手法

について，精度と速度を比較する．精度の評価では，Instance Accuracy を採用する．速度の評価では，テスト時間の測定を採用する．Instance Accuracy の計算式をは式 1 に示す．この指標はモデルが各サンプルで正確に予測した割合を示す．

$$\mathrm{Acc} = \frac{Correctly\ predicted\ samples}{Total\ \ samples} \times 100\% \quad (1)$$

各手法を用いて精度とテスト時間の比較を行った結果を表 1 に示す．

表 1　各手法の比較

	PointNet	PointNet++	Ours
Instance Accuracy [%]	89.2	90.7	**91.7**
Testing Time [s]	**4.3**	19.6	4.4

PointNet は最も高速である一方で，精度は最も低いことが確認された．PointNet++は局所特徴解析を実施しない PointNet に比べて精度が向上しているが，これに伴いテスト時間も 4 倍ほどに増加している．一方，提案手法は，テスト時間は PointNet よりわずかに増加しているが，これら三つの手法の中で最も高い精度を示した．

4.2 対照実験

提案手法の構造が最適な組み合わせであるかを検証するために，対照実験を行った．一番目の対照実験では，残差 MLP と CBAM の接続順序を変更し，残差 MLP と CBAM の位置順序がモデルの性能に与える影響について検討する．二番目の対照実験では，提案の Res MLP + CBAM Block の適用個数を二つから三つに増やし，Res MLP + CBAM Block の適用個数がモデルの性能に与える影響について検討する．対照実験で使われる対照用構造は図 2 に示す．また，二つの対照実験の実験結果はそれぞれ表 2 と表 3 に示す．

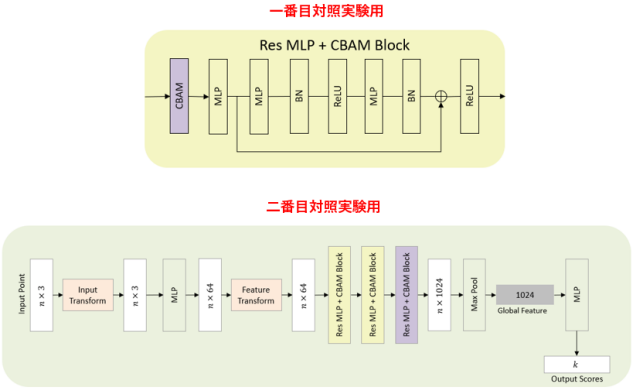

図 2 対照実験で用いる対照用構造

表 2 一番目の対照実験の結果

	Comparison	Ours
Instance Accuracy [%]	91.6	**91.7**
Testing Time [s]	4.4	4.4

表 3 二番目の対照実験の結果

	Comparison	Ours
Instance Accuracy [%]	**91.9**	91.7
Testing Time [s]	6.3	**4.4**

一番目の対照実験では，提案手法のモジュールと比較用モジュールの速度と精度がほぼ同等であることが確認された．

二番目の対照実験では，提案手法の適用モジュール個数が 3 個に増加しても，比較用モジュールと比較して速度が速く，精度はほぼ同等であることが分かった．

5. 結論

本研究では，局所情報の分析を実施しない PointNet に対し，残差 MLP と注意機構の CBAM をもとに構成した Res MLP + CBAM Block を導入した新たな点群分類手法を提案した．実験では，速度を維持しつつ精度を向上させることを確認でき，提案手法の構造の有効性を確認できた．

今後は，他の残差 MLP と注意機構の構造を導入する実験を行い，より計算負荷が少ない構造を考案する必要があると考えられる．

参考文献

[1] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, “PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space,” Neural Information Processing Systems (NIPS 2017), pp. 5105-5114, Dec. 2017.

[2] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “PointNet: Deep Learning on Point Sets for 3D Classification and

Segmentation," IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR 2017), pp. 652-660, Jul. 2017.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR 2016), pages 770–778, Jun. 2016.

[4] S. Woo, J. Park, J. Lee, and I. Kweon, "CBAM: Convolutional Block Attention Module," European Conference on Computer Vision (ECCV 2018), pp. 3-19, Sep. 2018.

[5] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, "3D ShapeNets: A Deep Representation for Volumetric Shapes," IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015), pp. 1912-1920, Jun. 2015.

2023 年度　卒業論文

残差 MLP と注意機構を導入した点群分類手法に関する研究

Research on Point Cloud Classification Method with Residual MLP and Attention Mechanism Integration

指導教員　渡辺 裕　教授

提出日：2024 年 1 月 30 日

早稲田大学　基幹理工学部　情報通信学科

1W202288-1

范 洪睿

目次

第１章　序論 3

１．１　研究背景 3

１．２　関連研究と問題点，および研究目的 3

１．３　本論文の構成 4

第２章　関連技術 5

２．１　まえがき 5

２．２　PointNet 5

２．２．１　T - Net 5

２．２．２　最大プーリング操作 6

２．３　残差モジュール 6

２．４　CBAM 7

２．５　むすび 8

第３章　提案手法 9

３．１　まえがき 9

３．２　提案手法 9

３．３　むすび 11

第４章　実験・実験結果 12

４．１　まえがき 12

４．２　有効性検証実験 12

４．２．１　データセット 12

４．２．２　学習方法 13

４．２．３　テスト及び評価方法 13

４．２．４　実験結果 14

４．３　対照実験 14

４．３．１　対照内容 14

４．３．２　対照実験結果 15

４．４　考察 15

４．５　むすび 16

第５章　結論と今後の課題 17

５．１　結論 17
５．２　今後の課題 17

謝辞 18

参考文献 19

図一覧 20

表一覧 21

研究業績 22

第１章　序論

１．１　研究背景

点群データは，物体や環境の 3 次元構造を表現するためのデータ形式で，現実世界の解析や認識において重要な役割を果たしている．特に，自動運転や，ロボットアームなど，複雑な処理が必要となる分野において，点群データの分類タスクは非常に重要である[1].

自動運転分野において，点群の分類は安全な自動車の運行において不可欠である．レーザレーダーにより生成された点群データは，道路，歩行者，障害物などの三次元情報を提供し，これにより自動車は周囲の環境を正確に理解し，適切な運転を行うことができる．分類タスクにおいては，異常検知や障害物の種別判別が含まれ，これによって安全性や運転の滑らかさが向上する．ロボティクス分野では，点群の分類はロボットのナビゲーションや環境認識において極めて重要である．例えば，サービス型ロボットが複雑な環境で特定のタスクを実行する際には，周囲の点群データから障害物を正確に識別し，適切な動きを計画する必要がある．このような分類タスクにおいて，点群データはロボットの行動の洗練度や安全性に直結している．三次元モデリング分野においては，点群の分類が高精度なモデルの生成に寄与する．建築，土地計画，文化遺産の保護などの分野では，点群データを用いて現実世界の物体や地形を正確にモデリングすることが必要である．分類タスクでは，これらのモデリングにおける物体の種別識別や詳細な特徴の抽出が要求される．

点群の分類タスクの進展は，これらの技術がより安全かつ効果的に機能し，現実の複雑な環境に適応する能力を向上させることに繋がる．現在の分類タスクでは，処理速度と分類性能のトレードオフが課題となっている．一部の手法は高い分類性能を示しているが，その一方で処理速度が低下する傾向がある．そのため，このトレードオフを解消する新しい手法の提案や，既存の手法の改善が今後の研究の重要な課題となる．より高速で効率的な点群の分類手法を確立することで，様々な分野において新たな革新をもたらすことができる．

１．２　関連研究と問題点，および研究目的

従来の点群分類手法として PointNet++[2] があるが，複雑な局所特徴の解析を行うため，膨大な計算量が必要となり，処理時間の大幅な増加を招くという問題がある．この問題が，大規模な点群データに対する高速で効率的な処理の障害となっている．

本研究の目的は，現在の点群の分類タスクにおける処理速度と分類性能のトレードオフの問題を解決し，大規模かつ複雑な環境においても優れた性能を発揮できる新たなモデルを提供することにある．提案するモデルは PointNet[3]を基本的なフレームワークと

して採用している．PointNet は局所的な特徴情報を分析せずに全体の点群を処理する特徴があり，これにより高速な処理が可能である．本論文では PointNet の枠組みをベースにすることで，既存の高速で効率的な点群データ処理の利点を継承しつつ，更なる高精度を達成することを目指している．

１．３　本論文の構成

本論文の構成を以下に示す．

第１章　本章であり，本研究の背景，関連研究と問題点および研究目的について述べる．
第２章　本研究で用いる従来の点群分類手法および関連技術について述べる．
第３章　本研究の提案手法について述べる．
第４章　本研究における実験の方法，結果および考察について述べる．
第５章　結論と今後の課題について述べる．

第２章　関連技術

２．１　まえがき

本章では，バックボーンの点群分類手法である PointNet について述べる．また，関連技術である残差ネットワーク ResNet および注意機構である Convolutional Block Attention Module（CBAM）について述べる．

２．２　PointNet

PointNet は，点群データに対して，深層学習を適用するためのエンドツーエンド手法である．従来の手法は，主に画像に対して分析を行ってきたが，点群データのような順序のない不規則な 3 次元データに対しては効果的な処理手法がなかった．そこで，PointNet が誕生した．PointNet のモデル構造を図２．１に示す．

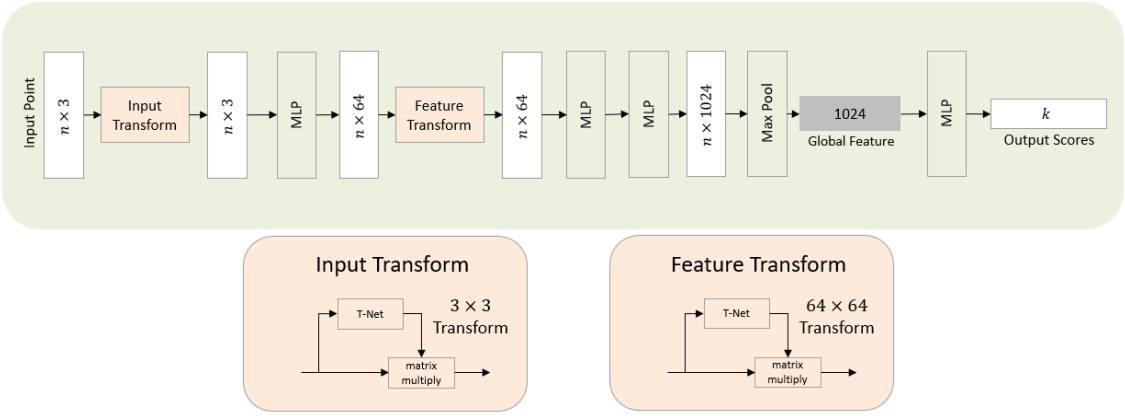

図２．１　PointNet のモデル構造

PointNet のネットワーク構造は T-Net, MLP と最大プーリング操作から構成されている．T-Net は，入力された点群データの変換行列を学習するためのネットワーク構造である．PointNet のアーキテクチャでは，MLP は点群内の各点の特徴を抽出するために使用される．これらの MLP 層は各点の元々の特徴に対して非線形変換を行い，ネットワークがより高度な表現を学習できるようにする．PointNet はグローバルな特徴を学習することで，点群の構造と情報を学習することができ，点群データの分類タスクとセグメンテーションタスクを実現する．以下に，特に重要な構成仕組みである T-Net と最大プーリング操作について説明を行う．

２．２．１　T - Net

T-Net は，点群の座標変換を学習するメカニズムであり，ネットワークが異なる座

標系や配置に対して頑健な特徴を抽出できるようになる．これにより，PointNet は入力の点群が異なる座標系や配置に対して不変であることが保証される．具体的には，T-Net は，入力として点群を受け取り，その点群の座標変換行列を学習する．この変換行列は，個々の点の位置を調整するために使用され，点群全体に対して一貫した変換を施す．これにより，入力の点群が異なる位置や向きにある場合でも，ネットワークはこれを補正し，同一の表現に変換することができる．

２．２．２　最大プーリング操作

最大プーリング操作は，入力データに対する特定の領域から最大の値を選択する効果的な手法である．点群データの無秩序な性質から，操作の際には点のデータの順序が変更されても同じ結果を得る必要がある．最大プーリング操作は，この状況に応えるために採用されている．特徴ごとに最大値を選択することで，点の順序が変更されても最大値は不変となり，正確な情報が記録される．これにより，点群の無秩序な性質に柔軟に対応しつつ，PointNet が安定して代表的な特徴を抽出できるようになっている．

２．３　残差モジュール

深層学習において，残差モジュール[4]はネットワークの訓練を安定化させ，学習効率を向上させる重要な手法となっている．残差モジュールは，通常，ショートカット接続と呼ばれる構造を含んでおり，この構造によって入力データと残差（差分）を学習する．ショートカット接続は，モデルが学習する際に直接的な情報の伝達を可能にするため，残差モジュールはモデル全体において勾配の安定性を保ちつつ，学習を行うことができる．これにより，ディープニューラルネットワークでも効果的に情報を伝達し，逆伝播中に勾配が適切に伝搬されるため，学習プロセスが円滑に進行する．

また，残差モジュールは入力データと残差を直接学習することで，ネットワークが特定の特徴やパターンを抽出しやすくなる．これにより，モデルはより複雑な表現を学び，非線形性を捉える能力が向上する．残差モジュールの構造を図２．２に示す．

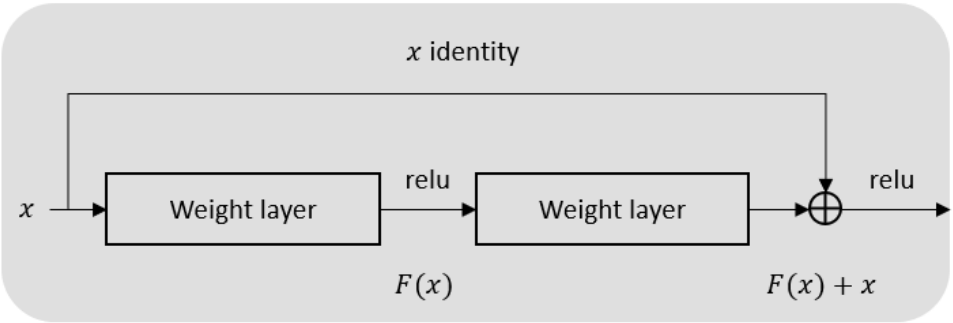

図２．２　残差モジュールの構造

２．４　CBAM

注意機構は，ニューラルネットワークにおいて，重要な特徴を強調し，モデルの表現力を向上させるための手法である[5]．Convolutional Block Attention Module（CBAM）[6]は，一種の注意機構で，Channel Attention（チャンネルアテンション）と Spatial Attention（空間アテンション）により，チャンネルごとおよび空間的な情報を同時に考慮することができる機構を提供する．

Channel Attention はチャンネルごとの重要性を学習するための機構である．前層の出力に対して，チャンネルごとに重みを掛けて重要な情報を強調する．これにより，異なる特徴マップの重要性を考慮し，有益な情報を抽出できる．Spatial Attention は空間的な特徴の重要性を学習するための機構である．畳み込み層の出力に対して，位置ごとに重みを掛けて重要な領域を強調する．これにより，モデルは重要な領域を注視できる．CBAM はこれらのアテンションメカニズムを組み合わせ，ニューラルネットワークが入力データの中で重要な情報を効果的に捉えるための優れた特徴抽出ができるようになっている．CBAM の構造を図２.３に示す．

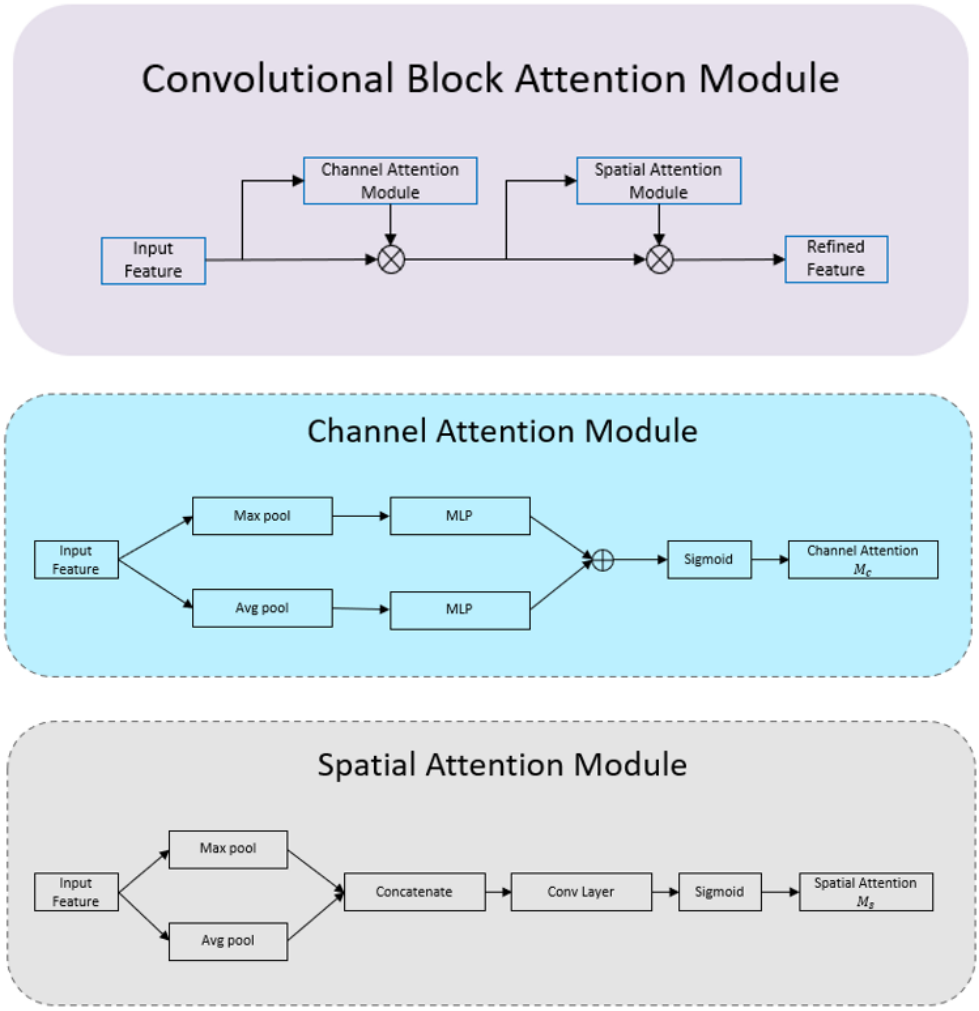

図２.３　CBAM の構造

２．５　むすび

本章では，バックボーンである PointNet のモデル構造が，T-Net, MLP と最大プーリング操作からなることを示した．また，関連技術である残差モジュール，CBAM について Channel Attention と Spatial Attention が含まれることを述べた．

第３章　提案手法

３．１　まえがき

本章では，PointNet のモデル構造に，残差 MLP と注意機構である CBAM を導入した新たな点群分類手法を提案する．提案手法では，局所特徴を探索しないため，速度を確保しながらモデルの認識能力が向上させることができる．

３．２　提案手法

従来の点群分類の手法では，PointNet++のように局所的な領域の特徴を分析することで精度向上を図るアプローチが主流であった．しかし，この手法には膨大な計算量が伴い，処理時間の大幅な増加を招くという問題がある．そのため，局所的な情報分析を実施しない PointNet をもとに, 処理速度を維持しつつデータ構造を捉える能力を向上させる新たな手法を提案する．提案手法のモデル構造を図３.１に示す.

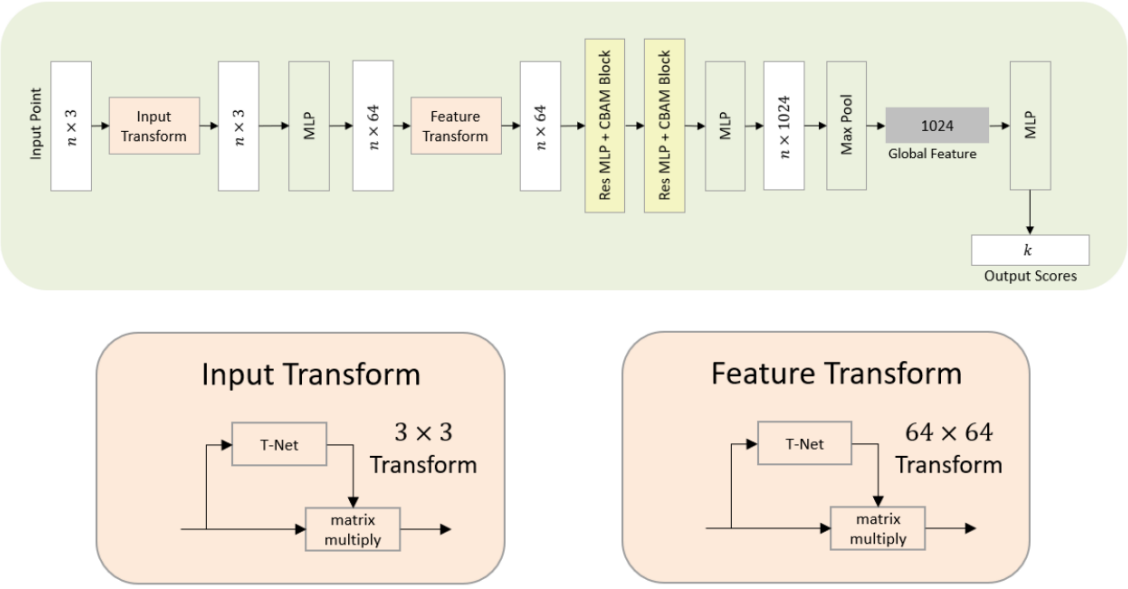

図３.１　提案手法のモデル構造

まず，入力データを Input Transform ブロックに入力し，T-Net を使用して入力点群に対する線形変換を行う．これにより，モデルは入力データの回転や平行移動に対して不変性を獲得する．次に，MLP を使用して特徴を抽出し，Feature Transform ブロックに入力し，再び T-Net を適用して得られた特徴に対して線形変換を行う．これにより，モデルはさらに特徴から学習したグローバル特徴に対して不変性を獲得する．Input Transform と Feature Transform の構造を図３.２に示す.

その後，前の層の結果を提案する Res MLP + CBAM Block に入力する．Res MLP +

CBAM Block のモデル構造を図３.２に示す．次に，前の層の結果を再び MLP 層に入力し，新しい特徴を抽出する．そして MaxPooling 操作により全ての特徴を集約し，最も顕著な特徴を抽出し，ネットワークが点群データをより包括的に理解できるようにする．最後に，ネットワークの出力を全結合層に接続して分類を行う．

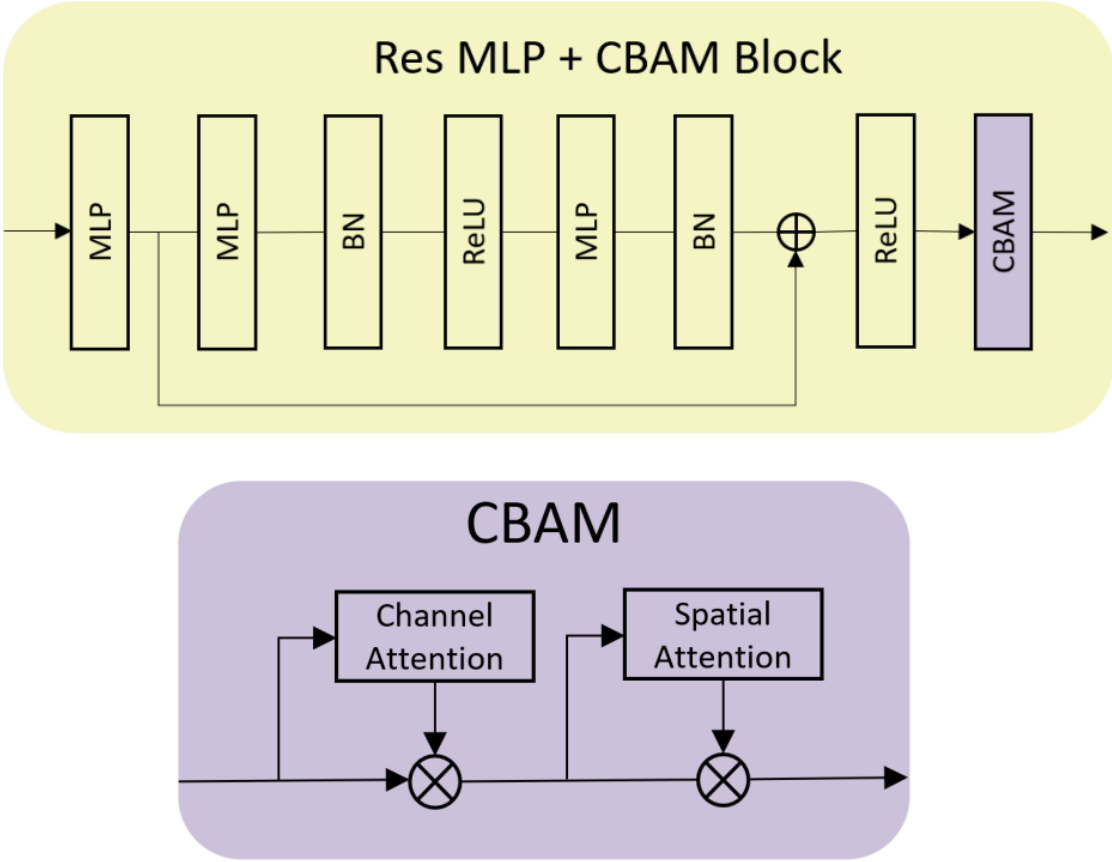

図３.２　Res MLP + CBAM Block の構造

Res MLP + CBAM Block は残差 MLP フレームワークと注意機構の CBAM を直列に結合した構造となっている．この構造は実験と分析に基づいて選択され，より優れた性能を発揮し，表現力の向上が見込まれる．

Res MLP + CBAM Block で採用した残差モジュールは，情報の伝達を最適化し，ネットワークの性能向上を目指している．このモジュールは，多層パーセプトロン（MLP），バッチ正規化（BN），および整流線形ユニット（ReLU）を組み合わせた構造を持っている．モジュールの流れは，左から右へと進み，MLP，MLP，BN，ReLU，MLP，BN，ReLU となる．一番目の MLP の後から残差接続が始まり，最後の ReLU の前に到達する．この構造により，モデルは勾配を効果的に伝達し，残差情報を学習しやすくなり，ネットワークのトレーニング効率と全体的な性能が向上する．

Res MLP + CBAM Block で採用した CBAM は，チャネル方向の Channel Attention と，空間方向の Spatial Attention を直列に結合した構造となっている．チャンネル注意力と

空間注意力を組み合わせ，入力データの重要なチャンネルと空間情報に効果的に焦点を当て，入力の抽象的な表現を向上させるので，モデルの表現力がさらに向上している．

３．３　むすび

本章では，本研究の提案手法である残差 MLP と注意機構を考慮した点群分類手法について述べた．具体的には，局所的な情報分析を実施しない PointNet のモデル構造に，残差 MLP と CBAM を導入した．残差 MLP の導入により，モデルは点群の複雑な特徴をより効果的に抽出し，モデルの表現力を向上させる．また，CBAM の導入により，チャネルと空間双方向の情報に対する注意を高め，重要な特徴に焦点を当てることで，モデルの精度向上に寄与する．

第４章　実験・実験結果

４．１　まえがき

本章では, 本研究で用いたデータセット, 実験方法, 実験結果, 考察について述べる. 具体的には，まず点群分類モデルの学習に用いたデータセット，実験方法，性能について述べる．その後，評価実験として，従来手法と提案手法の実験結果の比較を行う.

４．２　有効性検証実験

４．２．１　データセット

データセットには，ModelNet40 データセット[7]を使用する．ModelNet40 データセットは，点群の分類に使用されるデータセットであり，40 の異なるカテゴリを含んでいる．詳細なカテゴリを表４．１に示す．トレーニングセットには 9843 個の点群データがあり，テストセットには 2468 個の点群データが含まれている．データ形式は txt ファイルで，各 txt ファイルは特定のカテゴリオブジェクトを示す．点群データは 10000 点から成り，ModelNet40 データセットでは各点が 6 次元の情報を持っている．これらの次元は[x，y，z，nx，ny，nz]で，(x, y, z)は点の空間座標を，(nx, ny, nz)は点の空間中の法線ベクトルを表す.

点群可視化ソフトウェアの MeshLab を使用して, 点群データを視覚化することができる．これらのデータの可視化の例を図４．１に示す.

表４．１ 各条件における分類性能の分散

Class	Category	Class	Category	Class	Category	Class	Category
0	airplane	10	cup	20	laptop	30	sofa
1	bathtub	11	curtain	21	mantel	31	stairs
2	bed	12	desk	22	monitor	32	stool
3	bench	13	door	23	night_stand	33	table
4	bookshelf	14	dresser	24	person	34	tent
5	bottle	15	flower_pot	25	piano	35	toilet
6	bowl	16	glass_box	26	plant	36	tv_stand
7	car	17	guitar	27	radio	37	vase
8	chair	18	keyboard	28	range_hood	38	wardrobe
9	cone	19	lamp	29	sink	39	xbox

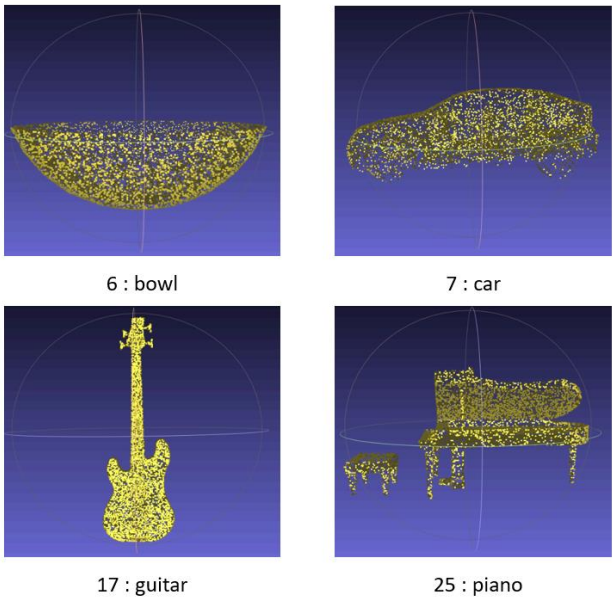

図４.１　ModelNet40 のデータの可視化の例

４．２．２　学習方法

本実験では，GPU として GeForce GTX 1080 Ti を搭載した PC を使用する．学習では ModelNet40 データセットのトレーニング用セットを使用する．これには合計 9843 のデータが含まれ, 40 の異なるクラスが存在する. この実験では法線情報は使用せず, 各データは 1024 の点でサンプリングされて学習される．トレーニング速度と性能のバランスをとるため，バッチサイズを 24 に設定し，200 エポックのトレーニングを行った．初期学習率は 0.001 とし，Adam オプティマイザを選択した．

４．２．３　テスト及び評価方法

本実験の評価では，ModelNet40 のテスト用データセットを使用する．これには合計 2468 のデータが含まれ，40 の異なるクラスが存在する．提案手法，バックボーンの PointNet，および局所特徴を分析する代表的な点群分類手法である PointNet++の 3 種類の手法について，精度と速度を比較する．精度の評価では，Instance Accuracy を採用する．速度の評価では，テスト時間の測定を採用する．

$$\text{Instance Accuracy} = \frac{Number\ of\ correctly\ predicted\ samples}{total\ number\ of\ samples} \times 100\% \qquad (4.2)$$

Instance Accuracy の計算式を式(4.2)に示す．この指標はモデルが各サンプルで正確に予測した割合を示す．具体的には，各点群データに対して，モデルの予測結果と実際のラベルを比較する．モデルが入力サンプルを正確にクラス分類できた場合，そのサンプルを「正しく予測された」と見なす．Instance Accuracy は，「正しく予測された」サンプルの総数に対する割合として計算される．

４．２．４　実験結果

各手法を用いて精度とテスト時間の比較を行った結果を表４.２に示す．PointNet は最も高速である一方で，精度は最も低いことが確認された．PointNet++は局所特徴解析を実施しない PointNet に比べて精度が向上しているが，これに伴いテスト時間も４倍ほどに増加している．一方，提案手法は，テスト時間は PointNet よりわずかに増加しているが，これら三つの手法の中で最も高い精度を示した．

表４.２　各手法の比較

	PointNet	PointNet++	Ours
Instance Accuracy [%]	89.2	90.7	**91.7**
Testing Time [s]	**4.3**	19.6	4.4

４．３　対照実験

４．３．１　対照内容

提案手法の構造が最適な組み合わせであるかを検証するために，対照実験を行った．対照試験は，残差 MLP と CBAM について異なる位置順序や個数がモデルの性能に与える影響をテストすることを目的としている．このため，注意機構は CBAM のままで，残差 MLP の構成も変更しない．

一番目の対照実験では，残差 MLP と CBAM の位置順序がモデルの性能に与える影響について検討する．対照用モジュールの具体的な構造を図４．２に示す．対照用モジュールでは接続順序を変更し，最初に CBAM を接続してから残差 MLP に接続する．

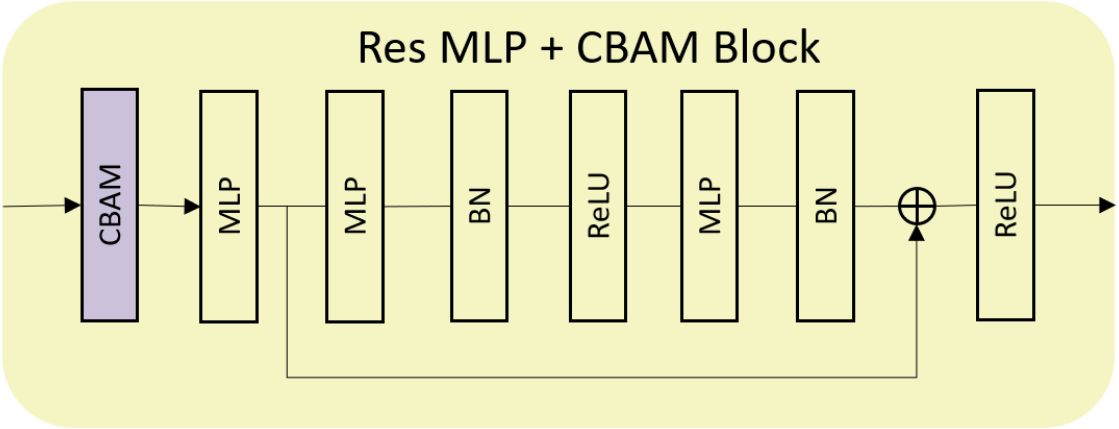

図４.２　一番目の対照実験で使われる対照用モジュール

二番目の対照実験では，提案の Res MLP + CBAM Block モジュールの個数がモデルの性能に与える影響について検討する．対照用モデルの具体的な構造を図４．３に示す．提案のモジュールの適用個数を二つから三つに増やし，実験を行った．

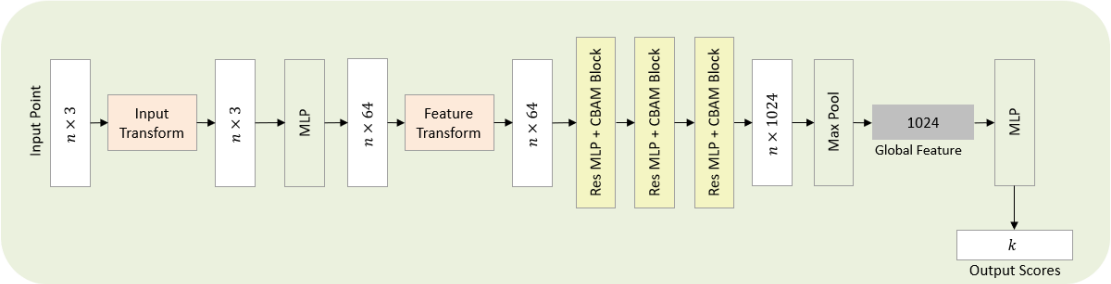

図４.３　二番目の対照実験で使われる対照用モデル

４．３．２　対照実験結果

モジュールの比較実験の結果を表４．３に示す．提案手法のモジュールと比較用モジュールの速度と精度がほぼ同等であることが確認された．

表４.３　対照実験１の比較結果

	Comparison	Ours
Instance Accuracy [%]	91.6	**91.7**
Testing Time [s]	4.4	4.4

モジュールの適用個数の比較実験の結果を表４．４に示す．提案手法の適用モジュール個数が３個に増加しても比較用モジュールと比較して速度が速く，精度はほぼ同等であることが分かった．

表４.４　対照実験２の比較結果

	Comparison	Ours
Instance Accuracy [%]	**91.9**	91.7
Testing Time [s]	6.3	**4.4**

４．４　考察

４．２では，ModelNet40 データセットを使用して，提案手法，バックボーンの PointNet，局所的情報の分析を実施する PointNet++の３種類の手法について，学習および評価実験を行った．実験結果から，提案手法は PointNet に比べて 2.5%の精度向上を達成していることが分かる．局所情報分析を使用する従来の PointNet++よりも 1%の精度向上が得られている．処理時間から見ると，提案手法は PointNet よりも 2.3%遅い程度である．しかし，PointNet++に比べては 77.6%速くなる．PointNet が既に速いことを考慮すると，提案手法は高速性を保ちつつ精度を向上させたと確認できる．

４．３では，提案手法が最も適した構成であるかどうかを検証する対照実験を行った．提案のモジュール内部の構造に関して，残差 MLP と CBAM をネットワークに接続する

順序を変更した．その結果，速度と精度はほとんど変わらず，精度では提案手法が対照用モジュールよりもわずかに高い（0.1%）ことが確認できた．また，モジュールの適用数を 2 個から 3 個に増やした対照実験では, 対照用モデルは精度がわずかに向上したものの，大幅な時間増加をもたらした．時間と精度の両方を考慮すると，提案手法はより高いコストパフォーマンスを持つことを言える．

本実験を通じて，残差ネットワークと CBAM の導入が点群の分類タスクの精度向上に寄与することが確認できた．残差 MLP はより複雑な特徴を捉え，点群の分類の難しさに対処できる．一方で，注意機構はモデルが重要な特徴に焦点を当て，不要な情報に対しては重みを低くすることで，モデルの注目すべき点を強調し，精度向上を実現できる．また，提案手法の導入により計算時間の増加も確認できた．したがって，速度と精度の両方を考慮する実際の状況では，提案モジュールの適用数に注意する必要がある．

４．５　むすび

本章では, 本研究の点群分類モデルにおける, 学習に用いたデータセット, 実験方法, 実験結果について述べた．また，対照実験として，提案手法のモジュール内部の構造を変更した場合と，提案手法のモジュール数を増やした場合の実験し，提案手法の実験結果との比較を行った．提案手法では，高速を維持しつつデータ構造を捉える能力を向上させることが達成でき，対照実験からもバランスの良い性能を示すことが分かった．

第５章　結論と今後の課題

５．１　結論

本研究では，従来の点群分類手法である PointNet に対し，残差 MLP と注意機構の CBAM をもとに構成した Res MLP + CBAM Block を導入した新たな点群分類手法を提案した．提案手法では，各残差ネットワークと CBAM が効果を発揮し，速度を維持しつつデータ構造を捉える能力を向上させることが可能となった．実験では，インスタンス精度とテスト時間の評価を行うことにより，提案手法の有効性を確かめた．具体的には，提案手法と従来手法の比較を行い，その有効性を検証した後，対照実験を行い，提案手法がバランスの良い構造であることを示した．定量的評価では提案手法は元のモデルに比べて処理速度が同じレベルの場合に精度が 2.5%向上することが確認できた．また，したがって，提案手法により残差 MLP と CBAM を導入することは，点群分類タスクにおいて有効かつ効率的であるといえる．

５．２　今後の課題

提案手法では，高速を維持しつつ精度の向上が達成できた反面，残差 MLP の構造と注意機構の構造は，それぞれ 1 種類しか使用しておらず，検討すべき構造の種類が不足している可能性がある．そのため，今回は行えなかった他の残差 MLP と注意機構の構造を導入する実験は必要性があると考えられる．さらに，提案のモジュールの適用数の増加に伴う計算時間の増加が大きいため，より汎用性が高く，計算負荷が少ない構造の考案の必要性があると考えられる．

謝辞

本論文の執筆に当たり，研究の方向性や問題点をご指導くださり，快適な研究環境を与えてくださった渡辺裕教授に感謝いたします.

本論文の作成にあたり，副査として適切なご助言をくださった早稲田大学大学院基幹理工学研究科の中山光典氏に感謝いたします.

毎週のゼミ活動において，研究の進捗や実験方法に関して丁寧なアドバイスをくだ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と早稲田大学大学院基幹理工学研究科の劉雲龍氏，小野蒼平氏，Phuong Thao Nguyen 氏に感謝いたします.

また，日頃から温かく接してくださり，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に感謝いたします.

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] 国立研究開発法人科学技術振興機構 研究開発戦略センター，“IoT 時代のセンサ融合基盤技術の構築　～センシング情報の高付加価値化に向けた多様なデータの取得と統合的処理～,” Mar. 2020, https://www.jst.go.jp/crds/report/CRDS-FY2019-SP-10.html

[2] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, “PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space,” Neural Information Processing Systems (NlPS 2017), pp. 5105-5114, Dec. 2017.

[3] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “PointNet: Deep Learning on Point Set s for 3D Classification and Segmentation,” IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR 2017), pp. 652-660, Jul. 2017.

[4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recogni tion,” IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR 2016), pages 770–778, Jun. 2016.

[5] M. Guo, T. Xu, J. Liu, Z. Liu, P. Jiang, T. Mu, S. Zhang, R. R. Martin, M. Chen g, and S. Hu, “Attention Mechanisms in Computer Vision: A Survey,” Springer, Co mputational Visual Media, Vol. 8, No. 3, pp. 331–368, Mar. 2022.

[6] S. Woo, J. Park, J. Lee, and I. Kweon, “CBAM: Convolutional Block Attention Module,” European Conference on Computer Vision (ECCV 2018), pp. 3-19, Sep. 2018.

[7] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, “3D ShapeNets: A Deep Representation for Volumetric Shapes,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015), pp. 1912-1920, Jun. 2015.

図一覧

図２.１　PointNet のモデル構造 5
図２.２　残差モジュールの構造 6
図２.３　CBAM の構造 7
図３.１　提案手法のモデル構造 9
図３.２　Res MLP + CBAM Block の構造 10
図４.１　ModelNet40 のデータの可視化の例 13
図４.２　一番目の対照実験で使われる対照用モジュール 14
図４.３　二番目の対照実験で使われる対照用モデル 15

表一覧

表４.１　各条件における分類性能の分散 12
表４.２　各手法の比較 14
表４.３　対照実験１の比較結果 15
表４.４　対照実験２の比較結果 15

研究業績

[1] 范洪睿, 渡辺裕：“残差 MLP と注意機構を組み合わせた点群分類手法の検討 (A Study on Point Cloud Classification Method Integrating Residual MLP and Attention Mechanism)”, 2024 年 電子情報通信学会総合大会, Mar. 2024. (発表予定)