

論文 / 著書情報  
Article / Book Information

題目(和文)	点群および波形情報を有するライダーデータを用いた地物分類手法の提案と点群のみのデータへの活用に関する研究
Title(English)	A STUDY ON A METHOD FOR CLASSIFICATION OF GEOSPATIAL FEATURES USING FULL WAVEFORM LIDAR DATA AND ITS APPLICATION TO POINT CLOUD ONLY DATA
著者(和文)	篠原崇之
Author(English)	Takayuki Shinohara
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11833号, 授与年月日:2022年3月26日, 学位の種別:課程博士, 審査員:松岡 昌志,山中 浩明,浅輪 貴史,藤田 康仁,大風 翼,小林 秀樹
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11833号, Conferred date:2022/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

点群および波形情報を有するライダーデータを用いた地物分類手法の  
提案と点群のみのデータへの活用に関する研究  
A STUDY ON A METHOD FOR CLASSIFICATION OF GEOSPATIAL  
FEATURES USING FULL WAVEFORM LIDAR DATA AND ITS  
APPLICATION TO POINT CLOUD ONLY DATA

東京工業大学環境・社会理工学院  
建築学系都市・環境学コース  
篠原崇之

博士論文

指導教員：松岡昌志教授

# Abstract

This thesis is titled "A Study on a Method for Classification of Geospatial Features Using Full Waveform LiDAR Data and Its Application to Point Cloud Only Data" and consists of the following six chapters.

Chapter 1 describes the background of this thesis, related research issues, research objectives, and the structure of the doctoral dissertation. This thesis proposes a method for automatic classification of point cloud and waveform data to realize the automatic creation of 3D maps. In order to achieve this goal, we are developing a method of geospatial feature classification for full waveform lidar data consists of point clouds and waveform data using deep learning techniques. In this paper, we propose a deep learning method for point clouds and waveform data that uses direct point clouds and waveform data as input. In this paper, we propose a deep learning method that uses point clouds and waveform data as input data. In this paper, we propose a method to estimate waveform data from point cloud data and convert it into pseudo point cloud and waveform data. Additionally, we propose a method for estimating waveform data using conditional GAN, one of the deep learning methods, and referring to the previous study of point cloud coloring method.

Chapter 2 presents a description of lidar in general and the characteristics of lidar for recording point clouds and waveform information, as well as a deep learning method for classifying point cloud data into geographic objects and a conditional adversarial generation network. First, the measurement method of lidar, which records point clouds and waveform information, is explained in principle, and the characteristics of the measured waveform data are described. The general measurement principles of the main point clouds and lidar that record waveform information is explained, regardless of their applications. Next, the principles of deep learning are explained, followed by a description of deep learning for general images, deep learning methods for point cloud data, and a description of the conditional adversarial generative network method used for waveform data estimation in this study. Next, we provide a comprehensive description of the main existing point cloud and waveform data geomorphology classification methods. At the end of the chapter, the relationship between the existing studies and the proposed method is described.

In Chapter 3, a new method of point cloud and waveform data land classification is proposed. In proposing the method, we assume that the classification of point clouds and waveform data can be divided into two categories: "geographic features for which individual waveform data is effective" and "geographic features for which global information, which is the spatial relationship between waveform data, is effective. For each feature, we apply a local block to classify the individual waveform data, and a global block to classify the feature by applying neighbor feature extraction hierarchically to the latter. As an effect of the proposed method, it does not depend on the analysis ability of the end-user because it uses deep learning and does not extract features based on human knowledge from waveform data, which is a common machine learning method. In addition, since the proposed method directly deals with waveform data and point cloud data, it is not affected by the loss of information on 3D structure when projected to 2D in the method of converting point cloud data and waveform data to images used in the conventional methods of point cloud and waveform data analysis using deep learning. An evaluation using a benchmark dataset shows the effectiveness of the combination of global and local blocks in the proposed

method.

In Chapter 4, a waveform estimation method from point cloud data is proposed. Since there are no previous studies on waveform estimation, we refer to the point cloud coloring method, which is a similar task in terms of estimating other information from point cloud data, and focus on individual waveforms as well as the geological classification of point cloud and waveform data. In this paper, we propose a new loss function that evaluates the estimated waveforms based on the frequency characteristics of their shapes. By using the new loss function, it is shown that the performance of estimating the waveform of the point cloud data can be improved.

In Chapter 5, we combine two proposed deep learning methods, a method for estimating waveform data from point cloud data and a method for classifying waveform data geodetic data from point cloud and waveform data, to create pseudo waveform data from point cloud data, and then compare the created waveform data with point cloud data (pseudo point cloud and waveform data) are input to the point cloud and waveform data land classification method to perform land classification. To evaluate the generalization performance of the waveform data estimation model, we used the Dublin dataset, which contains point cloud data and waveform data as a pair, and showed that waveform data can be estimated even for point cloud data from a different region than the benchmark dataset used for training. We applied the pseudo point cloud and waveform data classification method to the benchmark data and the archived data, and confirmed that the proposed pseudo point cloud and waveform data classification method has a higher F1Score in the benchmark data than the deep learning model that uses only point cloud data as input. In the case of archival data, the F1Score of the point cloud data is higher than that of the deep learning model. In the archived data, the proposed method is applied to the DALES dataset, which is a public dataset of point cloud data, and since the DALES dataset has true values, the performance of the proposed method is quantitatively evaluated using the true values. As a result, it was found that the pseudo point cloud and waveform data of the proposed method showed higher performance than simply classifying point cloud data. In addition, the effectiveness of the proposed method is confirmed by the results of land classification on Japanese archived data, focusing on the extraction performance of the land surface.

In Chapter 6, we summarize the results obtained in this study, show the scope of application of the proposed method, and discuss the future prospects.

## 要旨

本論文は、「点群および波形情報を有するライダーデータを用いた地物分類手法の提案と点群のみのデータへの活用に関する研究」と題し、以下の6章により構成している。

第1章「序論」では、本論文の背景、関連研究の課題、研究目的、博士論文の構成を述べている。本論文は、3次元地図の自動作成を実現するために点群・波形データを自動分類する手法の提案を行う。そのために、深層学習技術を用いた点群・波形データの地物手法の開発に取り組んでいる。点群・波形データの深層学習手法として、画像に変換する手法が存在するが、画像化の際に詳細な3次元情報が消失するため、直接点群・波形データを入力する深層学習手法を提案する。さらに、提案した点群・波形データの地物分類手法を活用する際には、アーカイブデータにおいて波形データが格納されない問題点を克服する必要がある。そこで、点群データから波形データを推定し、擬似的に点群・波形データ化する手法を提案している。その際に、深層学習手法のひとつである条件付きGANを用い、その先行研究である点群着色手法を参考に、波形データ推定手法を提案している。

第2章「点群および波形情報を記録するライダーと既往の深層学習手法」では、一般的なライダーに関する説明と点群および波形情報を記録するライダーの特徴、および、深層学習による点群データの地物分類手法と条件付き敵対的生成ネットワークを紹介している。まず、点群および波形情報を記録するライダーの計測方法を原理から説明し、計測された波形データの特性について説明している。この際、主要な点群および波形情報を記録するライダーについて、その用途を問わず一般的な計測原理を説明している。次に、深層学習に関して、その原理を説明し、一般的な画像に対する深層学習の説明を行い、次に点群データに対する深層学習手法、そして本研究において波形データ推定に用いられる条件付き敵対的生成ネットワーク手法に関する説明を行っている。次に、既存の主な点群・波形データ地物分類手法について包括的に説明している。章の最後に、既往研究と提案手法の関係性を述べている。

第3章「点群・波形データを用いた地物分類手法の提案」では、点群・波形データの地物分類を新たに提案している。手法の提案を行う際に、点群・波形データの地物分類を行う際に、「個々の波形データが有効な地物」と「波形データと波形データの空間的な関係性である大局情報が有効な地物」に大別できると仮定し、それぞれの地物に対して、前者には個々の波形データの地物分類を実施する局所ブロック、後者には近傍特徴抽出を階層的に適用して地物分類を実施する大局ブロックを適用し、それぞれを組み合わせている。提案手法の効果として、深層学習を使用しており、一般的な機械学習手法の波形データからの人間の知識による特徴量抽出を行わないため、エンドユーザーの解析能力に依存しない。また、提案手法は波形データと点群データを直接取り扱うため、深層学習を利用した点群・波形データ解析の従来手法で用いられた点群データと波形データを画像に変換する手法の2次元に投影する際の3次元構造の情報欠損の影響を受けない。ベンチマークデータセットを用いた評価によって、提案手法の大局ブロックと局所ブロックの組み合わせの効果を示している。

第4章「点群データからの波形推定手法の提案」では、点群データからの波形推定手法を提案している。その際、波形の推定における既往研究は存在しないため、点群データから別の情報を推定する点において似ているタスクである点群着色手法を参考にし、点群・波形データの地物分類と同様に、個々の波形に着目している。さらに、推定した波形を評価する損失関数として、その形状を評価する周波数特性にて評価する新たな損失関数を提案している。単に点群着色手法を適用するのではなく、新たな提案を用いることで、点群データからその点群データが観測された際の波形を推定する性能が向上することを示している。

第5章「波形データ推定手法と地物分類手法の効果検証」では、点群データからの波形デ

ータを推定する手法と、点群・波形データの波形データ地物分類手法の2つの深層学習手法を組み合わせることで、点群データから擬似的に波形データを作成し、その作成した波形データと点群データ(擬似的な点群・波形データ)を点群・波形データの地物分類手法に入力し地物分類を行っている。波形データ推定モデルの汎化性能評価には、点群データと波形データがペアとして存在するダブリンデータセットを用い、学習に使用したベンチマークデータセットとは別の地域の点群データにおいても波形データが推定できることを示した。そして、擬似的な点群・波形データの地物分類手法をベンチマークデータとアーカイブデータに適用し、ベンチマークデータにおいては、点群データのみを入力する深層学習モデルよりも提案手法である擬似的な点群・波形データ化は地物分類性能として高いF1Scoreを示すことを確認している。アーカイブデータにおいては、点群データの公開データセットである DALES データセットに対して提案手法を適用している。DALES データセットにおいては真値が付与されているため、その真値を用いて地物分類性能を定量評価し、その結果、単に点群データの地物分類を実施するよりも、提案手法の擬似的な点群・波形データ化により高い地物分類性能を示すことを明らかにしている。さらに、日本のアーカイブデータに対して地物分類を実行した結果、地表の抽出性能に着目して手法の有効性を確認している。

第6章「結論」では、本研究で得られた成果の総括と提案手法の適用範囲について示し、今後の展望について述べている。

## 謝辞

本研究を進めるにあたり終始あたたかいご指導と激励を賜りました指導教員の東京工業大学環境・社会理工学院松岡昌志教授に心から感謝の意を表します。また、本研究に関し、副査を引き受けてくださいました、東京工業大学浅輪先生、大風先生、小林先生 (JAMSTEC)、藤田先生、山中先生には、中間発表等の機会にて多大なるご指導をいただきました。深く感謝いたします。

松岡研究室の皆様にも感謝を申し上げます。同時期に博士後期課程に入学した修浩毅さんは同じような点群の研究をおこなっており手法に関する議論を熱心に行なっていただいた。大串文誉さんは、社会人博士の先輩として、道を示していただいた。同じく社会人博士の串山傳さんには応用よりの鋭い意見をたくさんいただいた。鄭閔隆さんは私が修士を卒業した頃からの付き合いがあり、彼との食事は良い息抜きになった。

所属会社の株式会社パスコの佐藤俊明さん、坂元光輝さんには、業務量の調整等をしていただき、体を壊すことなく博士課程を無事に終えることがきました。李勇鶴さんをはじめ、部署の皆様にも、エフォートを最大値で出せない中、業務を補助していただいた。

早稲田大学・オムロンサイニックエックスの千葉直也さんには点群に関する深層学習手法について様々な情報共有をいただきました。また、東京工業大学情報理工学院の林直輝さんには機械学習に関する理論的な議論をさせていただいた。幹事の皆川卓也さんをはじめとして関東コンピュータビジョン勉強会の皆様には国際会議の論文紹介などで情報交換をしていただきました。産業総合研究所の片岡裕雄さんをはじめとした [cvpaper.challenge](http://cvpaper.challenge) の皆様にはオンラインにて論文等の議論をさせていただいた。オンライン上での社会人博士のコミュニティでお互いに鼓舞し続けた皆さまに感謝いたします。

最後に、これまで私をあたたかく応援してくれた両親、昨年亡くなった祖母に感謝します。

# 目次

<b>1</b>	<b>序論</b>	<b>1</b>
1.1	はじめに	2
1.2	研究背景	3
1.2.1	3次元計測機器の普及	3
1.2.2	点群データの利活用	3
1.2.3	点群および波形情報を記録するライダー計測	5
1.3	既往研究の問題点	6
1.3.1	点群・波形データの地物分類	6
1.3.2	アーカイブ点群データの有効活用	6
1.4	研究目的	8
1.5	博士論文の構成	9
<b>2</b>	<b>点群および波形情報を有するライダーと既往の深層学習手法</b>	<b>11</b>
2.1	はじめに	12
2.2	点群および波形情報を記録するライダーの原理と解析方法	13
2.2.1	航空機によるライダー計測の原理	13
2.2.2	点群および波形情報を記録するライダー計測	18
2.2.3	点群および波形情報を記録するライダーの種類	24
2.2.4	波形データからの点群データの作成方法	24
2.3	深層学習手法	31
2.3.1	深層学習とは	31
2.3.2	一般的な深層学習手法の説明	32
2.3.3	畳み込みニューラルネットワーク	38
2.3.4	点群データに対する深層学習手法	44
2.3.5	条件付き GAN	51
2.4	点群・波形データに対する深層学習手法の適用	55
2.4.1	古典的な点群・波形データの解析方法	55
2.4.2	深層学習を利用した点群・波形データの地物分類	57
2.4.3	条件付き GAN の点群・波形データへの適用	59
2.5	既往研究と提案手法との関係	61
2.5.1	点群・波形データの地物分類を行う深層学習モデルの提案	61
2.5.2	点群データから波形データを推定する深層学習手法の提案	61
<b>3</b>	<b>点群・波形データを用いた地物分類手法の提案</b>	<b>62</b>
3.1	はじめに	63
3.2	点群・波形データの地物分類の問題設定	64
3.3	点群・波形データの地物分類を行う上での仮説	65
3.4	局所・大局モデルの提案	68
3.4.1	提案する局所・大局モデルの概要	68
3.4.2	局所ブロック	68
3.4.3	大局ブロック	70
3.4.4	損失関数	79

3.4.5	最適化アルゴリズム	79
3.4.6	各ブロックの出力の組み合わせ方法	79
3.5	ベンチマークデータセットによる地物分類実験	81
3.5.1	データセット	81
3.5.2	実験の設定	83
3.5.3	実験結果	85
3.5.4	2つのブロックの出力と最終予測結果との関係	88
3.5.5	波形データの効果	88
3.5.6	局所・大局モデルにおける入力データに対する感度分析	89
3.5.7	判断根拠の可視化	92
3.5.8	点密度の影響	96
3.6	第3章のまとめ	97
<b>4</b>	<b>点群データからの波形推定手法の提案</b>	<b>98</b>
4.1	はじめに	99
4.2	波形推定の問題設定	101
4.3	波形推定に点群着色手法を適用する際の仮説	104
4.4	波形推定モデルの提案	106
4.4.1	波形推定モデルの概要	106
4.4.2	大局生成器の層設計	106
4.4.3	識別器の層設計	111
4.4.4	損失関数	114
4.4.5	最適化アルゴリズム	116
4.5	ベンチマークデータによる波形推定実験	119
4.5.1	データセット	119
4.5.2	実験の設定	120
4.5.3	実験結果	120
4.5.4	大局生成器の影響	123
4.5.5	各工夫点の効果	124
4.5.6	物理モデルとの比較	125
4.5.7	点密度の影響	127
4.6	第4章のまとめ	128
<b>5</b>	<b>波形推定手法と地物分類手法の効果検証</b>	<b>129</b>
5.1	はじめに	130
5.2	提案手法	131
5.2.1	使用する学習済みモデル	131
5.2.2	アルゴリズム	131
5.3	波形推定の汎化性能評価	133
5.3.1	ダブリンデータセット	133
5.3.2	波形推定結果	133
5.4	ベンチマークデータセットへの適用	136
5.4.1	データセット	136
5.4.2	実験結果	136
5.5	アーカイブデータへの適用実験	140
5.5.1	アーカイブデータ	140
5.5.2	DALES データセット地物分類実験	143
5.5.3	日本のアーカイブデータにおける検証実験	147
5.6	第5章のまとめ	149

<b>6</b>	<b>結論</b>	<b>150</b>
6.1	博士論文全体のまとめ	151
6.2	詳細なまとめ	152
6.2.1	第3章: 点群・波形データを用いた地物分類手法の提案	152
6.2.2	第4章: 点群データからの波形推定手法の提案	153
6.2.3	第5章: 波形推定手法と地物分類手法の効果検証	154
6.3	手法の貢献点と適用範囲	155
6.3.1	貢献点	155
6.3.2	適用範囲	156
6.4	将来の展望	158
6.4.1	点群・波形データの普及	158
6.4.2	点群・波形データの地物分類性能の向上	158
6.4.3	点群データからの波形推定性能の向上	158
6.4.4	より波形データを活かした地物の地物分類	159
6.4.5	航空機で観測した点群・波形データ以外への拡張	159
	<b>付録</b>	<b>161</b>
A	在学中の採択論文	162
B	地物分類へ深層学習モデルの層設計が与える影響	163
C	点群データに対する条件付き GAN の予備実験	166
C.1	点群着色手法の問題設定	166
C.2	改良した点群着色手法	167
C.3	実験結果	172

## 図一覧

1.1	点群データへ地物情報を付与することによる 3 次元地図の作成とその応用先 . . .	4
1.2	本研究の構成 . . . . .	10
2.1	都市域における複数反射の例 . . . . .	17
2.2	フットプリントの影響 . . . . .	19
2.3	様々な状況における波形データ . . . . .	20
2.4	疎な点群を記録するライダーと点群および波形情報を記録するライダーの点群 データの作成方法比較 . . . . .	25
2.5	波形データ分解の概念図 . . . . .	25
2.6	波形データのモデル化の概念図 . . . . .	28
2.7	ジオリファレンス処理 . . . . .	30
2.8	深層学習モデルの学習と予測 . . . . .	32
2.9	パーセプトロン . . . . .	33
2.10	Rectified Linear Unit (ReLU) . . . . .	33
2.11	マルチレイヤーパーセプトロン . . . . .	33
2.12	二変数の場合の確率的勾配降下法の例 . . . . .	35
2.13	シグモイド関数 . . . . .	36
2.14	1D CNN による時系列信号の処理 . . . . .	38
2.15	2D CNN による画像データの処理 . . . . .	39
2.16	画像の畳み込み演算の概念図 . . . . .	40
2.17	全結合層と CNN の違い . . . . .	41
2.18	畳み込みとプーリングによる局所受容野 . . . . .	41
2.19	画像に対する深層学習手法を用いた主要なタスク . . . . .	43
2.20	点群データに対する深層学習手法を用いた主要なタスク . . . . .	44
2.21	PointNet の概念図 . . . . .	47
2.22	PointNet++ の概念図 . . . . .	48
2.23	点群データの近傍特徴抽出と画像の近傍特徴抽出の違い . . . . .	49
2.24	PointNet++ のダウンサンプリングとアップサンプリングの概念図 . . . . .	49
2.25	深層学習手法における受容野の概念図 . . . . .	50
2.26	GAN の概念図 . . . . .	52
2.27	GAN で使用される尺度である KL ダイバージェンスの正規分布を用いた時の例 . . . . .	53
2.28	条件付き GAN の概念図 . . . . .	54
2.29	点群・波形データ解析における深層学習手法 . . . . .	58
2.30	条件付き GAN を点群データに適用した先行研究の例 . . . . .	60
2.31	条件付き GAN を波形データに適用した先行研究の例 . . . . .	60
3.1	点群・波形データの地物分類の概念図 . . . . .	63
3.2	点群・波形データに対する地物分類の問題設定 . . . . .	64
3.3	提案する局所・大局モデルの概念図 . . . . .	68
3.4	局所ブロックの波形データに対する特徴抽出 . . . . .	69
3.5	局所ブロックの層設計 . . . . .	69
3.6	波形データに対する近傍特徴抽出の仕組み . . . . .	72
3.7	階層的な特徴抽出を行うためのダウンサンプリングとグループ化 . . . . .	74

3.8	最遠点サンプリング (FPS) の二次元平面上の例	75
3.9	デコーダにおける 3 次元内挿	76
3.10	大局ブロックの層設計	78
3.11	予測時の局所ブロックと大局ブロックの組み合わせ方法	80
3.12	学習・検証・テストデータの意味	81
3.13	ベンチマークデータセット	82
3.14	ベンチマークデータセットにおける波形データの例	82
3.15	学習済みモデルによる地物分類結果	87
3.16	ノイズを加えた波形データの例	90
3.17	ノイズを加えた点群データの例	91
3.18	隠蔽感度分析手法の概念図	93
3.19	勾配可視化手法の概念図	94
3.20	波形データの判断根拠の可視化結果	95
4.1	点群データからの波形推定手法の概念図	100
4.2	点群データと波形がペアになった学習データを利用した大局生成器の学習	102
4.3	点群データからの波形推定の問題設定	103
4.4	強度の差を計算する損失関数の問題点	105
4.5	FFT を適用する利点	105
4.6	個々の点群データに対する特徴抽出の仕組み	108
4.7	点群データに対する近傍特徴抽出の仕組み	109
4.8	波形推定モデルの大局生成器の層設計	110
4.9	大局識別器の層設計	112
4.10	局所識別器の層設計	113
4.11	波形推定モデルで推定した波形と観測値の例	122
4.12	FFT を適用した波形推定モデルで推定した波形と観測値の例	122
4.13	各工夫点の波形推定への効果	124
5.1	本章で提案する波形推定手法と点群・波形データの地物分類手法を組み合わせる手法の概要	130
5.2	波形推定モデルの汎化性能の検証実験のためのダブリンデータセット	134
5.3	ダブリンデータセットに対する波形推定結果	135
5.4	推定した点群・波形データによるベンチマークデータセットに対する地物分類結果	138
5.5	ベンチマークデータにおける波形推定の失敗例と成功例	139
5.6	真値付き点群データの DALES データセット	142
5.7	日本のアーカイブデータ	142
5.8	DALES データセットに対する地物分類結果	144
5.9	DALES データセットにおける提案手法の効果	145
5.10	日本のアーカイブデータに対する地物分類結果	147
5.11	日本のアーカイブデータにおける地表の抽出結果の例	148
B.1	全体横断モデルで使用する演算	164
B.2	全体横断モデルの層設計	165
C.3	点群着色手法の概念図	166
C.4	点群着色手法の問題設定	167
C.5	点群着色手法における生成器の層設計	169
C.6	点群着色手法における識別器の層設計	171
C.7	航空写真と航空機に搭載されたライダーで観測された点群データによる着色データセット	173
C.8	点群着色結果	174

## 表一覧

1.1	点群・波形データの地物分類に関する既往研究	6
1.2	波形推定に関する既往研究	7
2.1	リモートセンシングに用いられる代表的な機材	13
2.2	点群および波形情報を記録するライダーの長所と短所	23
2.3	深層学習と従来型の機械学習手法の比較	31
2.4	点群データに対するシーン分類, 物体検出, セグメンテーションのための公開データセットの概要	46
3.1	訓練データとテストデータ	83
3.2	局所クラスと大局クラス	84
3.3	定量評価指標の計算に使用する項目	85
3.4	ベンチマークデータにおける定量評価結果	86
3.5	最終出力に対する局所ブロックからの寄与率	88
3.6	局所・大局モデルにおける入力データが地物分類結果に与える影響	89
3.7	局所・大局モデルにおける波形データに対する感度分析結果	90
3.8	点群データに対する感度分析結果	91
3.9	点群・波形データの点密度が学習済み局所・大局モデルへ与える影響	96
4.1	推定された波形と観測値の波形の平均絶対誤差	121
4.2	生成器の層設計が波形推定に与える影響	123
4.3	提案した各工夫点の効果検証結果	124
4.4	物理モデルと提案手法の定量比較結果	126
4.5	点群データの点密度が学習済み大局生成器へ与える影響	127
5.1	ベンチマークデータにおける定量評価結果	137
5.2	DALES データセットとベンチマークデータセットとの対応	140
5.3	日本のアーカイブセットとベンチマークデータセットとの対応	141
5.4	DALES データセットにおける定量評価結果	143
5.5	大局点群モデルにおけるベンチマークデータの学習時のチューニングによる効果	146
5.6	日本のアーカイブデータセットにおける定量評価結果	148
6.1	点群データの標準仕様である LAS 形式で定義される地物分類対象	160
B.1	ベンチマークデータにおける深層学習モデル構造の比較結果	165

# 用語集

本博士論文で用いる用語をあらかじめ定義する。

用語	定義	英語表現
点群および波形情報を有するライダー		Full Waveform LiDAR
点群・波形データ	点群と波形で成り立つライダーデータ	Full Waveform LiDAR Data
点群データ	x,y,z 座標を持つ点の集合	Point Cloud
波形データ	照射したレーザの反射強度の時系列かつ連続的な記録	Waveform Data
生の	点群データや波形データを画像などに変換しない状態	Raw
サンブル	波形データの各反射強度値	Sample
エコー	照射したレーザが地物に反射してセンサに戻る成分	Echo
深層学習	ニューラルネットワークを多層化（隠れ層が1つ以上）した手法全般	Deep Learning Layer
層	ニューラルネットワークの層	
深層学習モデル	あるタスクを実行可能な多層化したニューラルネットワーク	Deep Learning Model
層設計	深層学習モデルの構造	Network Architecture
学習可能な重み	最適化で重みの値を獲得可能	Learnable
画像畳み込み	学習可能な重みを用いたある画素の値とその近傍画素の値の積和演算	Image Convolution
波形畳み込み	学習可能な重みを用いた波形データのある強度値とその前後の時間に観測された強度値の積和演算	Waveform Convolution
情報集約	ある画素・点の値とその近傍画素・点の値のを最大値等で集約する演算	Pooling
特徴抽出	量み込みや情報集約で入力されたデータの特徴量を得ること	Feature Extraction
近傍特徴抽出	一定範囲内の波形データや点群データに対して個別に特徴抽出した後に情報集約すること	Pooling
特徴量	量み込み・近傍情報集約で得られる値	Feature
特徴マップ	各層における特徴量が抽出された結果	Feature Map
コンテキスト	車は道路の上にある等の物体間の関係性	Context
受容野	ある画素・点を予測する際に畳み込み・近傍情報集約で集約される範囲	Receptive Field
局所的	点群・波形データの波形データを個別に言及する際	Local
大局的	点群・波形データの波形の広範囲なコンテキストに言及する際	Global
最適化	深層学習モデルの重みを訓練データを用いて獲得すること	Optimization
データセット	最適化や評価に用いるデータ全体	Dataset
パッチ	データセットをある一定範囲の大きさに切り出したデータ	Patch
学習データ	深層学習モデルの重みを最適化するために用いるパッチ	Training Data
学習済みモデル	深層学習モデルが学習データで最適化された状態	Trained Model
テストデータ	学習データで最適化した学習済みモデルによって推論を行うデータ	Test Data
アーカイブデータ	オープンデータサイトや納品されたデータ	Archived Data

# 第 1 章

## 序論

## 1.1 はじめに

まず本章の構成を述べる．最初に本研究の研究背景を説明する．研究背景としては点群データ計測の普及と，その点群データを利活用する応用先の説明，そして点群データを取得するための点群および波形情報を記録するライダー計測手法の説明を行う．次に，既往研究の問題を整理する．まず，点群・波形データを用いた地物分類の既往研究における問題点を整理し，さらに，点群・波形データを用いた地物分類をオープンデータ等のアーカイブ化された点群データへ活用する際の問題点を整理する．そして，本研究の目的を述べる．最後に，博士論文の章構成を述べる．

## 1.2 研究背景

### 1.2.1 3次元計測機器の普及

近年、3次元計測を行う機器や手法として様々な方式が普及した [257]. 3次元計測を実現する手法として、ステレオ写真による測量が古くから行われている. さらにステレオ写真の計測手法を拡張し、様々な角度で撮影した写真を用いて測量を行う Multi View Stereo(MVS) も普及している. これら技術は航空機で観測する方式の他に、Unmanned Aerial Vehicle(UAV) で測量を行う方式もある. 計測機器として写真を用いる手法の他に、ライダー (Light Detection and Ranging, LiDAR) を用いたレーザ計測が普及している. ライダーはセンサと地物間の距離をレーザの反射によって直接計測することができ、その成果物は点群データとして表現される. ライダーの機器そのものや、搭載されるプラットフォームが異なれば、様々な計測時の特性が現れる. 例えば、点密度はその計測方法によって、10 点/m<sup>2</sup> 未満から 5000 点/m<sup>2</sup> まで様々である [233].

### 1.2.2 点群データの利活用

点群データを用いることで我々の住む現実空間の幾何的な3次元情報を取得し、それをコンピュータ上で可視化し形状を確認可能である. この点群データの可視化によって、判読者は工事の進捗確認や、施設の経年変化の確認が可能である. 可視化は最新のコンピュータグラフィックスによるレンダリング手法とグラフィックス用のハードウェアである Graphics Processing Uni(GPU) を用いると、この確認作業は即時的に可能である. また、可視化する際には GPU は判読者の手元の計算資源にある必要はなく、別のサーバ上でレンダリングされた結果をスマートフォンやタブレットなどの低スペックな端末に配信し表示することもできる. 取得された点群データの用途として、可視化の他には高さデータの利用も可能である. 例として、建物の高さや地表面の形状等を取得可能であるため、点群データからデジタル地表モデル (Digital Surface Model, DSM) やデジタル地形モデル (Digital Terrain Model, DTM) が作成できる. さらに、建設分野においては、点群データを同じ地点で複数回計測することで、土量の変化等が計算でき工事の進捗管理に使用される. また、可視化や DSM/DTM 作成のような直接的な高さ情報の取得だけでなく、計測した点群データから特定の地物の情報を抽出した結果をシミュレーションに用いることも可能である [205]. 例えば、ある都市の点群データから建物を取得し、仮想的に都市の建物が再現された場合、それを用いて太陽光をシミュレートし日照解析 [259] が可能である. 同様に植生を点群データから抽出し都市の植生を仮想的に再現して熱環境の解析 [165] などが可能になる.

上記のシミュレーションのような応用を目指す場合、単なる幾何的な情報による高さの確認や可視化だけでなく地物情報と点群データを紐づけること、つまり点群データの地物分類が重要である. この点群データの地物分類を行なった点群データを3次元地図 (図 1.1) と定義し、3次元地図は様々な分野に応用が期待されている. 乗用車の自律走行においては、走行可能領域である道路や、障害物となりうる樹木や建物の地物情報のある3次元地図が重要となる [272]. また、UAV の自律走行 [268] においては、建物や樹木の他に電線や鉄塔などの障害物の地物情報が含まれた3次元地図 [157] が必要になる. なお、これらの乗用車や UAV の自律走行に関する3次元地図は ISO 規格化が検討 [267, 86] されており、3次元地図の国際的な利活用が期待されている. 乗用車や UAV の自律走行だけでなく、強靱で持続可能な都市を構築するためにも3次元地図が必要である. 例えば、空中に張り巡らされた送電線やそれを支える送電鉄塔などの3次元地図は都市のインフラ網を管理する場合において有効である. また、農業の観点では生育する植物の成長状態を監視することが重要である. 環境の観点からは二酸化炭素の排出量を監視するために、点群データから建物、植生、道路などの地物情報を含む3次元地図が都市を包括する規模が必要である. 二酸化炭素排出に関する具体的なサービスとして、アメリカのニューヨーク市における街路樹のモニタリングを行うサービス NYC' s Street Trees[162] がある. NYC' s Street Trees では、点群データから街路樹を抽出しており、この情報を元に二酸化炭素の吸収量を評価している. 国内においても同様に、3次元地図の活用が期待されている [261]. 具体例として、建物に限定してはいるが、国土交通省は

Project PLATEAU という建物の 3 次元モデルをオープンデータとして公開する仕組みを作成した [174]. 似た仕組みとして自治体でも行われており, 東京都は 3 次元可視化事業を開始して, 西新宿エリアや都庁前駅の建物モデルを作成している [2]. このように, ある特定の地物だけでなく複数の地物の情報が付与された 3 次元地図の整備は自動車やロボットの自律的な移動や, 都市環境の評価・将来予測を行うための基盤情報となるため重要になる.

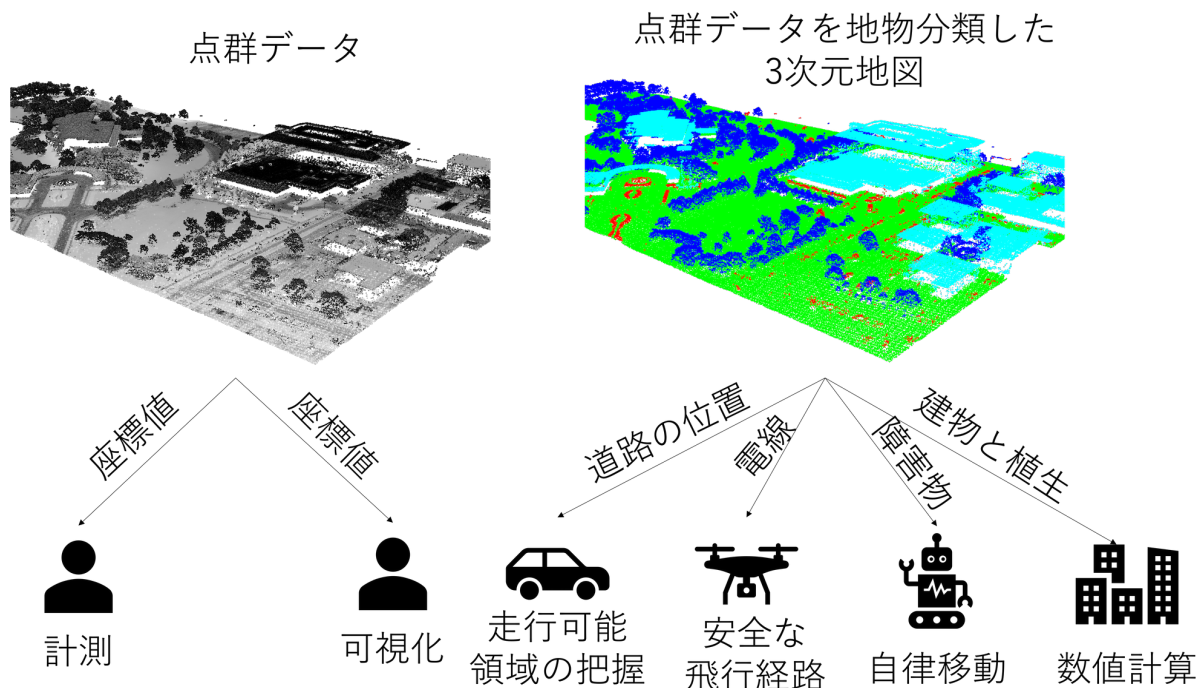


図 1.1: 点群データへ地物情報を付与することによる 3 次元地図の作成とその応用先. 道路の情報が点群データに付与されていれば, 乗用車の自動運転に使用でき, 電線等の空中の障害物の情報が付与されていると, UAV の自律移動にも使用できる.

このような理想に反して, 3 次元地図の整備は現在あまり行われておらず, 点群データの利活用において可視化や高さ情報の利用が依然として主流であり続けている. 3 次元地図の作成は大きく分けて計測・手作業による点群データの地物分類・3 次元モデル化のステップで構成されている. 計測は自動化が不必要であり, 3 次元モデル化は多くの商用ソフトウェアにて自動化が達成されている. しかし, 3 次元地図作成時の手作業による点群データの地物分類は未だ手作業に依存しており, 3 次元地図を作成・維持するためのコストが 3 次元地図によって受ける利点を大きく上回る. 具体的には点群データの計測によるデータ収集コストに対して, 3 次元地図作成時の手作業による分類で生じるコストがおおよそ 10 倍ほど高く [137, 83, 59], ある電力会社の例では電線に属する点群データを手作業で抽出するために年間 5 万時間を要している [175]. そのため, 3 次元地図の普及を促進させるためには, 手作業による労働集約的でない点群データの地物分類手法の開発が急務である. さらに, 3 次元地図の整備においては, 情報の鮮度が重要であり, 自動運転の 3 次元地図においては 1 ヶ月に 1 度更新される必要がある [262]. そのため, 現状の手作業による 3 次元地図の作成ではこの頻度を達成できない. したがって, 点群データの地物分類手法の実用化は更新頻度を高くする上でも必要である. そこで, 本研究では, 3 次元地図作成の中核となる手法, すなわち点群データの地物分類手法の開発を行う. とりわけ, 都市を包括するような広範囲な 3 次元地図を作成するには広範囲を計測可能な航空機に搭載されたライダーによって得られた点群データが重要である. したがって, 本研究では航空機に搭載されたライダーで得られた点群データの地物分類を対象にする.

### 1.2.3 点群および波形情報を記録するライダー計測

航空機に搭載されたライダーはセンサと地球上の地物や地形との間の距離をレーザによって直接計測する能動型のリモートセンシング技術である。このような距離計測の結果は点群データとして表現され、その点群データが持つ位置精度は 0.1 m 以下と高い精度を達成できる。この高精度な位置精度は、過去 20 年間に全球測位衛星システム (Global Navigation Satellite System, GNSS) 技術と慣性計測ユニット (Inertial Measurement Unit, IMU) が導入されたことに起因する。

ライダー開発の歴史は 1970 年代初頭、航空機や人工衛星に搭載された能動型センサーが設計されたことが始まりである。これらの初期の能動型センサは、シングルパルスのレーザを照射し、直下視に沿った 1 次元の反射のみを記録した。これに対し、レーザの照射を直下だけでなく入射角を変えられる機材が開発され、平行に重なった多数のストリップを取得可能になった。さらに、センサ技術がより発達し、計測時にレーザを照射された地球上の表面の形状によっては 1 つの照射パルスに対して複数の反射エコーが発生する現象も記録可能になった。このようなシステムは、マルチパルス式ライダー (本研究では 疎な点群を記録するライダーと定義) と呼ばれており、初期のセンサと比較すると高密度の点群データが得られるようになった。写真測量による高さ計測では最高点の点群データしか得られないが、疎な点群を記録するライダーは、樹木表面の高さとその下にある地形の標高の両方を一度に測定できるため、森林地帯の計測においては特に有効である。このように、疎な点群を記録するライダー計測は地形情報が取得可能であるため、森林のモニタリングや管理に使用されるようになった [5]。

上記で説明した一般的な疎な点群を記録するライダーは、ある地物の位置情報を点群データとして提供する。しかし、疎な点群を記録するライダーでは対象物からの反射信号が疎に記録され、照射した 1 つのレーザパルスに対して、最大 4 つまでの反射エコーしか得ることができない。そのために、地表面から反射した弱いエコーは検出が困難である。また、検出できる反射エコー間の距離が固定されているために、隣接した地物からの反射エコーは検出できない。このような疎な点群を記録するライダーの欠点に対して、照射したレーザの反射の強度を連続的に取得して、波形データとして密に記録し、その波形データから点群データを作成可能な「点群および波形情報を記録するライダー」が主流となった。

点群および波形情報を記録するライダーは、ここ 10 年の間に商業機材の発達とともに普及した。点群および波形情報を記録するライダーは、照射したレーザパルスごとに地上の地物に当たった反射の強度の信号を連続的に記録するもので、その記録された結果は波形データと呼ばれる。波形データは一定の時間間隔で反射の強度をサンプリングするため (時間分解能は GHz オーダー)。そのため波形データは反射の強度を等間隔に並べた時系列かつ離散的なサンプルの集合と定義でき、そのデータは「波形」のような形状を有する。この波形データの形状は、照射されたレーザが地面や建物や樹木等の反射面と相互作用する際の状況を表している。したがって、点群および波形情報を記録するライダーによって記録された波形データは地球上の地盤や地物の基本的な幾何学的表現以上の豊富な情報をもたらすことが可能である。

波形データは照射したレーザが各地物から反射した「エコー」で構成され、そのエコーのピークが地物のある位置に対応しており、そのピークの位置から点群データが計算できる。点群および波形情報を記録するライダーは多くの地物からの反射エコーを記録可能であるため、従来の疎な点群を記録するライダーより高密度の点群データが得られる。さらに、波形データと点群データ (点群・波形データと呼ぶ) を組み合わせることで、幾何学的な情報だけでなく照射されたレーザの進行方向に含まれる物体の特徴も得られる。具体的には、植生であれば波形データの中にピークが複数存在するなど、地物によって典型的な波形データの形状を示す場合がある。さらに、地物と被照射面の物理的な後方散乱特性に関する追加情報が波形データから提供される。したがって、点群・波形データの解析者は点群データのみを使用した場合よりも詳細な情報を波形データを併用して判読できるようになる。なお、本研究では現在の商用点群および波形情報を記録するライダーにおける主流である点密度の高い小さなフットプリントのものを対象にする。

## 1.3 既往研究の問題点

### 1.3.1 点群・波形データの地物分類

点群データのみを用いた地物分類よりも点群・波形データを用いる地物分類は高い分類性能を示す [47, 146, 159, 182, 57]. 例えば, 点群・波形データに対してルールベースの決定木が分類に用いられる [220, 4]. 他にも, 非線形分類が可能なサポート・ベクトル・マシン (SVM) 分類器などに対して, 人間が手で設計した特徴を入力する機械学習に基づく方法もある [77, 10, 249, 140, 11]. 上記のアルゴリズムの多くは, 統計的分類法や単純な機械学習アルゴリズムに入力するデータの作成方法, つまり点群・波形データからの手作業による特徴量設計方法に強く依存する. そのため, 点群・波形データの解析を担う人間の能力や知識に地物分類性能も依存する. さらに, 測量会社におけるデータ生産フローにおいては解析者が判断し適切なルールベースアルゴリズムのパラメータや機械学習モデルに入力する特徴量の作成は困難であることが多い. 解析者の能力に依存せずより高い分類性能を達成するためには, 入力されたデータから特徴量をアルゴリズムの内部で自動的に抽出する深層学習手法が効果的であると考えられる. 特に深層学習手法は 1 つのモデルで複数地物の分類が可能になるため, データ生産時における解析者の判断が削減できる利点がある.

深層学習を利用した点群・波形データの地物分類手法として, Zorzi ら [255] は点群・波形データを画像化し, 画像に対する深層学習手法を適用することで地物分類を行った. しかし, 元の点群・波形データが有する 3 次元的情報を 2 次元に投影しているため, 詳細な 3 次元情報が欠落する問題が生じる. なお, このような詳細な 3 次元情報の欠落は, 古典的な機械学習手法における, 特徴量の設計においても同様に発生する. そこで, 点群・波形データを別の形式へ変換することなく, 点群データと波形データが持つ詳細な情報を欠落させず, 生の点群・波形データを直接深層学習モデルで扱える手法が必要である.

変換を必要としない手法としては, 点群データを直接入力可能な深層学習手法が提案されている. PointNet[177] の登場以降, 点群データに対する様々な深層学習手法 [68, 3, 134] が研究されてきた. 直接点群データを扱う深層学習手法を用いることで, 点群データの地物分類性能において従来の人間が設計する特徴量を用いる機械学習手法を上回った. さらに, 一つの深層学習によって複数の地物を同時に分類可能であるため, 複数の手法を組み合わせる等の複雑な処理が必要ないという利点がある. しかしながら, 表 1.1 に示すように点群・波形データを変換することなく直接深層学習モデルで扱う手法は存在しない現状にある. そのため新たに点群・波形データに対する深層学習手法を提案する必要がある. この際, 使用する点群・波形データは波形データのファーストパルスから作成した点群データがペアになっているものを用いる. この理由は, 3 次元地図においてフットプリント内に占める面積が小さく弱い反射エコーを観測する地物も対象にするため, 反射の強度に依存しない点群データの作成方法としてファーストパルスに限定したためである.

表 1.1: 点群・波形データの地物分類に関する既往研究

	特徴量化	画像化	直接入力
ルール	✓	✓	-
従来型機械学習	✓	✓	-
深層学習	-	✓	-
			研究対象

### 1.3.2 アーカイブ点群データの有効活用

オープンデータとして公開された点群データに対して点群・波形データに対する地物分類手法の学習済み深層学習モデルを適用する状況を考える. 具体的なオープンデータのサイトとして, 日本においては, 3 次元の情報を基盤とする国土交通データプラットフォーム [260] が存在する. さらに, 産業技術総合研究所においても点群データ等を検索するプラットフォーム

ームが構築されている [1]. また, Shizuoka Point Cloud DB(2022 年 1 月に閉鎖) や東京大学が運営する G 空間情報センターで航空機に搭載されたライダーで観測された点群データが公開されつつある [69, 196]. G 空間情報センターで公開されている点群データはオープンデータ概念が普及した 2015 年以降の計測データが多い. このようなオープンデータ等で公開された点群データをアーカイブ点群データと定義する. これらの Web サイトにおいて, 実際に使用可能なオープンデータや納品物において波形データは公開されておらず点群データのみ公開されている. この理由として, 点群データの利用が地盤面の高さの測量に使用されることが多く, この際波形データは点群データの生産時にのみ使用され点群データの納品時には波形データが不必要であることが挙げられる. また, オープンデータの Web サイトにおいては, 波形データが点群データの最大で 15 倍程度のデータ容量があり, ストレージを圧迫する運用上の要素がある. つまり, 一般的に入手可能なアーカイブ点群データに対して点群・波形データに対する地物分類手法の学習済み深層学習モデルを活用する際に波形データが利用できない. 提案した点群・波形データに対する地物分類手法の学習済み深層学習モデルを利活用するため, 点群データから波形を推定し, 擬似的に点群・波形データ化して点群・波形データの地物分類モデルへ入力する. この際の前提条件として, 商用の点群および波形情報を記録するライダーが普及し, かつアーカイブデータのサービスが登場した 2015 年以降の計測データを対象にする. とりわけ, 今後点群データのオープンデータ化が促進すると考えられるため, 技術革新によって大きく観測方式が変わった過去の点群および波形情報を記録するライダーは対象外とする. また, アーカイブデータにおける入手性の高さからファーストパルスの点群データを対象にする.

点群データから波形を推定するタスクは「点群データが本来持っている情報を推定するタスク」と定義できる. このようなタスクとして, 深層学習を利用した点群データの色を推定する点群着色手法が研究されている [21, 131]. また, 波形データに対しては, 低いサンプリングレートの波形データから高いサンプリングレートの波形を推定する手法が研究されている [129, 130]. しかしながら, 表 1.2 に示すように, 点群データから波形を推定する手法は存在しないため, 新たに提案する必要がある.

表 1.2: 波形推定に関する既往研究

		出力	
		色情報	波形データ
入力	波形データ	-	✓
	点群データ	✓	-

研究対象

## 1.4 研究目的

本研究では、3次元地図作成を行う際の点群データの目視判読を自動化するため、点群および波形情報を記録するライダー計測で得られた点群・波形データの地物分類手法を提案する。そのために、点群・波形データに対する深層学習モデルを用いた地物分類手法を提案する。地物分類手法の対象は、アーカイブデータ等にて入手可能な点群および波形情報を記録するライダーで観測された点群・波形データであり、分類地物数は標準的な点群データや点群・波形データの仕様であるLASにて定義される15種類のうち代表的な地面・植生・建物・電線・鉄塔・道路に限定する。また、対象のセンサを現在のオープンデータにて入手可能な点群データや点群・波形データを観測したライダーと同程度のスペックを有するライダーを対象にする。また、提案する地物分類手法の適用範囲をオープンデータ等に広げるために、波形データが削除された状態のアーカイブ点群データを用いる必要がある。そこで、点群・波形データの地物分類手法をアーカイブ点群データに対しても活用するために、点群データからその点群データが計測された際の波形データ情報を推定する手法を提案する。これらの提案手法を組み合わせ、アーカイブ点群データから波形の推定と、その推定された波形データを用いた擬似的な点群・波形データに対して地物分類手法の活用を行う。

本研究では上記の目的を達成するために、手法の提案とその提案手法の適用を行う。提案手法は、「深層学習を利用したデータ変換を必要としない点群・波形データの地物分類手法」と「深層学習を利用した点群データからの波形推定手法」の2つである。手法の適用では、「波形推定手法と点群・波形データの分類手法の組み合わせによるアーカイブ点群データの地物分類」である。

## 1.5 博士論文の構成

この博士論文は本章を含めて6つの主要な章(図 1.2)で構成される。本章の後に、第2章では関連研究として、点群および波形情報を記録するライダーの観測方法を述べ、次に、深層学習手法についての基本的な説明を行う。第3章では、都市域を対象にした点群・波形データに対する深層学習手法を用いた点群・波形データの地物分類手法を提案し、ベンチマークデータによる検証実験を行なった結果を述べる。第4章では、アーカイブ点群データの有効活用のための波形データ推定方法とその検証実験を行った結果を述べる。次に、第5章では第3章で行なった深層学習を用いた点群・波形データの地物分類手法と第4章にて行なった波形データ推定手法を実応用した結果を述べる。具体的には、第4章で推定した波形データを第3章で作成した点群・波形データ地物分類モデルへ入力し、擬似的な点群・波形データとして地物分類を行いその分類結果を述べる。最後に、第6章では、本研究のまとめと課題点の整理と今後の展望について述べる。以降でそれぞれの章について詳しく説明する。

第2章では、一般的なライダーに関する説明と点群および波形情報を記録するライダーの特徴、および、深層学習による点群データの地物分類手法と条件付き敵対的生成ネットワークを紹介する。まず、点群および波形情報を記録するライダーの計測方法を原理から説明し、計測された波形データの特長について説明する。この際、主要な点群および波形情報を記録するライダーについて、その用途を問わず一般的な計測原理を説明する。次に、深層学習に関して説明する。まず、深層学習の原理を説明し、一般的な画像に対する深層学習の説明を行い、次に点群データに対する深層学習手法、そして本研究において波形データ推定に用いられる条件付き敵対的生成ネットワーク手法に関する説明を行う。次に、既存の主な点群・波形データ地物分類手法について可能な限り包括的に説明する。章の最後に、既往研究と提案手法の関係性を述べる。

次に手法の提案を担う章について説明する。第3章では、深層学習を利用した、点群・波形データの地物分類手法を提案する。まず、点群・波形データの地物分類についての問題設定を定義する。次に、深層学習モデルを提案する際に用いる仮説について述べる。その際に波形データの形状が分類に寄与する地物と波形データの空間的な関係性である大局特徴が分類に寄与する地物が存在すると仮説を立てた。そのため、それぞれの地物に特化した深層学習モデルを組み合わせて地物分類を行う。実験の結果、提案した点群・波形データの地物分類手法は既存の深層学習モデルを拡張した手法に比べ、高い地物分類性能を示した。

第4章では、点群データから波形を推定する深層学習手法を提案し、その学習済みモデルによる波形の推定を行う。その目的は、アーカイブ点群データには波形データが存在しないことが多く、提案した点群・波形データの地物分類手法を有効活用するために波形を推定し擬似的に点群・波形データとして地物分類を行う。まず、波形推定手法を提案する際に用いる仮説について説明する。次に提案手法を説明する。波形推定に関する検証実験を行なった結果、定量・定性評価において波形の推定性能が高いことが示された。

第5章では、第3章と第4章の内容を組み合わせ、アーカイブ点群データの波形を推定し、その推定された波形データを用いた擬似的な点群・波形データの地物分類を行う。具体的には、点群データから波形を推定する深層学習モデルが第4章にて提案され、また、点群・波形データに対する地物分類を行う深層学習モデルも第3章にて提案された。これらの訓練済みモデルを用いて波形を推定し、その推定された波形データと点群データを擬似的に点群・波形データとして地物分類モデルに入力する。実験の結果、推定された波形データを用いることで単に点群データのみを使用する場合と比較すると、地物分類性能が向上した。

最後に、第6章で博士論文の結論を説明する。具体的に、全体のまとめと各提案手法のまとめを行い、提案手法の貢献と適用範囲について説明を行う。最後に、提案手法の適用先と点群・波形データの利活用に関する可能性についての展望を述べる。

第 1 章	序 論
第 2 章	点群および波形情報を有するライダーと既往の深層学習手法
第 3 章	点群・波形データを用いた地物分類手法の提案
第 4 章	点群データからの波形推定手法の提案
第 5 章	波形推定手法と地物分類手法の効果検証
第 6 章	結 論

図 1.2: 本研究の構成. 本章の後に, 第 2 章では関連研究として, 点群および波形情報を記録するライダーの観測方法およびその観測データである点群・波形データの地物分類方法について述べ, 次に, 深層学習手法についての基本的な説明と深層学習を利用した点群・波形データの地物分類の既存手法を紹介する. 次の章からは提案手法の紹介を行う. 第 3 章では, 都市域を対象にした点群・波形データに対する深層学習手法を用いた点群・波形データの地物分類手法を提案し, ベンチマークデータによる検証実験を行なった結果を述べる. 第 4 章では, アーカイブ点群データの有効活用のための波形データ推定方法とその検証実験を行った結果を述べる. 次に, 第 5 章では第 3 章で行なった深層学習を用いた点群・波形データの地物分類と第 4 章にて行なった波形データ推定手法を実応用した結果を説明する. 具体的には, 第 4 章で推定した波形データを第 3 章で作成した点群・波形データ地物分類モデルへ入力し, 擬似的な点群・波形データとして地物分類を行いその分類結果を述べる. 最後に, 第 6 章では, 本研究のまとめと課題点の整理と今後の展望について述べる.

## 第 2 章

点群および波形情報を有するライダーと既往の深層学習手法

## 2.1 はじめに

本章では、博士論文に関係する分野として、点群および波形情報を記録するライダー計測の原理と深層学習について説明し、深層学習を用いた点群・波形データの地物分類方法を述べる。そして、最後に、これらの関連研究と本研究との関係を述べる。

点群および波形情報を記録するライダー計測の原理では点群および波形情報を記録するライダーに限らず一般的なライダーによる計測原理に関して説明し、その後、点群および波形情報を記録するライダーの原理を説明する。続いて、点群および波形情報を記録するライダーのハードウェア構成とデータ形式を紹介する。次に、商用の点群および波形情報を記録するライダーの代表例の説明と、計測された波形データを処理し点群データを得る手法を説明する。さらに、観測された波形データの反射の強度が持つ特性など、いくつかの一般的な説明を行う。

深層学習に関する説明では、提案手法である点群・波形データの地物分類と、点群データからの波形データ推定では、点群データに関連する深層学習手法を説明する。まず、深層学習の一般的な説明を行い、その歴史背景から、基本的な積和演算や活性化関数の説明を行う。次に、画像データと点群データに対する深層学習手法を説明する。さらに、波形データの推定において深層学習手法のひとつである条件付き敵対的生成ネットワーク（条件付き GAN）を用いるため、条件付き GAN に関する一般的な説明を行う。

深層学習を点群および波形情報を記録するライダーへ適用する場合の説明では、まず、点群・波形データの地物分類に関する説明を行う。次に、波形データと点群データを合わせた点群・波形データの地物分類を行うための古典的な手法を説明し、深層学習を利用した地物分類を説明する。さらに、条件付き GAN を点群データや波形データへ適用する場合の説明を行う。最後に、上記で紹介した手法と本研究との関係を述べる。

## 2.2 点群および波形情報を記録するライダーの原理と解析方法

### 2.2.1 航空機によるライダー計測の原理

本節では点群および波形情報を記録するライダーに限らず、一般的な航空機によるライダー計測に関して説明する。一般的な航空機によるライダー計測は、表 2.1 に示すようにライダーと地表との間の距離を直接計測する能動型のリモートセンシング技術のひとつである。リモートセンシングによる都市域の計測は、近年の測量技術の進歩により計測性能が向上しつつあり、建物や道路、植生や電線等の複雑な構造物で構成される都市域の解析にも用いられるようになった。光学衛星画像の空間分解能は民間向け商用衛星の WorldView では 30 cm になり、非商用で誰もが自由に使用可能な衛星データである Sentinel-2 においても 10 m 解像度となった。これまで広く用いられた Landsat の 30 m 解像度と比較すると地上解像度としては詳細な情報が得られるようになった。光学衛星画像は空間解像度だけでなく、スペクトルの解像度も高くなり、最新のハイパースペクトルセンサは、100 以上のバンドを収集可能になった [54]。また、10 cm 以下の空間分解能を持つ航空機による光学画像計測が普及し、より微細な構造物の検出など、新たな研究分野や応用の可能性が広がった [15]。さらに、全天候型の合成開口レーダ画像 (SAR) は、ALOS-2/PALSAR-2 では最大で 3 m、Sentinel-2 では 5 m の空間分解能を有し、都市域の分析にも十分に貢献可能である [199]。

このような高解像度のリモートセンシングデータを都市域で解析し、様々な情報が抽出可能になる [158]。建物などの人工建造物、道路、植生などの都市構造をリモートセンシングデータから検出した結果は、地理情報システムにとって必要な基本情報になる。また、建物、道路、樹木等の各地物の抽出結果や、土地被覆分類や土地利用分類などの分類結果から自動的に地図を作成することも可能となる。しかし、地物分類を行う際には光学衛星・航空画像を用いる場合には太陽光を利用した受動型センサであるため天候に左右され、同じ構造物でも撮影された時期や時間によって写り方が変化し、地物分類を行う上で堅牢なアルゴリズムを作成することが難しい。また、能動型センサである SAR を用いる場合においても、レーダの反射を目視判読することには判読者の慣れが必要であり、さらにその目視判読に必要な知識を地物分類に反映させることは困難である。

一方、これらの欠点に対し、航空機に搭載されたライダー計測技術は多くの利点を有している。航空機に搭載されたライダー計測では、航空機によって迅速に広範囲の点群データを高密度に取得できる。さらに、近年の航空機に搭載されたライダー計測は、高さ方向に 10 cm 程度の計測誤差で点群データを取得可能である。また航空機に搭載されたライダー計測は光学画像から 3 次元情報をステレオ計測にて推定する必要がなく、対象物の幾何学的な特性を直接計測可能である。そのため、航空機に搭載されたライダー計測は光学画像や SAR 画像と比較すると、影等による色の変化などの外乱が少なく堅牢な地物分類に特に適している。さらにこの 10 年間で、航空機に搭載されたライダーで得られた点群データに対して、道路の抽出 [270]、橋や送電線の検出 [198, 96]、DTM の作成 [17]、沿岸地域の管理 [85]、さらには材料計測 [56] や文化遺産の保全 [44] など、多くの研究が行われてきた。特に、乗用車や UAV の自律走行に必要な障害物や、都市のインフラ管理に必要な情報は電線やポールなどの 3 次元的な形状が重要になる物体であり、これらは航空写真や衛星画像の 2 次元的な計測では情報が取得できないため、ライダーで観測したデータの解析が重要になる。

表 2.1: リモートセンシングに用いられる代表的な機材

	受動型	能動型
形式	光学	SAR ライダー
メリット	人間の見た目に近い	全天候 位相がわかる
デメリット	天候に左右される	距離がわかる 判読しにくい 点群データの処理が重い

まず、航空機によるライダー計測原理の概要を説明する。ライダーと物体間の距離は、ライダーの搭載された航空機から高いサンプリング周波数でレーザを照射し、照射されたレー

ザが地上の物体に反射した後方散乱信号が受信機方向に帰ったものが測定されるまでに経過した時間から求める [12]. そして、距離を元に衛星測位システム (Global Navigation Satellite System, GNSS) と慣性計測装置 (Inertial Measurement Unit, IMU) で求めたライダーの中心位置を用いることで地球上の実座標である緯度経度高さ情報を計算でき点群データが得られる [13, 209]. この座標を計算する簡単な説明は以下に示す (詳細は次節で説明する).

- GNSS と IMU の両方を使用したハイブリッドシステムを用いて、航空機の位置姿勢とライダーの絶対方位を計算する. この位置姿勢の計算はリアルタイムで処理するのではなく、計測終了後のオフライン後処理によって平滑化された航空機の飛行経路を算出する [75].
- 後方散乱信号から地表にある物体からの反射を検出し、その反射が帰ってきた時間を計測する. そして、航空機の位置姿勢情報とライダーの絶対方位を元に、極座標で地物までの距離を表現する.
- その後、ジオリファレンス処理により、極座標で表現されたライダーと地物間の距離の値が、 $x, y, z$  座標として表現される.

上記の処理を行うために、航空機に搭載されたライダー機器は、レーザ送信機と受信機 (反射したパルスを受信)、GNSS/IMU、ハードディスクなどの記憶媒体、信号のデジタル化とデータ収集を担うコンピュータなどで基本的に構成される..

次に、航空機によるライダー計測のレーザ照射方式 [194, 218] を説明する. 航空機によるライダー計測に使用されるレーザの照射方式としては、パルス方式と連続波方式のいずれかが使われている. パルス方式では、レーザから物体までのレーザの往復時間を単純に測定する. 一方、連続波方式では、送信信号と受信信号の位相差を測定して測距を行う. 現在使用される航空機によるライダー計測に用いられる機器の多くはパルス方式である. パルス方式におけるレーザは高いエネルギー・短いパルス・高いサンプリングレート・狭いビーム幅で照射される. 現在、水深計測などの特別な用途を除くと、航空機によるライダー計測で主に使用されるレーザの光は単色の近赤外線で位相はコヒーレントである.

次に照射されたレーザについて説明する. レーザを照射するパルス繰返し周期 (Pulse Repetition Frequency, PRF) は、観測条件の設定に状況に依存するが、通常は 100 kHz よりも高速である. 各パルスの照射は、前に照射したパルスの記録が完了した後に実行される. この PRF は高度つまり地表との距離によっては遅くすることもでき、飛行高度 3 km で観測する場合においては、PRF が 50 kHz であれば、前のパルスが戻る時間が確保可能である. ただし、機材によっては、前のパルスを記録する前に次のレーザを照射する機能も有しており、レーザの照射回数に比例して観測可能な点も増加する [186]. 照射されるレーザは指向性があるため、必然的にその形状は扇形になる. なお、扇型の角度は 0.4~0.8 mrad である. そのため、1つのレーザが照射されると、そのレーザが通過した 3次元円錐領域内の異なる位置にある物体から反射して、複数の多重反射によるエコーが発生する可能性がある. なお、この多重反射の数は、物体の特性だけでなく、航空機の観測条件にも依存する. 次に、この照射されるレーザを数式を用いて定義する. レーザの直径  $D$  (物体に平行な正面を想定した場合) は、次のように定義できる.

$$D = 2H \tan \frac{\beta}{2} \quad (2.1)$$

ここで、 $H$  はライダーまでの距離であり、 $\beta$  は発散角度を示す. たとえば、1,000 m の距離で  $D = 0.5$  mrad である. 航空機によるライダー計測ではレーザが物体の表面と衝突すると、フットプリントと呼ばれるレーザが照射される面が発生する. このフットプリントは、物体の表面 (つまり、レーザと物体の角度の関係と材料の特性) によって異なる. レーザの発散が大きくなると、フットプリントと 3次元円錐領域が大きくなり、放出されたパルスごとに複数のエコーが得られる確率が高くなる. このような複数個のエコーを観測可能なライダーは、ある地表に複数の地物があった場合にも最上部と最下部を同時に観測できるため、樹木等があったとしても地面の高さを計測可能という利点がある. 次に、レーザのパルス幅  $p$  について説明する. 一般的に使用される、 $p = 4$  ns の典型的なパルス幅は、光速でレーザが進行する

場合、約 1.2 m のパルス幅に相当する。パルス幅は、複数の反射を検出または区別する際に、大きな影響を与えるため、ライダー機器において重要な特性である。具体的には、二つの物体間の距離  $d$  がパルス長  $l_p$  の半分より大きい場合は 2 つのエコーを区別可能である。

$$d \geq \frac{1}{2}l_p \text{ with } l_p = \frac{c}{n}\tau_p \quad (2.2)$$

具体例を示すと、 $p = 4$  ns の場合、物体間の距離が少なくとも 0.6 m より大きい場合、2 つの物体が識別される可能性がある。

このようなレーザーを使用して実際に距離を計算する方式について説明する。光がある媒体中を移動する際に生じる Time-of-Flight(TOF) の飛行時間 ( $t$ ) を測定することで、ライダーと物体の間の距離 ( $d$ ) を以下の式で簡単に表現できる。

$$d = \frac{v_g}{n} \frac{t}{2} \quad (2.3)$$

ここで、 $v_g$  は光速であり、 $n$  は媒体の屈折率（空気中では  $n = 1.00025$  であり、多くの場合 1 に等しいとされる）を表す。レーザーの送信機は、赤外線領域（波長 800 nm～1,550 nm）で短時間（通常は数ナノ秒）のパルスを照射する。したがって、物体サイズは波長よりもはるかに大きくなる。照射されたレーザーは、波長に応じて大気中の粒子と相互作用である吸収または散乱が発生するが水蒸気とエアロゾルがない場合は無視できる範囲であるため、主に地球上に存在する自然物または人工物にのみ相互作用が発生すると言える。したがって、レーザーの伝播速度はほとんどの場合一定であると想定される。

ここからは、ライダー計測を定式化するライダー方程式について説明する。ライダー方程式では検出器とレーザーの当たる物体の特性を考慮して、測定過程を定義する。例えば、送信する信号と反射し戻る信号の電力を関連付けることが古典的な定式化である [89]。また、実世界の条件をより詳細にモデル化するために、詳細な方程式も提案されている [39, 22]。しかし、これらのモデル化は単一の物体、または平坦な表面に対してのみ有効である。樹木などの空間的に広く分布している物体の場合、ライダーへ帰る信号は様々な物体にあたり反射し、それぞれで反射した小さなエコーの重ね合わせで観測値は表現できる。また、受信電力  $P_r$  は積分として表すことができる。

$$P_r(t) = \frac{D^2}{4\pi\lambda^2} \int_0^H \frac{\eta_{sys}\eta_{atm}}{R^2} P_t(t - \frac{2R}{v_g})\sigma(R)dR \quad (2.4)$$

ここで、 $t$  は計測時間、 $D$  は受信光学系の開口径、 $P_r$  は受信電力、 $P_t$  は放射電力、 $\lambda$  は波長、 $H$  は計測高度、 $R$  は LiADR システムから物体までの距離である。また、 $\eta_{sys}\eta_{atm}$  はそれぞれ、大気とシステムの透過率、 $v_g$  はレーザーの速度、 $\sigma(R)dR$  は見かけの有効微分断面積 [222] である。この「見かけ」とは、与えられた距離で信号を反射する物体が遠くの物体を遮る可能性があるため、仮定する断面は「見かけ」と呼ばれる。上記の定義は一つの物体からの反射しか考慮していないという問題がある。そこで、受信電力は、独立する特性を持つ  $N$  個の物体からの反射の合計と見なすライダー方程式も開発された。

$$P_r(t) = \sum_{i=1}^N (P_{r,i} * \eta_{sys} * \eta_{atm})(t) \quad (2.5)$$

ここで、 $*$  は畳み込み演算子であり、 $P_{r,i}(t)$  は次のように表される  $i$  番目の物体からの反射エコーである。

$$P_{r,i}(t) = \frac{D^2}{4\pi\lambda^2} \int_{R_i-\Delta R}^{R_i+\Delta R} \frac{1}{R^4} (P_t * \sigma'_i)(t) \quad (2.6)$$

ここで、 $R_i$  は平均距離、 $[R_i - \Delta R, R_i + \Delta R]$  は空間的な広がり、 $\sigma'_i(R)$  は  $i$  番目の反射と関連した物体の有効な微分後方散乱断面積である。つまり、反射信号は、送信されたパルスと有効な微分断面積の間の畳み込みとして見るることができる。通常であれば、現在商用で使用

される機材の場合、 $R$ は0.1~1.5 mの間である。したがって、 $\Delta R \ll R$ となり、結果として、次のように定義できる。

$$P_{r,i}(t) \approx \frac{D^2}{4\pi\lambda^2 R_i^4} (P_t \star \sigma'_i)(t) \quad (2.7)$$

ここで、 $\sigma'_i$ は、各範囲間隔 ( $m^2$ ) 内のレーザが照射された領域の見かけの断面である。受信電力は、最終的に次のように表すことができる。

$$P_r(t) = \sum_{i=1}^N \frac{D^2}{4\pi\lambda^2 R_i^4} \underbrace{P_t \star \eta_{sys}}_{system\_contribution} \star \underbrace{\eta_{atm} \star \sigma'_i}_{environment\_contribution} \quad (2.8)$$

したがって、受信波形データは、照射したレーザと表面応答  $\sigma'_i$  の間の畳み込みであり、後方散乱断面積 (Back Scatter Coefficient, BCS) として定義できる。 $\eta_{atm}$  は時間に対してほぼ一定であるが、主に  $R$  の影響を受ける。 $\eta_{atm} = \exp(-2R/a)$  [203, 94] または [78] の場合は  $\eta_{atm} = 10^{\frac{-2R/a}{10000}}$  になる。ここで、 $a$  は dB/km で表され、先行研究 [203] のレーザ波長と視程距離の関数として表される大気減衰係数である (Koschmieder の式に基づく経験的スペクトル近似)。実際には、 $\eta_{sys}$  は、時間の経過とともに、特定のライダーに対して一定であると想定されている。見かけの断面  $\sigma'_i$  は、すべての物体特性を収集し、次のように定義される。

$$\sigma'_i = 4\pi \frac{\rho A_t}{\Omega} \quad (2.9)$$

ここで、は物体の立体角 ( $\Omega = \pi$ )、 $A_t$  は物体領域 ( $A_t = \frac{\pi R^2 \beta^2}{4}$ )、 $\rho$  は物体の反射率である。ランバート散乱物体  $\rho = \rho_m \cos \alpha$  と仮定する。ここで、 $\rho_m$  は材料の反射率、 $\alpha$  はレーザ光と物体法線の間角度である。したがって、以下のように表記できる。

$$\sigma'_i = \pi \rho_m \cos \alpha \beta^2 R_i^2 \quad (2.10)$$

最終的に、上記の式をまとめると受信電力は次のように書き直すことができる。

$$P_r(t) = D^2 \eta_{sys} \rho_m \cos \alpha \sum_{i=1}^N \frac{\exp(-2R_i \alpha)}{4R_i^2} P_t(t) \quad (2.11)$$

初期の航空機に搭載されたライダーは、放出されたレーザに対して単一の反射エコーのみを計測した。このような機材を用いる場合、レーザが照射される経路内に単純な物体ひとつがある場合は、単一のエコーの記録で十分である。しかし、レーザのフットプリントが小さい場合でも、レーザの経路内に多くの物体が存在する可能性がある。これらの個々の物体からの反射エコーは、レーザが照射された物体ごとに作成される。一方疎な点群を記録するライダーでは、1つのレーザ照射から最大6つの反射エコーを区別できるようになった [209]。このような疎な点群を記録するライダーは単一の反射を記録する時代と比較すると主に、植生地域では複数の反射が発生するため有利になる。一般的には植生があまり密集していない場合、疎な点群を記録するライダーの最初のエコーは樹冠の上部に相当し、最後のパルスは地面に相当すると見なされる。しかしながら実際には、この条件が常に当てはまるとは限らない。都市域では、特定の視野角でレーザが建物の端に当たると、2つのエコーが生成される可能性がある。最初のパルスは屋根の一部から反射し、2番目のパルスは地面に当たるレーザに相当する。図 2.1 に示す青色は1回の反射、緑色や黄色で示した箇所が複数反射の具体例である。十分な高さの上部構造 (主に煙突) が存在する場合、建物の屋根内でも同じ動作が見られる場合がある。なお、この現象はフットプリントが大きい場合に強く作用するが、近年の機器を用いた場合においても、フットプリントは 30 cm 程度が一般的な大きさであるため、依然として境界部の複数反射は発生する。

このような計測原理によって、疎な点群を記録するライダーの欠点が存在する。疎な点群を記録するライダーの場合、反射パルスの検出は後方散乱信号に依存する。検出器は、観測された時系列の反射パルスからピークを検出し、いくつかのタイムスタンプ付きパルスに変換

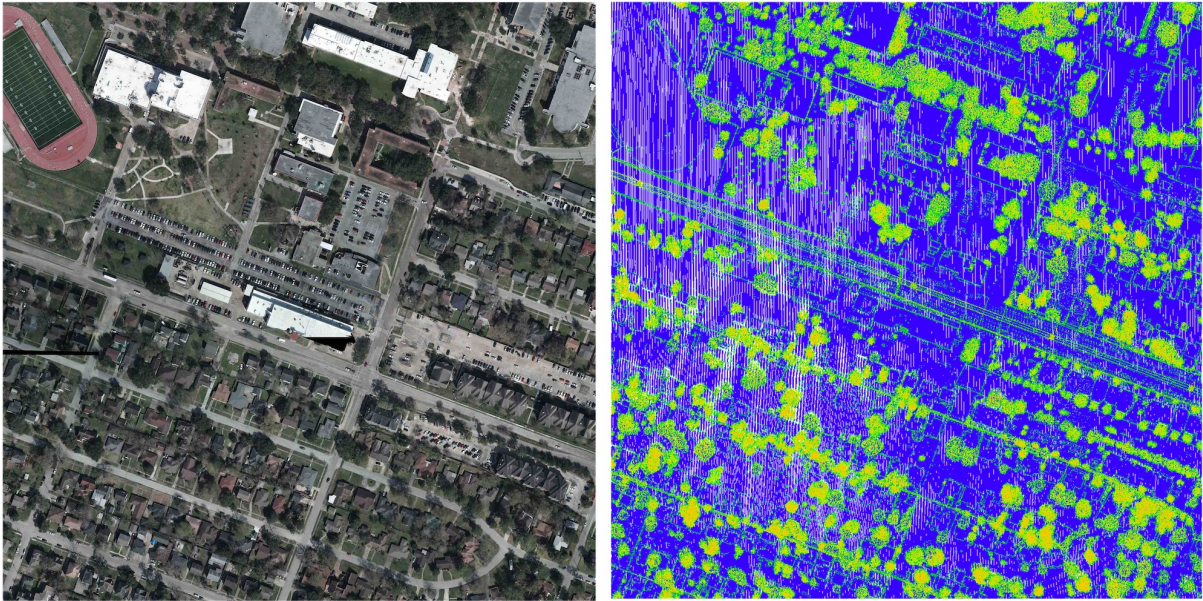


図 2.1: 都市域における複数反射の例. 青色は反射回数が1回, 緑色や黄色やオレンジ色で示した箇所が複数反射の具体例.

し, 個々の物体の位置を計算する. この処理過程におけるピーク検出において, 多くのピーク検出方法が提案されているが, 商用のライダーメーカーはハードウェアシステムに実装されている検出方法に関する情報を提供していない. また, エコーの形状に関する情報を提供しないため記録されるパルス数と計測時間はピークの検出方法に大きく依存する [221]. 反射信号は, 放出されたレーザ, 放出されたパルスと物体の間に入射角, 物体の粗さと反射率の動作 (双方向反射率分布関数で表現), 物体までの距離 ( $R^2$  減衰), およびオクルージョンとその影響で表現できる. その結果, 単純な距離計算方法は, 以下の状況下における欠点が存在する.

- パルス位置の誤推定  
物体の特性とレーザの入射角に応じて, レーザは減衰・伸長・傾斜する場合がある. これは, 先行研究 [221] の領域であっても, ピーク検出方法のパフォーマンスに影響を与える可能性がある. この原因で発生する最終的な点群データの誤差は, 主に高さ方向 (場合によっては 0.3 m 以上) である.
- エコーの検出  
ライダー機器は, 検出可能なピークに上限がある. この上限を超えると, 必然的にその他の反射エコーが失われる. これは主に複数の反射が起きる植生や鉄塔などの地物で起こる.
- パルスの重複  
標準的なピークの検出技術では, 複数のピークが近接している場合, それぞれの信号を分離された反射エコーに切り分けられない場合がある. 次に, これら2つのパルス間でパルス位置が推定され, 誤った点群データが生成される可能性がある [172]. 具体的には, 都市部では次のような状況で発生する.
  - 低い植生・建物の上部構造  
地面または建物のエコーに小さな隆起を引き起こす可能性がある.
  - 樹冠の反射による植生領域  
枝などに反射して多数のエコーが生じる.
- 弱いエコー  
5個の反射が発生したとすると, 反射された信号電力全体の大部分が最初の2つのエコ

一に含まれている。したがって、3つまたは4つを超えるパルスを実タイムで検出するには、ノイズ内の低強度信号を検出する必要がある。したがって、弱いエコーが見落とされる可能性がある。これらの状況は以下に対応する。

- 樹木の中の「エコー」と、下の地面で反射した「エコー」
- 屋根のエコーが素材の強い吸収や大きな入射角によって減衰する場合

### 2.2.2 点群および波形情報を記録するライダー計測

本項では、前述した疎な点群を記録するライダーの欠点を補うために開発された点群および波形情報を記録するライダーに関する説明を行う。点群および波形情報を記録するライダーは送信された各レーザの反射エコーを受信した信号を時系列かつ連続して反射の強度として記録可能である。このような反射の強度の時系列的な記録は波形データと呼ばれる。以降では、照射したレーザが地物に反射する際の振る舞いと、波形データの記録方法と、点群および波形情報を記録するライダーの長所短所を説明する。

点群および波形情報を記録するライダーデータの波形データとは、照射されたレーザが物体に当たり生じた反射エコーの重ね合わせであると定義できる。このような波形データの形状に依存するものは機材のスペックであり、点群および波形情報を記録するライダーのサンプリングレート・スキャン角・フットプリントの大きさが要素である。市販されている多くの点群および波形情報を記録するライダーシステムは、計測高度に依存するが、一般的にフットプリントが小さく（直径0.2~0.5 m）サンプリングレートが高い。フットプリントが小さく、サンプリングレートが高いシステムは、点密度と正確な高度情報を提供することで、高密度の点群データを作成可能である。この点群および波形情報を記録するライダーシステムは、フットプリント内の1つまたはいくつかの物体の表面にレーザが当たり、その当たった表面に対応する明確なエコーを持つ波形データを生成する。そのため、都市部のような幾何学的に明確な反射を観測するには、小さなフットプリントが必要となる。都市部ではフットプリントサイズが非常に小さい（0.5 m以下）ため、電線や電柱や生垣、建物と地面の境目などの不連続な部分を除いて、観測される波形データは実質的には1つのピークをもつ波形データである。このようなケースは、地面や建物の屋根などの地物に対応する（図2.1で複数反射が起きていない部分や図2.2の地面の部分に相当する）。一方、樹木のような分散した物体の場合、図2.3に示すような複数の反射が発生する。

樹木等の中では、フットプリントの小さい点群および波形情報を記録するライダーシステムの最大の欠点は、密集した植生の下の地面に到達したかどうかを判断するのが難しいことである。その結果、地面や樹木の高さを正確に推定することが困難である[46]。一方、大きなフットプリント（直径10~70 m）を持つライダーは樹冠構造の特徴を明らかにすることを目的としている。このシステムの利点は、レーザが地面に当たる確率を高め、小さなフットプリントのバイアスを回避可能である。さらに、照射したレーザの反射波形データから、より広い範囲で遮断された表面の垂直方向の分布が記録されるため、簡単なフットプリントで代表的な樹冠の体積を推定することができる。その結果、フットプリントを広げると点密度が下がり高さ情報が集約され、平均的な反射が得られる。しかし、そのぶん平均的な地面の高度しか取得できないという欠点もある。

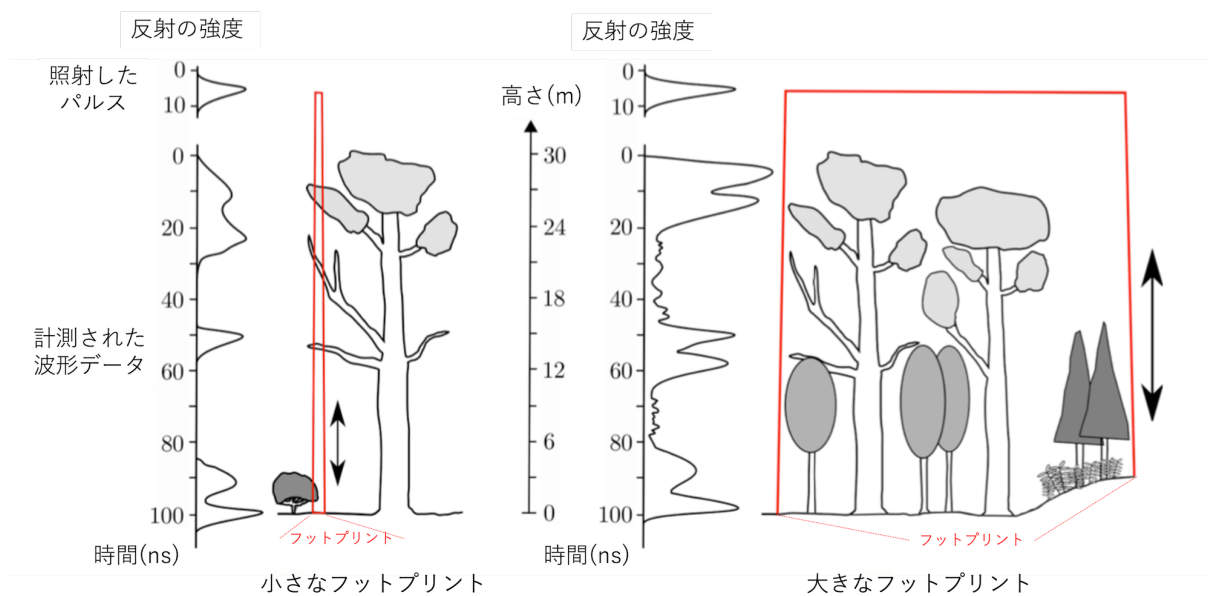
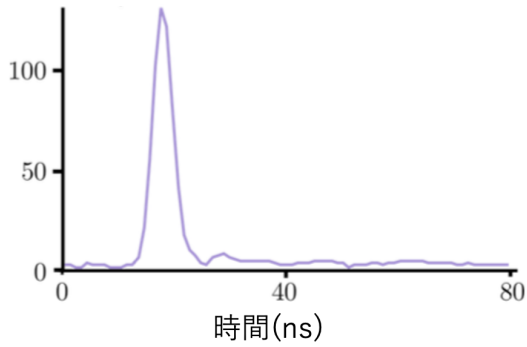


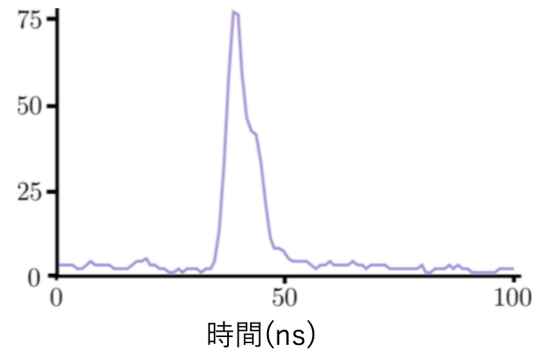
図 2.2: フットプリントの影響. 左図は小さなフットプリント, 右図は大きなフットプリントの波形データである. 大きいフットプリントでは平均的な反射を記録する. 図 LAS 仕様書 [7] より引用し加筆.

反射の強度



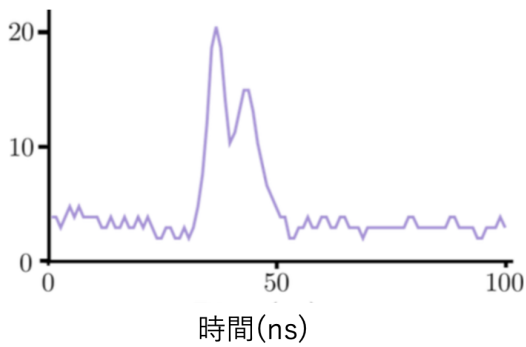
建物の屋根では1つのピーク

反射の強度



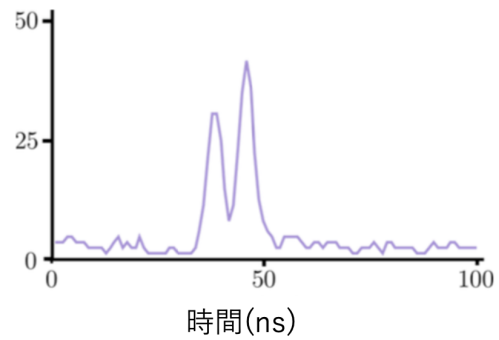
建物に上部構造物がある場合は  
ピークに乱れが発生

反射の強度



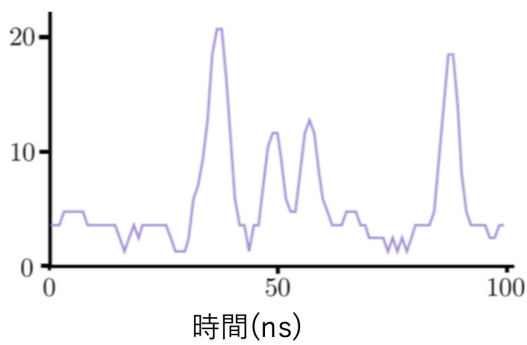
生垣の場合は生垣と地面でピークが発生

反射の強度



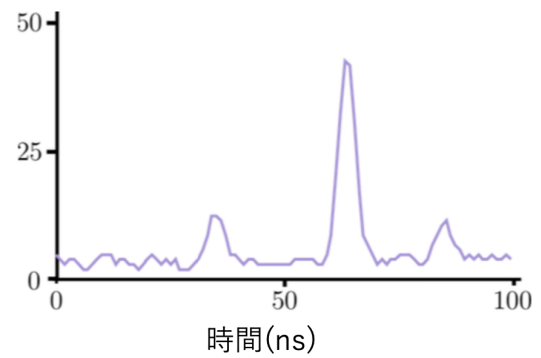
密な植生

反射の強度



疎な植生の場合

反射の強度



高木と低木が混ざった場合

図 2.3: 様々な状況における波形データ. 地物等の影響で観測される波形データには違いが生じる. 図は [145] の Figure 2.4 を参考に作図.

次に、波形データの記録方法について説明する。アナログ・デジタル・コンバーターを用いて信号を記録するための2つの主要な技術が[93]で紹介されている。まず、受信したエコー（照射したレーザの後方散乱光である）を受信光学系で集め、アバランシェフォトダイオードで電気信号に変換する。次に、出力された光電流が増幅された後、 $n$ ビットの量子化器とサンプリング周期（市販のデバイスでは1 ns）を用いてデジタル化される。このときの出力が波形データの振幅値になる。なおこの値には単位はなく、Digital Number(DN)で表されることが多い。

波形データを記録するために、ライダーの製造メーカーは既存のシステムにデジタル化端子を追加したり、記憶容量の大きいハードディスクを搭載して対応した。波形データは通常、8ビット・12ビット・16ビット等でデジタル化される。データ量は、同じエリアの点群データよりも非常に多くなる（通常は5倍以上）。点群および波形情報を記録するライダーを使って地域を調査する際の主な制限は、ストレージ容量の消費量である。現在の課題は、より大量のデータが記録されるようになったため、データの取り扱いと管理にあると言える。デジタル化のサンプリング周期は一定で、1~10 nsの間が用いられる。大規模なストレージの問題を避けるために、波形データ全体（レーザがライダーから物体に到達し、戻ってくるまでの時間履歴）はすべて記録されず、固定された最小値を超えた後に所定の長さだけ記録される。このような記録方式はライダーの仕様で規定されており、測量する対象エリアによってはその計測方式をその場に合わせて変更する。以下に詳細な説明を述べる。

- レンジゲート  
レンジゲートとは、デジタイザーが信号に反応しない時間帯のことで、次のような目的がある。大気で散乱した信号がデジタイザーを起動させないように、レンジゲートを設けている（例：Optech 3100では100 m）。
- ドロップアウト  
ドロップアウトを使用する場合は、後方散乱信号が閾値を超えないためセグメントは記録しない。
- 最大サンプル数  
記録される波形データは、あらかじめ定義された最大サンプル数に制限される。例えば、Optechでは、各パルスに対して最大440サンプルを保存することができる。これは、垂直方向に66 mの離散的な区間に相当する。TopEye MarkIIシステムでは、最初のパルス以降（最初のパルスから127サンプル）または最後のパルス以前のいずれかの定義された観測モードに従って、128サンプルを保存する。樹木の高さがシステムの最大記録長よりも高い場合、点群および波形情報を記録するライダーでは、樹冠からのエコーと地上からのエコーの両方を1つの波形データ内に記録できない可能性がある。
- 複数段階の記録  
この記録できない問題を克服するために、システムは1回のレーザ照射に対して複数段階に分けた記録を用いることもある。波形データの最後のサンプルの振幅が一定の閾値を超えると、一定の長さの新しいセグメントが記録される。波形データセグメントの数は、地形によって異なる。平坦で開けた地形では1つのセグメントが生成されるが、複雑な物体ではより多くのセグメントが生成される。記録可能な上限は、Riegl LMS-Q560システムでは2つ、Optech 3100システムでは7つとなっている。しかし、実際の計測においては、2つのセグメントがあれば十分である。なお、閾値を超えたデータに加えて、最初のサンプルが閾値を超える前と、最後のサンプルが閾値を下回った後のサンプルも記録される。

波形データのデジタイザーは、従来の疎な点群を記録するライダーよりも弱い信号を検出することはできない。実際、デジタル化された信号のSNR（信号対雑音比）は、量子化やその他のデジタイザーによるノイズのために低くなる。しかし、複数の波形データを平均化することで、より高い感度を実現可能である。最後に、レーザが照射されると、そのエネルギーのごく一部がビームスプリッターによってフォトダイオードに導かれ、その信号またはス

ターゲットパルスによってタイミングデバイスが起動される。これを「射出/透過パルス」または「T0パルス」と呼ぶことがある。なお、実際に放出されたレーザーエネルギーのうち、記録される割合はエンドユーザーには不明である。

ここからは、表 2.2 に示すように実際に利用する場合における点群および波形情報を記録するライダーで得られた点群・波形データを使用する際の長所と短所について詳しく説明する。これらの結論は客観的と点群・波形データを利用する際に気づいた主観的な実用上の問題点にも基づいている。

- 長所

- 最小パルス間隔の制限が無いこと  
点群および波形情報を記録するライダーは、距離を計算するシステムの上の必然的な限界である、「照射方向に数メートル以下しか離れていない反射を分離できない離散的なシステムの問題」を克服した。重複するピークが見つかることもあり、結果としてより密度の高い点群データが得られる。
- 地面の反射の検出  
大きな背景ノイズがあると、ライダーのハードウェア・アルゴリズムでは弱いエコーを見逃すことがある。波形データ処理を行うことで、このような弱いエコーの検出能力を高めることができる。
- 有用な情報の追加  
計測後のオフライン処理により、生の信号からの情報抽出を促進することができる。また、ピークの抽出だけでなく、波形データの形態的な解析も可能である。
- データの精度向上  
適切なピーク検出アルゴリズムを設計することで、エンドユーザーは距離推定を向上させ、抽出された点群データの品質を向上させることができる。
- メタデータの可用性  
放射されたパルスごとに、放射された信号と後方散乱信号に加えて、GPS タイムスタンプ、スキャン角度、ライダーの軌跡、位置と姿勢、レンジファインダーのデータ（範囲+振幅）が得られる。これらの情報は、点群データを生成するために必要である。メタデータを利用することで、対応するのピーク振幅をより均質にすることができる。

- 短所

- 低いパルス周波数  
波形データをデジタル化して記録すると、その処理を待つ時間を稼ぐために、ライダー機器はパルスの周波数を下げざるを得ない。高密度の点群および波形情報を記録するライダー計測では、より多くのストリップを重ねる必要があるが、これには測量コストの増加とストリップの追加登録が必要になるという 2 つの欠点がある。
- 信号対雑音比 (SNR) の低下  
波形データの SNR は離散的な疎な点群を記録するライダーに比べて低い傾向がある。そのため、信号のダイナミクスが悪くなり、稀にデータセットに欠損が発生することもある。具体的には、弱いパルスは観測される強度がハードウェア上にてあらかじめ定義された閾値以下であるため、波形データが記録されないことがある。
- 波形データ記録方法  
何百もの無駄な強度情報の記録を避けるために、記録数は制限されている。高い建物の縁や下草の無い植生地域では、レーザーが地面に到達する前に、その間に十分な振幅のサンプルが存在しないため、波形データ記録が停止することがある（いわゆるデッドゾーン）。これは、何百もの無駄な強度情報の記録を避けるために、ハードウェア上にて記録数が制限されているためである。

- 大きなデータサイズ  
現在、点群および波形情報を記録するライダーによる計測では、放射された信号と後方散乱された信号の完全な時系列データに加え、メタデータも記録されている。そのため、点群および波形情報を記録するライダーによる計測ではバイナリ形式であっても、疎な点群を記録するライダーの5~15倍のファイル容量になる。したがって、波形データやジオリファレンスの情報を抽出するためには、大容量のデータ処理が必要となる。そこで、[112]は、ウェーブレットベースの波形データの非可逆圧縮方式を提案し、21%のデータ削減を実現した。しかし、依然として波形データのデータ量が多い事実は変わらない。
- 計測時間の制限  
莫大なデータ量を有する点群および波形情報を記録するライダーの計測はハードディスクの容量によって制限がある。140 GBのハードディスクは、Optech 3100システムの50 kHzのPRFで1.6時間分のデータ取得に相当する。
- 専用ソフトが存在しない  
現在のところ、エンドユーザーが生波形データを1次元や3次元で可視化したり、加工したり、ジオリファレンスしたりすることができる標準的な点群および波形情報を記録するライダーデータ形式や専用ソフトウェアは普及してしない。ASPRS Las file format version 1.3では波形データデータの保存が可能だが、商用ソフトにはまだ採用されていない。さらに、ライダーの開発企業は、加工ツールまで含めた表示ソフトウェアを顧客に提供しているが、独自のファイルフォーマットに限定されている（例：Riegl社のRiAnalyze）。いくつかのフリーまたはオープンソースのソフトウェアが存在するが、測量会社等における生産ラインでは運用されていない[26, 148]。

表 2.2: 点群および波形情報を記録するライダーの長所と短所

	波形データ情報の付与	評価
長所	最小パルス間隔の制限なし	+++
	地面からの反射の特定	+++
	取得した追加情報	+++
	データ精度の向上	++
	メタデータの利用	++
	強度データの利用	++
	幾何的な情報へ変換	++
	低いPRF	-
短所	低いSN比	--
	波形データ全体の利用	-
	大きなデータサイズ	--
	限られた計測時間	-
	ソフトウェアの未普及	---

### 2.2.3 点群および波形情報を記録するライダーの種類

航空機に搭載されたライダー機器は、1970年代初頭に開発された。当時は直下視のシングルパルスの値しか取得できなかったが、近年では、回転ミラーやプリズムなどの機構により、レーザの照射角度を変化させながら観測できるため150 mから600 mの広い観測幅を持つようになった。このような広い観測幅により、様々な入射角で観測できるようになり、高密度の点群データが得られるようになった。現在、航空機によるライダー計測は、広範囲を迅速に、かつ高密度で正確な（計測誤差として通常、高度方向では0.1 m未満、平面方向では0.4 m未満）の点群データを取得可能である。そして現在では、点群および波形情報を記録するライダーが主流となった。点群および波形情報を記録するライダーは、1980年代に水深測定を目的として設計された[65]ことが始まりである。また、地形を計測する用途の点群および波形情報を記録するライダーは、1990年代半ばにNASAの実験で登場した。水深計や実験用システムは、いずれも特定の用途のために開発されたものであるが、ライダーメーカーは特定の用途を目的とせずにデジタル化端末を開発してきた。商用点群および波形情報を記録するライダーはフットプリントが小さく、サンプリング周波数が高いのが特徴である。なお、航空機に搭載する点群および波形情報を記録するライダーは2004年から商業的に利用可能になった。

最初に、開発当初の点群および波形情報を記録するライダーについて説明する。点群および波形情報を記録するライダーは開発当初において水深計として設計された。正確な海面深度を測定するために設計されたこの装置は、波長帯の違う2つのレーザで構成されている。緑色のレーザ(532 nm)は、空気と海面の境界を通過し、わずかな減衰を伴いながら海底まで水中を伝搬する。赤外線レーザ(1,064 nm)は水で反射し、海面までの範囲を測定する。近赤外波長では吸収係数が大きい、レーザのフットプリントは十分に大きく(2.5 m以上)、小波による反射で信号が受信機に戻る。したがって、水深計の波形データは2つのピークで構成されている。これらの波形データを処理するには、2つの主要な信号の最大値を見つけ出し、距離の値を導き出す必要がある。本研究では、陸地を対象にするため、水深測定システムの詳細については触れないが、より詳しい情報は[66]を参照されたし。

次に、商用の機器について説明する。2004年に最初の商用航空機点群および波形情報を記録するライダーが発表された(Riegl LMS-Q560 レーザスキャナをベースにした Lite Mapper-5600) [82]。ライダーメーカーは、Riegl社(LMS-Q560/680, VQ-380/460), Toposys(Falcon シリーズ), TopEye/Blom(MarkII), Optech(ALT3100), Leica(ALS-50II 60)などがある。また、Riegl社のように、地上レーザスキャニングやモバイルマッピングシステム用の点群および波形情報を記録するライダーを開発する会社も存在する。

### 2.2.4 波形データからの点群データの作成方法

ここでは、実際に観測された波形データから点群データを得る過程を説明する(図2.4参照)。点群および波形情報を記録するライダーの波形データから点群データを作成するために、現在では観測波形データを一定の要素に分解する手法が検討されている。照射したレーザの経路上にあるさまざまな物体を区別し観測された波形データを要素(航空測量分野ではエコーと呼ばれる)に分け、そのエコーを合計した値が観測で得られた波形データであると仮定する方法である。このようにエコーに分離する目的は、地物に関連するピークの検出率を最大化し、より密度の高い点群データを波形データから作成することである。なお、点群データの数を増やすことは、林業への応用や、都市部の建物のエッジを検索する際に役立つ。さらに、生の波形データ信号からの特徴量の作成を行うことで、点群データの分類時にその特徴量を追加することでその分類能力を向上することも期待される。また一般的には、より多くの情報を抽出することは、森林地帯でも都市部でも、地物分類の目的に役立つ可能性がある。波形データから点群データを作成する処理の流れは、「波形データ分解(デコンポジション)によるピーク位置決定」、「波形データモデル化」、「ピーク位置への位置情報の付与(ジオリファレンス)」で構成される。以降ではそれぞれの工程について述べる。

まず、波形データ分解(デコンポジション)について述べる。送信されるレーザが既知の

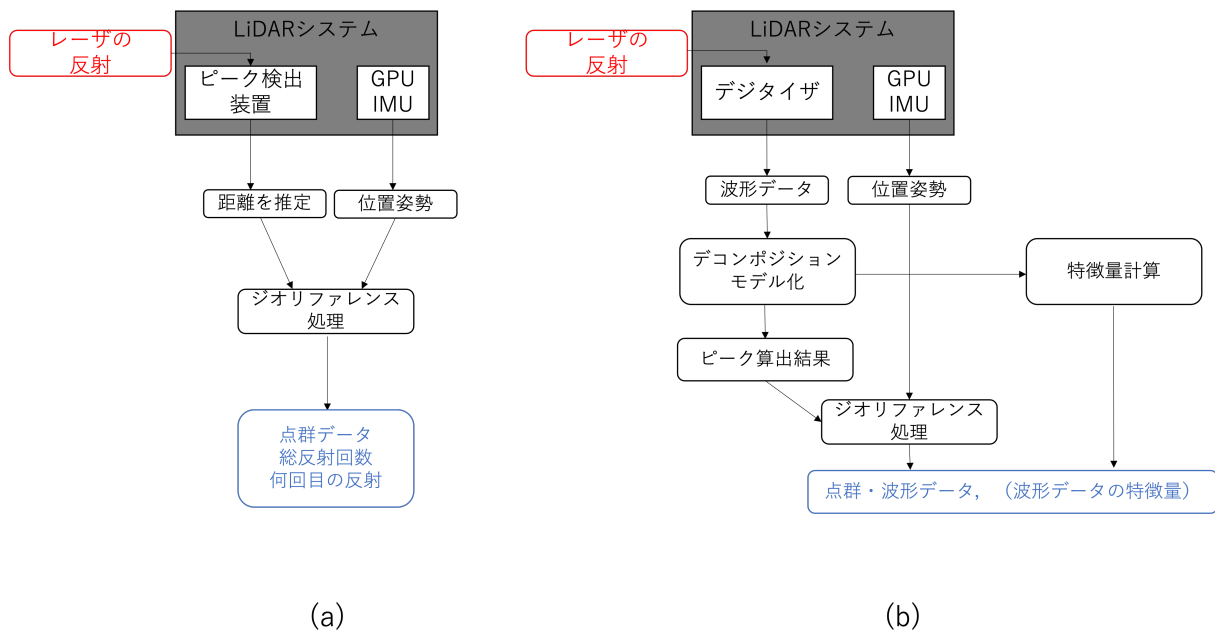


図 2.4: 疎な点群を記録するライダーと点群および波形情報を記録するライダーの点群データの作成方法比較. (a) 疎な点群を記録するライダーの場合, (b) 点群および波形情報を記録するライダーの場合. 疎な点群を記録するライダーでは観測時に直接ピークをハードウェアで検出しその距離を算出する. 点群および波形情報を記録するライダーでは, 観測された波形データから分解 (デコンポジション) によるピーク算出とモデル化を実行し, その値を元に点群データを作成する.

場合, 物体に当たり反射してライダーへ帰った波形データに対して波形データ分解 (図 2.5) を実行して, 記録された波形データに対するその形状の影響を取り除き点群データを作成することが可能である. しかし, この問題は, 制約となる予備知識がない場合には逆問題となることが大きな課題である. そのため, Fourier-quotient 法などの標準的な信号処理手法では満足いく波形データ分解結果が得られない. 以降, この問題を解決するために提案されてきた波形データ分解手法を紹介する.

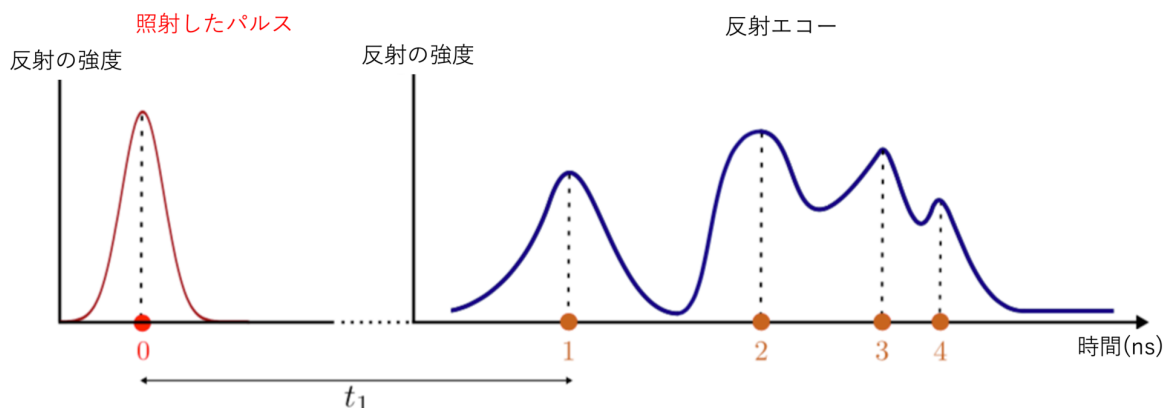


図 2.5: 波形データ分解の概念図. 波形データ分解では各エコーに分解し, ピークの位置を算出する. 図の例の場合では, 検出されたピークは4つである. 図は [145] の Figure 3.1 を参考に作図.

- ウィナーフィルタリング

[95] は、パルスの伝搬と照射面との相互作用の物理的な理解に基づいて、非常に近接した範囲 ( $< 0.15\text{m}$ ) にある異なる反射波を識別するためのアルゴリズムを提案した。まず、後方散乱されたパルスの受信波形データをライダー方程式を用いて計算する。これは、送信パルス (ガウスノイズでランダムに変調されたガウス関数でモデル化されている)、送信パルスの空間エネルギー分布 (これはレーザ装置をモデル化している)、表面反射 (与えられた反射率と速度を持つ) に依存する。その他に、表面の応答 (反射率と幾何学的特性、ここでは2つの異なる高さの鏡面とする)、大気の透過率、受信機の効率などがある。

エコーはまず、ノイズに依存した閾値を用いて受信波形データ内で検出される。次に、送信パルスと受信波形データのフーリエ領域での波形データ分解を行い、ウィナー・フィルターを用いて表面関数を推定する。なお、ウィナー・フィルターは、変調された送信波形データと背景ノイズから推定される実関数で、ここでは主に周波数領域でのノイズ除去に用いられている。最後に、Levenberg-Marquardt法を用いて、表面関数をモデル化関数として用いて波形データのフィッティングを行う。さまざまな種類の表面を対象とした実験によると、 $0.15\text{ m}$  以下の距離に相当する表面を分離できた。ウィナー・フィルターに代わるものとして、リチャードソン・ルーシー・アルゴリズム [160] が考えられた。リチャードソン・ルーシー・アルゴリズム [160] の論文では、ウィナー・フィルターは収束が遅く、計算量も多く、反復処理中にノイズを増幅する可能性がある。

- EM アルゴリズム

[167, 168] では、[58] で採用された Expectation-Maximization (EM) アルゴリズムベースの手法が採用されている。この問題は、ベイズ推定問題として定式化され、与えられたエネルギー関数の最大事後 (MAP) を推定する。この関数は、波形データに関する事前の知識を含んでおり、波形データのノイズに関する仮定 (ウィナー・フィルター手法のように明示的または暗黙的に) を必要としない。反復 EM アルゴリズムを用いて、MAP 推定量の近似解を求める。この手法の大きな特徴は、検出されるピークの数が多いことと、誤検出の確率が高いこととの間のトレードオフを調整するパラメータが用意されている。この手法は、空港の障害物調査において、電柱やアンテナなどの物体の検出を改善するために開発されたものである。[264] では、EM アルゴリズムに使用する際のフィッティング手法を変更することで、上記の手法よりもピークの検出数を増加させた。

- 正則化付き最小二乗法

[227] は、時間領域での波形データ分解の手法を提案した。3つの変数 (送信・受信波形データ+後方散乱断面積) を区分けされた線形関数とみなす。そして、波形データ分解問題を最小二乗法による連立方程式の解とすることができる。さらに、ノイズが存在するため解が不安定になる可能性があり、正則化技術が必要となる。そこで、平滑性制約を最小二乗法の問題に適用する。この手法は効率的ではあるが、データ内のノイズレベルが既知である必要がある。合成シミュレーションと様々なライダーからの実際の波形データでの結果を示す。

- B スプライン

[183] では、B スプラインを用いた手法が提案されている。この方法では、初期化ステップや放出パルスと散乱体の形状を仮定することなく、線形最小二乗法で問題を解くことができる。実際、次数3のB スプラインは、連続的な曲率を持つ関数に相当する。送信波形データと受信波形データはB スプラインとしてモデル化される。そのため、後方散乱断面積もB スプライン曲線として推定することができる。B スプライン曲線を観測データにフィッティングすることは、線形最小二乗問題である。実際に実施した結果 [183] は、既知の断面を持つ合成データセットと実際の波形データの両方で良好な結果を示した。

次にエコー抽出時の工夫点について説明する。波形データ分解で得られる個々のエコーに対する検出性能を向上させるための、ハードウェア的な工夫点とソフトウェア的な工夫点が

ある。

- 相関を利用した手法

[208]では、パルスと連続波の両方の測距システムを用いてパルスを検出する。著者ら [208]は、単一の閾値を用いたピーク検出では、SNRに依存した一定の誤差が生じるが、放出されたパルスが既知の場合には、適応的な閾値を用いることができることを示した。両システムから送られてくるエコーの相関関係を測定することで、ノイズを大幅に減らし、ピーク検出を改善することができる。波形データ内のエコーの数と位置を推定するために、[184]は Average Square Difference Function (ASDF) 法を採用した。ASDFは、送信したレーザと記録された波形データとの相関関係に基づいて時間遅延を推定する手法である。2つの波形データのずれを表す関数が定義できる。シングルエコーはこの関数の極小点に位置するが、複数のエコーが現れることもあるので、局所的な極小点でも大局的な極小点でも対応できる。

- 弱いパルスの検出

[204]では、ライダーと物体の間で部分的にオクルージョンが発生した場合や表面の後方散乱特性が悪い物体に対応する弱いパルスを検出することができると示した。[204]は、連続した波形データの間近傍関係を確立することで、波形データを積み重ねる技術である。これにより、相互情報が蓄積され、「物体のグローバルな散乱」が得られる。したがって、標準的なアルゴリズムでは検出されない新しいエコーを特定できる。

また、[126]では、特にピークの重なりに着目してこの問題に取り組んだ。著者らは、非対称または非ガウスの挙動を示すパルス形状内のエコーを検出するアルゴリズムを開発した（例：地上の低層植生）。主に見えるピークの両側で、十分な振幅と距離を持つ変曲点を見つける。その後、ガウス関数を適用し、小さなピークを残差の最大値の中から反復的に抽出する。この手法により重なり合ったパルスが非常によく分解され、その結果、近い物体を区別できることが示された（この例では 0.6 m 未満）。

次に図 2.6に示す波形データのモデル化について説明する。波形データモデル化の利点は3つある。まず、上述の方法と同様に、弱いエコーが検出される。森林地帯では、例えば、最大2倍多くの弱いパルスを検出することができる [181],[27]。次に、波形データ処理を行うことで、複雑な表面でも物体の範囲判定が改善される [88]。森林地帯では、樹冠と地面の高さの両方の推定値が改善される可能性があるが [49]、この結果は計測高度に依存する [27]。また、都市部では、距離推定精度を向上させることができる [102]。最後に、エコーをモデル化することで、ルールベースアルゴリズムによる点群および波形情報を記録するライダーデータの分類を実施する際に有用な追加特徴量が得られる。

波形データ内のエコーをモデル化する際、現在用いられている手法では常にパラメトリックなアプローチが選択されている。信号内で検出された各ピークに対して、ある関数のパラメータが推定される。これらのパラメータは、物体の特性（形状や反射率）に関する追加情報を提供し、波形データ処理機能を拡張する。信号処理技術によって抽出された統計的要素は、有意なピークの数、ライダーまでの距離、およびそれらのモデル化関数のパラメータである。波形データのすべてのエコーをモデル化するために、常に1つの関数が使用される。波形データ  $y = f(x_i)$  を  $n$  個の成分の和に分解する場合を考える。

$$y_i = \sum_{k=1}^n \phi_k(x_i) + b \quad (2.12)$$

ここで、 $\phi_k$  はパラメータ  $\theta$  を持つ関数、 $f$  は波形データモデル ( $f = \sum_k \phi_k$ )、 $x_i | i = 1, \dots, N$  は  $N$  個の等間隔点列、 $y = y_i | i = 1, \dots, N$  はサンプリングされた波形データ、 $b$  は背景雑音である。

以降の節で紹介する点群および波形情報を記録するライダーの地物分類で紹介するように、モデル化で得られたパラメータが点群データの地物分類に使用される。与えられたパラ

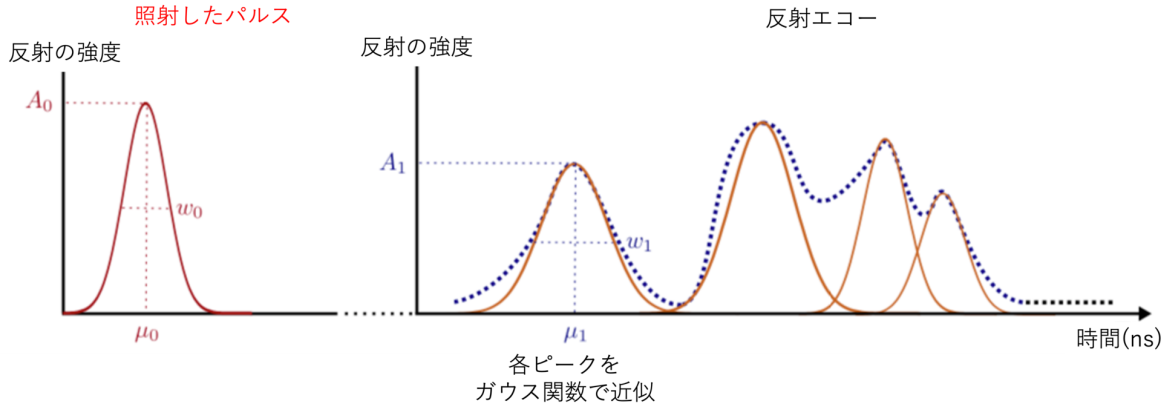


図 2.6: 波形データのモデル化の概念図. エコーをモデル化し詳細なパラメータを計算する. 図の例では, 青色で示す観測値の波形データに対して, 4つのピークそれぞれに対してガウス関数によるフィッティング (赤色) を行った結果である. 図は [145] の Figure 3.1 を参考に作図.

メトリックモデルを用いたエコーのフィッティングについては, 多くの文献で取り上げられている. 現在, 主に使用される2つのステップと関連する文献を以下に示す.

まず, ガウス関数による波形データのモデル化について説明する. 波形データとは, ガウス関数で表現可能な照射したレーザと, ガウス関数とみなされる「表面」散乱関数 (後方散乱断面積) との間の畳み込みであると定義できる [222]. そして, 受信信号は, ガウス関数の足し合わせであると仮定できる. その式は以下のように定義できる.

$$\phi_k(x) = A_k \exp\left(-\frac{(x - \mu_k)^2}{2w_k^2}\right) \quad (2.13)$$

ここで,  $\mu_k$  はパルスの位置,  $A_k$  はパルスの強度,  $w_k$  はパルスの幅を示す. さらに,  $\theta_k = \mu_k, A_k, w_k$  である.

ほとんどの用途において, 特に大きなフットプリントのライダーデータではガウス関数で十分である [88], [222]. しかし, 大気による散乱により波形データが変形した場合, ガウス関数は不正確であることが指摘されている [16]. その上, 小さなフットプリントの場合, このモデルは必ずしも現実に即してはいない. また, 都市部では, 多くのピークが歪んで観測される. 都市域においては, ほとんどの観測波形データは, 幾何学的特性 (屋根の傾斜など) とラジオメトリックな物体の特性 (異なる種類の道路や屋根の材質など) の混合した影響を受けている. そのため, 反射のピークの特性が大きく異なる場合が多数存在する.

次に, フィッティングについて説明する. 単一の関数で波形データを表現するために, 主に2つの方法が実施されている. 2つの手法は, 非線形最適化手法と Expectation-Maximization (EM) アルゴリズムによる最尤推定である.

- 非線形最小二乗法

[79] は, 海底地形 [230], 衛星 [16], 地上 [95], 航空機 [49], [181] に搭載された点群および波形情報を記録するライダーで使用されている非線形最小二乗法の一般的な説明をしている. この問題は,  $m \times n$  個の未知のパラメータを持つ  $N$  個の観測値である.  $m = \text{card}\theta$  はモデル化関数のパラメータの数,  $n$  はエコーの数である.  $n$  は, 標準的なピーク検出で述べた手法を用いて推定される. 結果の品質は, 変数  $\xi$  によって決定される. この処理では, データを所定の精度  $\epsilon$  でフィッティングさせる.

$$\xi = \sqrt{\frac{1}{N} \sum_{i=1}^N (f(x_i) - y_i)^2} < \epsilon \quad (2.14)$$

このフィッティングは、非線形最小二乗法である Levenberg-Marquardt (LM) 法 [149] を用いて解く。  $\phi_{kk}$  関数は線形ではないため、この問題を直接解くアルゴリズムは存在しない。 LM 法はロバスト性が高いことが知られているが、良い初期値が必要である。そこで初期値は、伝統的なパルス検出方法によって提供され、そのような方法では検出されない弱いパルスやオーバーラップするパルスを考慮して改良されることが多い。例えば、間違っただ初期化ステップの問題を解決するために、[79] では、 $\xi$  が所定の閾値よりも大きくなるまで、最小二乗法において、振幅に応じて（最高のピークから最低のピークへ）段階的にピークを追加する戦略を用いる。 [126] では、パラメータ推定値に制約を加えるために、非線形最適化手法として信頼領域法を使用することを提案した。

- 最尤法：EM アルゴリズム

[172] は、Expectation-Maximization(EM) アルゴリズム [38] に基づいたパルス検出法を開発した。 EM アルゴリズムは、最尤解を求めるための 2 段階の反復最適化手法である。 EM は、尤度の期待値を計算する期待 (E) ステップと、E ステップで求めた期待値を最大化してパラメータの最大尤度を計算する最大化 (M) ステップを交互に実行する。このステップでは、各サンプルについて、信号を分解する  $k$  個の分布のうちの 1 つに属する確率を計算する。 M ステップで求められたパラメータは、次の E ステップでも使用され、この M ステップと E ステップが繰り返される。これらの手法では、反射してライダーへ戻った観測波形データがガウス関数の和で表現可能であると仮定している。なお反射波形データに合わせて関数を変えることも可能であるが一般的にはガウス関数で統一することが多い。ここでは、エコーの数が事前にわからないため、モデル選択のステップが必要となる。

なお、上記の波形データ分解と波形データモデル化の出力は、 $p$  個の追加特徴を持つ点群データである。多くの研究例では  $p = 3$  とし、特徴量は一般化ガウスモデルのパラメータ  $A, w, \alpha$  が用いられている。

最後に、各ピークに対する位置情報の付与（ジオリファレンス）による点群データの作成について説明する。まず、照射したレーザに関する説明を行う。エコーのピーク位置  $\mu_i$  は、照射したレーザが物体から反射するまでの経過時間に相当する。この値から、送信されたパルス ( $T0$  パルス) の位置である  $\mu_0$  差し引く必要がある。  $T0$  パルスはわずかに形状が変化し、その振幅と幅は後の正規化に必要なので、一般化ガウス関数によるフィッティングも行う。波形データフィッティングとの主な違いは、 $T0$  パルスには見つけやすい 1 つのピークしか存在しないことが既知である点である。

次に、点への位置情報付与（ジオリファレンス処理）に関する説明を行う。地物の各座標は、図 2.7 のようにライダーの座標フレーム（レーザー座標系 (las)）において極座標で表される。極座標を地球楕円体の  $x, y, z$  座標にする、いわゆるジオリファレンス処理が行われる。ジオリファレンスされたデータの  $x, y, z$  座標を計算するために、図 2.7 に示すように航空機の GNSS(GPS) とライダーの位置・姿勢値を用いる。これらの位置姿勢情報によって、極座標の位置を任意の地図投影法（平面座標系）に変換することが可能になる。基本的な変換式を以下に示す。

$$c_P^{gc} = c_P^{gc} + \mathcal{R}_{gps-gc}(\mathcal{T}_{las-gps} + \mathcal{R}_{las-gps}c_P^{las}) \quad (2.15)$$

ここで  $c_P^{gc}$  は地球楕円体における点  $P$  の座標である。  $c_P^{gc}$  は地球楕円体における GPS 座標系である。  $\mathcal{R}_{gps-gc}$  は、GPS フレームから地球楕円の回転である。  $\mathcal{T}_{las-gps}$  は、レーザー座標系から GPS 座標系への並進である。  $c_P^{las}$  は、レーザー座標系におけるレーザー点の座標である。  $\mathcal{R}_{las-gps}$  は、レーザー座標系から GPS 座標系への回転である。なお変換式は、ライダーごとに異なるため、詳細に表現することはできないため、上記の式は簡易的な変換式である。また、オフセット値は観測機器によって異なり、ライダー機器、GNSS(GPS)、Inertial Measurement Unit(IMU) の位置関係によって決まる。

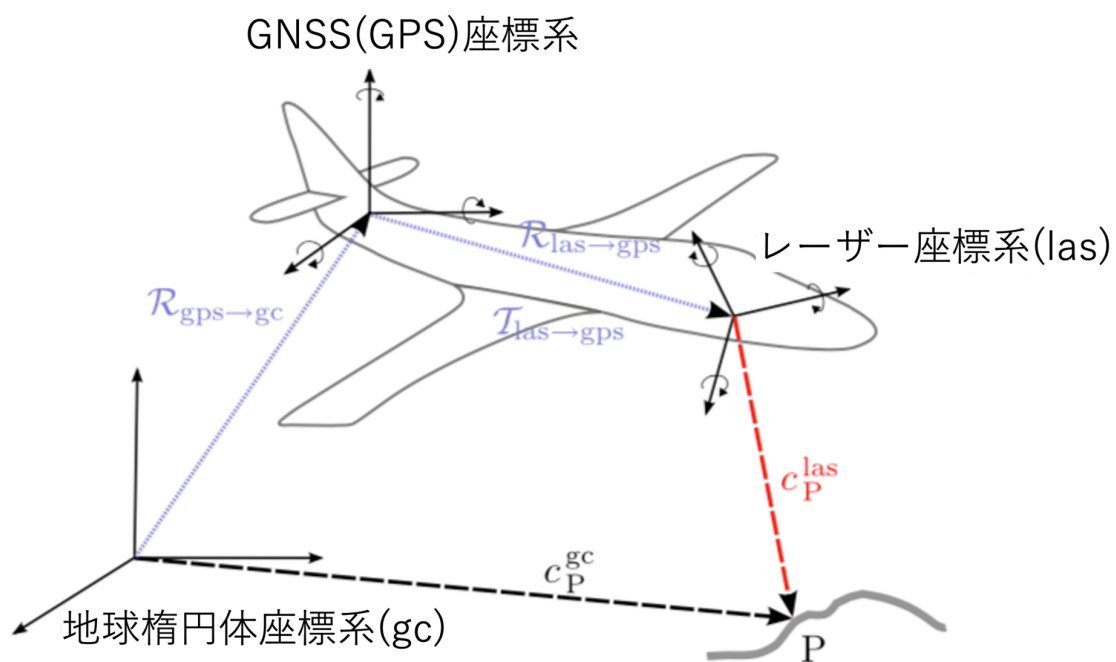


図 2.7: ジオリファレンス処理. 飛行機の位置姿勢を GPS/IMU で計算しているため, 波形データから求めた距離と照射した角度の情報から地球上の座標に変換できる. 図は [145] の Figure 3.5 を参考に作図.

## 2.3 深層学習手法

### 2.3.1 深層学習とは

深層学習はニューラルネットワーク (NN) を参考に発展した分野である。深層学習は、任意の活性化関数と学習可能な重みを持つニューロン (ユニット) で構成される NN を多層化したものである。一般的に深層学習手法は、入力データ (リモートセンシング分野では衛星画像や点群データなど) を出力データ (土地利用・土地被覆図・地物分類など) を予測する際に、入力データのエッジなどの簡単な特徴から、特徴と特徴の関係性などの抽象的な特徴を段階的に抽出する。この段階的な特徴の学習は、入力と出力の間のユニットの集合である層 (隠れ層と呼ばれる) で行われる。深層学習は、最も伝統的な形 (つまり 2 つ以上の隠れ層を持つ NN) をより多層にした形と定義できる。

このような深層学習の考え方は、生物学的なニューロンの結合をモデルとした人工知能 (AI) に基づくもので、1950 年代から存在している。その後、計算ハードウェアが進歩し、ラベル付けされたデータセットが大量に入手できるようになったことで、主要な科学分野でその関心が再び高まった。リモートセンシングの分野では、2010 年代半ばから深層学習が注目されているが、これは深層学習が画像地物分類タスクで大きな成功を取めたことに由来する。

リモートセンシング分野にて広く用いられる深層学習手法の多くは画像処理であり、教師付き学習と呼ばれる画像とその画像に対応する真値を用いた学習手法を行う。これらの入力画像における特徴を複数の演算を適用することで抽出し、さらにその特徴を組み合わせて、最終層で目的の予測値を返す。以上の説明では、深層学習以前の純粋な NN と変わらないが、その違いを強調すると、以下のように定義できる [116]。深層学習とは、複数の表現レベルを持つ表現学習である。表現学習とは深層学習に生データ (通常は画像、文章、ビデオなどの非構造化データ) を与えて、そのデータが持つ表現を深層学習モデル自身によって自動的に発見することである。また、深層学習が発展した経緯としては、「計算ハードウェアの性能向上」、「大量のデータ収集が可能になったこと」が挙げられる。

深層学習と従来型の機械学習のアルゴリズムは、どちらもある統計的なモデルを学習して数値や地物などを予測する方法である。従来型の機械学習では、まずデータから手作業でそのデータを端的に表す特徴を抽出する必要があり、その特徴を用いて入力データを説明または予測するモデルを学習する。一方、深層学習では、関連する特徴を抽出する手作業の工程を省略可能である。その代わりに、データは直接深層学習のアルゴリズムに与えられ、地物分類などを行う。つまり、深層学習は機械学習の一種ではあるが、データを直接扱い、より複雑なモデルであることが多い。従来の機械学習と深層学習のどちらを選択するかは、まず大量のラベル付きデータと高性能な計算機 (一般的には GPU) が利用できるかどうか重要である。実際には、深層学習を成功させるためには、信頼性の高い結果を得るために少なくとも数千の訓練データが必要であり、モデルの学習には強力な計算資源が必要であると言われている [62]。しかし、深層学習には精度の面で多くの利点があり、また、どの特徴が対象物の最適な表現であるかを人間が理解する必要がなく、直接深層学習モデルから最適な表現が抽出されるためである。したがって、従来型の機械学習と深層学習の選択は、データの量と解決しようとしている問題の複雑さに依存する。深層学習と従来型の機械学習の比較を表 2.3 にまとめる。

表 2.3: 深層学習と従来型の機械学習手法の比較

	深層学習	従来型の機械学習
訓練データ数	多い	少ない
人間による特徴量の作成	不必要	必要
学習可能な重み数	多い	少ない
訓練時間	長い	短い

最後に本研究で使用する教師付き学習による深層学習モデルの学習と予測について説明する。深層学習手法は、図 2.8 に示すように、入力データ (図 2.8 中では画像を例とする) と真値がペアとなった訓練データを利用して、深層学習モデルが持つ学習可能な重みを推定する。こ

の重みの推定には、予測結果と真値の差を損失関数として定義し、その損失関数を最小化するように最適化を実行する。最適化の実行後に得られた重みを持つ深層学習モデルを用いて、未知のデータに対して予測を行い出力を得る。詳しくは次節にて説明する。

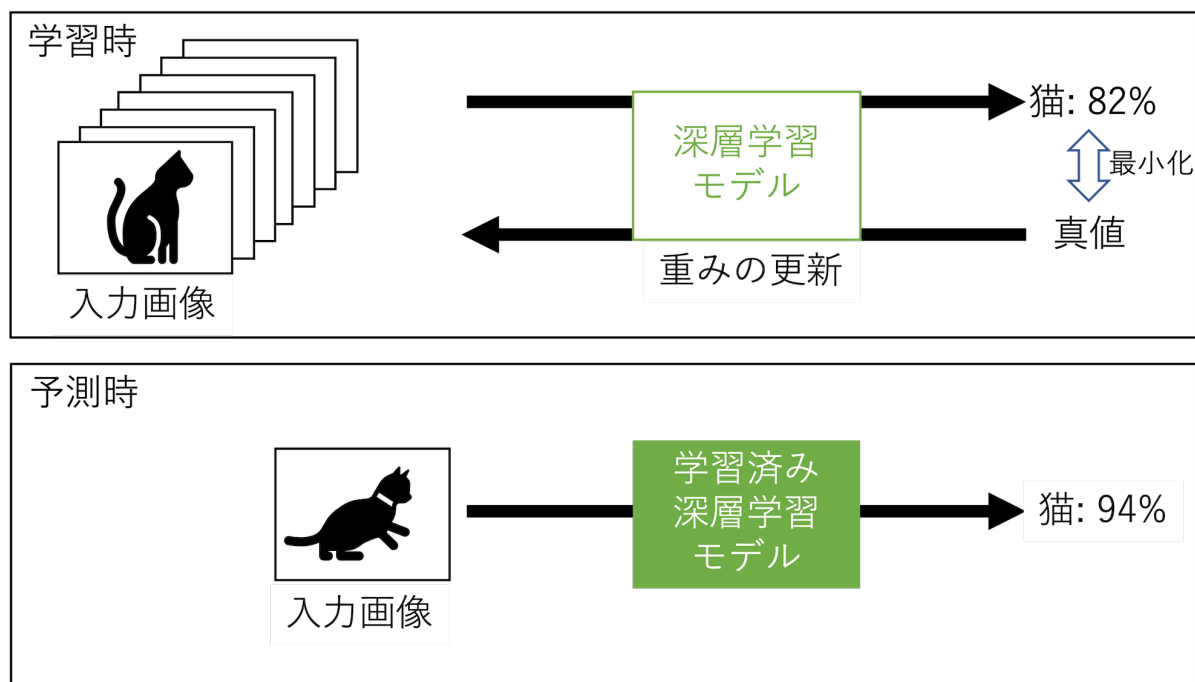


図 2.8: 深層学習モデルの学習と予測

### 2.3.2 一般的な深層学習手法の説明

まず、最も単純なニューラルネットワークであるパーセプトロンを説明する。図 2.9 に示すようにパーセプトロンは入力データの各値  $x$  と重み  $w$  の積和演算を行うニューロンで構成される。この積和演算は入力されたデータの全てに対して重みをかけて足し合わせる。そしてその後、活性化関数が適用される。活性化関数は、ニューロンが活性化され次の層へ送るかどうかを決定するために用いられる。これらの活性化関数は内在的なパターンの学習に寄与する重みを得る。つまり、各ニューロンが他のニューロンとどのような相互作用があるかを自動的に獲得することが深層学習の主要な考え方の 1 つである。一般的に適用される活性化関数には、線形、シグモイド、ハイパボリックタンジェント、Rectified Linear Unit (ReLU)、および ReLU の亜種である Leaky ReLU, Exponential Linear Unit (ELU), Parametric Rectified Linear Unit (PReLU) などがある。ReLU は現在では最も使用される活性化関数として知られており、図 2.10 に示すようにある変数  $x$  のすべての負の値に対して 0 の値を定義している。ReLU が用いられる理由には複数個存在する。この ReLU は、他の活性化関数と比較して計算コストがあまり高くないこと、消失勾配問題をうまく扱うことができること、データをより疎に表現することができることなどである。勾配消失問題とは、深層学習モデルの最適化を行う際に使用する勾配情報が入力層に近い層では消えることであり、ニューラルネットワークの普遍的な解決課題である。次にマルチレイヤーパーセプトロンについて説明する。図 2.11 に示すようにマルチレイヤーパーセプトロンはパーセプトロンを隠れ層にてニューロンを多数作成し、さらに複数層適用する。図 2.11 の隠れ層における各青丸印がパーセプトロンに相当する。このように、ある隠れ層のニューロンが、一つ前の層のニューロン全てと結びつく層を全結合層と呼ぶ。なお、このような層を深くつなぎ合わせることを深層学習の「深層」と呼ばれる所以である。深層学習と呼ばれる以前では多くても 3 層であったが、現在では計算機性能の向上や工夫点によって多層化が可能になった。

活性化関数の他に、深層学習が従来の NN と比較して高性能になる理由を説明する。それ

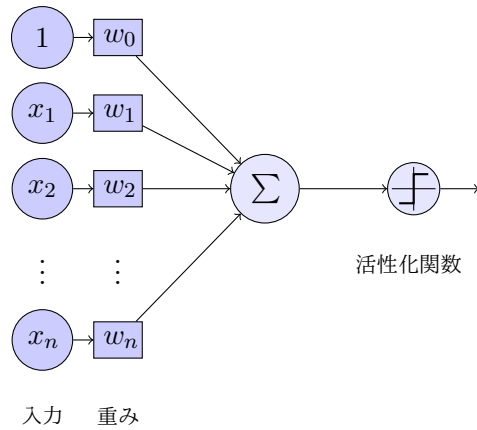


図 2.9: パーセプトロン. 入力データに重みをかけ, 足し合わせた後に活性化関数を適用する.

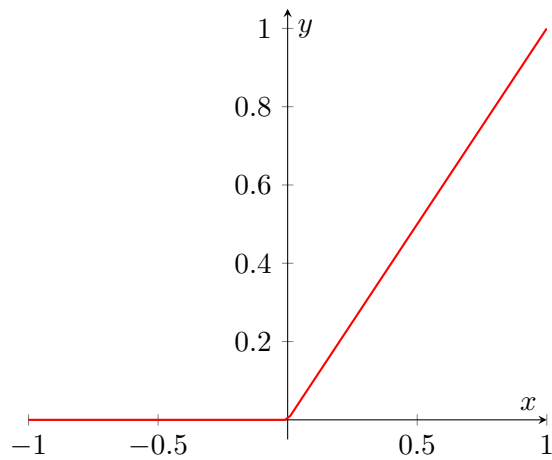


図 2.10: Rectified Linear Unit (ReLU)

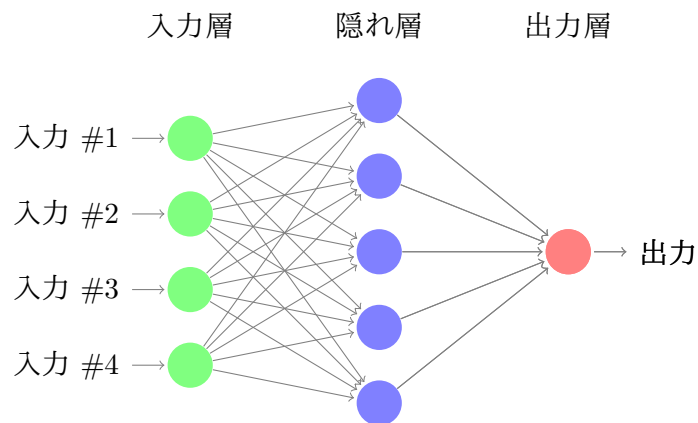


図 2.11: マルチレイヤーパーセプトロン. パーセプトロンの積み重ねで成り立ち, あるニューロンの出力に対して一つ前の層のすべてのニューロンが結びついた状態を全結合と呼ぶ.

は, ドロップアウト層, バッチ正規化層, 畳み込み演算, Max Pooling などの演算や, エンコーダ-デコーダ構造といった層設計の工夫に関連している. 畳み込み演算と Max Pooling, エンコーダ-デコーダ構造は後述するため, ここではドロップアウト層とバッチ正規化層を中心に説明する. ドロップアウト層は, ある層  $L$  のニューロンと層  $L+1$  のニューロン間の接続を一定の確率でランダムに選択して「ドロップ (削除)」するため, 深層学習モデルに正規化を

持たせるために重要な役割を果たす。これにより、訓練データへの過剰適合を低減し、汎化性能を向上させる。また、ドロップアウト層によって、多数のモデルを学習していることと同義であるため、予測時にはアンサンブル予測をしていることに等しい。バッチ正規化層は、最適化を行う際の勾配の流れを滑らかにすることで、汎化性能を向上させる。また、バッチ正規化層は、特徴量マップ内の共分散シフトの問題を解決するためによく使用される。

深層学習モデルは上記のように最適化によって得られる重みを持つ演算で構成される。したがって、最適化アルゴリズムの選択が重要になる。深層学習分野において最適化アルゴリズムとして最もよく使われる方法は誤差逆伝播に基づいた、Adam, モメンタムアルゴリズム, 確率的勾配降下法 (Stochastic Gradient Descent, SGD), Root Mean Squared Propagation (RMSprop) などである。深層学習モデルの最適化は非凸最適化であるため、局所解が存在する。そのため、深層学習モデルとその目的に応じて最適化アルゴリズムを適切に選択することが重要である。SGD は最もシンプルな最適化手法で、1 ステップごとに1つのサンプルを計算することで各層の重みを推定する。モメンタムアルゴリズムは、SGD に対し時間的な概念を加えることで、局所解に留まるといった問題を解決するものである。勾配ベースの最適化技術である RMSprop は、モメンタムと Adaptive Gradient Algorithm (AdaGrad) と呼ばれる別のアルゴリズムの両方を組み合わせて、勾配の指数関数的に減衰する平均値を計算する。Adam は現在最も使用されている最適化アルゴリズムであるが、普及した理由はモメンタムと適応学習率の両方を使用できることに起因する。次に具体的な最適化手法について述べる。深層学習モデルを用いて予測を行う場合、入力  $x$  を起点として中間層の出力  $y$  を経て  $z$  を得る。

この計算は、順伝搬と呼ばれ、情報が入力層から隠れ層を通して出力層に流れる。最適化の過程において、学習可能な重み  $\theta$  を持つ深層学習モデルにおける順方向計算は誤差  $J(\theta)$  を生成する。一方、逆伝播では、誤差からの情報を第1層まで遡ってモデルに取り込み、各重みを更新するのに必要な勾配を単純に計算する。次に一般的な関数  $f$  の勾配  $\nabla \times f(x, y)$  を推定する方法を説明する。 $y$  は関数の出力となる変数のベクトルで、 $x$  は微分が必要な入力データである。深層学習では、重みとバイアス  $\nabla_{\theta} J(\theta)$  に関する損失関数の勾配を計算することが求められる。ここで説明する方法は、単純のため関数の出力がスカラーである場合に限定されるが、複数の出力を持つ関数にも容易に拡張できる。逆伝播は、微積分の連鎖法則を用いて勾配を求めるアルゴリズムである。連鎖律は、導関数が既知の他の関数の合成によって形成される関数の導関数を計算するために利用される。 $x$  はスカラー、 $f$  と  $g$  はともに実数を入力し実数を返す関数、 $y = g(x)$ ,  $z = f(g(x)) = f(y)$  とすると、連鎖律では次のように書ける。

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} \frac{\partial y}{\partial x} = \frac{\partial f(y)}{\partial x} \frac{\partial g(x)}{\partial x} \quad (2.16)$$

$x$  と  $y$  がベクトル ( $y = g(x)$ ,  $z = f(y)$ ) の場合に一般化すると、次のようになる。

$$\nabla_x z = \left(\frac{\partial y}{\partial z}\right)^T \nabla_y z \quad (2.17)$$

ここで、 $\frac{\partial y}{\partial z}$  は  $g$  のヤコビ行列である。勾配は、ヤコビ行列に勾配を乗じることで計算できる。逆伝播アルゴリズムは、この積を計算グラフの各演算に対して実行する。

次に、最適化の際に使用されるミニバッチについて説明する。勾配を正確に計算するには、訓練データセット全体のすべてのサンプルに対して深層学習モデルを評価する必要があるため、計算コストがかかる。しかし、訓練データセットから少数の例をランダムにサンプリングし、それらの例を平均して期待値を算出することで、期待される勾配を計算することができる。この方法の利点は、すべての最適化アルゴリズムが、正確な勾配をゆっくりと計算するのではなく、その推定値を用いて頻繁に勾配を更新することで、より速く収束することである。最適化アルゴリズムは、データセットの使用目的に応じて、3つのグループに大別される。

- すべての学習サンプルを利用して正確な勾配を計算する最適化アルゴリズムは、バッチ法または決定論的勾配法と呼ばれる。これらの手法では、データセット全体を同時に処理する。
- 一方、一度に1つのサンプルに対して勾配を計算する最適化アルゴリズムは、確率的またはオンライン法と呼ばれる。

- また、1つ以上の学習サンプルを利用する最適化アルゴリズムをミニバッチ法またはミニバッチ法と呼ぶ。ミニバッチ法は、深層学習モデルで最も一般的に使用される最適化手法である。

バッチサイズを選択は、通常これらの要素に影響される。

- 勾配のより正確な推定値を得るためには、大きなバッチサイズに頼る必要がある。
- GPU のアーキテクチャは、非常に小さなバッチサイズでは十分に活用できない。
- 一般的には、バッチ内のすべてのサンプルが並行して処理される。この場合、タスクに必要なメモリの量は、バッチサイズに比例する。
- ハードウェアの種類によっては高速に学習する条件もある。例えば、GPU を使用する場合、バッチサイズとして2の累乗を使用するのが一般的である。

ミニバッチから期待される勾配の推定値を計算するためには、それらのサンプルが独立していることが必要である。つまり、ミニバッチは無作為に選択されなければならない。この操作は、データセット全体を入れ替えてからミニバッチに分割することで簡単に行うことができる。

ここで、勾配を用いた学習を行うために最も一般的な最適化手法である図 2.12 に示すような確率的勾配降下法 (Stochastic Gradient Descent, SGD) について詳細に説明する。確率的勾

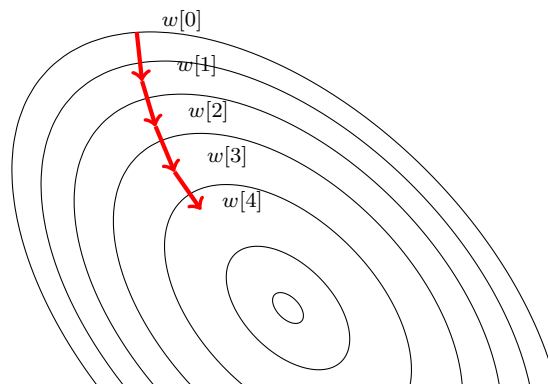


図 2.12: 二変数の場合の確率的勾配降下法の例。初期値の重み  $w[0]$  から徐々に勾配情報を使用して重みを更新する。

配降下法 (SGD) は、学習の反復ごとに重み  $w$  を更新し、以下の手順を利用する。なお、初期の重みを  $w[0]$  とする。まず、訓練データセットから  $m$  個の入力データ (画像であれば  $m$  枚の画像) とそれに対応する真値を集めたミニバッチを用意する。次に、勾配推定値を計算する。最後に重みを更新し、 $w[1]$  にする。これを繰り返すことで、 $w[0], w[1], w[2], w[3], w[4]$  のように徐々に重みを更新していく。SGD アルゴリズムで最も重要なハイパーパラメータは重みの更新量を決める学習率  $\epsilon$  であり、通常、学習エポックの間に学習率を徐々に下げていく必要がある。一般的には、次のようなルールで、反復回数  $\tau$  まで学習率を直線的に減少させる。

$$\epsilon_k = (1 - \alpha)\epsilon_0 + \alpha\epsilon_\tau \quad (2.18)$$

ここで、 $\epsilon_k$  は反復  $k$  における学習率を表し、 $\alpha = k/\tau$  であり、反復回数 (訓練データを何周するか) の指標であるエポックが用いられる)  $\tau$  の後は  $\epsilon$  を一定にするのが一般的である。学習率の値は、時間の関数としての学習曲線を監視しながら、試行錯誤によって選択される。主な問題点は、 $\epsilon_0$  をどのように設定するか依存する点である。 $\epsilon_0$  が大きすぎると、損失関数が大幅に増加し、学習曲線が激しく振動する可能性がある。また、学習率が低すぎると、学習手順の進行が遅くなる。他にも、一般的な説明で述べたように、Adam など最適化中に学習率を

自動的に調整する最適化アルゴリズムは数多く存在する。なお、深層学習モデルの重みを得る最適化においては、非凸の多変数最適化であるため、一般的に局所解を持つ。この最適化における非凸性は活性化関数の性質と層設計における多層性の二つが要因である。一方、訓練データを用いた最適化によって得られる局所最適解は最適解に近いと予想されている [136]。

次に最適化に使用される損失関数を説明する。深層学習モデルの最適化においては、最適化アルゴリズムだけでなく、モデルの性能を評価するために定義された関数である損失関数(目的関数またはコスト関数とも呼ばれる)が重要である。損失関数は、モデルに入力されるデータを単一のスカラー値へ写像するモデルの能力を表す。最適化問題はこの損失関数を最小化するように深層学習モデルの重みを調整する方法を見つけることに帰着する。また、深層学習モデルが地物分類問題を扱っているのか、回帰問題を扱っているのかなどのタスクにも損失関数は依存する。そのため、深層学習モデルが出力する値の性質そのものに関係していることから、適切に損失関数を定義することが深層学習モデルの性能上重要である。深層学習モデルの出力の例として、地物分類の場合を説明する。二つの地物分類の値を予測する必要があるタスクは多く、例えば、2つの地物を区別するために入力データにラベルを割り当てる必要がある場合などである。例えば、入力画像が建物であるかどうかを認識するように学習された深層学習モデルがあるとすると、この場合、出力値はそれぞれのクラスに属する確率を予測する。つまり、この例では、与えられた入力画像の中の建物が存在する確率のみを予測する。これが有効な確率分布であるためには、この数値が区間  $[0,1]$  に存在しなければならない。モデルの出力に関わらず、区間  $[0,1]$  へ変換するために図 2.13に示すシグモイド関数を使用される。ここで、入力画像内に存在する物体を地物分類するようにネットワークを最

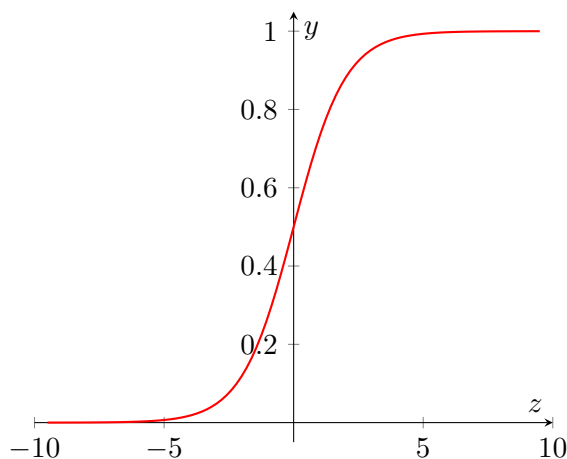


図 2.13: シグモイド関数

適化する場合を考える。この場合、ネットワークの課題は離散変数の確率分布を表現することである。すなわち、望ましい出力は、各地物に属する確率であり、0-1の値域であることが期待される。その際に使用されるのがソフトマックス関数である。

$$\text{softmax}(\mathbf{z}) = \frac{\exp(z_i)}{\sum_j^n \exp(z_j)}, \quad i = 1, \dots, n \quad (2.19)$$

なお、地物分類のタスクを解決するためには、クロスエントロピーや Kullback-Leibler(KL) ダイバージェンスなどの損失関数を使用することも可能である。また、連続値の比較などの回帰関連のタスクでは、平均二乗誤差 (MSE)、平均絶対誤差 (MAE)、平均絶対率誤差 (MAPE)、平均二乗対数誤差 (MSLE) などに基づく損失関数が一般的に実装されている。これらの損失関数についての詳細は、Goodfellow[62]の資料に記載されている。

次に、学習済みモデルの評価方法を説明する。学習済みモデルの性能を評価するために、様々な定量指標が採用されている。地物分類タスクでは、正答率や Recall, Precision が一般的に使用される指標だが F1 指標, Receiver Operating Characteristics(ROC) 曲線下の面積、

Intersection over Union(IoU)などの指標を比較することもある。回帰関連のタスクでは、MSE, MAE, Mean Relative Error (MRE), Root Mean Squared Error (RMSE), 相関係数などの指標も使用される。これらの指標は、モデルの予測値とラベル付けされた真値（リモートセンシング分野では Ground Truth とも呼ばれる）との関係を定義するために重要である。特に、手法の提案を行う際には、先行研究とデータセットに対してこれらの指標を適用することで既存手法と提案手法間の比較を客観的に実行可能である。

上記のように、深層学習モデルを改良するために、さまざまな種類の層や演算、最適化アルゴリズム、損失関数などが提案された。しかし、今日の深層学習の普及の背景には、深層学習モデルの学習に利用できるデータ量の増加も関係している。経験則では、1つのカテゴリーにつき最低 5,000 個の真値付きのデータセットが推奨されていた [62]。リモートセンシング分野においても同様に、様々なデータセットが公開 [271] されつつあり、深層学習モデルを利用した研究が加速している。

### 2.3.3 畳み込みニューラルネットワーク

現在広く用いられている深層学習手法は畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) である。CNN は、近年のコンピュータビジョン分野で普及した深層学習の考え方である。CNN は 1998 年に LeCun によって初めて提案され、現在では多くのコンピュータビジョンタスクに対する解決策となっている。CNN は特に、規則正しいグリッド状の形状を持つデータの処理に適している。通常、CNN で処理されるデータの例は以下の通りである。

- 1D データ：音声信号などの時系列データで、各要素が特定の時間間隔を持つ信号値に対応している。図 2.14 のように 1 次元のフィルタを適用する。
- 2D データ：2 次元の格子状の画素で表現される画像データ。RGB 画像は、各色（赤、緑、青）ごとに 3 つのチャンネルで構成される。図 2.15 のように 2 次元のフィルタを適用する。また、RGB だけでなく、RGB 近赤外やハイパースペクトルのように多チャンネルにすることも可能である。

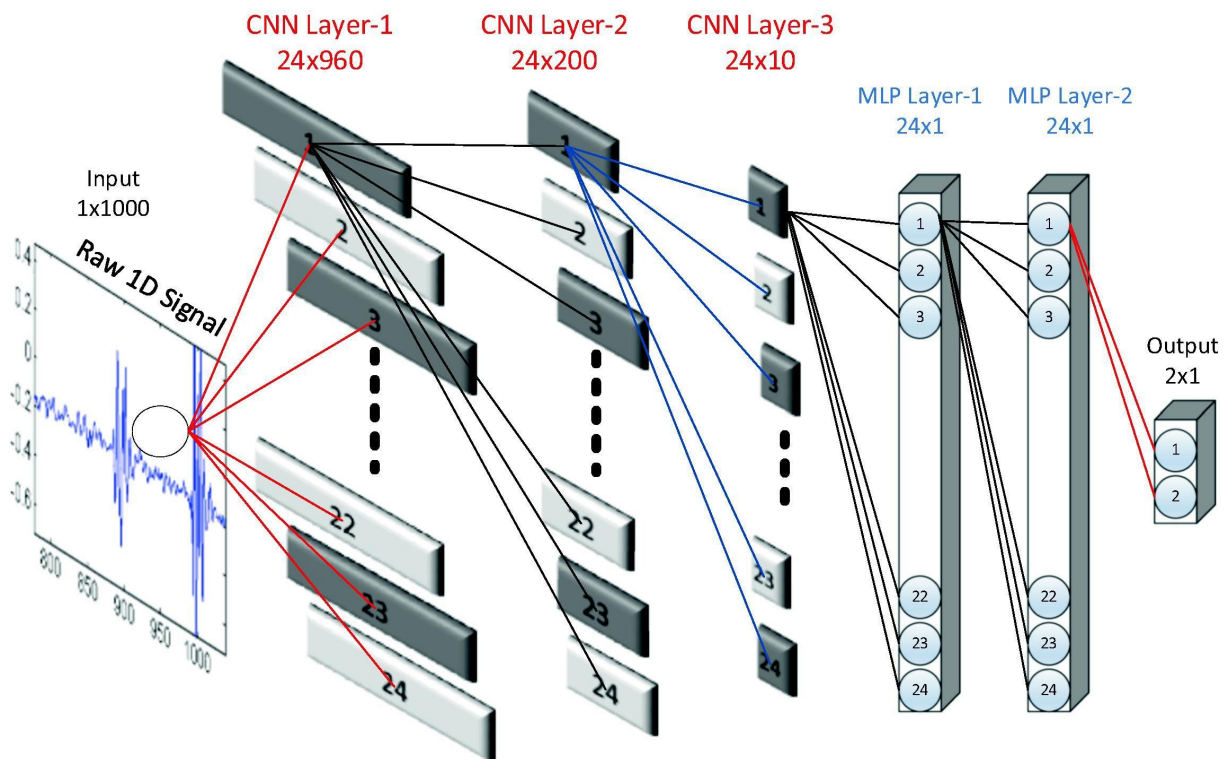


図 2.14: 1D CNN による時系列信号の処理。図は [101] の Fig.2 から引用。

CNN の畳み込みという名称は、数学的演算の畳み込みに由来する。畳み込みとは、2 つの関数の間の演算であり、次のように定義できる。

$$s(b) = \int_{-\infty}^{+\infty} x(a)w(b-a)da \quad (2.20)$$

この操作は通常、 $s(b) = (x w)(b)$  と表され、第一の関数  $x(a)$  は入力、第二の関数  $w(b-a)$  はカーネルと呼ばれることが多い。また、畳み込み演算による出力は特徴マップと呼ばれることもある。コンピュータビジョン分野でのアルゴリズムでは、一度に複数の軸の畳み込みを行うことが多い。例えば、画像を処理するモデルの場合、畳み込みカーネルは、行と列に沿ってスライドさせる必要がある。2 次元の画像  $I$  を入力とし、2 次元のカーネル  $K$  を与える

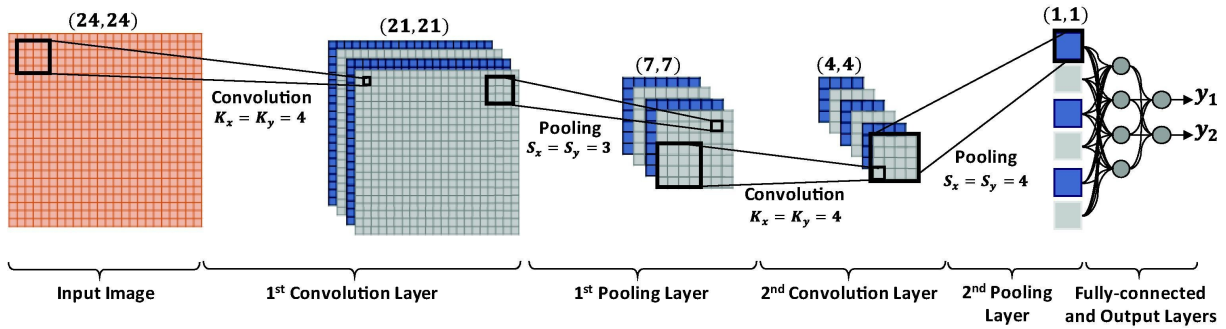


図 2.15: 2D CNN による画像データの処理. 図は [101] の Fig.5 から引用.

と、畳み込み演算は次のように定義できる.

$$S(i, j) = (K \star I)(i, j) = \sum_m \sum_n I(i - m, j - n) K(m, n), \quad (2.21)$$

図 2.16は画像を入力した際の畳み込みの例である. 入力画像の画素  $(i_1, i_2, i_5, i_6)$  に対して, 重み  $k_1, k_2, k_3, k_4$  を持つ畳み込みフィルタ (カーネル) を適用する. この際の出力は,  $i_1 * k_1 + i_2 * k_2 + i_5 * k_3 + i_6 * k_4$  となる. この演算を, 畳み込みフィルタを適用する場所を変えながら繰り返し適用することで, 入力された画像に対する特徴が計算できる. CNN の成功は, 深層学習モデルを向上させることができる, 畳み込み演算に関連した 3 つの重要な考え (疎な相互作用, 重み共有, 等変量表現) によるものである.

次に, CNN と全結合型の従来のニューラルネットワークとの違いを説明する. マルチレイヤーパーセプトロンなどの全結合型のニューロンを持つ従来のニューラルネットワークは, 図 2.17の左図のように, ある層のニューロンの全てに対し重みを適用し次の層のあるニューロンへ接続される. 図 2.17の左図において, 赤丸で示すニューロンは青丸で示すニューロンに接続される状態である. 入力層の各ニューロンが隠れ層のニューロンに接続されている全結合型の従来のニューラルネットワークとは対照的に, CNN では入力層のニューロンのうち, 隠れ層のニューロンに接続できるのは小さな領域 (図 2.17の右図) だけである. 図 2.17の右図において, 赤丸のニューロンは, 青丸の三つだけに接続される. これらの領域は受容野と呼ばれている. 一回の畳み込み演算とプーリング処理では, ごく限られた狭い領域の受容野しか得られないが, これらの演算を繰り返し適用することで広い範囲の受容野を獲得する. 広い受容野は, ある画素を予測する際に使用すると特徴の範囲が広くなり, 広い範囲の画素と画素の関係性を用いて分類可能である. そのため, 主にセグメンテーションにおいて広い範囲の受容野の獲得は重要な概念になる. CNN においては, ある層において適用される畳み込みの重みは入力データ全体に同じ値が適用され, 入力層から隠れ層のニューロンへの特徴量を作成する. これは重み共有と呼ばれており, この特性は, メモリ消費量と計算時間を削減させる. つまり, 保存する重みの数が少なくなり, 出力を計算するために必要な演算の数も少なくなる. 図 2.17と図 2.18は, 疎な相互作用と局所的受容野の概念を示している. 図 2.17において, 左図が全結合であり, ある赤丸の値は次の層に影響を与えるのは青丸のすべてのニューロンである. 一方, CNN では, 右図に示す赤丸の値が次の層に影響を与えるのは青丸の限られた範囲のニューロンである. 図 2.18の破線の矩形の範囲内を  $2 \times 2$  でプーリングを適用した例である.  $2 \times 2$  の範囲内の値で, 代表値がひとつだけ次の層へ送られる.

次に, CNN で獲得される特徴量について説明する. CNN における重みの共有という考え方は, 重み行列の各要素が入力の 1 つの要素にのみ適用される全結合層のニューラルネットワークとは対照的に, 畳み込みの各要素が入力のすべての位置に適用される. これにより, 入力位置ごとに個別の重みを用いるのではなく, 各畳み込み層ごとに少ない重みの数を用いる. 重みを画像全体に共有することで, フィルタは入力の異なる領域で同じパターンを探ることができるようになり, 「並進に対する等変性」と呼ばれる特性が得られる. 例えば, 画像を畳み込み層で処理した場合, 結果として出力されるのは, 入力サンプル中の特定の特徴が現れる場所を記述した特徴マップである. 入力で示されたオブジェクトが移動した場合, 出力表

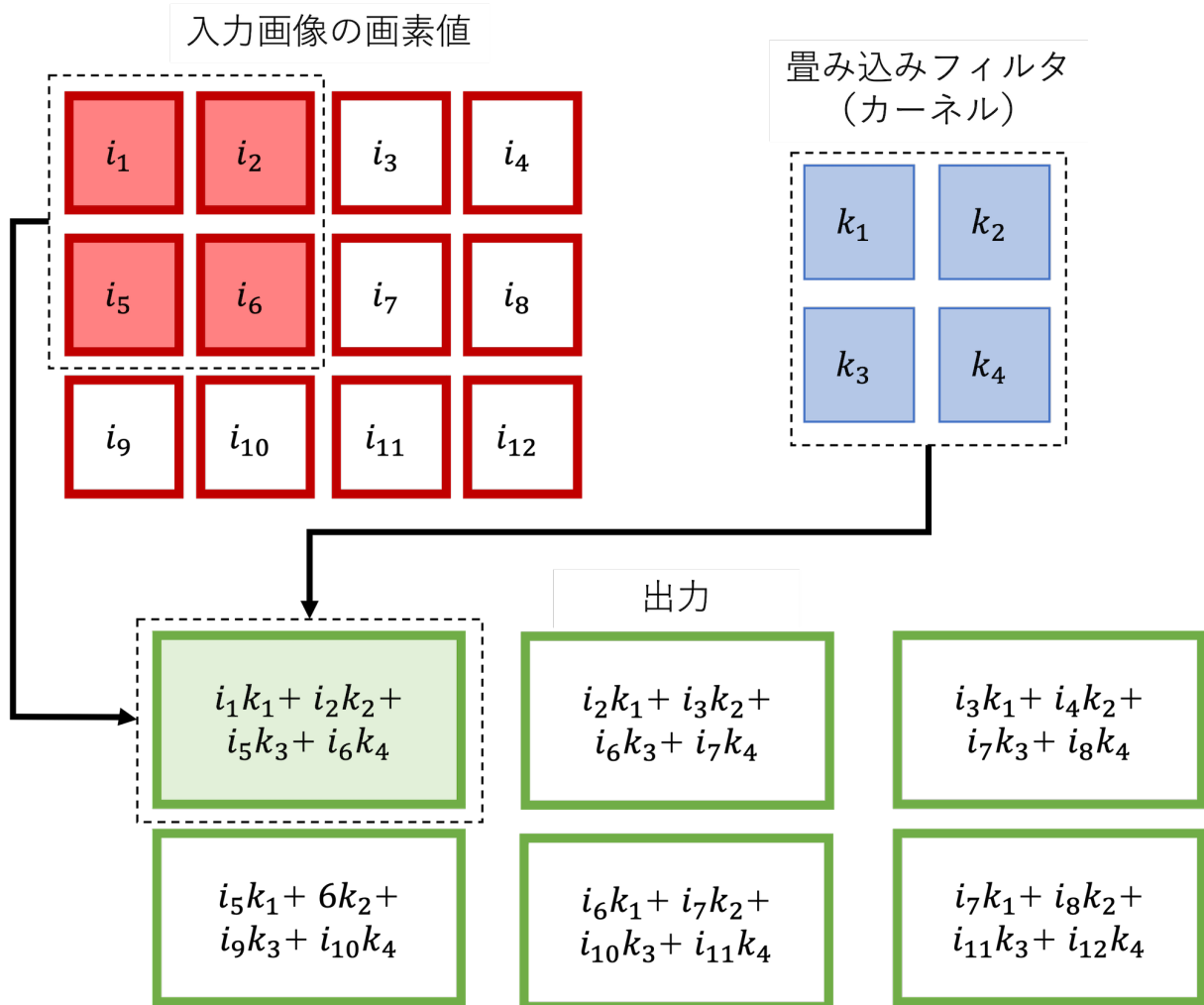


図 2.16: 画像の畳み込み演算の概念図。画像の畳み込み演算はある画素に対して、一定サイズのフィルタを適用する。このフィルタは最適化によって得られる重みで構成される。

現は同じ量の画素だけ移動する。これは、画像のいたるところに現れる可能性のある同じエッジ、コーナー、パターンを検出する場合に役立つ。このような畳み込み演算では、入力画素の近傍を考慮するため、フィルタのサイズに応じて異なるレベルの空間的自己相関関係を探索できる。

次に、畳み込み層の構築方法について説明する。一般的に、畳み込み層は3つの要素で構成されている。最初の段階では、層は複数の畳み込みを並行して行い、入力の線形表現を生成する。続いて、非線形活性化関数が実行される。最後に、プーリングが実行され、ある位置にある畳み込み層の出力を、隣接する出力の統計値に置き換える（図を参照）。なお、プーリングは、畳み込み層の出力表現を、入力の小さな変換に対して不変なものにするのに役立つ。プーリングの典型的な例としては、矩形内の最大値を保持する Max Pooling がある。プーリングは、ニューロンの小さな領域の出力を1つの出力に集約することで、特徴マップの次元を減らす。これにより、後続の層が単純化され、深層学習モデルの重みの数が減るだけでなく、広い範囲の特徴が次の層へ送られる。プーリングは、さまざまなサイズの入力を扱う際にも不可欠な処理である。例えば、様々なサイズの画像を地物分類するためには、最終的な地物分類を行う分類層のサイズを一定にする必要がある。これは、プーリング領域間のオフセットのサイズを変更することで実現でき、画像の寸法に関わらず、分類層が常に同じ数の総和特徴を受け取るようになる。要約すると、数十から数百の隠れ層で畳み込みとプーリングを組み合わせることで、CNN は画像のさまざまな特徴を検出することを学習できる。隠れ

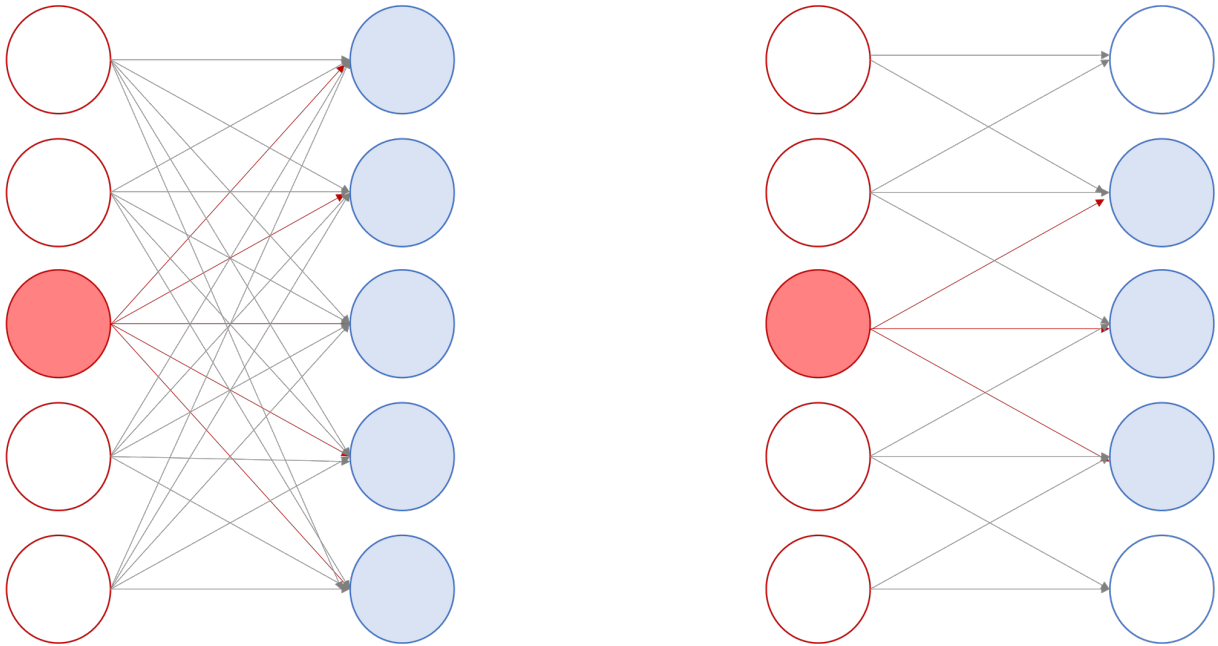


図 2.17: 全結合層と CNN の違い. 全結合はすべてのニューロンが次の層へ繋がる. しかし, CNN の場合は, ある一定範囲のニューロンが次の層へ繋がる.

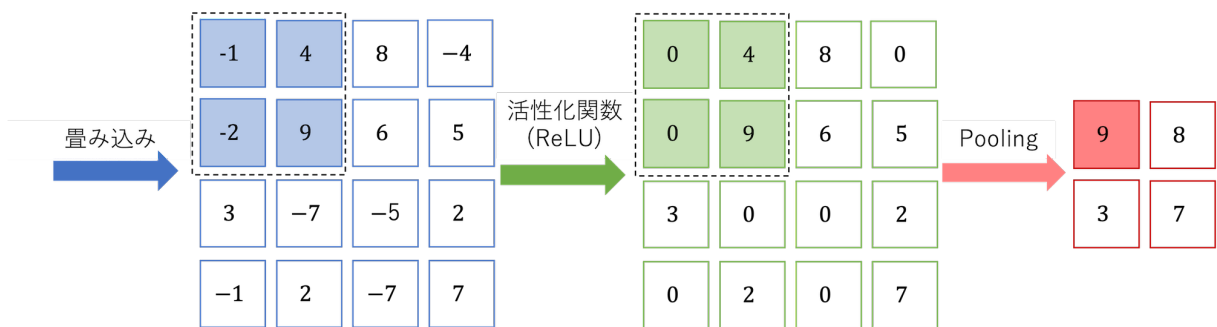


図 2.18: 畳み込みとプーリングによる局所受容野. 局所受容野は畳み込みとプーリングによって獲得できる特徴がどれくらいの範囲から得られたかを示す.

層が増えるごとに, 学習する画像特徴の複雑さが増していく. たとえば, 最初の隠れ層はエッジの検出方法を学習し, 最後の隠れ層はより複雑な形状の検出方法を学習する. 最後の層では, 全結合層と同様に, 最後の隠れ層からのすべてのニューロンを出力ニューロンに接続して, 最終的な出力を生成する.

次に, CNN の最適化方法について説明する. CNN を最適化するには 3 つの方法がある. 1 つ目は, CNN をランダムな初期値から最適化する方法である. この方法は分類精度が高いが, 何十万ものラベル付けされた画像と膨大な計算資源が必要となるため, 最も困難な方法でもある. 2 つ目の方法は転移学習に頼るもので, ある種の問題の知識を使って似たような問題を解決することができるという考えに基づいている. 例えば, 自動車とトラックを区別する新しいモデルを最適化する際に, 動物を認識するように訓練された CNN の重みを初期値として最適化できる. この方法では, 前述の方法よりも少ないデータと少ない計算資源で済む. 3 つ目の方法は, 事前に学習した CNN を使って, 別の機械学習モデルを最適化するための特徴量を抽出する方法である. たとえば, 画像のエッジを検出する方法を学習した隠れ層は, さまざまなドメインの画像に関連していることが期待される.

この CNN を用いて, 画像分類コンペティションで 2012 年に Krizhevsky ら [107] が従来の

人間の手で設計する特徴量を画像分類モデルに入力する大差をつけて優勝した手法を発表したことで、画像処理用の CNN に関心が集まった。AlexNet として知られるようになったこのネットワークは、8 層で構成されている。最初の 5 層はすべて畳み込み層で、いくつかの層の後には Max Pooling 層が続き、最後に 3 つの全結合層がある。これらはすべての層において ReLU 活性化関数を使用している。このコンペティションで優勝した手法は、現在ではシンプルな深層学習モデルと考えられているが、その画像分類性能は層の深さに関係している。

次に、リモートセンシング分野に対する CNN の応用について説明する。CNN は元々、複数の配列でデータを処理するように設計されており、この特性は、画像の画素が規則的に配列されているためマルチバンドのリモートセンシング画像を扱うのに特に適している。その結果、CNN を用いた層設計は現在最も利用される深層学習モデルの一つとされており、リモートセンシング分野では画像を用いた応用において実証されている [265]。画像形式で表現されるリモートセンシングデータを深層学習ベースのアルゴリズムで処理した場合、シーン分類、セグメンテーション、物体検出のタスクに分けられる (図 2.19)。まず、シーン分類について説明する。リモートセンシングにおけるシーン分類とは、入力された画像に対して農業のシーン、海岸のシーン、都市のシーンなど、画像 (またはパッチ) にラベルやテーマを関連付ける手法を指す。このようなシーン分類が活躍する機会はリモートセンシング分野における応用では稀である。その代わりに、ほとんどの応用事例では、物体検出や画素単位の地物分類が可能なセグメンテーション手法の方が仕様される。シーン分類では、画像のクラスラベルのアノテーションのみが必要だが、物体検出のような他のタスクでは、画像内のすべての物体に対するの矩形を真値として作成する必要があり、データセットの構築にコストがかかる。また、セグメンテーションでは、真値の作成者が物体の各画素を含む領域をマスク画像としてを描く必要があり、真値作成作業にはより注意と精度が求められる。そのためデータセットの構築コストが莫大になる。これらのコストに対して、シーン分類タスクは真値作成のコストにおいては優位性があるため、実利用する際はコストと得られる情報量の多さを考えて実行する必要がある。

次に、セグメンテーションの説明を行う。セグメンテーションは一般的に画素単位の地物分類問題として定義されている [152]。セグメンテーション手法は大きく分けて、学習済み CNN で作成した特徴マップに対してアップサンプルを行う FCN 型と、学習済みモデルにとられず、畳み込み-Pooling を複数回階層的に行うエンコーダと、エンコーダによって作成された特徴マップから徐々に入力解像度までアップサンプリングするデコーダを有するエンコーダ-デコーダ型に分かれる。FCN 型は Fully Convolutional Network [135] によって提案され、自然画像に対するセグメンテーションや [127]、車載動画像に対するセグメンテーション [33] で効果を発揮した。エンコーダ-デコーダ型のネットワークとして U-net [185] が幅広く用いられ、リモートセンシング分野においも U-net を適用した例が多い [241]。U-net はエンコーダによって畳み込み-プーリングを繰り返し入力された画像の解像度から徐々に特徴マップの解像度を落とし広い範囲から情報を集約する。また、デコーダでは解像度の落ちた特徴マップから  $2 \times 2$  の学習可能なアップサンプリングやバイリニア補間によって解像度を上げる。その後、エンコーダ側で対応する解像度の特徴マップからスキップコネクションを用いて特徴マップをデコーダ側の特徴マップと結合する。そして、通常の畳み込みを適用することで特徴マップを新たに作成する。このようなセグメンテーション向けの層設計を用いて、リモートセンシング分野においては、画素単位の建物抽出 [266] や道路の抽出 [248] 等の特定クラスの抽出や、土地被覆地物分類 [256] などが行われている。また、リモートセンシング分野では、これらのタスクの研究を後押しするために公開データセットの整備も行われている [271]。

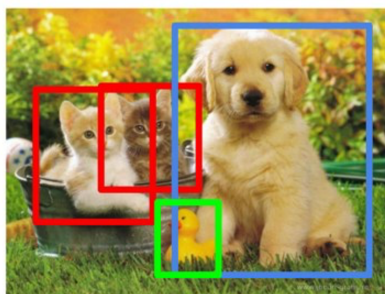
近年においては、リモートセンシング分野において、画像だけでなく点群データに対する深層学習手法も提案され実用化されつつある。この点群データに対する深層学習手法を次の節で説明を行う。

### シーン分類



画像に映る物体のクラス

### 物体検出



猫の位置  
犬の位置  
アヒルの位置

### セグメンテーション



猫の画素  
犬の画素  
アヒルの画素

図 2.19: 画像に対する深層学習手法を用いた主要なタスク. 入力された画像に何が写っているかを計算する「シーン分類」, 画像のどこに何が写っているかを計算する「物体検出」, 画像画素単位に分類を行う「セグメンテーション」がある. 図は [99] の Fig.1 より引用し加筆した.

### 2.3.4 点群データに対する深層学習手法

本研究では、コンピュータビジョン分野にて研究されている深層学習を利用した点群データの地物分類手法に基づいて点群・波形データの地物分類手法や波形データ推定手法を提案している。そのため、まず、コンピュータビジョン分野における点群データの解析に関する背景を述べる。航空機ライダーで観測した点群データだけでなく、RGB-D カメラ (Kinect, RealSense), スマートフォン (iPhone 12) などによって点群データが取得でき、点群データは取得しやすくなった。これらのセンサで取得した点群データは、空間的な形状やスケールの情報を豊富に含んでいる。点群データは、画像計測では観測できない物体の裏側の情報を補うことが可能であり、リモートセンシングだけでなく、自律走行、ロボティクス、医療など様々な分野で応用されている。通常、点群データを用いて3次元形状を表現するが、それ以外にも深度画像やメッシュ、Voxelなどのデータ形式で表現されることもある。点群データによる3次元形状の表現は、離散化することなく3次元空間の元の幾何学的情報を保持可能なため、一般的に広く使用されている形式である。このような点群データに対しても同様に深層学習を使用した解析手法が提案されている [68, 3, 134]。

点群データに対する深層学習の研究は、ModelNet [232], ScanObjectNN [212], ShapeNet [23], PartNet [154], S3DIS [6], ScanNet [34], Semantic3D [70], ApolloCar3D [201], や KITTI Vision Benchmark Suite [60, 14] などのデータセットの公開によって支えられている。深層学習手法を使用した点群データの解析手法として、画像と同様に、地物分類、物体検出、セグメンテーションが行われている [84, 3, 234, 180]。また、2つの観測間での点群データ位置合わせや、6自由度の物体姿勢推定、3次元再構成など、点群データ処理に関する様々な問題に対応する手法が提案されている [51, 243, 71]。本節では、深層学習手法の基本となるタスクの説明とそのデータセットの説明、および、本研究で対象とするセグメンテーション手法を説明する。



図 2.20: 点群データに対する深層学習手法を用いた主要なタスク。タスクは画像と同様にシーン分類・物体検出・セグメンテーションがある。図は [123] の Fig. 2 より引用し加筆した。

まず、以降の関連研究紹介にて利用されている学習用の点群データのデータセットについて述べる。様々な点群データに対する深層学習アルゴリズムの研究を補助するため、もしくは、手法の性能を定量的かつ平等に評価するために数多くのデータセットが公開されている。表 2.4は、図 2.20に示す点群データに対する主要な問題設定である、シーン分類、物体検出、およびセグメンテーションに使用される典型的なデータセットを示している。シーン分類のためのデータセットには、CAD モデルから合成した擬似的なデータセット (CAD データセット) [232, 23] と、実世界で実際の物体を計測したデータセット (実世界データセット) [34, 212] の2種類ある。CAD データセットで対象となる物体は、オクルージョンや背景がない理想的な状態である。一方、実世界データセットに含まれる物体は、様々な状況でオクルージ

ョンが発生し、さらにノイズが含まれている。物体検出には、部屋のレイアウトなどの屋内計測のデータセット（屋内データセット）と、屋外の都市域を計測したデータセット（屋外データセット）の2種類がある。屋内データセットに含まれる点群データは、高密度の深度画像から変換されたものか、3Dメッシュからサンプリングされたものである。屋外データセットは様々な物体が空間的にまばらであり、これらの点群データは疎であることが多い。セグメンテーション（本研究における地物分類に相当）のために使用される公開データセットは、その計測する範囲や対象によって様々な機器で計測された点群データを使用する。例として、自動運転用の3次元地図を作成する状況を想定した場合には Mobile Laser Scanners (MLS) が使用され [14, 193, 188], ある都市を包括的に広範囲の3次元地図を作成する際には航空機ライダーで計測した点群データが使用され [187, 216], より高密度な3次元地図を作成する用途では地上設置型 (TLS) が使用され [70], 屋内では取り回しの良い RGB-D カメラ [34] や屋内計測用の特殊なレーザスキャナ [6] が使用される。上記のデータを計測対象ごとに分けると、屋内データセットは S3DIS [34, 6], 屋外データセットは Semantic3d[70], Paris Lille 3D[188], Toronto3D [207], DALES[216], ISPRS[187], semantickitti [14] のようになる。上記のように、現状の研究においては、点群データの解析に興味はあるが、波形データに着目したデータセットは無い。そのため、本研究のような波形データと点群データを組み合わせた深層学習手法による地物の地物分類を提案することはリモートセンシング分野特有の考え方と言える。

表 2.4: 点群データに対するシーン分類, 物体検出, セグメンテーションのための公開データセットの概要

シーン分類向けデータセット									
データセット名	公開年	点数	クラス数	訓練数	テスト数	種類	データの表現		
McGill Benchmark [197]	2008	456	19	304	152	合成	メッシュ		
Sydney Urban Objects [37]	2013	588	14	-	-	実世界	点群データ		
ModelNet10 [232]	2015	4899	10	3991	605	CAD	メッシュ		
ModelNet40 [232]	2015	12311	40	9843	2468	CAD	メッシュ		
ShapeNet [23]	2015	51190	55	-	-	CAD	メッシュ		
ScanNet [34]	2017	12283	17	9677	2606	実世界	RGB-D		
ScanObjectNN [212]	2019	2902	15	2321	581	実世界	点群データ		
物体検出向けデータセット									
データセット名	公開年	シーン数	クラス数	真値付きフレーム数	矩形の数	種類	センサ		
KITTI [60]	2012	22	8	15K	200K	Urban (Driving)	RGB & LiDAR		
SUN RGB-D [200]	2015	47	37	5K	65K	Indoor	RGB-D		
ScanNet V2 [34]	2018	1.5K	18	-	-	Indoor	RGB-D & メッシュ		
H3D [170]	2019	160	8	27K	1.1M	Urban (Driving)	RGB & LiDAR		
Argovse [24]	2019	113	15	44K	993K	Urban (Driving)	RGB & LiDAR		
Lyft L5 [98]	2019	366	9	46K	1.3M	Urban (Driving)	RGB & LiDAR		
A*3D [173]	2019	-	7	39K	230K	Urban (Driving)	RGB & LiDAR		
Waymo Open [206]	2020	1K	4	200K	12M	Urban (Driving)	RGB & LiDAR		
nuScenes [20]	2020	1K	23	40K	1.4M	Urban (Driving)	RGB & LiDAR		
セグメンテーション向けデータセット									
データセット名	公開年	点数	クラス数 <sup>1</sup>	スキヤンの数	広さ	色情報	センサ		
Oakland [156]	2009	1.6M	5(44)	17	-	N/A	MLS		
ISPRS [187]	2012	1.2M	9	-	-	N/A	ALS		
Paris-rue-Madame [193]	2014	20M	17	2	-	N/A	MLS		
IQuibus [213]	2015	300M	8(22)	10	-	N/A	MLS		
ScanNet [34]	2017	-	20(20)	1513	8×4×4	Yes	RGB-D		
S3DIS [6]	2017	273M	13(13)	272	10×5×5	Yes	Matterport		
Semantic3D [70]	2017	4000M	8(9)	15/15	250×260×80	Yes	TLS		
Paris-Lille-3D [188]	2018	143M	9(50)	3	200×280×30	N/A	MLS		
SemanticKITTI [14]	2019	4549M	25(28)	23201/20351	150×100×10	N/A	MLS		
Toronto-3D [207]	2020	78.3M	8(9)	4	260×350×40	Yes	MLS		
DALES [216]	2020	505M	8(9)	40	500×500×65	N/A	ALS		

次に、点群データに対する深層学習を用いたセグメンテーション手法について述べる。点群データに対する深層学習手法は、不規則に分布した点群データに対して画像等に変換することなく直接適用できる利点がある。なお、点群データは順番が定義されてなく、なおかつ画像のように空間的に構造化されていないため、標準的な CNN を直接適用することはできない。そこで、点群データに対する深層学習手法として、図 2.21 に示すような PointNet[177] が提案された。PointNet は、点群データを直接入力可能な手法であり、対称関数を用いて点群データの入力順に対する順列不変性を実現した研究である。具体的には、PointNet は複数の全結合層層で点ごとの特徴を独立して学習し、Max Pooling 層で入力順に依存しない大局的な特徴を抽出する。さらに、PointNet と同時期に発表された Deep Sets[242] は、すべての点の特徴量を合計し、非線形変換を適用することで入力点群データに対する順列不変性を獲得する。しかしこれらの手法では、点ごとに独立して特徴を学習するため、近傍点の関係性を捉えることが不可能である。そのため、画像の分野で使用される CNN のような、近傍情報を階層的に学習することができない。そこで、Qi ら [178] は、図 2.22 にあるような各点の近傍の空間的構造の階層性を捉えるために、階層的なネットワーク PointNet++ を提案した。PointNet++[178] は、サンプリング層、グルーピング層、PointNet[177] ベースの特徴抽出層の 3 つの層で構成されており、入力された点群データに対し、近傍情報の特徴抽出を階層的に行い、徐々に抽象化された特徴を大局的に獲得できる。

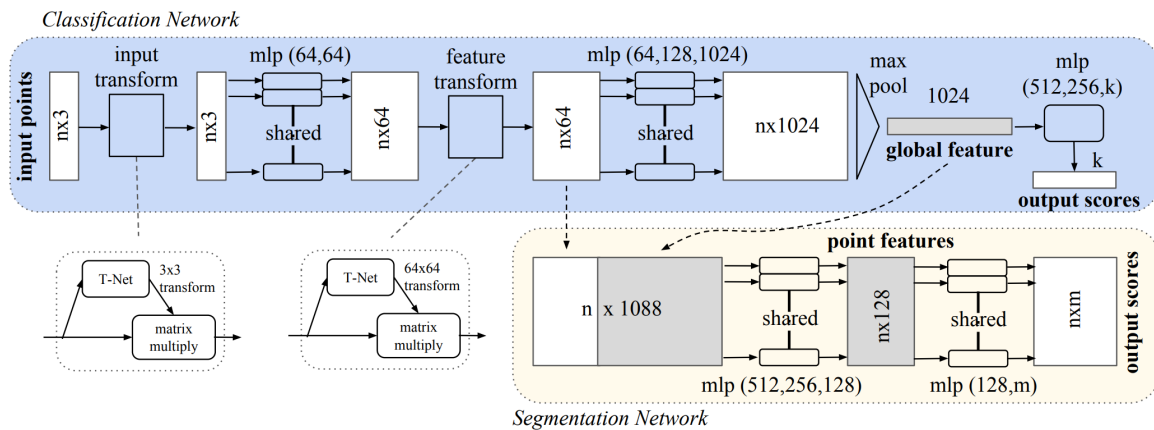


図 2.21: PointNet[177] の概念図。各点に独立して重みを共有した全結合層を適用し、最後に最大値等を求めて情報を集約する。図は [177] の Fig.2 より引用。

次に、点群データに対する深層学習手法の工夫点について説明する。PointNet のような全結合層ベース手法では、計算効率の良さから重み共有全結合層を用いることが多い。しかし、重み共有全結合層で抽出された全点に対し独立した点単位の特徴量では、点群データの局所的な形状や点間の相互作用を捉えることができない。各点のより広い範囲の大局的な特徴を捉えつつ近傍構造を学習するために、近傍特徴のプーリングなど点群データ専用ネットワークが導入された。以降で、これらについて説明する。全結合層ベース手法では、計算効率の良さからネットワークの基本単位として重み共有全結合層を用いることが多い。しかし、重み共有全結合層で抽出された全点に対し独立した点単位の特徴では、点群データの局所的な形状や点間の相互作用を捉えることができない。各点のより広い背景情報を捉えつつ近傍構造を学習するために、近傍特徴集約が使用される。近傍特徴集約では、ある点に対する近傍点の幾何学的パターンを捉えるために、近傍点からの情報を集約して各点の特徴を学習する。PointNet++[178] では近傍点の定義をグループ化と呼ばれる手法を用いている。このグループ化では、図 2.23 のようにある点に対して球 (図 2.23 の青丸で示した範囲) を仮定してその球内に入る点をグルーピング (図 2.23 の赤点に対する周辺の青点がグループ化された点) する。このグループ化された領域において図 2.23 の緑の  $w$  で示す重み付き演算を適用する。このグループ化は、画像処理における畳み込み演算のある画素に対する上下左右の画素に相当する。また、PointNet++[178] は点を階層的にグループ化し、小さな近傍点から大きな近傍

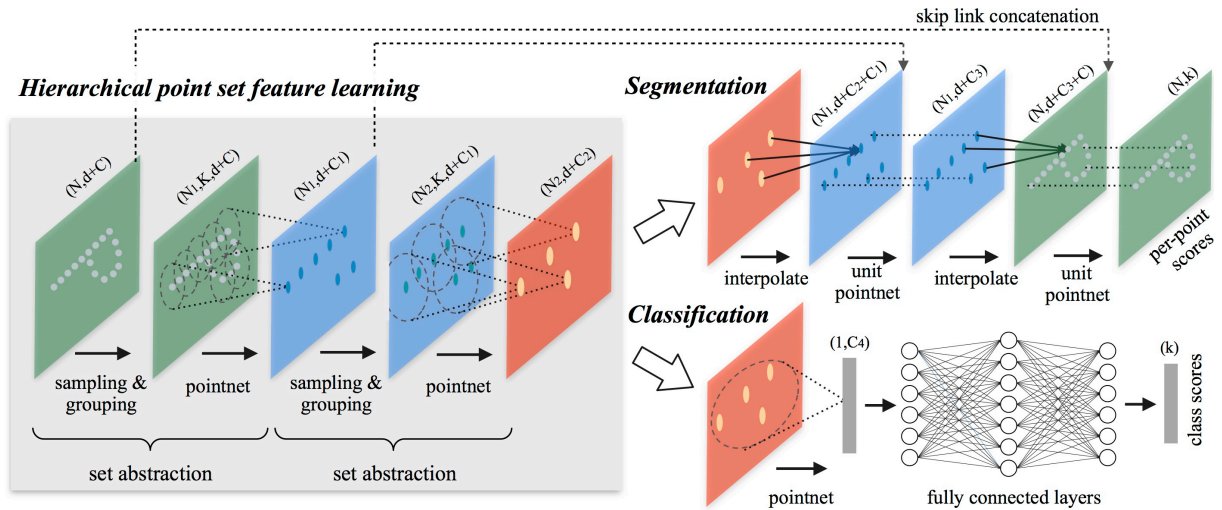


図 2.22: PointNet++[178] の概念図. ある点の周辺をグループ化し, そのグループに対して PointNet の演算を適用する. 図は [178] の Fig.2 より引用.

点へ徐々に特徴抽出を行う. 階層的なグループ化の際には, ダウンサンプリングによる代表点の抽出を適用することで, 図 2.24 のように徐々に点数を少なし, 近傍点として定義される点をそれに応じて広い範囲にする. これにより, ある一点を地物分類するために使用する範囲が広がる効果がある. また, 図 2.24 のように, ダウンサンプリングされた点から徐々に元の点数に戻すアップサンプリングを適用して広範囲から抽出された特徴を元の点数に伝搬させる. これによって, 大局的な特徴を元の点に付与可能になる. また, 点群データの不均一性や密度の変化に起因する問題を解決するために, グループ化の際に複数スケールグループ化や複数解像度グループ化が提案された. PointNet++[178] の登場後, 様々な手法が提案された. Jiang ら [91] は, 向き情報を扱うため PointSIFT モジュールを提案した. このモジュールは, 8 つの空間的方位からの情報を, 3 段階の順序付き畳み込みによって連結し符号化する. Engelmann et al. [52] は, PointNet++ で使用されているグループ化技術 (すなわち, ボールクエリ) とは異なり,  $K$ -means クラスタリングと KNN を利用して,  $x, y, z$  の実空間座標と特徴空間の 2 つの近傍領域を別々に定義した. また, 同じクラスに属する点は特徴空間内で近いと考えられることから, 得られる特徴量をさらに正則化するために, ペアワイズ距離損失とセントロイド損失を導入した. 異なる点の間の相互作用をモデル化するために, Zhao ら [247] は, 局所的に完全にリンクされた近傍点の関係を密に構築することで, 近傍領域内のすべての点のペアの関係を探索する PointWeb を提案した. Zhang ら [246] は, 同心円状の球から得られる統計情報に基づいて, Shellconv と呼ばれる順列不変の畳み込みを提案した. この手法では, まずマルチスケールの同心円球の集合に問い合わせを行い, 統計量を要約するために異なる球の中で Max Pooling 操作を行う. そして, 最終的な出力を得るために全結合層と 1 次元畳み込みを使用する. Hu ら [81] は, 大規模な点群データセグメンテーションのために, RandLA-Net と呼ばれる効率的で軽量なネットワークを提案した. この RandLA-Net は, ダウンサンプリング時にランダムサンプリングを利用して, メモリ消費と計算量の面で著しく高い効率性を実現した.

このような階層構造や特徴集約が提案された理由には, 画像と同様に受容野の広さとセグメンテーションの関係が存在する. 点群データに限らず画像においても同様に, ある点を分類する際に受容野を広く設定すると, ある点とその周辺点の相対的な特徴量を用いた分類可能になる. 例として, 図 2.25 に示すように建物の点を分類する際には, 受容野を広くして大局的な特徴を得て道路の上には建物が存在するなどの特徴を使用して分類ができる. しかし, 狭い受容野であれば, 建物を分類する際に, 同じような平坦性を持つ道路なのか地面なのかの区別ができない. このように, 深層学習手法を用いて地物を分類する際には, 受容野を広く設定し大局特徴を用いて分類することが望ましい.

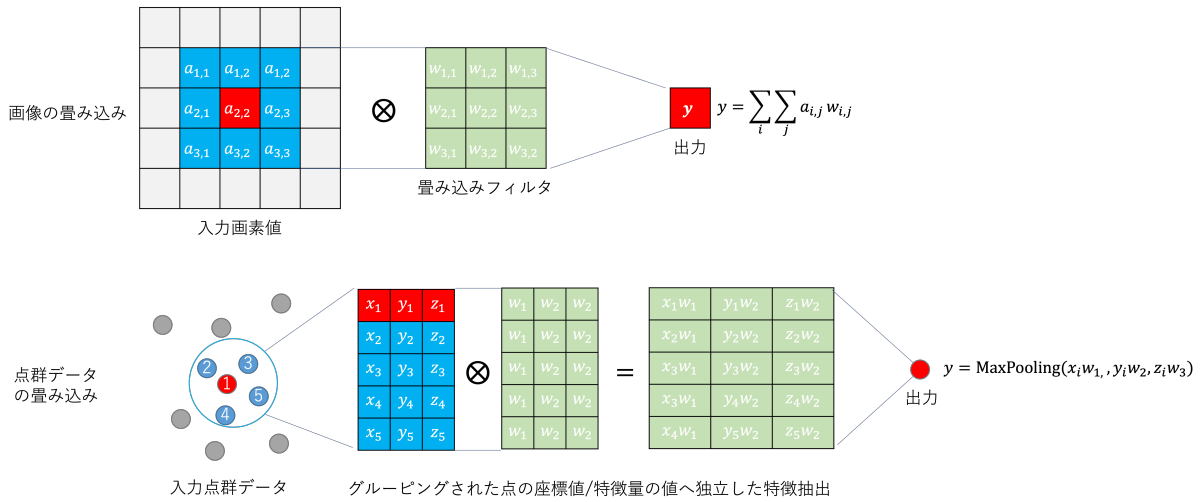


図 2.23: 点群データの近傍特徴抽出と画像の近傍特徴抽出の違い。画像の場合はある画素の上下左右が任意に決定できるが、点群データは位置関係の定義をアルゴリズム上で定義する必要がある。

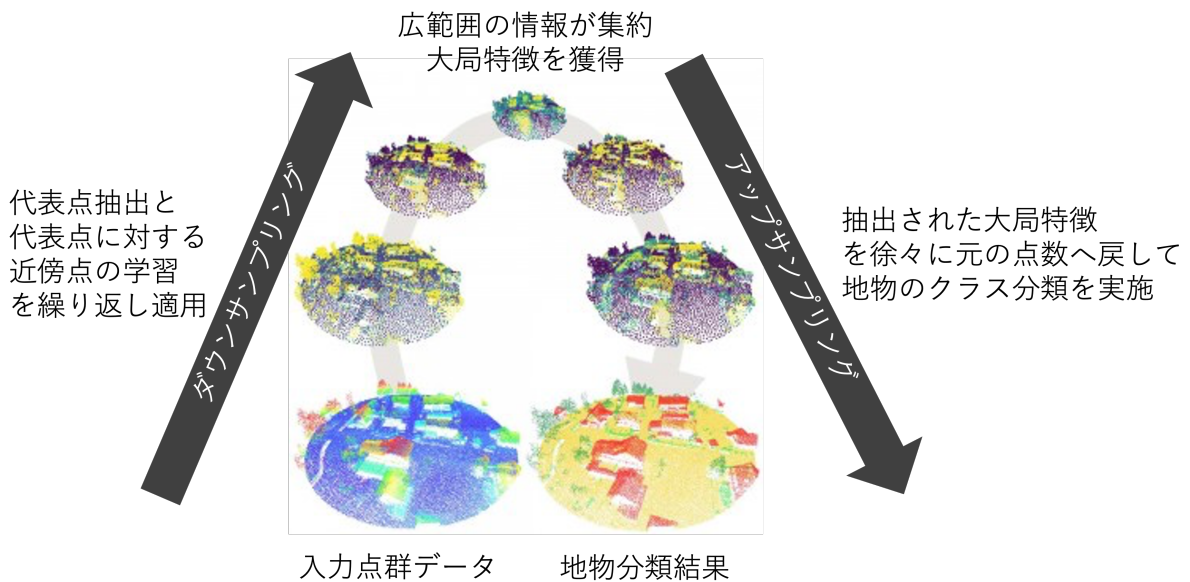


図 2.24: PointNet++ のダウンサンプリングとアップサンプリングの概念図。徐々にダウンサンプリングを繰り返し、広い範囲の情報を抽出可能である。図は [229] のデータを加筆し作成。

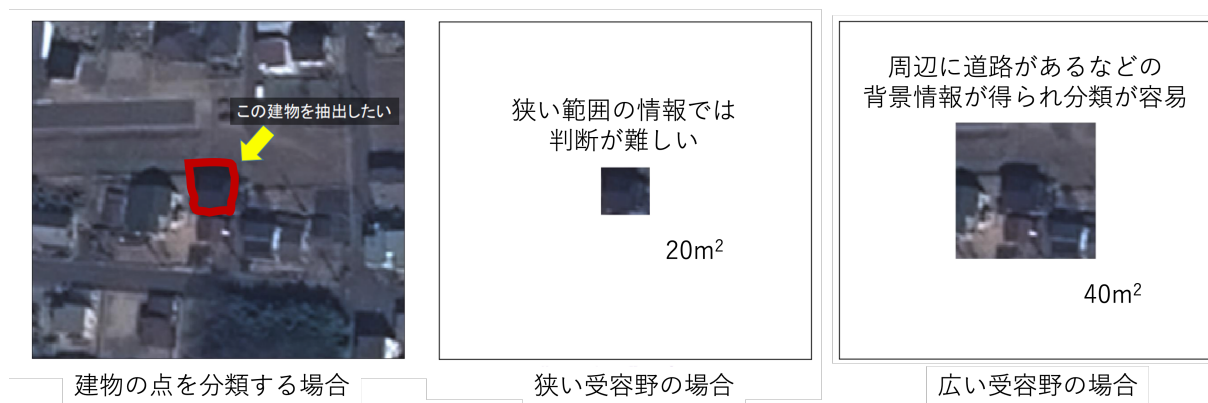


図 2.25: 深層学習手法における受容野の概念図. 点群データに限らず深層学習手法の基本戦略は広い受容野によって得た特徴を用いて分類する.

### 2.3.5 条件付き GAN

本研究では、深層学習技術を利用して点群データからその点群データが観測された際の元の波形データを推定する手法を提案する。この際に参考にするのが、条件付き敵対的生成ネットワーク (Generative Adversarial Network, GAN) と呼ばれる手法である。条件付き GAN の説明に入る前に、その基本となる GAN について説明する。

- GAN

敵対的生成ネットワーク (Generative Adversarial Network, GAN) は、深層学習の分野で注目を集めている技術の一つである [63, 133, 190, 211, 231, 166]。GAN が用いられる分野としては、コンピュータビジョン [50, 141, 217, 235, 163, 124, 251, 115]、自然言語処理 [55, 238, 36, 90]、時系列合成 [43, 73, 53, 120, 19]、セグメンテーション [254, 138, 45, 179, 202] など多岐にわたる。

GAN は、機械学習における生成モデル、つまり、データが持つ確率密度関数の推定が可能なモデルの一つである。GAN 以外にも変分オートエンコーダーなどの他の深層学習手法を利用した生成モデルが提案されているが、それらと比較して GAN は多峰性のある確率密度関数を扱うことができる。具体的な GAN の特性を以下に示す。

1. GAN は他の深層生成モデルよりも優れた出力が得られる。可変オートエンコーダー (VAE) と比較すると、VAE が平均的なデータしかできないのに対し、GAN はあらゆるタイプの確率密度を生成することができる [61]。
2. GAN の考え方をを用いると、様々なタイプの生成ネットワークでも学習可能である。他の深層生成モデルでは、生成器の出力層がガウス分布であるなど、生成器の事前分布が必要であるが、GAN には出力層において、分布の仮定が必要ではない [61, 41, 100]。
3. 生成器に入力する潜在変数の次元数に制限がない。一般的には 100 次元を入力するが、任意の数を設定可能である。

この特性により、GAN は特にコンピュータビジョンの分野で使用される。例えば、尤もらしい画像の生成 [97, 225, 176, 30, 244]、画像の変換問題 [253, 133, 252, 210, 132, 87, 31, 142]、画像の超解像 [118, 226, 144, 45, 143]、画像の傷修復 [240, 239, 42, 29, 122]、画像の着色 [113] などである。

GAN の仕組みの説明を具体的にするため、画像生成の学習を例にアルゴリズムの説明を行う。GAN は、図 2.26 に示すように、真値の画像と生成画像を分類する識別器 ( $\mathcal{D}$ , Discriminator) と、識別器を欺くために画像を生成する生成器 ( $\mathcal{G}$ , Generator) の 2 つの深層学習モデルで構成される  $\mathbf{z} \sim p_z$  という分布が与えられたとき、 $\mathcal{G}$  はサンプル  $\mathcal{G}(\mathbf{z})$  の分布として確率分布  $p_g$  を定義する。GAN の目的は、真値の画像のデータの確率分布  $p_r$  に近似した生成器で作成したデータの確率分布  $p_g$  を学習することである。

まず、GAN の最適化戦略について説明する。GAN の最適化は、 $\mathcal{D}$  と  $\mathcal{G}$  の両者に対して行われる。

$$\begin{aligned} & \mathbb{E}_{x \sim p_r} [\log \mathcal{D}(x)] + \mathbb{E}_{z \sim p_z} [\log(1 - \mathcal{D}(\mathcal{G}(z)))] \\ & \arg \min_{\mathcal{G}} \max_{\mathcal{D}} \mathcal{L}_{cGAN}(\mathcal{G}, \mathcal{D}). \end{aligned} \quad (2.22)$$

次に GAN の層設計について説明する。オリジナルの GAN 論文 [63] では、生成器と識別器の両方に全結合層のみを使用している。この層設計は、MNIST [117]、CIFAR-10 [106]、Toronto face dataset などの単純な画像データに適用された。次に具体的な最適化を説明する。学習の設定で  $\mathcal{D}$  の最適化が完了すると、識別器のオーバーフィッティング対策のために  $\mathcal{D}$  の最適化を  $k$  ステップ行い、 $\mathcal{G}$  の最適化を 1 ステップ行うことを提案している。実際には、式(2.22)において生成器を最適化する際に損失関数の勾配が消失する問題を引き起こす可能性があり、は  $\mathcal{G}$  の学習において  $\log \mathcal{D}(\mathcal{G}(z))$  を最大化している。こ

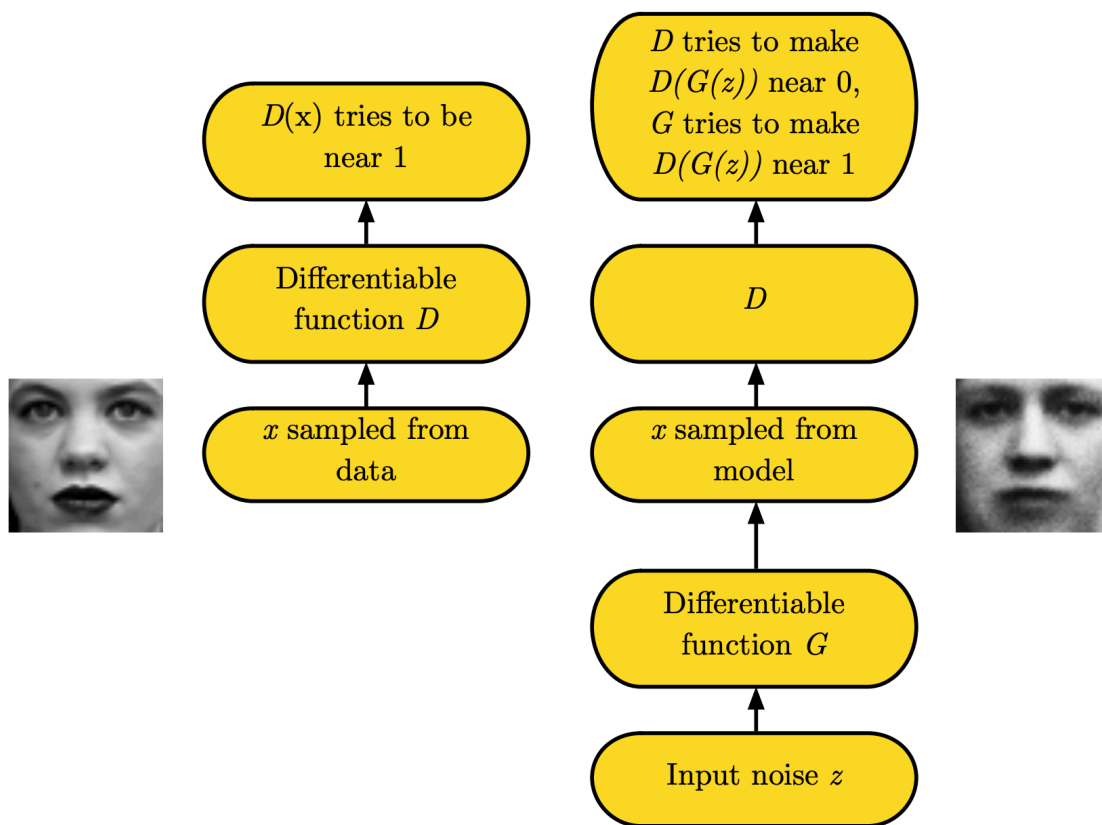


図 2.26: GAN の概念図. GAN は訓練データの確率分布を推定する手法である. 生成器  $D$  は乱数  $z$  から生成画像 ( $x = G(z)$ ) を出力する. 識別器  $D$  は生成画像 ( $x$  sampled from model) か真値の画像 ( $x$  sampled from data) かを分類する. この際識別器  $D$  は生成画像と真値の画像の確率分布を計算していることに等しく, これを最小化すると両者の確率分布が一致する. 図は NIPS2016 チュートリアル論文 [61] の Fig.12 から引用.

の最適化は, 図 2.27 に示すような  $G$  に対して  $p_g$  と  $p_r$  の間の Kullback-Leibler (KL) ダイバージェンスを最小化と定義できる. なお, KL ダイバージェンスは 2 つの  $p_g$  と  $p_r$  間の分布が一致した場合に 0 になる.

このようにオリジナルの GAN では, 生成器と識別器の組み合わせによって理論的に入力データの確率分布を推定できることを示した. しかし, 実際の複雑な画像に対しては優れた画像生成の性能を示さなかった.

この問題は GAN の学習は 2 つの深層学習モデルを同時に最適化するので非常に不安定であることに起因する. 学習が困難であることは, 生成器と識別器はお互いに非協力状態であることに起因する. これは, GAN の損失関数の設定が, 生成器と識別器で逆の最適化問題, つまり生成器は損失関数を最小に, 一方, 識別器は損失関数を最大にする Min-Max 最適化を行うためである. 及び生成器がデータセット分布全体をうまく学習できないことに由来する. 前者は生成器と識別器の両者が非協力ゲーム状態にあるため理論上回避することは難しいが, 後者はモード崩壊問題と呼ばれており, モード崩壊を回避する手法が提案されている [104, 35, 103, 125].

- 条件付き GAN

GAN は教師なし学習によって, 入力されたデータの確率分布を獲得する手法である. しかし, 獲得した確率分布からデータを生成する際に明示的に制御することが困難であった. そこで, ラベル情報等の条件を一緒に生成器へ入力することで, 生成結果を制御す

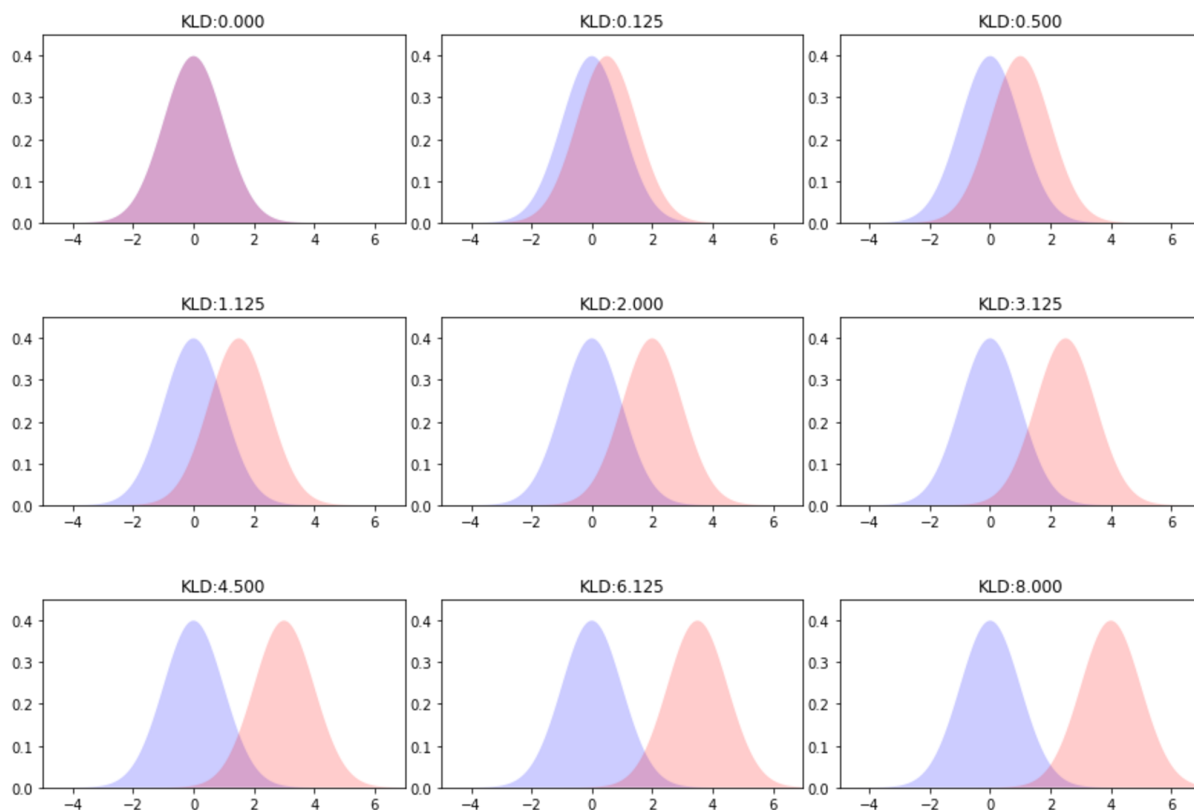


図 2.27: GAN で使用される尺度である KL ダイバージェンスの正規分布を用いた時の例. 完全に一致した際には 0 になる.

る手法が提案された. その手法は条件付き GAN と呼ばれ, 実際の学習時には識別器と生成器の両方にクラスラベルを与えて条件付けし, 推論時には生成器にのみクラスラベルを入力する [153].

条件付き GAN は, 学習時において, 追加情報  $\mathbf{y}$  ( $\mathbf{y}$  はクラスラベルや他の形式の画像などのデータでもよい) を識別器と生成器の両方に与える. なお,  $\mathbf{y}$  は, 通常, 生成器と識別器の内部で符号化された後, 符号化された  $\mathbf{z}$  と符号化された  $\mathbf{x}$  と結合される. 例えば, 原著の MNIST 実験では,  $\mathbf{z}$  と  $\mathbf{y}$  は, それぞれ 200 層と 1000 層の隠れ層に写像された後, 生成器で互いに結合されている (結合された層の次元は  $200 + 1000 = 1200$ ). このようにすることで, 条件付き GAN は識別器の識別能力を高めることができる. 条件付き GAN の損失関数は, 通常の GAN とは若干異なり,  $\mathbf{x}$  と  $\mathbf{y}$  を  $\mathbf{z}$  で条件付けている (式 (2.23) 参照).

$$\min_{\mathcal{G}} \max_{\mathcal{D}} \mathbb{E}_{\mathbf{x} \sim p_r} \log[\mathcal{D}(\mathbf{x}|\mathbf{y})] + \mathbb{E}_{\mathbf{z} \sim p_z} \log[1 - \mathcal{D}(\mathcal{G}(\mathbf{z}|\mathbf{y}))]. \quad (2.23)$$

条件付き GAN は, あらかじめ与えられた  $y$  情報のおかげで, 単一的な画像データセットだけでなく, 自然画像のようなデータセットも扱うことが可能になった. 著者らは, 公開されているデータセットである手書き数字の MNIST とインターネット上の画像で作成したデータセットである Yahoo Flickr Creative Common 100M (YFCC 100M) で cGAN を実験した. MNIST では, SGD を用いて, ミニバッチサイズ 100, 初期学習率 0.1, 減衰係数 1.00004 で指数関数的に  $1e - 6$  まで減少させてモデルを学習した. 生成器と識別器の両方に対して, 0.5 の確率でドロップアウトを行った. クラスラベルはワンホットベクトルとして符号化され,  $\mathcal{G}$  と  $\mathcal{D}$  の両方に与えた. なお, YFCC 100M の実験では, 学習用ハイパー重みは, MNIST の実験で設定したものと同じである. 条件

付き GAN は符号化されたラベルを導入することで識別器の能力を高めているが、生成されたラベルの中には画像との関連性を失っているものもある。

上記の条件付き GAN は画像生成において、数字画像の生成であれば数字のように生成する対象のラベル情報を入力した。この入力条件を図 2.28 に示すように画像として与え、入力とは別の画像に変換する条件付き GAN が提案された。図 2.28 では、入力した画像  $x$  を  $G$  によって入力画像に対して色を推定した結果を推定している。そして、推定された着色画像を  $D$  に入力して、着色された画像であるか真値の色の画像かを分類する。これは Pix2Pix[87] と呼ばれ、白黒画像から着色画像、低解像度画像から高解像度画像、ボケ画像からボケなし画像などの情報が欠如した画像から元の理想的な状態を推定可能な手法である。これらのことを入力された画像を別の画像に変換する画像変換と Pix2Pix では定義しているため、ここでもそれに従う。この欠如した情報を推定する手法は問題設定不良であるため、1:1 に定まるものではなく、1:多の写像である。そのため、機械学習・深層学習を用いて統計的にとりうる数値を推定する手法が用いられる。Pix2Pix は、入力画像を元に変換された画像と真値の間の画素単位の回帰を実行する損失関数に加えて、敵対的な損失関数を利用して出力が「実際の」画像と区別できないようにした。その目的関数 ( $\mathcal{L}$ ) は以下のように定義できる。

$$\mathcal{L} = \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L_1}(G). \quad (2.24)$$

ここで、 $\mathcal{L}_{cGAN}(G, D)$  は GAN と同様の損失関数、 $\lambda \mathcal{L}_{L_1}(G)$  は生成画像と実際の画像における画素の差を示す。

Pix2Pix は、強力な画像変換手法であり、それに基づいて改良された多くの画像間変換手法に影響を与えている。また、Pix2Pix の考え方は、画像だけにとどまらず、音声や言語のほか、後述する点群データの分野においても点群データの超解像や着色に応用されている。

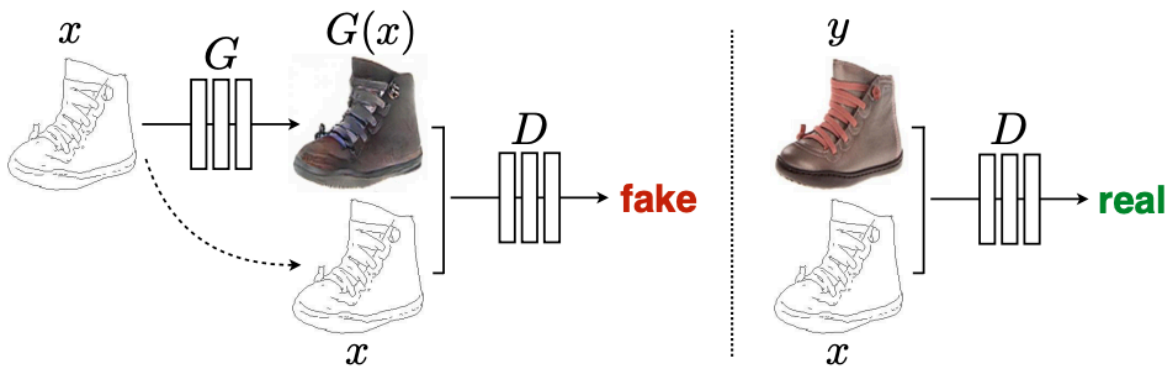


図 2.28: 条件付き GAN の概念図。条件付き GAN の例としては Pix2Pix[87] による画像変換が主な例である。Pix2Pix[87] では、入力された画像  $x$  から別の情報 (ここでは色) を生成器  $G$  によって推定することができる手法である。そして、生成器  $D$  で推定された画像 ( $G(x)$ ) か本物の画像 ( $y$ ) かを分類する識別器  $D$  を併用し、推定結果の質を向上させた。図は Pix2Pix[87] の Fig.2 より引用。

## 2.4 点群・波形データに対する深層学習手法の適用

### 2.4.1 古典的な点群・波形データの解析方法

本節では、点群・波形データの古典的な解析手法について説明する。一般的には、波形データのモデル化で得られた点群データの点単位の地物分類を実行する。古典的な手法においては、各点は幾何学的特徴と波形データ特徴を有している。幾何学的特徴は、 $xyz$ 座標値そのものや点群データに依存する空間的な特徴量（曲率や法線ベクトルなど）であり、一方、波形データ特徴は波形データモデル化から抽出された振幅やエコー幅等である。以降では、森林領域、DTMの作成、都市域に関する手法を説明する。

- 森林領域

点群および波形情報を記録するライダーデータは、森林領域の分析に広く利用されている。森林領域においては、樹木の枝や幹からの弱いエコーを検出できるため、より詳細な樹木の計測が可能になる [258]。日本国内においては、波形データのピークから樹冠内の高密度な点群データを作成した後に、主に林相分類を実施するために使用される [269]。森林領域において点群および波形情報を記録するライダーデータが使用されるパターンは以下の3つに類型化できる。

- バイオマスや森林構造の推定をする際には、従来の疎な点群を記録するライダーでは低木の検出能力に限界があるため、森林パラメータの推定を向上させるために、高密度の点群データを利用すること。
- 個々の樹木の領域推定と樹種分類のために、点群データによる幾何学的情報と波形データから推定された放射特性（後方散乱断面）を同時に使用することに取り組んだ研究である。
- 森林のパラメータが波形データに与える影響を理解するためのモデル化を行う研究である。

次に、植生を対象にした具体的な研究例について説明する。植生分類の問題は、オブジェクトベースの手法 [189] が用いられることが多い。まず、波形データの特徴量と点群データにおいて、エコー幅に基づいた領域拡張アルゴリズムを用いて、同じような特徴を含む波形データをまとめて小領域に分ける。次に、各領域について、領域内の各点の特徴（振幅、エコー幅、幾何学的属性）について、基本的な統計（最小、最大、標準偏差など）を計算する。これらの特徴を、分類木アルゴリズムに入力する。これにより、最良の特徴量とそれに関連する閾値を自動的に選択できる。そして、従来手法 [47] で報告されたものよりも若干高い分類性能（90%以上）を達成することができた。また、改良されたエコー比特徴量を新たに提案した研究 [76] では、より少ない計算量でより良い植生の分類を可能にしている。この研究 [76] は、非常に高い点密度、例えば  $50 \text{ 点}/\text{m}^2$  を扱う場合に有効である。また、非植生領域のデータも対象にしており、特にエコーの振幅と幅を使用した新たなルールベースの地物分類を作成し、建物の壁、屋根の縁、送電線などの非植生地物を分類した。

- DTMの作成

ここでは、デジタル地形モデル（DTM）の作成について、波形データの効果を述べる。一般的な利点として、疎な点群を記録するライダーよりも弱いエコーを検出できるため、地盤からのエコーを検出可能になる [263]。波形データ内に存在する各エコーのパルス幅は、物体表面の垂直方向の広がりに関する情報を保持しているため、重要な特徴量になりうる。フットプリントの小さいライダーでは、エコー幅は主に低植生（シダ、小さな茂みなど）や地面の上にある低い物体（道路上のポールや生垣）によって変化する。また、低植生や生垣はフットプリント内での高さの変化が大きいことも特徴的である。したがって、これらの物体からのエコー幅は照射された元のエコー幅よりも大きくなる。逆に、地形の傾斜による高さの違いは、一般的にパルスの長さ（4 ns のパルスで 1.2 m 相当）よ

りもはるかに小さいため、観測されるエコーは地形の傾斜に強く影響されない。このように、エコー幅を利用することで、地盤からのものではないエコーを識別することができる。この特性を考慮した DTM 作成のアルゴリズムを改良した手法を以下で紹介する。

まず、植生下の特徴に着目することで DTM の作成を高性能にした手法を紹介する [44]。この手法では、経験的に設定された閾値以上の反射エコー幅を持つ点を、最後のエコー点群データから削除する。この閾値による削除は、実際には下層植生に相当する。次に、残りの点に対して従来の地面/非地面分類によるフィルタリング処理を行う。このフィルタリング処理では、地形の変化ではない計測誤差や複数反射などの影響で発生した小さな凸凹を除去することで、DTM の作成を改善する。なお、このような反射エコーの幅を考慮する手法は波形データだけでなく疎な点群を記録するライダー向けの DTM 作成アルゴリズムに対しても同様に使用できる。疎な点群を記録するライダーの反射エコー幅を地面/非地面の識別特徴として使用するよう拡張した研究も存在する。例えば TIN ベースの DTM 作成手法 [9] に疎な点群を記録するライダーのエコー幅の特徴を導入した手法がある [128]。点群・波形データを用いた研究としてはエコーの幅に応じて点群データを 3 つのクラスに分ける手法がある [114]。その対象のクラスは、「オーバーラップするパルス」、「タイプ B」、「タイプ A」である。「オーバーラップするパルス」は、エコー幅の推定が信頼できない近接したエコーに対応する。「タイプ A」は、幅が所定の閾値よりも大きい粗い表面のパルスである。「タイプ B」は滑らかな表面に対応する。さらにこの「タイプ A」は、波形データ内の位置（単一または複数のリターン）に応じて、「Open smooth」と「Obscured smooth」の点に分割される。このクラス分け戦略が DTM 作成において、低い植生を除去することができ、局所的な地形表面の過小評価や過大評価を修正できる可能性があることを示した。この手法を元に「タイプ A/タイプ B」の閾値が決定され、続いて地上/非地上点の分類が行われる。決定木による分類は、近隣の点のラベルに応じて点のラベルを正則化することで改良される。また、弱い振幅のエコーや非常に強い振幅のエコーに対しては、パルス幅を注意して使用しなければならないと結論づけている。

しかし上記の手法は実行速度に問題がある。そこで、DTM 作成のアルゴリズムを高速化した手法が提案された [155]。高速化では、DTM 作成アルゴリズム [105] に経験的な重みを導入した。この重みによって、地面に属する確率が高い点で直接地形面を決定できる。重みの設定は、パルス振幅に対するパルス幅のヒストグラムに分布関数を当てはめることで行う。この重み演算によって、DTM 作成の処理の改善と高速化が図られた。

#### ● 都市域

都市部では、照射したレーザーが建物の端に当たったときや、生垣、電線等の地物に当たった際に複数のエコーが現れる。波形データから推定されたエコー幅等の特徴量を用いることで都市域における地物分類性能を向上できる。

生の点群・波形データにおけるパターン認識の性能向上という問題に対する取り組みが行われている。その例として、レーザーが当たった都市の物質を波形データを用いて区別可能であると示した研究がある [93]。この論文 [93] では、物体を認識するためには、より詳細な波形データの処理が必要であると結論付けた。そこで、点群・波形データを分類するために、波形データと波形データ間の近傍性という概念を使用した。さらに、連続した波形データのピークの空間的・時間的な相互相関をとることで、点群データの分類性能を改善できた。

点群データを用いた建物境界の検出は、波形データ情報を併用することでその検出性能が向上した [92]。波形データから求めたパルス数、振幅、地物までの距離を用いた領域拡張方法により、建物を抽出した後に、「境界」と「建物内部」に分割する。その結果、各領域（内側と外側）でレーザーの反射をモデル化することができる。これにより、レーザーの中心がどこにあるのかを特定し、境界線がどこにあるのかを見つけられる。しかし、これらの結果は、内部の反射率が同様であると仮定したシミュレーションデータに基づいており、モデル化の問題を単純化している。そのため、現実的に存在する複雑な建物においては適用が難しい。

次に複数地物の境界線の抽出手法について説明する [92]. 具体的には, 各エコーのパラメータを解析し, 隣接する波形データの関係を考慮することで, 植生と人工構造物 (建物など) を区別する. さらに人工構造物に対してその輪郭などの特徴的な直線を抽出する. アルゴリズムは, まず, 不規則なボクセル空間  $(x, y, t)$  の時空間計測を行い, 近傍関係を示す  $x$  (前進方向),  $y$  (スワス方向),  $t$  (時間) を生成する. このようなボクセルを 2D で解析し, 2.5 次元の情報と必要な関係性を得るために,  $x-t$  平面上で垂直に切り取る. これにより, 入射信号に垂直な面がある状況 (例: 建物の壁面) を想定して, この立方体内の 3 次元領域を検出する. その後, 切り取り処理で得られた強度画像に対して線検出器 (ここではハフ変換) で直線を検出する. また, 点群データをボクセルではなく, 領域分割適用後に各領域ごとに同様の手順で分類する手法も研究されている [189].

以上の内容では, 波形データの特性に関して解析者が波形データに対する前提知識を利用して特徴量を作成し, ルールに基づいて地物分類する手法を説明した. しかし, 近年では機械学習, とりわけ深層学習を使用した点群・波形データの地物分類手法が用いられることが多く, 次節にて深層学習を使用した方法を詳しく説明する.

## 2.4.2 深層学習を利用した点群・波形データの地物分類

波形データから得られる多くの特徴量 (エコー番号や地上からの高さなど) は物理的な意味を持っており, 単純な閾値で地物分類できる可能性があるため, ルールベースの手法が用いられることが多かった. 例えば, 決定木を用いることで多数の地物を分類でき, 木, 低木, 草, 屋根, 道路の分類を点群・波形データの特徴量から達成できた [47]. 一方で, このようなヒューリスティックな手法では, 解析対象の地域やライダーに応じた閾値が必要となる [4]. また, 決定木は少ない特徴量で容易に構成できるが, 高い分類精度は得られない. 逆に多い特徴量では, 最適な分割順序を見つけるのが容易ではなく非常に複雑な木になる可能性がある. そのため, このような問題に対処するために, 教師あり学習の機械学習手法が選択されている. このような機械学習手法では, 入力する特徴量に加えて, 地物のラベルが関連付けられたデータ (真値) が必要となる. このデータセットは訓練データと呼ばれ, 分類器の重みを推定するために用いられる. ここで重要な前提は, データセット全体が特徴の分布という点で訓練データと似ているということである. これは汎化性能に関連し, 分類器が良い分類を行うためには, 学習において類似した特徴を観察していなければならないことを意味する. [25] では, 期待値最大化アルゴリズムを用いたガウス混合モデルが採用しており, [67] では, ランダムフォレスト分類器が採用された. さらに, 統計的学習理論に基づく分類器は, 高次元データと限られた学習データの両方に対応する顕著な能力を示しているため, サポートベクターマシン (SVM) は都市部における点群・波形データの分類のために研究されてきた [191]. この SVM 手法は, アルゴリズムの複雑さがデータ次元に依存しないため, 高次元の特徴空間を扱うのによく適している. さらに, SVM はノンパラメトリックな分類手法に属し, パラメトリックな確率密度関数を必要としない. さらに, 学習ステップは凸最適化問題に還元され, 最終的にはカーネルトリックにより, ある特徴空間における線形分類器を見つけることができる.

まず, 点群・波形データに対する一般的な機械学習手法について説明する. 解析手法は, 基本的には人間が作成する特徴量ベースの方法と, データから特徴量を得るデータ駆動な方法に分けられる. まず, 特徴量ベースの手法について説明する. 人間が作成する特徴量ベースの手法は点群データだけでなく波形データから作成した特徴量を組み合わせることで点群データのみの場合よりも地物の認識性能を向上させると言われている [57]. 例として, ルールベースアルゴリズムによる地物分類手法が古くから研究されている [220, 4]. 他の手法として, support vector machine (SVM) 地物分類器 [147] 等の非線形な地物分類器を用いた機械学習手法も提案されている [77, 10, 249, 140, 11]. さらに, 土地被覆地物分類を行うために, Wang らは [223] 波形データの空間的な特徴の作成が重要であることを示した. また, Lai [111] はアンサンブル手法を用いることで高性能な地物分類を達成した. 複数のセンサを組み合わせるマルチモーダルな研究もなされており, ハイパースペクトル画像と点群・波形データの組み合わせる手法が提案された [224, 139]. しかしながら, これらの手法は人間が設計する特徴量に依存するため, 地物分類性能上の限界があると言われている.

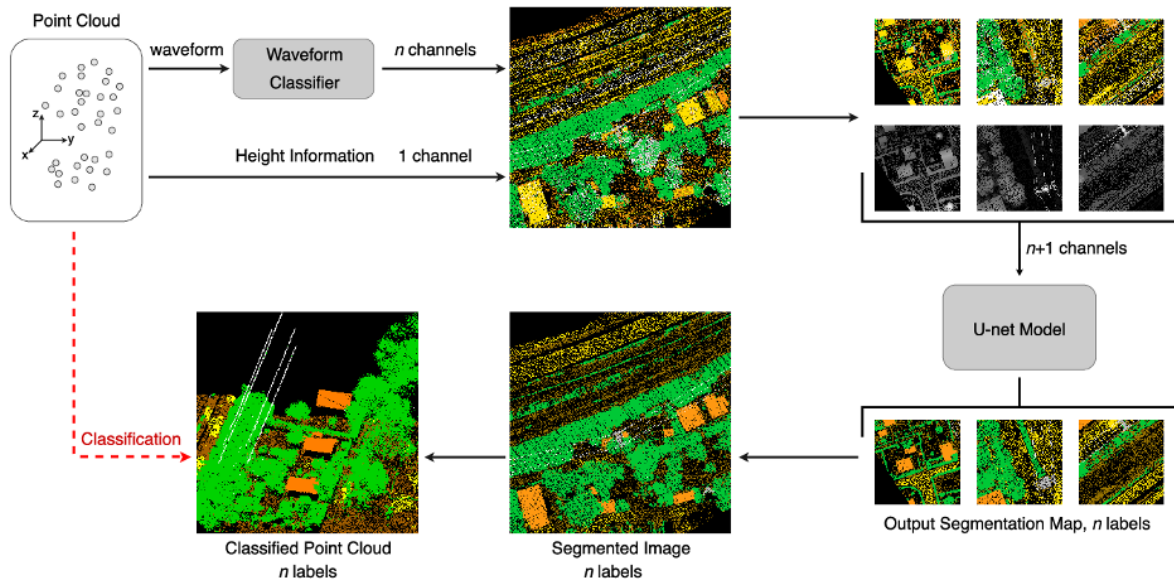


図 2.29: 点群・波形データ解析における深層学習手法. 点群・波形データを画像に変換する深層学習手法が唯一の先行研究 [255] である. 図は [255] の Fig.1 から引用.

次に、データ駆動の手法について説明する. データ駆動な手法は特徴量作成の依存性を克服している. 特徴量の作成を必要としない初期の手法は Maset ら [150] によって提案された. Maset ら [150] は Self-Organization Map により個別の波形データの教師なし地物分類学習 (低層植生・樹木・道路) を行った.

さらに近年では深層学習手法のひとつである畳み込みニューラルネットワーク (CNN) を用いた点群・波形データの教師あり 6 クラス (地面・植生・建物・電線・電線・道路) 地物分類が提案された [255]. 図 2.29 に示すように, CNN による手法は 1D CNN と 2D FCN から成り立つ. 1D CNN は入力された波形データに対して個別にクラス地物分類を行い各地物に属する確率値を求める. 次に, クラスに属する確率値と高さ情報を画像化し, 画素単位のクラス地物分類を行う Fully Convolutional Network (FCN) によるセグメンテーションを行う. なお, FCN の層設計としては U-Net を使用している. 画像化する 2 段階の手法を用いる理由として, 波形データを独立して扱う 1D CNN では空間的な特徴を抽出することができないためである. したがって, 個々の波形データでは分類が困難な地物においては分類性能が悪くなる.

また, 地物分類以外の目的に深層学習手法が使用される場合がある. 波形データから点群データを作成する波形データ分解とモデル化の工程に使用されることもある [8].

波形データを用いない航空機ライダー計測で得られた点群データに対して深層学習手法を適用した研究を紹介する. Winiwarter ら [229] PointNet++ [178] を用いて公開データだけでなく, 実際の航空機ライダーで観測した点群データでも性能を評価した. さらに, タスクに特化した手法も提案されており, Dynamic Graph CNN [228] を用いた点群データの地表面フィルタリング手法も提案された [245]. 深層学習手法を点群データに適用した結果の網羅的な比較を行なった論文 [121] を参考にすると, PointNet++ が安定した地物分類性能を示した. これらの手法では, 点群データという幾何的な情報のみしか使用できないため, 幾何的な形状が類似した地物の地物分類には課題が残っている. この課題に対しては, 点群データだけでなく波形データも利用することで解決が可能である.

深層学習手法を点群・波形データに適用する際には以下の問題点が未解決である.

- 波形データ単位の局所的な波形データ情報で地物分類可能なクラスと, 波形データだけでは地物分類が困難なクラスが存在する.
- 大局的な特徴を取り込むために, 画像に変換し, 画像のセグメンテーションタスクとして解く必要があるため, 詳細な 3 次元情報が欠落する.

- 3次元的な特徴抽出と波形データの特徴抽出の相互作用を考えていないこと。

### 2.4.3 条件付き GAN の点群・波形データへの適用

点群・波形データへ条件付き GAN を適用した研究の紹介の前に、点群データに対する条件付き GAN の適用例を紹介する。点群データにおいて、上記で説明した画像の分野にて研究されている条件付き GAN を用いて失われた情報の復元を行なった例は存在する。しかしその例としては、図 2.30 に示すような点群データに対する着色を実現する 2 つの手法が提案された程度にとどまる [131, 21]。これらの手法は、条件付き GAN を用いた pix2pix[87] に基づいている。生成器は PointNet[177] を用いて各点の色を予測し、PointNet ベースの識別器は生成器からの色情報と真値の色情報を判断しようとする。これらの方法は、簡単な CAD モデルの着色で検証された。上記のように、点群データから波形データを作成する手法は提案されていないため、直接的な参考にはできない。しかし、このような条件付き GAN を用いた点群データ着色に関する取り組みを参考にできる。

次に点群・波形データへ適用した例を紹介する。現在、本研究で行うような点群データから波形データ推定を行なった手法は存在しない。しかし、点群・波形データに対しては条件付き GAN を用いた疎なサンプリング周波数の波形データから密なサンプリング周波数相当の波形データを推定する「波形データの超解像」に関する研究は行われている [129, 130]。図 2.31 に示すように、波形データの超解像は疎な波形データを入力し、深層学習モデルによって得られる出力は密な波形データである。これらの入力と出力をつなぐ生成器には 1 次元の CNN を適用する。波形データは音声等と同様に時系列のデータと定義できるため、時系列の解析に使用される深層学習手法である 1 次元の畳み込みで構成される 1D CNN を適用している。また、作成された密な波形データに対して、真値の密な波形データとの差を計算し、それを最小化するように最適化を行う。さらに、識別器を用意することで、ここの波形データの強度だけでなく、波形データ全体の整合性に関する評価を行うことを可能にした。このように、波形データに対する超解像手法は存在するが、点群データから波形データを推定する手法は存在しない。しかし、このような波形データ単位に評価する識別器の考え方は参考になると考えられる。

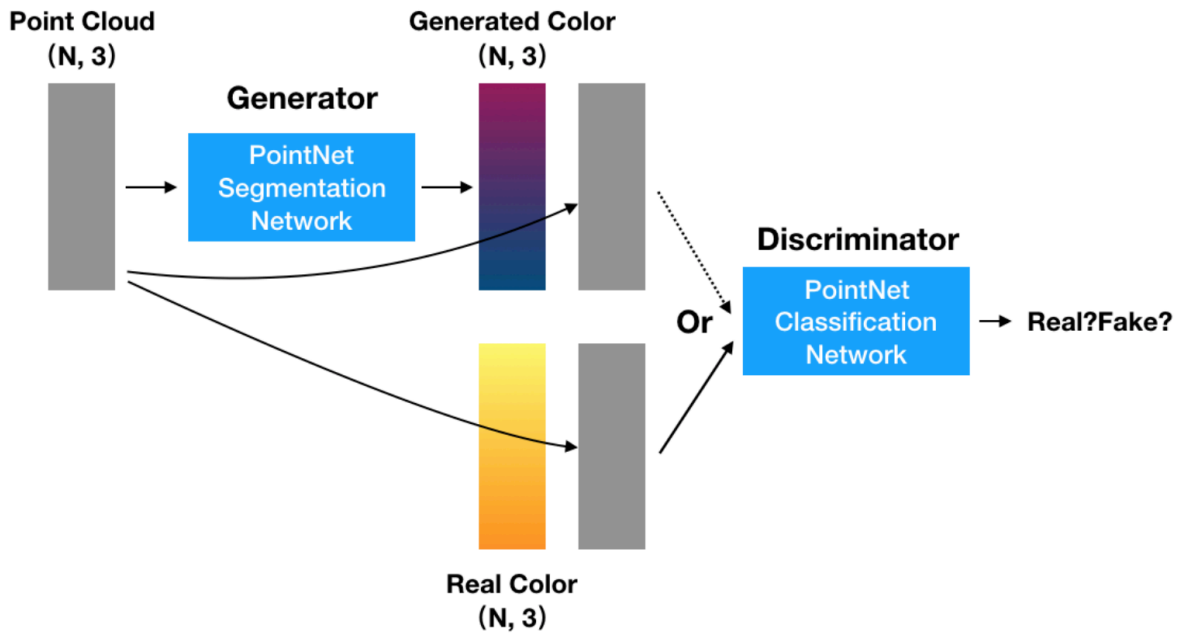


図 2.30: 条件付き GAN を点群データに適用した先行研究の例. 点群データの着色手法 [21] では点群データに対する深層学習を生成器 (Generator) として使い, 識別器 (Discriminator) においても同様に点群データに対する深層学習を用いている. 図は先行研究 [21] の Fig.5 より引用.

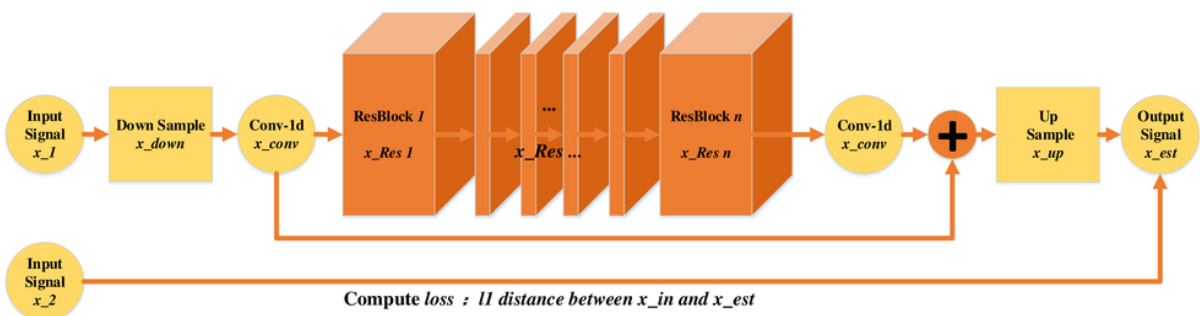


図 2.31: 条件付き GAN を波形データに適用した先行研究の例. 波形データの超解像手法 [130] では 1 次元の畳み込みを用いている. 図は先行研究 [130] の Fig.2.4 より引用.

## 2.5 既往研究と提案手法との関係

本研究では、「点群・波形データの地物分類を行う深層学習モデルの提案」と「点群データから波形データを推定する深層学習手法の提案」を行う。

### 2.5.1 点群・波形データの地物分類を行う深層学習モデルの提案

点群・波形データの地物分類に関する既往研究では、点群・波形データを特徴量に変換もしくは画像に変換し、その変換されたデータに対して地物分類を行なっている。画像に変換する際に点群・波形データが持つ幾何的な情報が欠落するため、直接点群・波形データを地物分類する手法が望ましい。しかし、点群・波形データを変換せず点群・波形データを直接扱った既往研究は存在しない。一方、点群データに対する深層学習手法は点群データを画像等に変換することなく直接地物分類が可能である。そこで、点群データに対する深層学習を参考した点群・波形データを変換せずに地物分類を直接実行できる手法を新たに提案する。

### 2.5.2 点群データから波形データを推定する深層学習手法の提案

点群データからの波形データ推定は「失った情報の推定」とであると定義できる。そこで、白黒画像からのカラー化や、低解像度の画像から高解像度の画像推定などの失った情報の推定に用いられる条件付き GAN に着目した。

点群データに対して条件付き GAN を適用した既往研究として、点群データからその点を持つ色情報を推定する手法が提案されている。また、波形データに対して条件付き GAN を適用した既往研究として、低いサンプリングレートで観測した波形データから高いサンプリングレートで観測した波形データ推定する手法が提案されている。しかし、本研究で対象にするような波形データを点群データから推定する手法は存在しない。そこで、点群データから波形データを推定する深層学習手法を新たに提案する。

## 第 3 章

### 点群・波形データを用いた地物分類手法の提案

### 3.1 はじめに

点群・波形データの地物分類として、点群・波形データから特徴量を設計しその特徴量を機械学習手法に入力する研究や、点群・波形データを画像化し画像に対する深層学習手法を適用する研究が行われている。しかし、点群・波形データを画像等に変換する際に詳細な3次元情報が欠落するため、直接点群・波形データを深層学習モデルで扱う手法が重要になる。そこで本章では、図 3.1 に示すような点群・波形データの地物分類を実現する深層学習手法を提案する。

次に、本章の概要を述べる。本章の流れは、まず点群・波形データ地物分類の問題設定を説明する。次に点群データに対する既存の深層学習手法を波形データに適用する際の仮説を立てる。仮説としては、波形データと地物の関係から波形データを2つクラスに大別できると仮定した。具体的には点群・波形データの特性に鑑み、地物を波形データの形状で分類が可能な地物（局所クラス）と波形データの空間的な関係性である大局情報を用いて分類が可能な地物（大局クラス）に大別する。次に、仮説に従い本研究における提案手法である点群・波形データに対する深層学習手法の提案を行う。提案手法は、局所クラスに特化した層設計と大局クラスに特化した層設計を組み合わせる手法（局所・大局モデル）である。次に、ベンチマークデータを用いた実験を行い、提案手法と先行研究の画像化する深層学習手法を比較する。ベンチマークデータによる局所・大局モデルの学習と学習済みモデルの検証を実施し、検証実験の結果、点群・波形データを画像に変換することなく、既存手法よりも高い地物分類性能を示した。さらに、提案手法における考察を実施する。具体的には、局所ブロックの効果、波形データの効果検証、ノイズに対する頑健性の評価、判断根拠の可視化、点密度に対する堅牢性の評価を実施する。最後に、本章のまとめを行う。

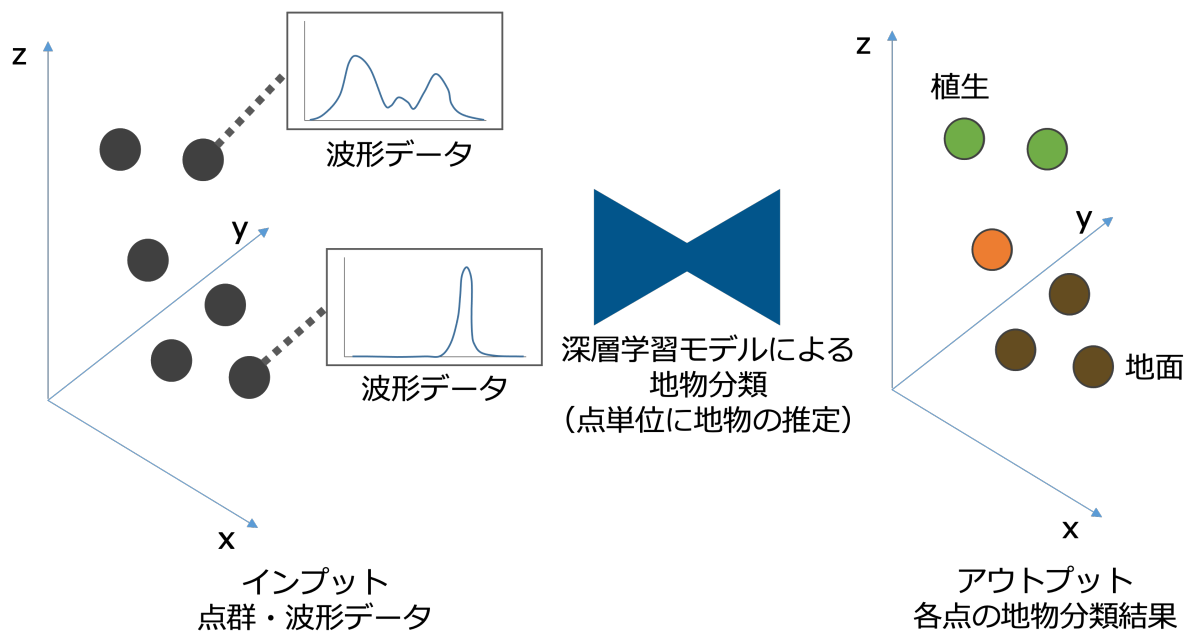


図 3.1: 点群・波形データの地物分類の概念図。点群・波形データを深層学習モデルへ入力し、各点に対する地物分類を実行する。

### 3.2 点群・波形データの地物分類の問題設定

まず、本章の点群・波形データを用いた地物分類についての問題設定を定義する。図 3.2 は本章における問題設定、つまり点群・波形データの地物分類タスクを示す。点群・波形データ  $\mathbf{P}$  が与えられたとき、 $\mathbf{p}_i \in \mathbb{R}^M$  を持つ  $\{\mathbf{p}_i\}_{i=1}^K$ 、入力データは  $\mathbf{P} \in \mathbb{R}^{K \times (3+M)}$  の行列として表現できる。ここで  $3+M$  は入力する特徴次元の数（つまり  $x, y, z$  と波形データの反射の強度が格納された数）、そして  $K$  は深層学習モデルに入力する点数である。入力データは二次元配列として表現され、行方向は  $x, y, z$  と波形データの反射の強度、列方向は一度に入力する点数である。地物分類を実行する深層学習モデルは、 $\mathbf{P}$  から  $\mathbf{S}_{seg} \in \mathbb{R}^{K \times C}$  への写像を獲得する。ここで、 $\mathbf{S}_{seg}$  地物分類の結果を示し、 $C$  はクラスの数である。

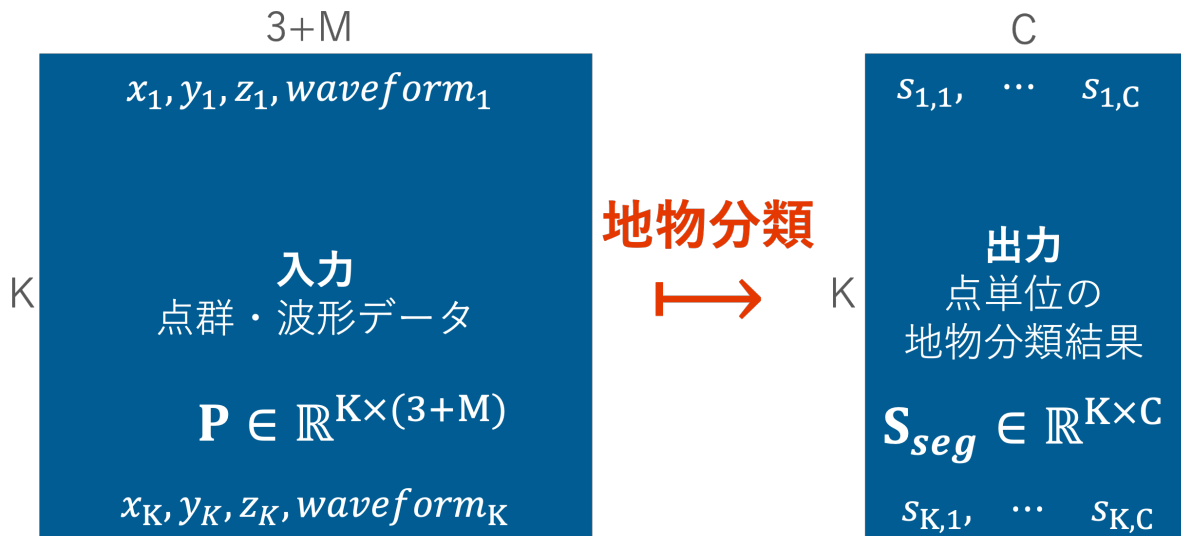


図 3.2: 点群・波形データに対する地物分類の問題設定。点群・波形データの地物分類タスクは点群・波形データ  $\mathbf{P}$  が与えられたとき地物分類の結果  $\mathbf{S}_{seg}$  を出力する。

### 3.3 点群・波形データの地物分類を行う上での仮説

点群・波形データの地物分類に特化した手法を提案するために、まず、点群・波形データの特性と地物の関係について整理する。いくつかの地物において、波形データの形状としてその地物の特性が顕著に現れることがある。一般的に、道路や草地に比べ、樹冠や耕作地では反射エコーの幅が広がる性質がある。また、草地や裸地では高い振幅が見られ、建物の屋根では屋根の材質に応じて振幅が変化する。波形データにおける物体の散乱特性を表す後方散乱係数などの属性は、地物の物理的特性を把握し、分類性能を向上させるのに有効である [219]。この際、エコー形状のパラメータを利用するためには、波形データの  $n$  番目のエコーの形状が  $n - 1$  番目のエコーの散乱特性に依存することを念頭に置く必要がある。 $n$  個目の地物の断面積は、その地物とセンサの距離に応じて減少することが知られている。しかし、現実の地物は複雑に配置されているため、どのように減少するかを予測することは容易ではない [220]。このように、点群・波形データの波形データを用いることで、地物の地物分類が容易になる場合があるが、個別の波形データだけでは複数の地物を地物分類することは難しい場合もある。道路や建物の屋根などは、場所によっては部材が変わるなどの状況が考えられ、材質がすべて同じであるとは限らないため、その波形データの特性も多種多様になることが多い。さらに、道路や建物の屋根においては、波形データから点群データを作成する際に抽出された特徴量は裸地等の自然物に近い特徴量を持つことが多いと報告されている [64]。その結果、経験的な閾値を用いたルールベースアルゴリズムなどの単純な地物分類アルゴリズムでは、個別の波形データに対しては高い頻度で誤った地物分類が行われる [48, 47]。このように個々の波形データだけでは誤分類が発生する地物に対しては、点群データ作成の際に計算される波形データの特徴量だけでなく、周囲の状況を考慮するルールベースアルゴリズム手法が提案されている。具体的には、円筒状の環境下で各点の強度値を用いて計算した共分散行列の固有値に基づいた手法が提案された [64]。このように、地物分類には波形データの特徴に加えて空間的な特徴が重要であり、広い範囲の大局的な波形データの特徴を用いることで地物分類が可能になる場合がある。

波形データの形状を変化させる要素については、これまでに複数の著者によって主にシミュレーションや近距離のデータを用いて行われてきた。Swedish Ministry of Defense Research Institute は、包括的なレーザ反射のシミュレーションモデルを提案した。このシミュレーションモデルには、検出器の特性、物体の形状、および反射率、大気の減衰などが含まれている。主な評価としては、物体の形状とラジオメトリの影響が混在しており、単純にシミュレーションはできないと結論づけた。この論文で得られた知見をもとに、影響を及ぼす要因について整理する。

- 入射角度  
照射されたレーザが物体へ当たる際の入射角が変わると、波形データの形状の変化が顕著になる。一般的に入射角度が小さいほど、観測される波形データのピークは狭く、そのピークの形状は対称的になる。
- 物体表面の粗度  
滑らかな表面と荒い表面とでは、異なる反応が見られる [214], [80]。滑らかな表面では不連続性がある場合は1つまたは2つのエコーが発生、荒い表面では樹木や植生の挙動に相当する、異なる深さでの複数の反射エコーが発生する。また、反射エコー幅は物体表面の荒さが影響している。
- 対象物の幾何形状  
照射されるレーザの反射エコーはレーザが当たる物体の幾何形状にも依存する。この反射について、様々な都市において存在しうる地物において、振幅の入射角依存性を測定した研究がある [108]。さらに、[203] や [22] では、単純な形状の物体を用いた実験を行い、反射エコーに対する形状と物体の反射率の影響を調査した。入射角が大きくなると、パルスの広がりとその振幅の減少が表面上で観察された。これにより、特に入射角が大きい場合に、距離を誤って推定してしまう可能性がある。また、平面とエッジの間では似たような波形データが観察される。その結果、2つの平面を1つの平面として分類す

る可能性がある。最終的には、近接した2つの平面を分離できるか否かを調べるために、段差のある物体でシミュレーションを行った。様々な形状で反射した波形データを比較すると、1つのパルスからの波形データを解析することで物体を分類できるというメリットを示した。ただし、これらの結論は、対象物がランバート反射を持つという大前提に基づいている。

- 物体の部材

様々な都市の素材がリターン波形データに与える影響を評価する実験を行った研究がある [93]。観測される波形データに対する素材の影響を調べるために、屋根材を対象として、石造り、波板、スレート板、瓦で反射をシミュレーションした。テストした4種類の屋根材は、それぞれ異なる挙動を示した。異なるクラスの物体間では、部材によって大きなばらつきが確認できるため、波形データだけでは植生・建物・道路を単純に分類できないことが確認された。この部材の影響に研究は、超高解像度の航空写真を用いた様々な都市域における部材（アスファルト、コンクリート、天然の石材など）の反射率の変化の研究 [109] と同様である。

- 距離

ライダーの計測原理より明らかではあるが、ライダーと物体の距離によって波形データの形状は変化する。受信電力  $\frac{1}{R^2}$  は、ライダーから物体までの距離を  $R$  とした関数である。このように、計測方法や現地の地形が受信波形データに影響を与える。

- その他の要因

その他の波形データに対して影響を与える要因として、大気と鏡面反射の影響は存在する。しかし、他の影響に比べて無視できる [22]。

上記のことを踏まえて、点群・波形データを深層学習モデルを用いて地物分類する際に、特徴的な波形データを有する地物に対しては波形データのみを用いて地物分類が可能になる場合と、点群データの幾何的な情報を有効活用し波形データと波形データの大局的な特徴を用いることで地物分類が有効となる場合が存在すると仮説を立てる。具体的には、地面や建物などの平面的な地物の波形データにおいてはひとつの強いピークが観測されるため、個々の波形データだけでは地物分類が困難であると考えられる。これを克服するには、広い範囲の波形データを大局的に扱い、波形データの空間的な関係性を用いて地物分類をする必要がある。例えば、建物の波形データと地面の波形データは形状が似ているが、平面な場所の上に浮いた独立した別の平面があればそれは建物である可能性が高い。このような相対的な位置関係を大局的に獲得すると地物分類が容易になる。次に、植生等の個々の波形データが特徴的な形状を示す場合を考える。この場合は、波形データのみを使用して地物の地物分類が可能であり、局所的な波形データに対する特徴抽出で充分となる。そこで、点群・波形データの地物分類を行う際に地物は2種類に大別する。

- 局所クラス

波形データが特徴的な形状を有しているため、個々の波形データのみで地物分類可能。

- 大局クラス

点群データが持つ座標値による空間的な情報と波形データを組み合わせることで、地物分類可能。

この、局所クラスと大局クラスそれぞれに特化した深層学習モデルを組み合わせた点群・波形データの地物分類手法を提案する。

このような波形データに対する仮説に対して、既存の点群データの深層学習手法をそのまま点群・波形データに適用した場合、以下のような問題点が生じる。点群データに対する深層学習手法は、大局的な特徴の関係性を用いて地物を分類する。そのため、個々の点群データに対応する波形データの特徴を有効活用できない可能性がある。これは、点群データに対する深層学習手法は屋内で取得した点群データを対象にすることが多く、建物の壁面や床を認識することが重要であるため、小さな物体よりも大きな物体を優先的に認識させるために

提案されてきたためである。そこで、既存の点群データの深層学習手法を適用するだけでなく、個々の波形データを用いた地物分類を実行可能な手法を組み合わせることにした。

### 3.4 局所・大局モデルの提案

本節では、点群・波形データの特性を考慮した深層学習モデル（局所・大局モデル）による地物分類手法を提案する。手法の提案する際に、個々の波形データで分類可能な地物（局所クラス）と波形データと波形データの空間的な関係性である大局的な特徴を用いて分類可能な地物（大局クラス）に大別できると仮説を立てた。そこで、大局クラスに対して既存の点群データに対する深層学習手法を点群・波形データに適用した大局ブロック、局所クラスに対して波形データの個々の特徴に着目する局所ブロックを組み合わせた深層学習モデルを提案する。以降で、提案する局所・大局モデルの層設計、最適化に用いる損失関数、最適化アルゴリズム、そして予測時のアルゴリズムを説明する。

#### 3.4.1 提案する局所・大局モデルの概要

図 3.3 は提案手法である点群・波形データの地物分類に特化した新たな深層学習モデルである局所・大局モデルの概念図を示す。提案する局所・大局モデルは、波形データに対して局所的な特徴抽出を適用して地物分類を実行する局所ブロックと、階層的な特徴抽出によって大局的な特徴抽出が可能なエンコーダデコーダ構造を持つ大局ブロックによって成り立つ。

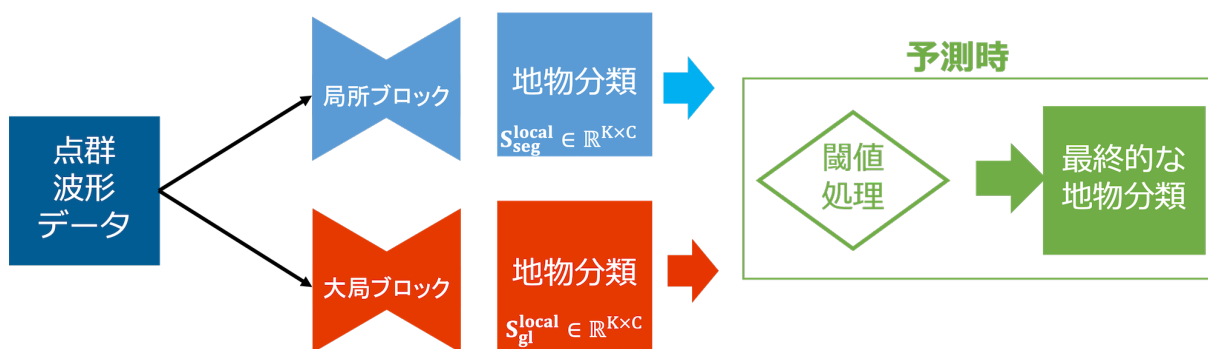
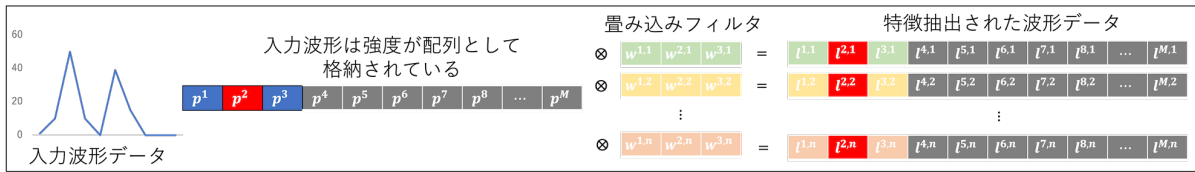


図 3.3: 提案する局所・大局モデルの概念図。局所・大局モデルは局所ブロックと大局ブロックによって構成される。局所ブロックは各波形データに対して独立した特徴抽出を実行し地物分類を行う。大局ブロックは波形データの近傍特徴抽出を階層的に実行する。予測時には局所ブロックと大局ブロックからの出力を組み合わせる。

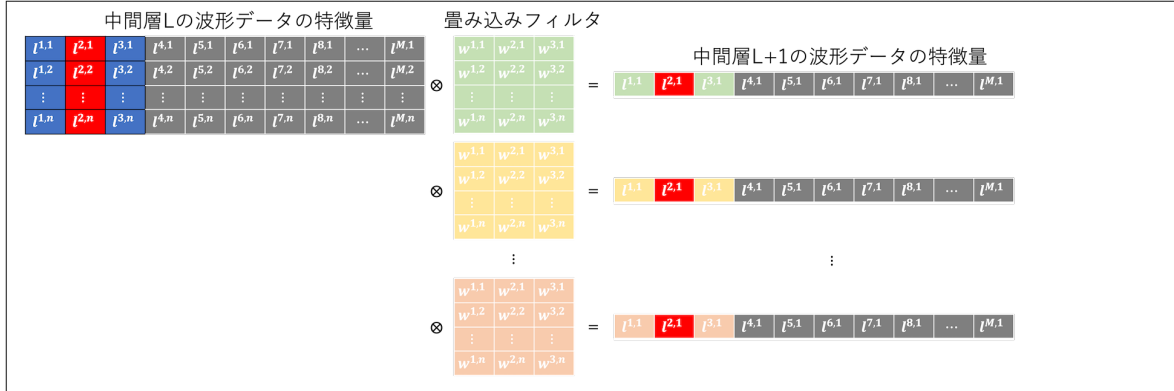
#### 3.4.2 局所ブロック

個々の波形データに対して特徴抽出し、その特徴を用いて地物分類を実施する局所ブロックを説明する。この局所ブロックでは、個々の波形データ単位に 1D CNN を適用する。この  $M$  次元の入力に対して図 3.4 に単純な 1 次元の畳み込みである  $1 \times 3$  の学習可能な重みを持つフィルタを複数回適用する。このフィルタ演算を図 3.5 の赤色の三角形で示すように複数回適用することで、波形データの特徴を段階的に抽出することが可能である。最終的には、図 3.5 の緑色で示された抽出された特徴量を用いて  $C$  個の地物それぞれに属する確率を計算する。このようにして得られる出力は、 $K \times C$  次元の地物分類結果 ( $S_{seg}^l$ ) である。なお、局所ブロックを設定するにあたり、音声データに 1D CNN を適用した先行研究手法 [236] を参考にした。

入力層の波形データに対する畳み込み



中間層の波形データの特徴量に対する畳み込み



分類層の演算 (6クラスの例)

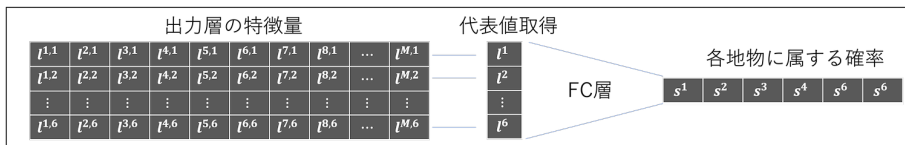


図 3.4: 局所ブロックの波形データに対する特徴抽出. 入力された波形データに対して  $1 \times 3$  の学習可能な重みを持つ畳み込みを適用する.

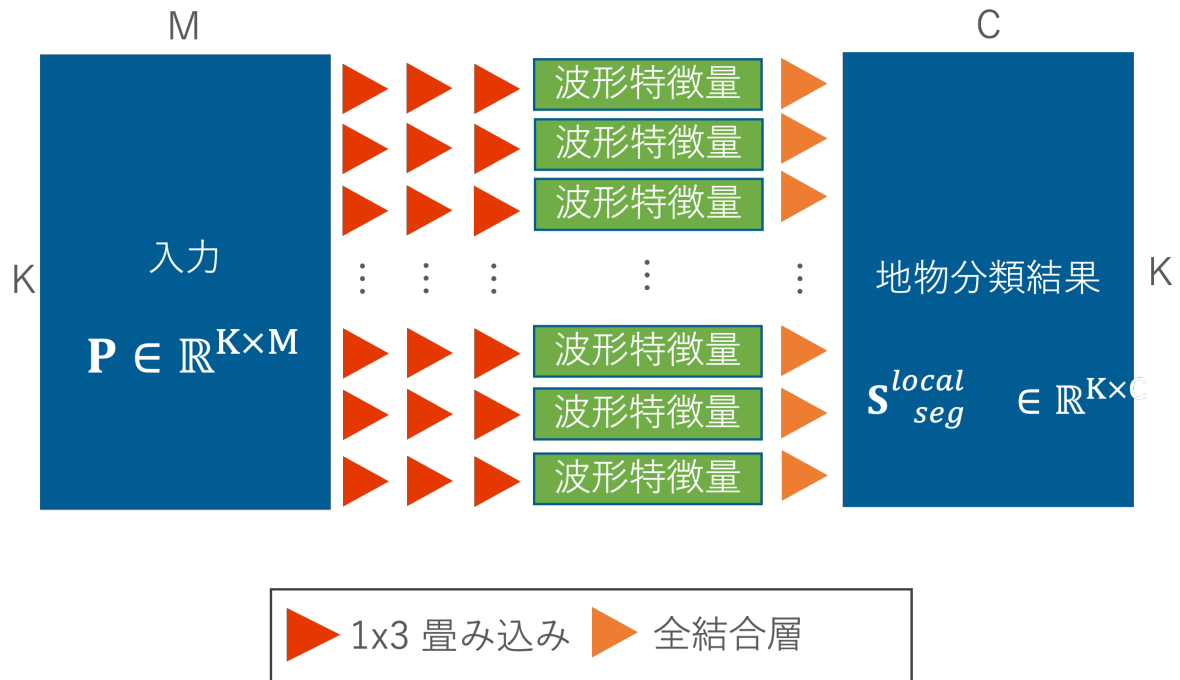


図 3.5: 局所ブロックの層設計. 波形データに対する畳み込み演算を複数回適用し, 入力された波形データの地物分類結果を出力する.

### 3.4.3 大局ブロック

ここでは、既存の点群データに対する深層学習手法である PointNet++[178] に対して点群・波形データを扱えるように拡張した大局ブロックを説明する。この PointNet++ を波形データを扱えるように拡張する際には、波形データに対する近傍特徴抽出と、近傍特徴抽出の内部で使用される個々の波形データの特徴抽出を定義する必要がある。そのため、まず、近傍特徴抽出と個々の波形データの特徴抽出を説明する。次に、近傍特徴抽出を階層的に適用する階層的な特徴抽出手法について説明する。最後に層設計の詳細な設定について説明する。

- 波形データに対する近傍特徴抽出

まず、波形データに対する近傍特徴抽出を説明する。波形データに対する近傍特徴抽出は、波形データと波形データの近傍関係の特徴抽出器として使用する。近傍特徴抽出手法として、点群データ処理に使用される演算を参考にしているため、まずは点群データの演算を説明する。点群データの近傍特徴抽出手法はある点に対してその近傍に存在する点群データをグループ化して、このグループ化された点群データにおける特徴を抽出する手法である。形式的には、点群データの近傍特徴抽出は次のように定義される。

$$(f * g)(\mathbf{p}_i) = \int_{-\infty}^{+\infty} f(\mathbf{p}_j) \cdot g(\mathbf{p}_i - \mathbf{p}_j) d\mathbf{p}_j, \quad (3.1)$$

モンテカルロ積分を用いることで、連続的な演算を離散化できる。

$$(f * g)(\mathbf{p}_i) \approx \frac{1}{N} \sum_{n=1}^N f(\mathbf{p}_n) \cdot g(\mathbf{p}_i - \mathbf{p}_n), \quad (3.2)$$

無限次元のカーネル関数  $g(\cdot)$  は学習可能な重みを持つ関数である。そしてこの関数はニューラルネットワーク ( $\mathcal{NN}$ ) で実装されることが多い:

$$g(p; \theta) = g(\mathcal{NN}(p; \theta)), \quad (3.3)$$

ここで、 $\theta$  は学習可能な重みを示し、 $g$  は max-pooling、 $\mathcal{NN}$  はニューラルネットワークである。具体的に、大局ブロックで参考にする PointNet++ においては、近傍特徴抽出手法を以下の方法でニューラルネットワークを用いて近似する。まず、与えられたある点に対する近傍の点群データに対して、独立した全結合型の演算  $h$  と活性化関数 ReLU によって非線形変換を実行する。その後、非線形変換された点に対して、max-pooling を適用し、入力全体の特徴量を算出して、近傍の点群データにおける代表値を得る。max-pooling は対称関数であり、入力点の数を  $K$  としたときに、 $K!$  の入力順番に変動があった際に、入力順序に依存せず同じ出力が得られる関数である。つまり、近傍の点群データというある限られた領域に対して、PointNet を適用して、特徴量を計算している。これを定式化すると、以下のように書ける。

$$p^{l+1} = \text{MaxPooling}(\text{ReLU}(h(p_0^l), (h(p_1^l), \dots, (h(p_{n-1}^l))))), \quad (3.4)$$

$p$  は入力点、点の添え字は入力する点の個数に対応、 $l$  は何層目であるか、MaxPooling は対称関数 (max-pooling) を示す。この演算によって各点で独立した特徴抽出だけでなく、入力データ全体を横断した特徴抽出も可能になる。

次に点群データに対する近傍特徴抽出の演算を波形データに対応できるように拡張する方法について説明する。点群データの近傍特徴抽出において、点単位の非線形な特徴抽出と max-pooling が用いられる。これを波形データに対して拡張するためには、max-pooling 演算は変更できる要素がないため、点単位の特徴抽出を波形データに対して適切な演算に変更することが重要になる。ここで、波形データに対する特徴抽出を提案する際に波形データというデータ構造に着目する。波形データは照射したレーザーが物体に反射して戻ってきた反射の強度が時系列に格納されたデータであることに着目す

る。時系列のデータであることから、音声分野と近い関係であることが推察される。そこで、音声処理で用いられる 1 次元の畳み込み演算を波形データに対する特徴抽出として用いることにした。図 3.6 に示す、波形データに対する 1 次元の畳み込み演算つまり 1D CNN は  $1 \times 3$  の学習可能な重みを持つフィルタとして実装する。このように点群データの座標値を使用してグループ化された波形データに 1D CNN を適用してその後グループ内の波形データから代表値を max-pooling にて獲得する。1D CNN の詳しい説明は次で説明する。

個々の波形データに対して適用される 1D CNN は畳み込み演算を使用し、図 3.4 のように定義できる。まず、入力された波形データに対して、 $1 \times 3$  の畳み込みフィルタを適用する場合を考える。

まず、入力された波形データに対してどのように畳み込みが適用されるかを説明する。入力層では、 $1 \times 1 \times M$  の配列として波形データは扱われる。これに対して、畳み込みを適用すると、ある波形データの値に対して、その両側の値を用いて畳み込みを行う。この際、端の値を畳み込みする際には、ゼロで埋める。このような畳み込みを全ての配列の値に対して畳み込みフィルタを動かしながら適用する。なお、様々な特徴を獲得するために、この畳み込みフィルタは 1 つだけでなく 64 個などある程度の量を設定する。この演算によって、入力された  $1 \times 1 \times M$  の波形データに対し、前述の 64 個の畳み込みフィルタである場合、 $64 \times 1 \times M$  の特徴マップが作成できる。

次に、入力ではなく、隠れ層における畳み込み演算を説明する。隠れ層  $l$  では、 $l-1$  層の畳み込み演算によって 128 個の特徴マップが作成された場合、 $128 \times 1 \times M$  の特徴マップである。これに対し、 $1 \times 3$  の畳み込みフィルタを適用する場合を考える。入力層とは違い、128 個の特徴マップがあるため、 $128 \times 1 \times 3$  の畳み込みフィルタになる。つまり、波形データの時系列方向と、特徴マップのチャンネル方向に対して演算を行うことと等しい。これは、様々な畳み込みフィルタで作成した特徴を混ぜつつ、波形データの特徴を抽出できる効果がある。このフィルタを例えば 256 個用意するとすると、 $l+1$  層に対しては、 $256 \times 1 \times M$  の特徴マップを送る。

なお、これらの畳み込みフィルタは学習可能な重みを持つため最適化によって得られる。つまり、地物分類に寄与するように波形データの特徴を自動的に抽出できることを示す。

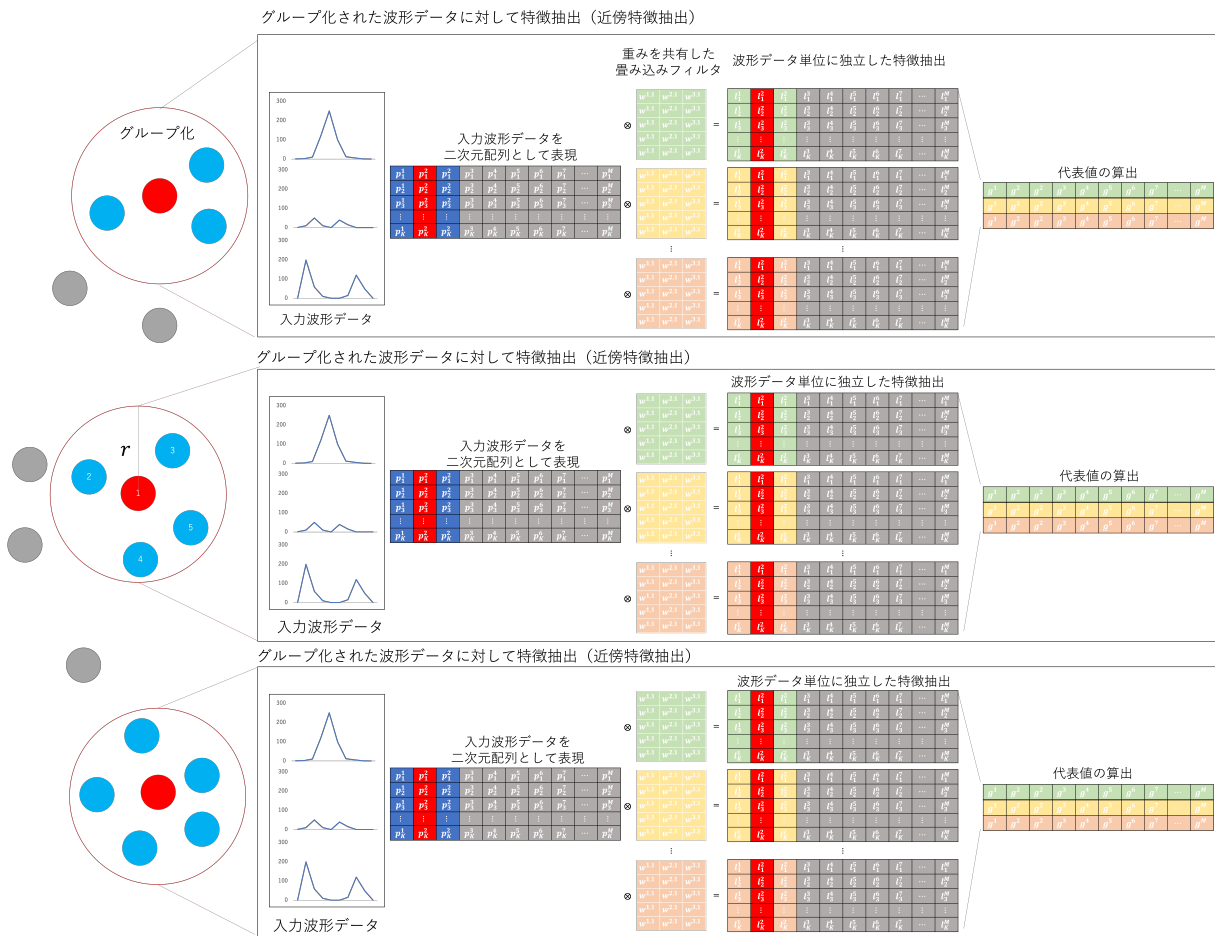


図 3.6: 波形データに対する近傍特徴抽出の仕組み 近傍特徴抽出では、ある波形データの近傍の波形データをグループ化し、波形データに対して独立して波形に対する特徴抽出 ( $1 \times 3$  の畳み込み演算) を適用し、グループ内の畳み込み演算の結果から代表値を得る。

- 階層的な特徴抽出

次に階層的な特徴抽出について説明する。階層的な近傍特徴抽出は、画像認識タスクにおける CNN が成功を取めた重要な技術である。そこで、点群データの深層学習手法に対しても CNN を做って同様に階層的な近傍特徴抽出が取り入れられた。CNN で用いられている階層型の近傍特徴抽出では、周辺画素から得られる特徴量の積み重ねによって、抽象化された高レベルの特徴量を大きな受容野から得られる。このような大局的な背景情報は相対的な波形データの関係性が重要な地物の分類を可能にするため、十分な受容野を持つ深層学習モデルを設計することが不可欠である。なお、プーリングは CNN のために広い範囲の情報を収集するためのダウンサンプリング手法である。提案手法における階層的な特徴抽出は、ダウンサンプリングによる代表点の選択、代表点の近傍点の波形データをグループ化、グループ化した波形データに対する近傍特徴抽出によって成り立つ。以降でそれぞれに対する説明を行う。

まず、ダウンサンプリング手法について説明する。ダウンサンプリング手法としては PointNet++ で使用されている [178] 最遠点サンプリング (FPS) を用いる。図 3.8 に示すように、FPS はある点を初期点として、点と点の間のユークリッド距離が最大になる点を選択することを繰り返して全点からの代表点を決定する手法である。FPS は、まず初期点をランダムに決定する (図 3.8 の初期点)。次に、その初期点から最も遠い点を選択する (図 3.8 の 2 点目)。この選択された点を起点として同様に最も遠い点を選択する (図 3.8 の 3 点目)。そして、この新たに選択された点を起点として最も遠い点を選択 (図 3.8 の 4 点目) する過程を繰り返し適用することで、ダウンサンプリングを行い全点からの代表点を得る。ダウンサンプリングを定式化する。入力された点群・波形データ ( $\mathbf{P} = \{\mathbf{x}_1, \dots, \mathbf{x}_K \mid \mathbf{x}_i = (x_i, y_i, z_i, \mathbf{p}_i)\}$ ,  $x_i, y_i, z_i$  は 3 次元座標,  $\mathbf{p}_i$  は波形データを示す) を扱う。各ダウンサンプリングからの出力は新しい集合 ( $\hat{\mathbf{P}} \subseteq \mathbf{P}$ , ダウンサンプリングされた点の数である。言い換えると、 $i$  番目のダウンサンプリング ( $K_i$ ) は一つ前の層の数  $K_{i-1}$  よりも少なくなる。新しい点の集合 ( $\mathbf{P}_i$ ) を得るため、FPS によるダウンサンプリングを行う。

次に、グループ化を説明する。このような FPS でダウンサンプリングを行って得た代表点に対し、グループ化を実行する。まず、代表点とダウンサンプリングする前の点を用いて代表点に対する近傍点をボールドクエリにて定義する。なお、このボールドクエリは半径  $r$  の球であり、この球内に入る点を近傍点とする。なお、この半径の値はハイパーパラメータとなる。また、グループ化する対象の数もハイパーパラメータとして与えられる。

次に、波形データに対する近傍特徴抽出を説明する。波形データに対する近傍特徴抽出はグループ化した波形データに適用する (図 3.7 の上図)。この近傍特徴抽出によって、各代表点にそのグループ内の波形データに対する特徴量が格納される。

これらの、FPS による代表点の算出、代表点の近傍点の波形データをグループ化、そしてグループ化した波形データに対する近傍特徴抽出を繰り返し適用することで、徐々に広い範囲の大局的な特徴を抽出できる。

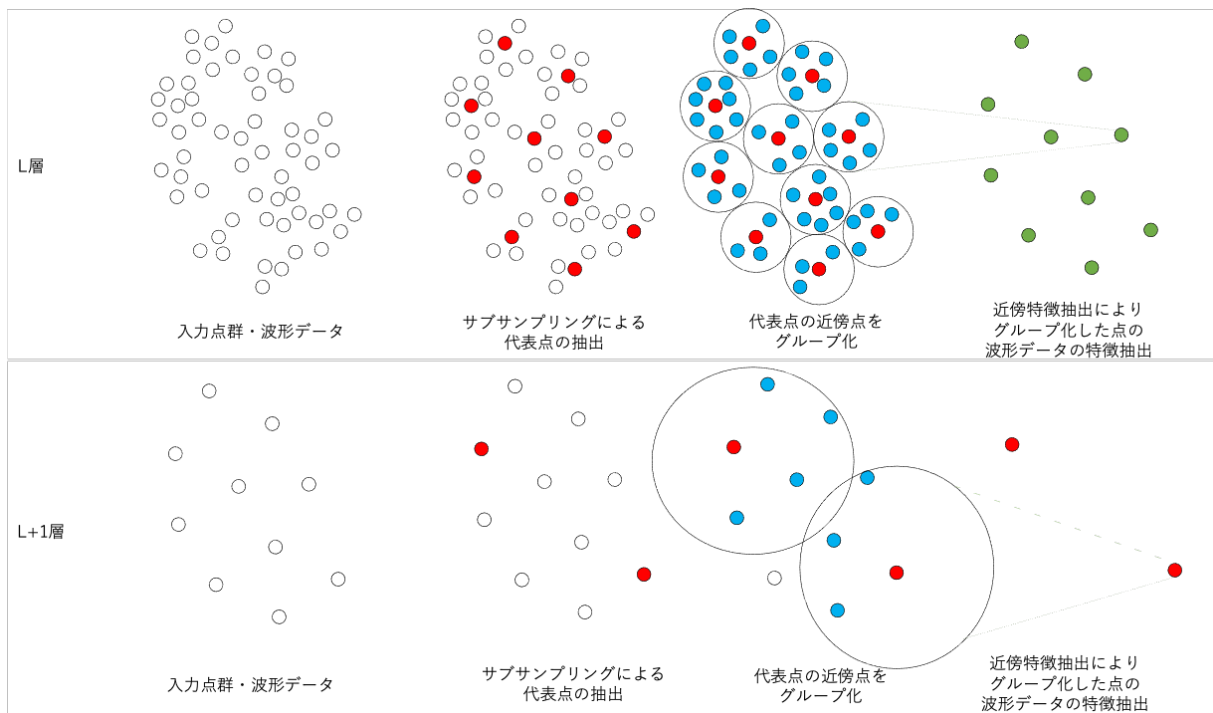


図 3.7: 階層的な特徴抽出を行うためのダウンサンプリングとグループ化。近傍特徴抽出を適用する際には、点群・波形データの代表点を決定する。そして代表点に対してその近傍の波形データをグループ化し近傍特徴抽出を行う。近傍特徴抽出によって、代表点に特徴量が集約され、その代表点以外の点を除去する。これを繰り返して適用することで、階層的に波形データの特徴抽出が可能になる。

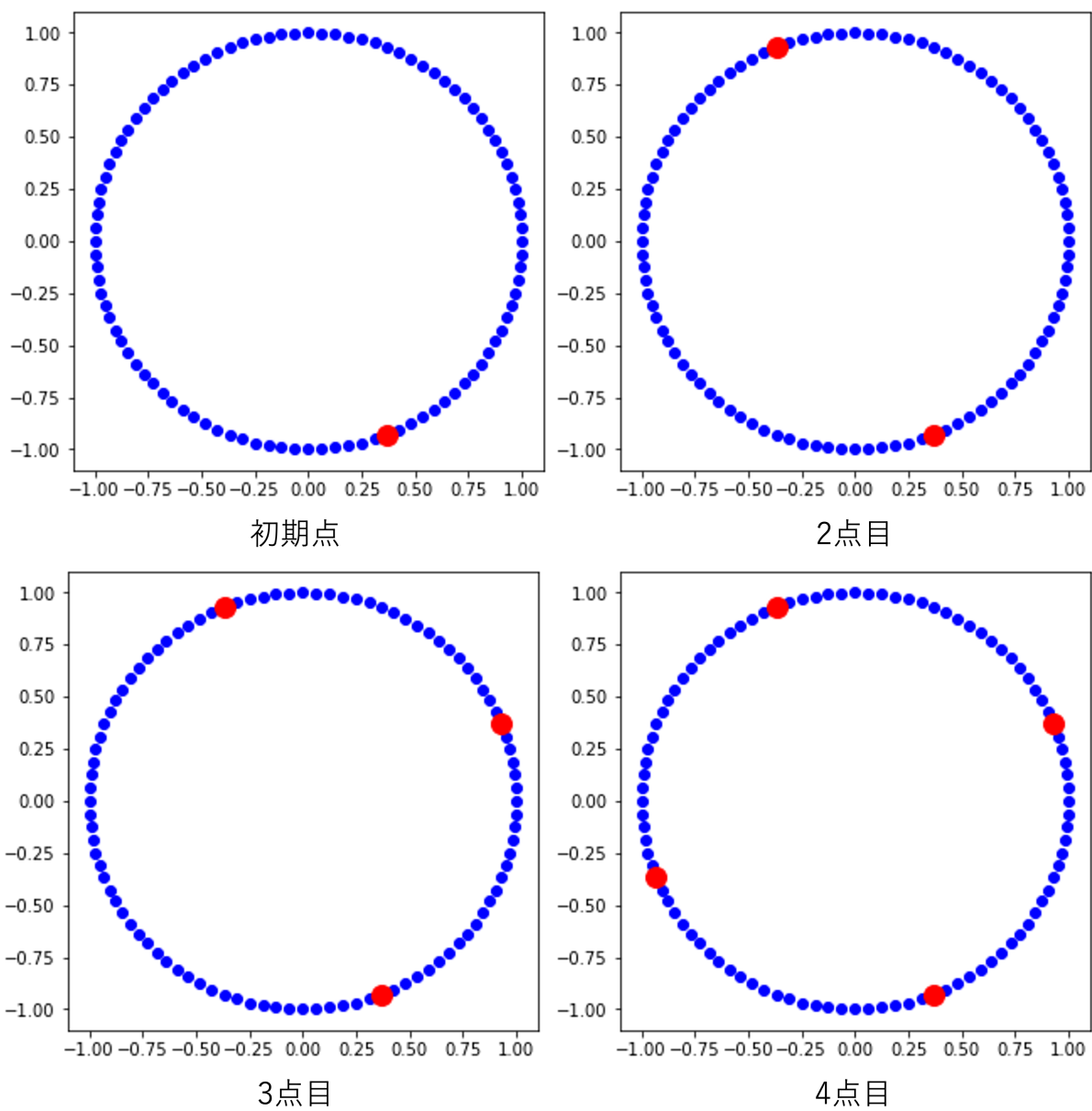


図 3.8: 最遠点サンプリング (FPS) の二次元平面上の例. ある点を初期値とし, その点から最も遠い点を設定した個数に到達するまで選択し続けるアルゴリズムである. 図の場合は 4 点を選択する場合を示す.

最後に、階層的なアップサンプリングについて説明する。階層的な特徴抽出により、大局的な波形データ特徴を獲得する機能を獲得した。しかしながら、代表点抽出を繰り返したことで入力した点数より少ない点数になったため、それを元の点数に復元する必要がある。そこで、図 3.9 のような特徴量の近傍点への内挿によるアップサンプリングを行う。ダウンサンプリングされた点群・波形データを入力し、そのダウンサンプリングが適用される前の点群・波形データをコピーする。そして、コピーした点群・波形データに対して、代表点における点群・波形データの特徴量を伝播させる。このアップサンプリングを定式化する。アップサンプリングにおいては二つの異なる次元数の集合を取り扱う。 $P_i$  で 2 つ目は  $P_{i-1}$  であり、それぞれ  $i$  番目と  $(i-1)$  番目のダウンサンプリングの出力である。対応するダウンサンプリングと同じ点数の点群データ ( $\bar{P} \supseteq P_i$ , この際  $P_{i-1}$  と同じ点数) を出力する。アップサンプリング手法としては、3次元内挿を行う。これは  $P_{i-1}$  から既知の  $P_i$  座標を用いることで可能となる。

このような階層的に点数下げつつ特徴抽出を実行し、その後入力点数へ特徴量を伝播させる層設計をエンコーダ-デコーダ構造と呼ぶ。

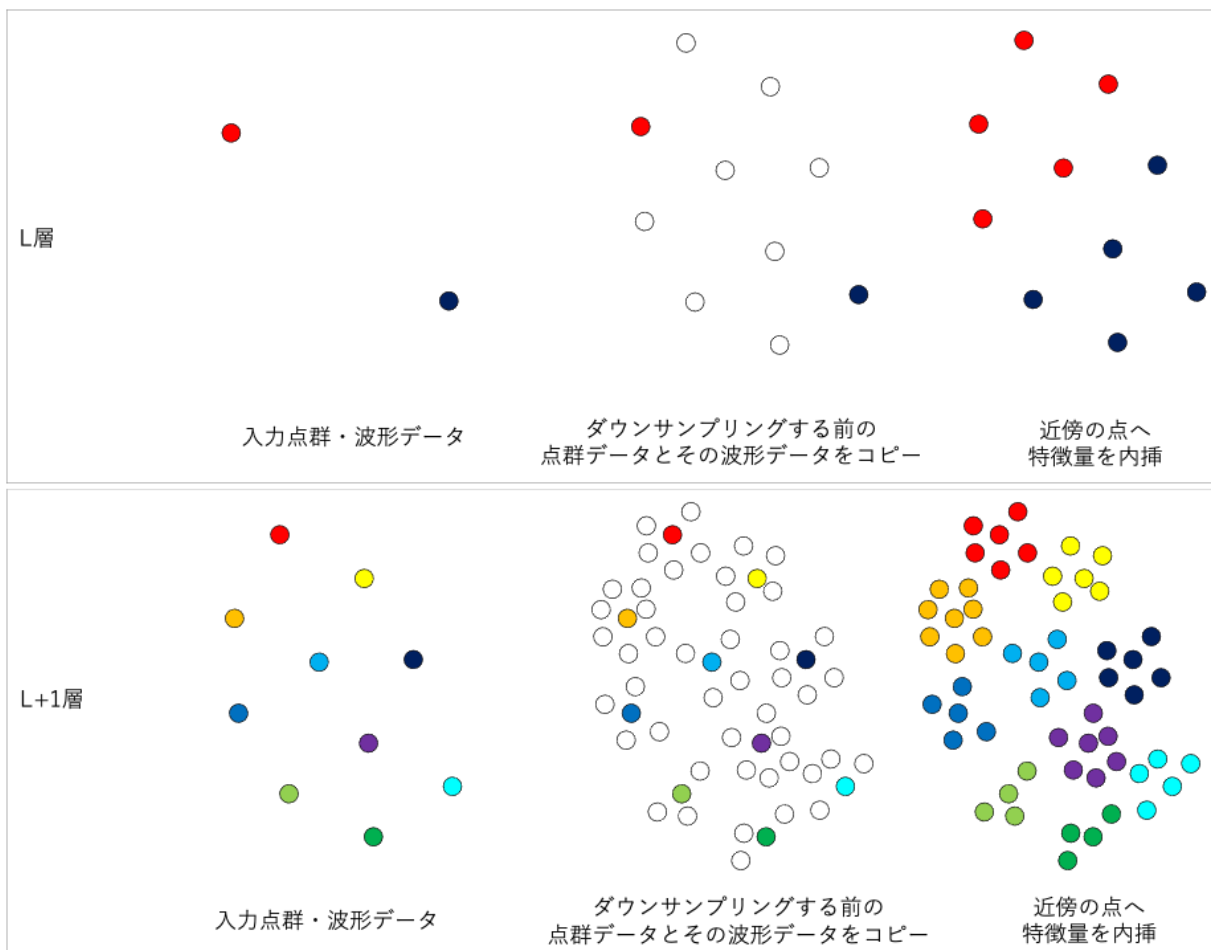


図 3.9: デコーダにおける 3 次元内挿。エンコーダで抽出された特徴を徐々に入力点群・波形データと同じ点数まで復元する。その際には、ダウンサンプリングする前の点群・波形データの情報を用いて各点群・波形データの近くにある代表点として選択されなかった波形データに特徴量を 3 次元内挿で伝播させる。

- 層設計

最後に大局ブロックの層設計について説明する。層設計とは、上記で説明した、近傍特徴抽出とダウンサンプリング、アップサンプリングの組み合わせ方法のことである。図 3.10 は大局ブロックの層設計を示す。大局ブロックは階層的な特徴抽出が可能なエンコーダ-デコーダ構造を持つ層設計であり、複数の波形データの空間的な関係性を学習しながら地物分類が可能である。図 3.10 において、近傍特徴抽出を繰り返し適用するエンコーダを図 3.10 の左側で表現している。また、デコーダは図 3.10 の右側で示している。

まず、エンコーダについて説明する。エンコーダには  $K$  個の点群・波形データを入力する。エンコーダの構造は図 3.10 の三角印で示すように 3 回の近傍特徴抽出で構成される。近傍特徴抽出の際には「FPS による代表点の決定」と「グループ化」が適用されるが、図 3.10 中においては簡略化するため単に近傍特徴抽出と記述している。FPS を適用する際の代表点の点数は 8,192 点、4,096 点、2,048 点へ段階的に点の数を減少させつつ大局的な特徴を得る。代表点に対するグループ化は近傍の点の定義を半径  $r$  の球で行う。 $r$  の設定は徐々に半径を大きくしながら、1 m、5 m、15 m とする。また、近傍点の数は浅い層から深い層へ 16、64、128 に設定した。近傍特徴抽出を行う際の特徴マップはダウンサンプリングでは浅い層から深い層へ 256、512、1,024 である。

そして、デコーダ側では、エンコーダで作成した特徴量を入力点数へ復元する。この際、復元する点数はエンコーダと対になり、2,048、4,096、8,192 と段階的に点数を復元する。また、内層によって得た特徴量に対して全結合層を適用し、その特徴マップ数は深い層から浅い層へ 1,024、512、256 である。さらに、スキップコネクションをエンコーダとデコーダの対応する点を繋ぐことでダウンサンプリング処理で消えた低レベルな特徴量を伝達する。最後に、デコーダの最終層では実際に地物分類を行うために全結合層を適用する。なお、このようなハイパーパラメータを設定するにあたり、点群データに対して深層学習手法を適用した先行研究手法 [121, 18, 28, 229] を参考にした。

このようにして得られる大局ブロックからの出力は、 $K \times C$  次元の地物分類結果 ( $\mathbf{S}_{seg}^g$ ) である。

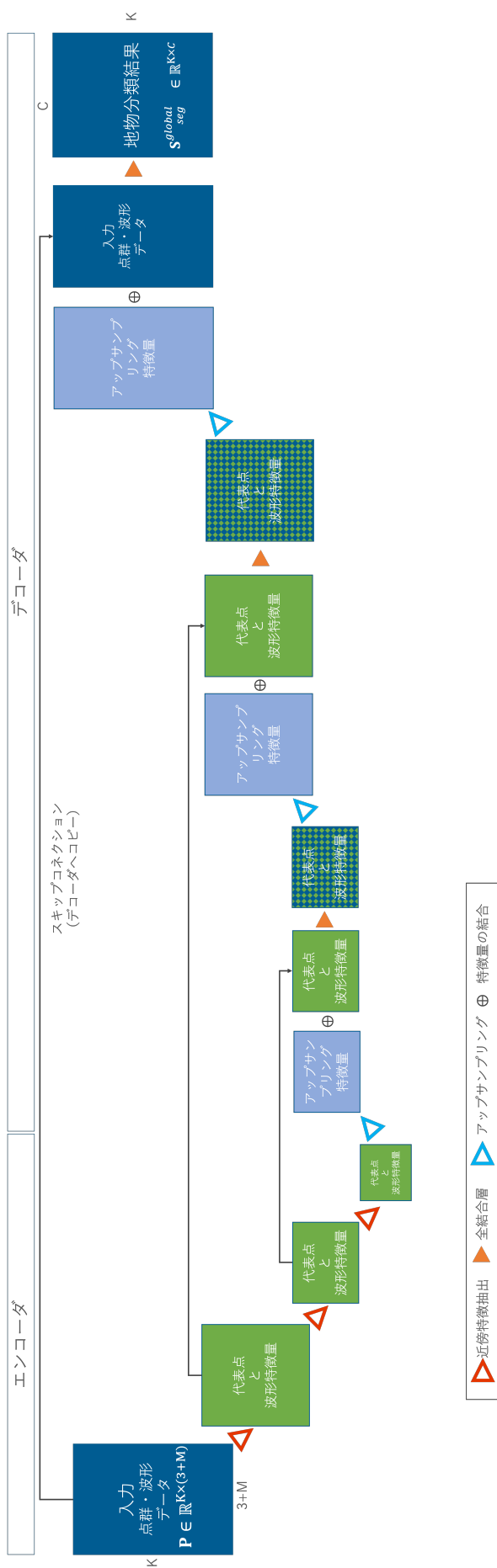


図 3.10: 大局ブロックの層設計. 近傍特徴抽出を行うダウンサンプリングによって成り立つエンコーダと, エンコーダで抽出された特徴を元の点数まで復元し地物分類を実施するデコーダによって成り立つ. ダウンサンプリングは3回適用し, それに応じてデコーダのアップサンプリングも3回適用する. なお, アップサンプリングの際には, 代表点として残らなかった情報をエンコーダ側からスキップコネクションによって受け渡す.

### 3.4.4 損失関数

最適化に使用する損失関数について説明する．一般的に，点群・波形データに限らず航空機ライダーで観測した点群データにおける地物分類では，各地物に属する点の数が不均衡になりやすいという問題が存在する．このような不均衡なデータセットで直接最適化を実行すると，深層学習モデルの全体的な性能に決定的な悪影響を及ぼす．この問題を解決するために，損失関数に各地物の存在比率を考慮した優先度係数を追加する．これにより，モデルが比率の少ない地物を学習することを補助する．各地物の優先度係数は，以下に示す各地物の割合によって決定される．

$$\lambda_c = \frac{1}{\ln\left(\alpha + \frac{\text{Num.}_c}{\sum_{c=1}^C \text{Num.}_i}\right)} \quad (3.5)$$

ここで， $\lambda_c$  は  $c$  番目の地物への優先度係数を表す．また， $\text{Num.}_c$  は  $c$  番目の地物に属する点の数， $C$  は地物数， $\alpha$  は負の数にならないための係数である．

少ない点数の地物に対する優先度係数をかけた損失関数は以下のように定義される：

$$\mathcal{L}_{seg} = \sum_{i=1}^K \sum_{c=1}^C \lambda_c t_{i,c} \log s_{i,c}, \quad (3.6)$$

ここで  $K$  は点の数， $t_{i,c}$  は真値， $s_{i,c}$  は  $i$  番目の点が  $c$  番目の地物に属する確率， $\lambda_c$  は  $c$  番目の地物に対する優先度である．

提案する局所・大局モデルでは，局所ブロックと大局ブロックからの2つの出力があるため，そのままでは適用できない．そこで，局所ブロックと大局ブロックにおいてそれぞれに損失関数を定義して，それらを線形に足し合わせたものを局所・大局モデルの最終的な損失関数として定義する．

$$\mathcal{L}_{total} = \mathcal{L}_{seg}^{local} + \mathcal{L}_{seg}^{global} \quad (3.7)$$

ここで， $\mathcal{L}_{seg}^{local}$  は局所ブロックの損失関数， $\mathcal{L}_{seg}^{global}$  は大局ブロックの損失関数を示す．

### 3.4.5 最適化アルゴリズム

本節では，提案する波形データに特化した深層学習モデルである局所・大局モデルのための最適化アルゴリズムを説明する．点群・波形データを入力し，局所ブロックから地物分類結果を出力  $\mathbf{S}_{seg}^l$ ，大局ブロックから地物分類結果を出力  $\mathbf{S}_{seg}^g$  する関数として表現できる．この関数について，局所ブロックを  $\text{MODEL}^l$ ，大局ブロック  $\text{MODEL}^g$  をと表現する．それぞれの学習可能な重みを，局所ブロックは  $\theta_{\text{MODEL}^l}^l$ ，大局ブロックは  $\theta_{\text{MODEL}^g}^g$  と表現する．そして， $\mathcal{L}_{seg}^{local}$  は局所ブロックの損失関数， $\mathcal{L}_{seg}^{global}$  は大局ブロックの損失関数を示す．

基本的には， $\theta_{\text{MODEL}^l}^l$  と  $\theta_{\text{MODEL}^g}^g$  の最適化は  $\mathcal{L}_{seg}^{local}$  と  $\mathcal{L}_{seg}^{global}$  を最小化するように  $\theta_{\text{MODEL}^l}^l$  と  $\theta_{\text{MODEL}^g}^g$  の値を調整することである．この際には，確率的勾配降下法を用いて最適化を行う．確率的勾配降下法による最適化を擬似的なアルゴリズムで示すと以下ようになる (Algorithm1)．まず，訓練データからミニバッチ  $\{x\}$  をランダムに抽出する．そして，局所ブロックと大局ブロックをそれぞれ最適化する．まず，局所ブロックを最適化する．ミニバッチを  $\text{MODEL}^l$  へ入力して， $\mathbf{S}_{seg}^l$  を出力する．この後， $\mathbf{S}_{seg}^l$  と真値の  $\hat{\mathbf{S}}_{seg}$  を用いて，損失関数の計算を行う．この損失関数の値を元に  $\text{MODEL}^l$  の重みに対する勾配  $g_w^l$  を計算して，その勾配を元に局所ブロックの重み  $\theta_{\text{MODEL}^l}^l$  を更新する．次に，大局ブロックを最適化する．ミニバッチを  $\text{MODEL}^g$  へ入力して， $\mathbf{S}_{seg}^g$  を出力する．この後， $\mathbf{S}_{seg}^g$  と真値の  $\hat{\mathbf{S}}_{seg}$  を用いて，損失関数の計算を行う．この損失関数の値を元に  $\text{MODEL}^g$  の重みに対する勾配  $g_w^g$  を計算して，その勾配を元に大局ブロックの重み  $\theta_{\text{MODEL}^g}^g$  を更新する．この動作を最大繰り返し回数まで実行する．

### 3.4.6 各ブロックの出力の組み合わせ方法

局所・大局モデルでは，予測時において局所ブロック ( $\mathbf{S}_{seg}^l$ ) と大局ブロック ( $\mathbf{S}_{seg}^g$ ) からの2つの地物分類結果が出力されるため，最終的な地物分類結果 ( $\mathbf{S}_{seg}$ ) を予測するために

**Algorithm 1** 局所・大局モデルの最適化アルゴリズム

**Require:** training data, MODEL<sup>g</sup>, MODEL<sup>l</sup>, max\_iteration, α: learning rate

**Require:** θ<sub>MODEL</sub><sup>g</sup>, θ<sub>MODEL</sub><sup>l</sup>

**Ensure:** trained model

```

1: for itr=1,...,max_iteration do
2:   Sample {x} a batch of Full Waveform LiDAR Data(Point Clouds and Waveform).
3:   Sample Sseĝ a batch from the ground truth corresponding to input batch.
4:   Ssegl = MODELl(x)
5:   gwl ← ∇w [Lseglocal(Ssegl, Sseĝ)]
6:   θMODELl ← θMODELl + α · Optimizer(θMODELl, gwl)
7:   Ssegg = MODELg(x)
8:   gwg ← ∇w [Lsegglobal(Ssegg, Sseĝ)]
9:   θMODELg ← θMODELg + α · Optimizer(θMODELg, gwg)
10: end for

```

は、これら2つの出力を統合する必要がある(図 3.3)の緑枠). その際に、各ブロックで得意不得意な地物を定義する. 図 3.11に示すように、局所クラスについては、波形データの特徴が地物分類に有効であると仮定し、ソフトマックス関数を介した局所ブロックからの出力が0.5以上であれば局所ブロックの結果を採用する. 図 3.11の例としては、地物 A, C, F が大局クラス、地物 B, D, E が局所クラスとした場合に、局所ブロックが推定した地物 B, D, E に属する確率が0.5を超えた場合において局所ブロックの値を最終出力に採用する. それ以外は大局ブロックの結果を採用する. 具体的に図 3.11の例として局所ブロックが推定した地物 B, D, E に属する確率が閾値を満たさない場合と、大局クラスの、地物 A, C, F は大局ブロックを採用する. 最後に、局所ブロックと大局ブロックを統合した結果から最大の確率を有する地物を最終的な予測結果とする.

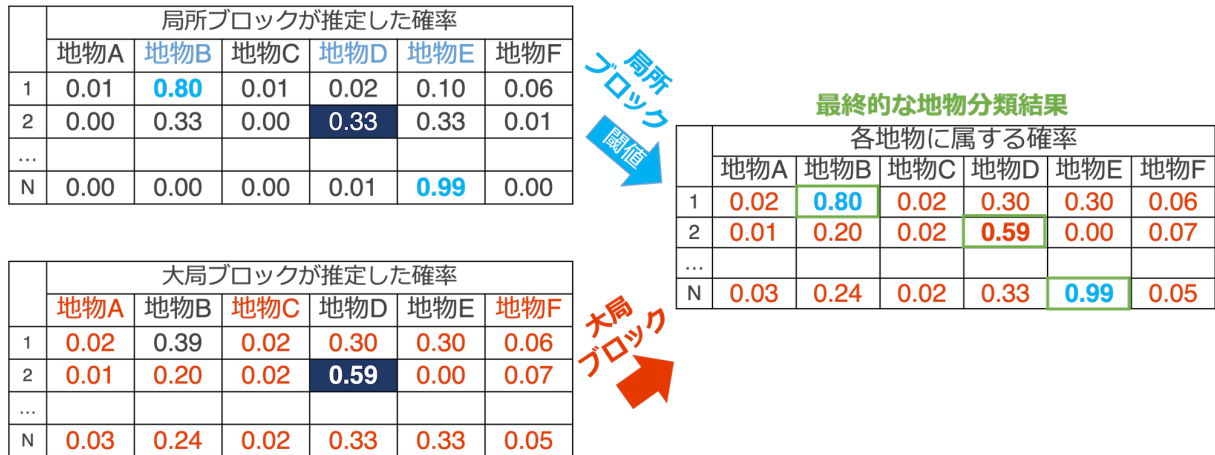


図 3.11: 予測時の局所ブロックと大局ブロックの組み合わせ方法. 学習済みモデルを適用する際には、局所ブロックと大局ブロックの2つの出力を組み合わせる必要がある. 組み合わせる方法は単純に局所ブロックが推定したそれぞれの地物に属する確率において局所クラスに属する確率が閾値を超えた場合には局所ブロックから採択する. それ以外は大局ブロックから採択する.

### 3.5 ベンチマークデータセットによる地物分類実験

#### 3.5.1 データセット

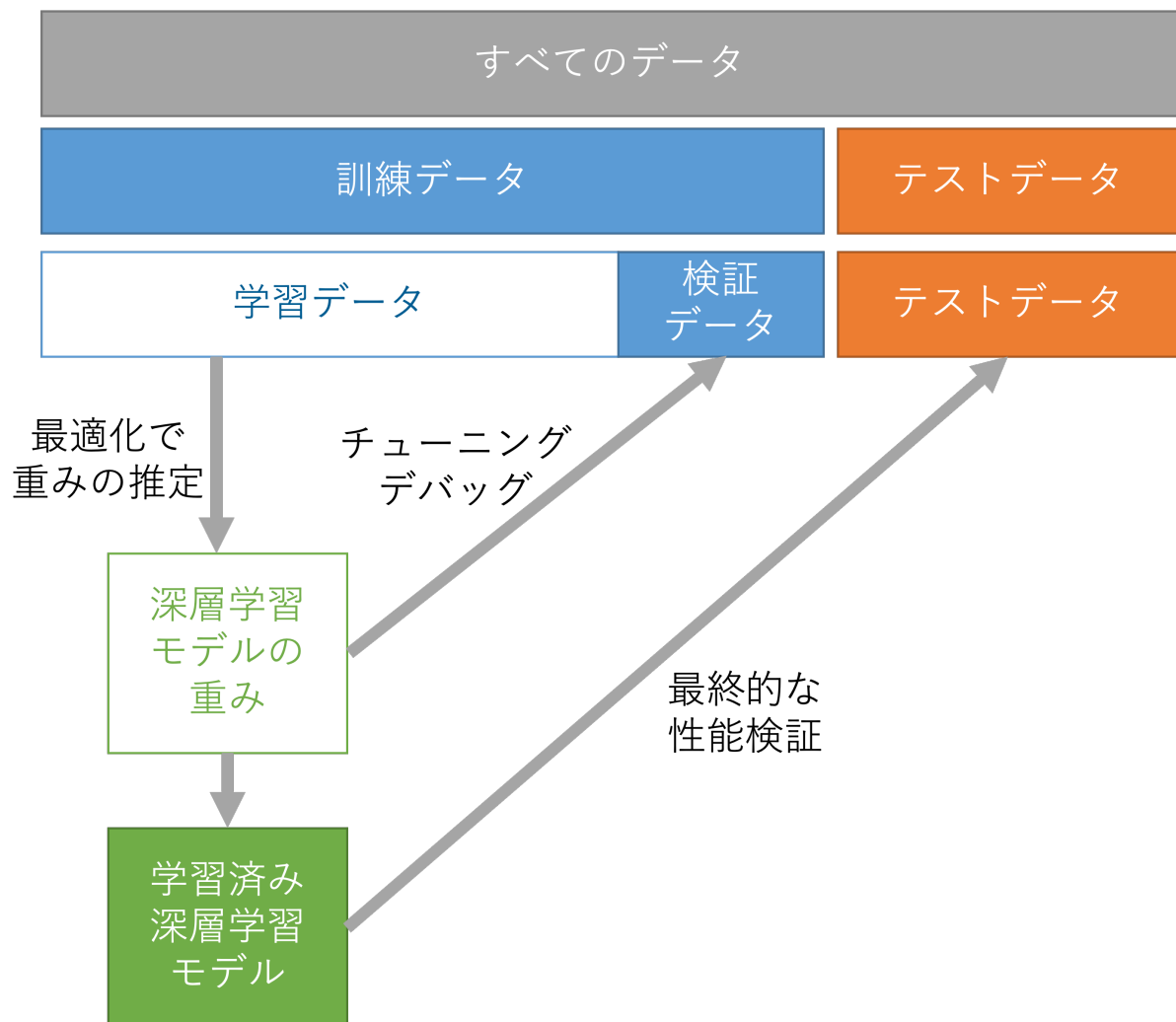


図 3.12: 学習・検証・テストデータの意味。 全てのデータを訓練データとテストデータに分割する。訓練データをさらに学習と検証データに分ける。訓練データは深層学習モデルの重みの推定に用い、検証データは深層学習モデルの提案を行う際の試行錯誤（チューニングやデバッグ等）に使用して現時点でのモデルの性能を評価する。テストデータは最終的に完成した手法の性能を評価するために使用する。

深層学習モデルの最適化およびその地物分類性能の検証には Riegl LMS-Q780 点群および波形情報を記録するライダーで取得されたベンチマーク用のデータセットを用いた [255]。このベンチマークデータを使用する理由は、本研究では点群・波形データに対する地物分類手法の提案・改善を実施するため、先行研究と同じデータセットにて平等に比較を行うことが重要であるためである。

ベンチマークデータセットにおける対象地域には自然物の他、人工的な地物が存在した (図 3.13)。波形データの値は 160 次元のベクトルで表現され、記録された反射の数が 160 に満たない場合は足りない部分を 0 で埋めた。なお、波形データの値は、Riegl LMS-Q780 の処理ソフトが出力したノイズ除去済みのデータを用いる。真値は各点に付与されており、分類対象の地物は 6 種類存在し、それぞれ、地面、植生、建物、電線、鉄塔、道路である。これらの地物は、点群および波形情報を記録するライダーで観測可能な空間解像度・点群データ密度

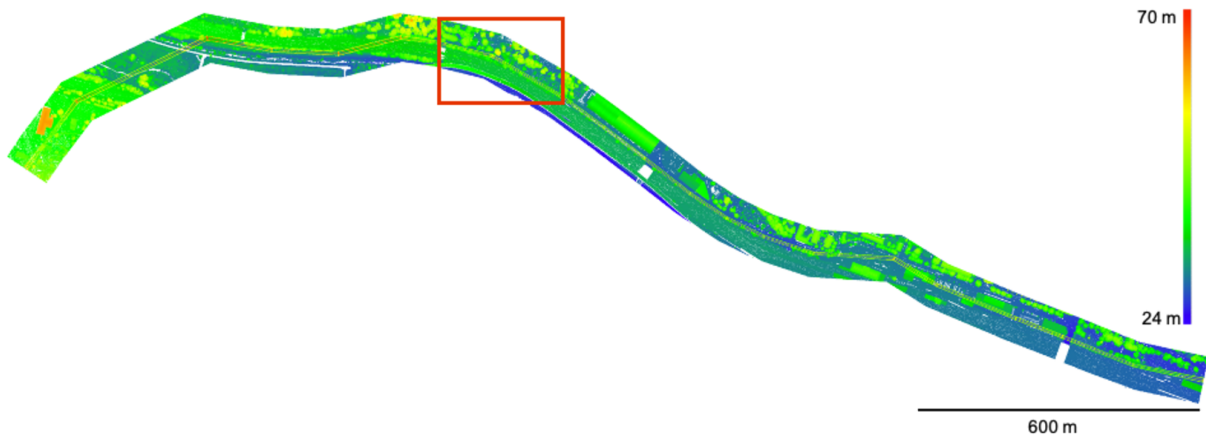


図 3.13: ベンチマークデータセット. 赤枠はテストデータとして設定した範囲を示す. 過去の手法と平等の条件で性能比較を行うため, ベンチマークデータを用いる.

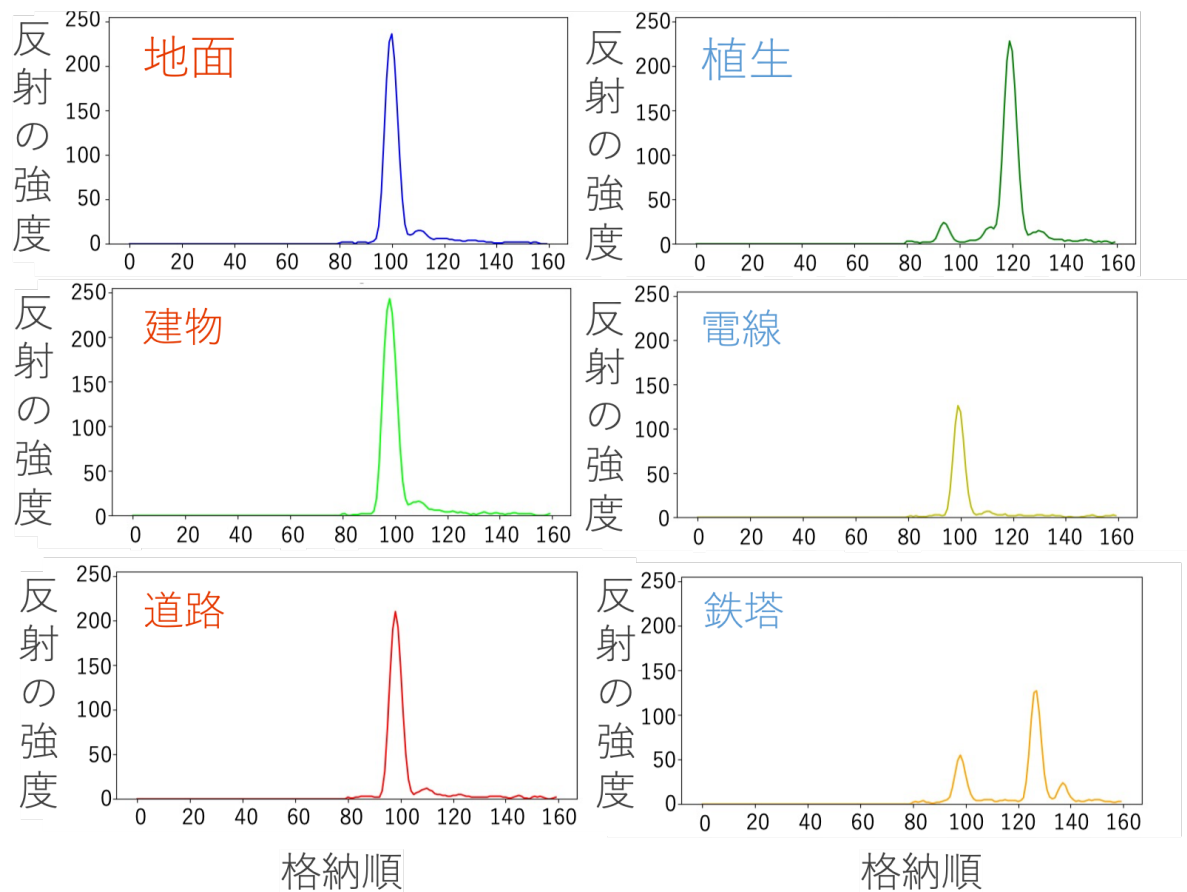


図 3.14: ベンチマークデータセットにおける波形データの例. 傾向を確認するため各地物においてランダムに 100 個の波形データサンプル抽出しその平均を求めた値. 赤字で示した地物は大局クラス (波形データの空間的な関係性である大局特徴が重要な地物), 青字で示した地物は局所クラス (個々の波形データが分類時に重要な地物) を示す.

等で分けすべき最低限の地物であり, 3次元地図を作成する際には有効な地物になる.

図 3.14 はそれぞれの地物からランダムに 100 サンプル抽出し, その平均値を求めた波形

データである。地面、建物、道路の地物はひとつの大きなピークという似た形状を持ち、それ以外の地物はそれぞれの地物に応じて特徴的な形状を持つ。このデータセットを用いて深層学習モデルの訓練を行う際には、図 3.12 に示すような定義に従って、訓練と性能評価を行うためのテストデータに分ける必要がある。そこで、ある小領域をテストデータとして分けた(図 3.13 の赤枠)。このテストデータは最適化には用いず、性能評価にのみ使用した。テストデータの設定は、分類対象の地物が全て含まれている領域を対象にした。最適化に用いるデータは赤枠以外の点群・波形データであり、訓練データと呼ぶ。訓練データをさらに、学習データと検証データに分ける。学習データは深層学習モデル(ここでは局所・大局モデル)が持つ学習可能な重みを推定するために用い、検証データは層設計を試行錯誤する際のチューニングに使用する。チューニングとは、学習済みモデルによる性能を検証データを用いて分類性能を判断し、各ハイパーパラメータや畳み込みの個数などの試行錯誤に使用する。この試行錯誤によって得られた最適な学習済みモデルをテストデータに適用する。また、このデータの特性として、表 3.1 では各地物の点の数が示す通り、非常に地物の偏りが大きい傾向を示した。これは 3 次元的な観測を行うため、体積に比例して点群データ数が多くなるため、大きな地物を小さな地物の間で差が大きくなるためである。

波形データの形状を元に、表 3.2 のように局所クラスと大局クラスの 2 つに分けた。局所クラスは個々の波形データという局所的な波形データの形状が地物分類に寄与すると考えられる地物である。一方、大局クラスは広い範囲の大局特徴を取り込むことで地物分類が可能になる地物である。図 3.14 が示す通り 植生、電線、鉄塔は波形データの形状によって地物の予想が可能である。しかしながら、地面、建物、道路は局所的な波形データのみを見ただけでは区別できないため、広い範囲の波形データ情報を総合的に見て判断する必要がある。従って、植生、電線、鉄塔を局所クラスに設定し、波形データ情報の特徴抽出を担う局所ブロックからの予測結果を優先的に採用し、地面、建物、道路を大局クラスに設定し、空間的な特徴抽出を担う大局ブロックからの予測結果を優先的に採用した。

また、点群データと波形データを紐付ける方法は、各波形データからファーストリターンの点のみを選択し、それを点群・波形データと定義する。これは、本研究にて使用している地物分類手法は点群データの深層学習手法を参考にしており、点群データの深層学習手法は物体表面に着目するため、ファーストリターン以降の点の特徴抽出が困難であるためである。

ベンチマークデータは、数百万規模の点群データを含む非常に大規模なものである。このような大規模なデータは、深層学習モデルの計算に使用する GPU のメモリが限られているため、モデルの学習過程で全てのデータを一度に直接扱うことが不可能である。とりわけ、点群データだけでなく波形データも扱うためメモリ消費は必然的に多くなる。そこで、大きな領域を小さなパッチに分割して学習データを作成した。学習用のデータを 50,000 点の固定点を持つパッチに分割した。なお学習時では、5 回のクロスバリデーションを行った。

表 3.1: 訓練データとテストデータ

波形データ情報が効果的	訓練データ		テストデータ		
	地物	点数	%	点数	%
	地面	1,787,352	20.4	193,070	18.1
✓	植生	4,719,634	53.9	765,327	71.7
	建物	1,514,486	17.3	49,138	4.6
✓	電線	71,978	0.8	8,151	0.8
✓	鉄塔	32,008	0.4	1,829	0.2
	道路	633,606	7.2	49,580	4.6

### 3.5.2 実験の設定

まず、訓練済みモデルによる予測方法について説明する。予測方法は以下の Algorithm2 を用いる。訓練済み局所・大局モデルとテストデータを使用する。なお、モデルの学習時に GPU のメモリ容量の関係から、入力次元数を限定したため局所・大局モデルは入力点数は 50,000

表 3.2: 局所クラスと大局クラス

局所クラス	植生
	電線
	鉄塔
大局クラス	地面
	建物
	道路

点である。そこで、テストデータを固定点の 50,000 点に抜き出す必要がある。したがって、テストデータ全体を予測するためには、一定範囲ごとに切り取りながら移動しつつ予測する。そこで、抜き出す中心点を決定し、一定範囲に区切られたパッチを作成する (getPatch)。抜き出したパッチを局所・大局モデルに入力し、入力された各点に対する地物分類結果を推定する (prediction)。このようなパッチ単位の処理を、50,000 点ごとに移動しながら実行する。地物分類時においては、各地物に属する確率が計算され、そのうち最大の確率を示した地物を選択する (getMaxClass)。なお、重複を許しながら、パッチを抜き出すため、重複部分では後から予測した結果を用いる。

---

**Algorithm 2** 学習済みモデルによる地物分類のアルゴリズム

---

**Input:** test data and trained model

**Output:** predicted data

```

model = loadModel(trained model)
data = loadData(test data)
patches, indexes = getPatch(data)
for i = 1 to len(patches) do
    pred prob = prediction(model, patches[i])
    predClassID = getMaxClass(pred prob)
    data[indexes[i]] = predClassID
end for
return data

```

---

次に、地物分類性能の評価方法について説明する。本研究では、学習済みモデルをテストデータに適用した地物分類結果の精度を定量的に評価する指標として、Precision, Recall, F1 Score を用いる。これらの指標は、主に画像では画素ごとに地物分類するタスクや点群データの点単位に地物分類するタスクの性能評価に広く用いられている。Precision は過検出の指標であり、Recall は抽出性能の指標である。F1 Score は、Precision と Recall のバランスを見ることが可能な指標である。一般的には、地物分類性能を定量的に評価する際には、F1 Score が重視される。Precision, Recall, F1 Score は下式のように定義される。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (3.8)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (3.9)$$

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (3.10)$$

ここで、表 3.3 に示すように、TP (true positive) は真値とモデルが予測した地物が一致していた場合、FP (false positive) は過剰抽出を示し、FN (false negative) は見逃しを示す。

最後に、実験に用いた実装や最適化の設定について説明する。本章における深層学習モデルの実装は Pytorch[169] を使用した。また、学習時のすべてのクロスバリデーションにおいて、Adam を使用した最適化を行い、学習率の初期値を 0.001 とした。なお、学習率は 10 エポ

表 3.3: 定量評価指標の計算に使用する項目

		真値	
		正	負
予測	正	true positive	false positive
	負	false negative	true negative

ックごとに 50% ずつ減衰させた。モデルの学習には東工大の TSUBAME3.0 の NVIDIA Tesla P100 GPU を 4 台用いた。バッチサイズは各 GPU で 1 とし、各層にはバッチ正規化を行った。さらに、表 3.1 に示すように、本節で使用するベンチマークデータセットにおいても、各地物の数が不均衡である。そのため、損失関数における優先度を表 3.1 の訓練データの点数を使用して計算した。過学習抑制のためドロップアウトを使用した。乱数で初期化された重みを用いて、モデルを End-to-End で学習した。

### 3.5.3 実験結果

深層学習モデルは、訓練データを用いて最適化を行なった。なお、最適化において固定点の 50000 点に切り出したパッチを使用し、パッチの総数は約 100 個であった。提案ネットワークの学習には、5 回のクロスバリデーションを用い、収束するまで学習を続けた。この学習過程には、1 エポック（訓練データを 1 周）あたり 6 時間を要した。なお、予測時において、テストデータを処理する時間は TSUBAME3.0 にて 30 分を要した。以降では、学習済みモデルによる地物分類論結果を紹介する。

まず、定性評価結果を述べる。学習済みモデルによる地物分類の結果を図 3.15 に示す。図 3.15(c) に示す提案手法（局所・大局モデル）による結果はテストデータの点について、図 3.15(d) に示す真値と比較すると正しい地物を予測した。一方で、境界部分のデータ（白丸で示した部分）の地物分類には失敗する傾向を示した。これは、大局ブロックで必要な広範囲の情報が不足しているために発生したと考えられる。具体的には、データの無い領域では十分に広い受容野が獲得できないことや、学習データに境界部のデータが少ないことなどが原因と考えられる。しかし、先行研究の評価と同等にするために境界部も予測対象にしているため生じた不具合だが、実運用では境界部は対象外もしくは隣接する図郭のデータを用いるため回避可能である。また、図 3.15(a), (b) はそれぞれ、局所ブロック結果、大局ブロックの結果を示しており、局所・大局モデルと比較すると分類性能が低いことが確認できた。そのため、両者を組み合わせる局所・大局モデルの有効性が示された。

次に、定量評価結果を述べる。定量評価を行うために、Precision, Recall, と F1 Score を各地物で計算した。表 3.4 は各地物の地物分類性能を示しており、提案手法の F1 Score は 6 つの地物で 80% を超えた。さらに、提案手法は画像化手法 [255] よりも殆どの地物において高い Recall を示し、画像等に変換することなく高い性能で点群・波形データの地物分類が可能であると示した。さらに点群・波形データをそのまま深層学習モデルに入力し地物分類を直接可能な手法が開発されたことで、画像化する手法はで中間生成物である点群・波形データを画像化した結果の品質を目視にて確認する手順などが必要になるが、本手法ではそれが省略可能になる。したがって、実用化を考える上で作業者の負担を減らすことに役立つ。次に、それぞれの局所ブロックと大局ブロックのみを学習した場合と局所・大局モデルを比較した。局所ブロックのみを用いた場合では、個々の波形データで分類可能と仮説を立てた植生・電線においてある程度の分類性能を示したが、全体として低い分類性能であった。これは、個々波形データだけでは分類が困難な建物等の特徴抽出も実行する必要がある、局所ブロックの持つ表現力では 6 つの地物を同時に分類することが困難であったことに起因する。さらに、波形データの特徴を考慮していない既存の点群データに対する深層学習手法を適用した大局ブロックと比較すると、提案手法の局所・大局モデルがより高い地物分類性能を示した。特に、大局クラスである地面・建物・道路では高い性能を示した。なお、Precision では植生と建物で大局ブロックと局所・大局モデルは同じ値を示したが、Recall において局所・大局モデルが高い性能を示した。したがって、我々の提案手法の波形データの局所的な地物分類と波形

データの大局的な関係性を用いた地物分類を組み合わせる局所・大局モデルの優位性が示された。また、基本とする深層学習モデルの層設計が分類性能へ与える影響は付録 B にて示す。

また、一般的に人間が手作業でデータにラベルをつける作業時においては3~6%の間違いが含まれている [161]。そのため、地物の抽出性能を示す Recall に着目した場合において、本提案手法である局所・大局モデルは人間と同程度の分類性能を地面・植生・建物・電線においては達成したと言える。これにより、従来の目視で行なっていた作業を自動化し、手作業の短縮が期待される。しかし、分類性能において鉄塔と道路の Recall は未達となったため、依然として分類性能の向上は必要である。抽出性能の高い建物や植生や電線に着目すると、これらは国土交通省で取り組まれているは Project PLATEAU[174] などの都市の 3 次元モデル作成において重要な地物である。本提案手法によって人間の作業と同等を達成した地物から自動的なフローに置き換えることによる 3 次元地図の更新頻度を向上させるなどの運用が考えられる。

表 3.4: ベンチマークデータにおける定量評価結果。太字は最高性能を示す。また、示している結果は 5 回のクロスバリデーションで得た学習済みモデルによる分類結果の平均値である。画像化手法の結果は元論文 [255] より引用した。

手法	評価指標	地面	植生	建物	電線	鉄塔	道路	平均
画像化 [255]	Recall	0.84	0.97	0.93	0.91	<b>0.88</b>	0.69	0.87
	Precision	0.51	0.96	0.14	0.26	0.05	0.29	0.37
局所ブロック	Recall	0.68	0.83	0.02	0.80	0.36	0.58	0.55
	F1 Score	0.59	0.89	0.04	0.40	0.08	0.38	0.40
大局ブロック	Precision	0.56	<b>0.98</b>	<b>0.99</b>	0.97	0.52	0.96	0.83
	Recall	0.93	0.85	0.85	0.85	0.66	0.60	0.83
	F1 Score	0.70	0.91	0.91	0.91	0.58	0.74	0.80
局所・大局モデル	Precision	<b>0.80</b>	<b>0.98</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.97</b>	<b>0.95</b>
	Recall	<b>0.94</b>	0.96	<b>0.95</b>	<b>0.99</b>	0.79	<b>0.71</b>	<b>0.89</b>
	F1 Score	<b>0.86</b>	<b>0.97</b>	<b>0.97</b>	<b>0.99</b>	<b>0.88</b>	<b>0.82</b>	<b>0.92</b>

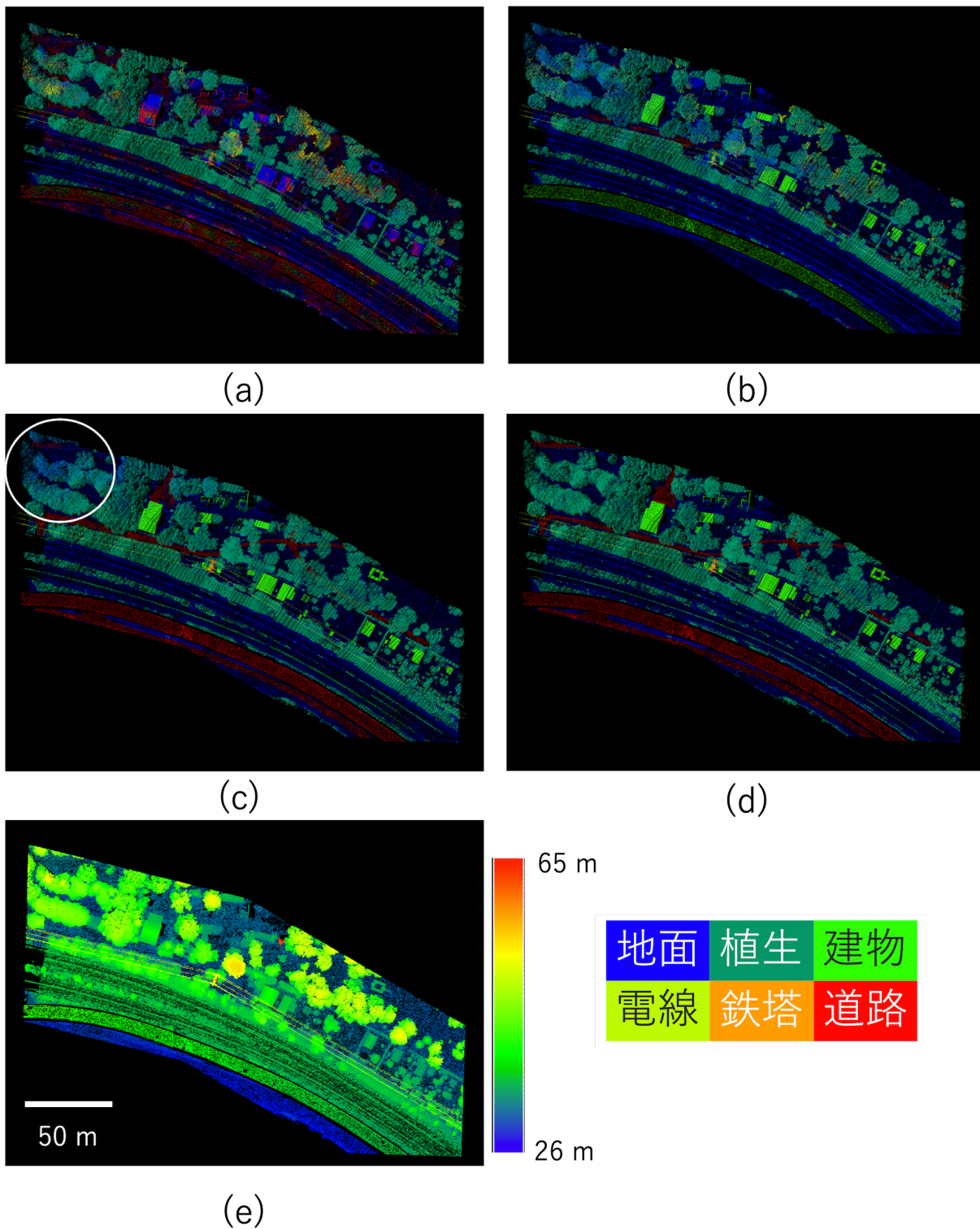


図 3.15: 学習済みモデルによる地物分類結果. (a): 局所ブロック, (b): 大局ブロック, (c): 局所・大局モデル, (d): 真値, (e): テストデータの高さ情報

### 3.5.4 2つのブロックの出力と最終予測結果との関係

本提案手法において、局所クラス（個々の波形データが分類性能に寄与する地物）とい大局クラス（波形データの空間的な関係性である大局特徴が分類に寄与する地物）に大別できると仮説を立てた。そこで、提案した局所・大局モデルは、局所ブロックと大局ブロックの2つの出力を統合して、テストデータの最終出力を得た。実験に使用したベンチマークデータにおいては、植生、電線、鉄塔の3つを局所クラスと定義した。これらの地物に対して、局所・大局モデルの局所ブロックがどれだけ最終出力に貢献したのかを調査した。

局所ブロックの貢献度を調べるために、最終出力に対する局所ブロックから得られる出力の寄与率を計算した。表 3.5 に示すように、局所ブロックは仮説通りに局所クラスとして設定した植生、電線、鉄塔で高い寄与率を示した。これにより、局所クラスと大局クラスに分けてそれぞれに特化した層設計を用いた有効性が確認できた。しかし、本稿で用いた局所ブロックと大局ブロックの統合方法は、局所ブロックから計算された各クラスに属する確率が閾値 0.5 を超えたら積極的に採用するため、想定していなかった地面、建物、道路も若干ではあるが局所ブロックから採用された。これは単純な閾値処理では対応できない状況が発生したためと考えられる。そこで、本論文で使用したような単純なルールベースの統合手法ではなくより高度なデータ駆動の統合手法などのより高度な統合方法が必要と言える。

表 3.5: 最終出力に対する局所ブロックからの寄与率。太字の箇所は局所クラスに相当する。

クラス	局所ブロックからの採用率
地面	0.15
植生	<b>0.82</b>
建物	0.09
電線	<b>0.89</b>
鉄塔	<b>0.96</b>
道路	0.13

### 3.5.5 波形データの効果

- 局所ブロックに対する波形データの効果

局所ブロックに対して波形データを入力する有効性を示すために、点群データのみを局所ブロックへ入力するモデルを開発した。点群データのみを局所ブロックへ入力するモデルは提案手法と同様にベンチマークデータセットを学習し、その際に点群データのみを与えた。このモデルの地物分類結果を表 3.6 1 行目に示す。表において、-を示している結果は損失関数が下がらず収束しなかったため、学習済みモデルを得られず分類ができなかったことを示す。点群・波形データを入力した局所ブロック (表 3.6 2 行目) は損失関数の値が下がり、学習済みモデルにおいて地物分類が実行できたことと比較すると、波形データの有効性が確認できた。
- 大局ブロックに対する波形データの効果

大局ブロックに対する波形データを入力する有効性を示すために、点群データのみを大局ブロックへ入力するモデルを開発した。点群データのみを大局ブロックへ入力するモデルは提案手法と同様に、ベンチマークデータセットを学習し、その際に点群データのみを与えた。このモデルの地物分類結果を表 3.6 3 行目に示す。点群・波形データを入力した大局ブロック (表 3.6 4 行目) と比較すると、点群データのみの場合には大きな変化が無い。これは、点群データ向けに開発された大局ブロックの構造を用いたため単に大局ブロックに点群・波形データを入力するだけでは問題があるためである。
- 局所・大局モデルに対する波形データの効果

局所・大局モデルにおける波形データを入力する有効性を示すために、点群データのみを局所・大局モデルへ入力するモデルを開発した。点群データのみを局所・大局モデル

へ入力するモデルは提案手法と同様に、ベンチマークデータセットを学習し、その際に点群データのみを与えた。このモデルの地物分類結果を表 3.6 6 行目に示す。点群・波形データを入力した局所・大局モデル(表 3.6 7 行目)と比較すると、点群データを入力した局所・大局モデルは F1 Score が低下した。波形データの形状を個々に分類可能と定義した植生・電線・鉄塔の局所クラスにおいて、波形データを入力した場合の局所・大局モデルと比較して大きく F1 Score が低下した。一方、地面・建物・道路の大局クラスにおいては局所クラスほどの大きな低下は見られないが、波形データを入力した局所・大局モデルの分類結果よりも低い F1 Score を示した。また、局所クラスにおける局所ブロックからの採用率は植生で 5%、電線で 9%、鉄塔で 2% であり、波形データを入力した局所・大局モデルにおいて植生で 82%、電線で 89%、鉄塔で 89% であったため、点群データのみの場合において低い局所ブロックからの採用率を示した。これらの実験結果により、点群・波形データの特性に適した層設計を提案できたと言える。

表 3.6: 局所・大局モデルにおける入力データが地物分類結果に与える影響

層設計	入力	地面	植生	建物	電線	鉄塔	道路	平均
局所ブロック	点群データ	-	-	-	-	-	-	-
局所ブロック	点群・波形データ	0.59	0.89	0.04	0.40	0.08	0.38	0.40
大局ブロック	点群データ	0.67	0.89	0.64	0.87	0.33	0.60	0.67
大局ブロック	点群・波形データ	0.62	0.87	0.51	0.79	0.49	0.61	0.65
局所・大局モデル	点群データ	0.62	0.87	0.51	0.79	0.49	0.61	0.65
局所・大局モデル	点群・波形データ	<b>0.86</b>	<b>0.97</b>	<b>0.97</b>	<b>0.99</b>	<b>0.88</b>	<b>0.82</b>	<b>0.92</b>

### 3.5.6 局所・大局モデルにおける入力データに対する感度分析

局所・大局モデルにおいて局所ブロックには、波形データを入力し波形データ単位の地物分類を行なっている。また、大局ブロックでは点群データと波形データを入力している。そこで、局所ブロックが有効に作用したかを確認するために、波形データ情報に変更を加えた外乱のあるデータを入力した際の挙動を実験的に調査した。さらに、大局ブロックも同様に、点群データの座標値に対して変更を加えて外乱のあるデータを入力した際の挙動を実験的に調査した。

- 波形データに対する感度

まず、波形データに対する学習済み局所・大局モデルの感度を調査する、特に局所ブロックは波形データのみを用いて地物分類しているため、実質的に局所ブロックに対する入力に対する感度の調査である。具体的には、学習済み局所・大局モデルに対して、実際の波形データに乱数を付与した波形データ、実際の波形データではなく単なる乱数を入力した際の地物分類性能や、波形データの値をすべて 0 にした際の地物分類性能と、実際の波形データを入力した際の結果を比較することで局所ブロックの有効性の検証を行う。

学習済み局所・大局モデルに入力するための乱数による波形データの作成方法を説明する(図 3.16)。入力する波形データはホワイトノイズを適用してベンチマークデータセットの 160 次元の強度のサンプル数を作成した。この際、全ての波形データを個別に乱数のシードを変えながら作成した。同様に、実際の波形データに対して乱数を付与したデータも作成した。また、0 の波形データは純粋にベンチマークデータセットの 160 次元の 0 で埋められた配列を作成した。

次に、実験結果について述べる(表 3.7)。まず、0 を入力した際の結果を説明する。表 3.7 のゼロでは 0 を入力した際の地物分類性能と、実際の波形データを入力した際の地物分類性能(Recall)を示している。この結果地物分類性能の低下が示された。したがっ

て、提案した波形データの学習を行なったことが実験的に示された。次に、乱数としてホワイトノイズを波形データとして入力した際の結果を説明する。表 3.7 のノイズ付与ではホワイトノイズを入力した際の地物分類性能と、実際の波形データを入力した際の地物分類性能を示している。この結果、ホワイトノイズを局所・大局モデルに入力した際には、全体的に低下した。これは、局所ブロックも大局ブロックも波形データを学習していることを示している。なお、幾何的な特徴が効果的であると仮定した「建物」「地面」「道路」では地物分類性能においても大きな低下が見られた。これは、波形データに対する特徴抽出も大局ブロックにて計算が行われており、波形データの空間的な相互作用を獲得していることに起因する。これらのことから、波形データの特徴を学習しているため、ノイズのような外乱に弱いことを示しており、学習データに依存している可能性がある。今後、波形データの観測過程のような制約を入れるような条件を深層学習モデルの内部に入れると外乱に対して堅牢な地物分類が可能であると期待できる。

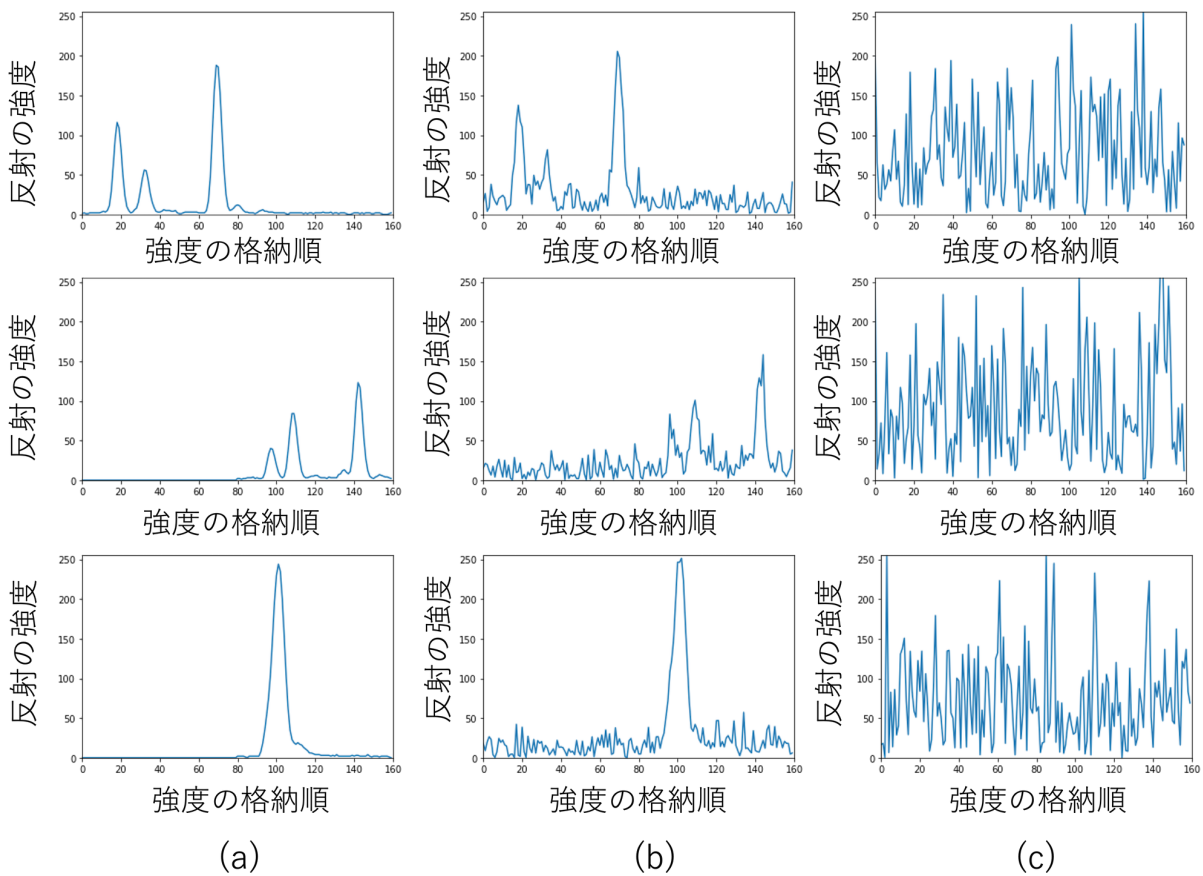


図 3.16: ノイズを加えた波形データの例。(a): 元の波形データ, (b): ノイズを付与した場合, (c): 完全に乱数を与えた場合を示している。

表 3.7: 局所・大局モデルにおける波形データに対する感度分析結果

	入力		指標	地面	植生	建物	電線	鉄塔	道路
	点群データ	波形データ							
実際の点群データ		ゼロ	Recall	0.01	0.01	0.01	0.99	0.01	0.11
実際の点群データ		ノイズのみ	Recall	0.01	0.88	0.01	0.18	0.92	0.01
実際の点群データ		ノイズ付与波形データ	Recall	0.81	0.92	0.91	0.95	0.79	0.78
実際の点群データ		実際の波形データ	Recall	<b>0.86</b>	<b>0.97</b>	<b>0.97</b>	<b>0.99</b>	<b>0.88</b>	<b>0.82</b>

- 点群データに対する感度分析

次に、点群データに対する学習済み局所・大局モデルの感度を調査する、局所ブロックでは大局ブロックは点群データと波形データのみを用いて地物分類しているため、実質的に大局ブロックに対する入力に対する感度分析である。具体的には、学習済み局所・大局モデルに対して、実際の点に乱数を付与した点群データ、実際の点群データの代わりに乱数を入力した際の地物分類性能や、点群データの値をすべて0にした際の地物分類性能と、実際の点群データを入力した際の結果を比較することで大局ブロックの有効性の検証を行う。

学習済み局所・大局モデルに入力するための乱数による点群データの作成方法を説明する(図 3.17)。入力する点群データは乱数としてホワイトノイズを適用してベンチマークデータセットの点群データと同じ次元の乱数を作成した。この際、全ての点群データを個別に乱数のシードを変えながら作成した。同様に実際の点に乱数を付与した点群データも作成した。なお、乱数の範囲はテストデータの座標が持つ最大値と最小値に設定した。また、0の点群データは純粹にベンチマークデータセットの点群データと同じ次元の0で埋められた配列を作成した。

次に、実験結果について述べる。まず、点群データの座標値として0を入力した際の結果を説明する。座標値を使用して近傍点を検索するグループ化の計算上、すべてゼロの値を持つ入力は想定されていないため、計算エラーとなった。そのため、表 3.8のゼロでは0を入力した際の地物分類性能として Recall は得られないので、-を記入した。次に、ホワイトノイズを入力した際の結果を説明する。表 3.8のノイズ付与ではホワイトノイズを入力した際の地物分類性能と、実際の点群データを入力した際の地物分類性能を示している。この結果、ホワイトノイズを局所・大局モデルに入力した際には、点群データが効果的であると仮定した「建物」「地面」「道路」において地物分類性能が低下した。また、波形データな特徴が効果的であると仮定した「植生」「電線」「鉄塔」では地物分類性能に大きな低下は見られなかった。低下した分は、大局ブロックからの採用されていた点が減ったために地物分類性能に影響したためである。このことから、提案手法における波形データに対して効果的であろうと仮定して提案した局所ブロックの効果が実験的に示された。なお、点群データの座標値を用いてグループ化を行うため、座標にノイズが入った場合には大局ブロックからの出力は期待された出力が得られない。

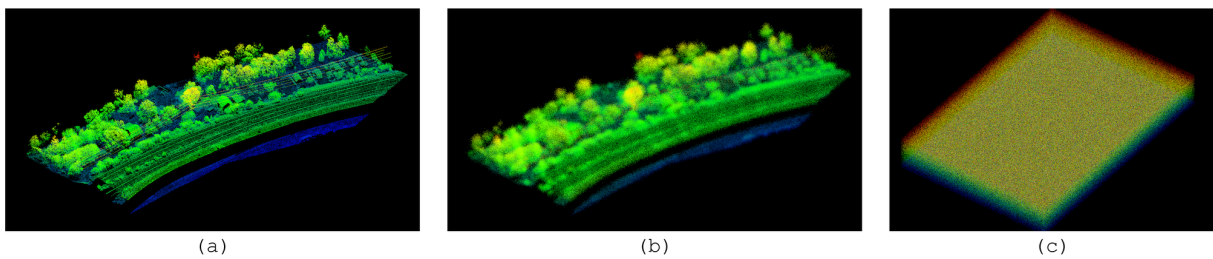


図 3.17: ノイズを加えた点群データの例, (a): 元の点群データ, (b): ノイズを付与した点群データ, (c): 完全に乱数を与えた場合を示している。

表 3.8: 点群データに対する感度分析結果

点群データ	入力		指標	地面	植生	建物	電線	鉄塔	道路
	点群データ	波形データ							
ゼロ	実際の点群データ	実際の波形データ	Recall	-	-	-	-	-	-
ノイズのみ	実際の点群データ	実際の波形データ	Recall	0.15	0.96	0.09	0.96	0.84	0.13
ノイズ付与点群データ	実際の点群データ	実際の波形データ	Recall	0.36	0.97	0.97	0.99	0.49	0.66
実際の点群データ	実際の点群データ	実際の波形データ	Recall	0.86	0.97	0.97	0.99	0.88	0.82

### 3.5.7 判断根拠の可視化

一般的に深層学習モデルはブラックボックスであるため、その判断根拠を知る必要がある。深層学習モデルの判断根拠を可視化する手法として、画像処理の分野で用いられる隠蔽感度分析と勾配可視化を参考にした。

画像処理に使用される隠蔽感度分析 (図 3.18) は直感的に理解しやすい深層学習モデルの判断根拠可視化手法である。まず、入力画像をいくつかの小さなグリッドに分割する。次に、訓練済みモデルによる予測を実行する。その際に、グリッドの1つをマスキングする (画素値をゼロに設定する)。各グリッドには、その「マスキングされている」ことが元のマスクしていない画像と比較して予測にどれだけ影響を与えたかに基づいて、寄与率が割り当てられる。直感的には、これらの寄与率はどのグリッドが予測に最も影響を与えているかを定量的に評価している。例として、図 3.18では画像を  $3 \times 3$  のグリッドで9つのグリッドに分割した場合を示す。この画像において、猫/犬の分類をする学習済みモデルがあった場合を考える。黒のグリッドが入力に加えたマスク (画素値が0) で、数字はマスクを加えた画像を学習済みモデルに入力した際の猫である確率の変化を示す。負の値は元の画像に比べて猫である確率が低下した場合を示す。この数字の絶対値が高いほど学習済みモデルに対する影響が大きいことを示す。この概念図の場合では、猫の胴体や足をマスクすると猫であると予測した確率が下がり分類結果への寄与度が大きく、学習済みモデルはそれらの箇所を判断材料にしていることがわかる。なお図 3.18では簡単のため  $3 \times 3$  という荒いグリッドであるが、よりグリッドを細かくすると猫の鼻や目、耳、尻尾に対するパーツ単位の評価が行うことが可能になる。

次に画像処理に使用される勾配 (図 3.19) による寄与度の計算による判断根拠の可視化について説明する。勾配による寄与度を用いた判断根拠の可視化は、深層学習モデルの判断根拠を隠蔽感度分析と同様に可視化する手法であるが誤差逆伝播の勾配情報を用いる。上記の隠蔽感度分析では、画像中のグリッドを1つ1つマスキングしたときに、出力の確率値がそれぞれどのように変化するかを考えた。そのため、ひとつの可視化結果を作成するために複数回の予測の計算が必要であった。一方、勾配可視化では、個々の画素が出力にどのような影響を与えるかを一度に計算できる。具体的には、深層学習モデルの最適化手法である誤差逆伝播で用いられる偏微分による勾配情報に着目し実現した。深層学習モデルの学習に一般的に用いられる誤差逆伝播では、真値と出力の誤差を用いて深層学習モデルのパラメータ (重み) に対する勾配を計算する。なお、勾配とは、誤差を各重みで偏微分した値である。この勾配を用いることで、深層学習モデルの学習時に各重みの値を更新できる。勾配は、各重みの値を含むベクトルで、その重みの小さな変化が出力にどれだけ影響するかを反映している。勾配可視化では、この勾配情報を元の入力画像に重畳 (オーバーレイ) することで画像を地物分類する上でどの入力画素が最も重要であるかを提示する。この勾配を可視化する方法として、Gradient-Weighted Class Activation Mapping (Grad-CAM) [192] が提案されている。Grad-CAMは勾配を各クラスに属する確率ではなく、最後の畳み込み層の各重みの値で取ると、解釈の質が向上することを発見した。つまり、特徴マップ上の勾配の大きさを可視化している。複数のクラスが画像に写っている場合の解釈性を高めるために、分類対象の地物に対する寄与度を計算する。手法名の「勾配による重み付けクラス活性化マッピング (Gradient weighted Class Activation Mapping)」とあるように、CNNの内部を流れる特定の勾配に重みをかけて、画像内の地物分類対象のいる領域を特定してヒートマップとして表示する。Grad-CAMは、画像地物分類、画像キャプション、Visual Question Answering への回答などのコンピュータビジョンの様々なタスクで使用可能である。特徴マップの勾配が分類結果への寄与率を示し、この可視化結果を確認することによって、これらのタスクをCNNが解く際に「どの部分を根拠として予測したのか」を解釈可能になる。

提案手法において、個々の波形データの形状が有効と仮定した地物において、局所ブロックからの分類結果を用いた。この分類時において、局所ブロックが機能しているかを確認するために判断根拠を可視化する。具体的には、これらの可視化手法を用いて、波形データに対して波形データのどの部分が局所・大局モデルの地物分類結果に寄与したかを可視化して確認する。この隠蔽感度分析を波形データに対して実行し、波形データのどの部分が地物分類に寄与したかを可視化する。波形データに適用する方法は、各波形データに対して160サンプルの強度データが格納されている。そこで、5サンプルずつその強度値をゼロにし

て、学習済みモデルに入力する。そして、ゼロにする前後での真値の地物に属する確率の変化を計算する。これを繰り返し適用することで、波形データのどの部分が局所ブロックの地物分類結果に寄与したのかを可視化できる。また、勾配可視化では、地物分類を行う直前の特徴マップと入力層の特徴マップにおいて Grad-CAM と同様に勾配を計算し、各層の勾配を足し合わせ元の波形データに重畳した。勾配は各波形データの最大値・最小値で正規化する。なお、複数回のプーリング処理によって入力よりも小さくなった特徴マップで勾配を計算するため、入力の 160 次元ではなく、40 次元になった特徴マップで計算した勾配を用いる。そのため、隠蔽感度分析と比較すると荒い可視化結果を示す傾向がある。

このような手法を用いて、局所・大局モデルの地物分類結果に寄与している箇所を可視化した結果を図 3.20 に示す。左図が入力波形データ、中図が隠蔽感度分析結果、右図が Grad-CAM の結果を示す。隠蔽感度分析は確率の低下度（寄与率）を示している。また、Grad-CAM は正規化された勾配を寄与率として示す。全体的な傾向として、波形データの形状が特徴的な場合においては隠蔽感度分析・Grad-CAM 共に波形データの強度が大きくなった箇所において反応が確認できた。このことから、局所・大局モデルにおいて波形データのピークを判断根拠として使用して地物分類していることが確認できた。

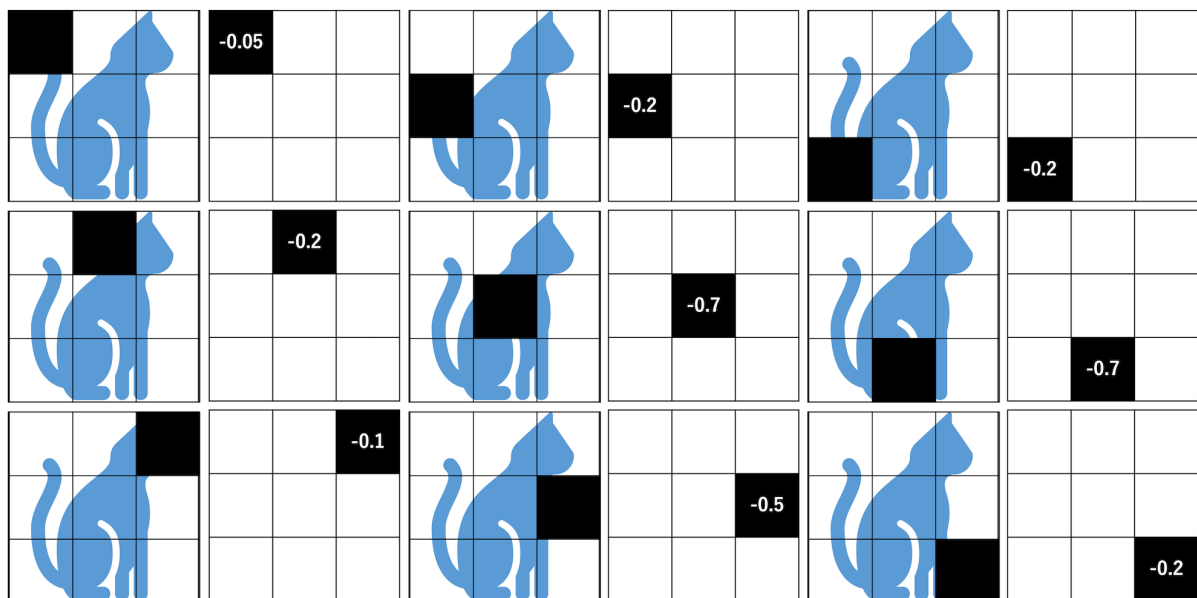
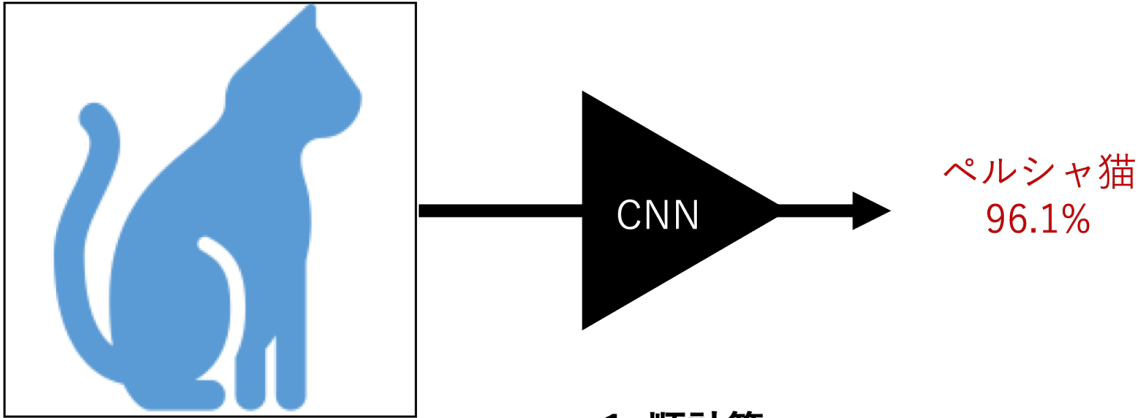
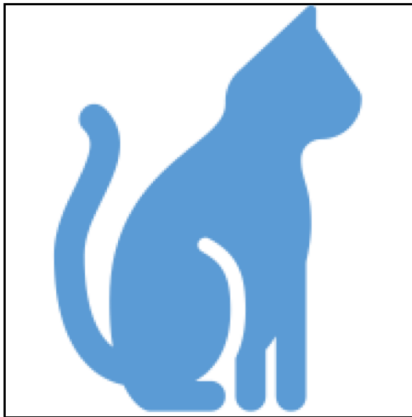


図 3.18: 隠蔽感度分析手法の概念図。画像に対して提案された学習済みモデルの感度分析手法であり、入力された画像に対してある一定範囲の大きさをマスクする。マスクした画像を学習済みモデルに入力し、その分類性能の低下（寄与率）を確認する。これをすべての領域に対して適用すると、画像のどの部分が分類に寄与するのかを確認可能になる。



### 1. 順計算



$\frac{\partial y}{\partial x_{1,1}}$	$\frac{\partial y}{\partial x_{1,2}}$	$\frac{\partial y}{\partial x_{1,3}}$
$\frac{\partial y}{\partial x_{2,1}}$	$\frac{\partial y}{\partial x_{2,2}}$	$\frac{\partial y}{\partial x_{2,3}}$
$\frac{\partial y}{\partial x_{3,1}}$	$\frac{\partial y}{\partial x_{3,2}}$	$\frac{\partial y}{\partial x_{3,3}}$

ペルシヤ猫  
96.1%

### 2. 誤差逆伝播で各画素に対する勾配計算

$\frac{\partial y}{\partial x_{1,1}}$	$\frac{\partial y}{\partial x_{1,2}}$	$\frac{\partial y}{\partial x_{1,3}}$
$\frac{\partial y}{\partial x_{2,1}}$	$\frac{\partial y}{\partial x_{2,2}}$	$\frac{\partial y}{\partial x_{2,3}}$
$\frac{\partial y}{\partial x_{3,1}}$	$\frac{\partial y}{\partial x_{3,2}}$	$\frac{\partial y}{\partial x_{3,3}}$

### 3. 入力画像に勾配を重畳

図 3.19: 勾配可視化手法の概念図. 画像分類モデルにおける勾配を可視化する手法として Grad-CAM[192] が提案された. Grad-CAM は入力された画像に対して学習済みモデルを適用して得られた中間の特徴マップにおいて真値と比較した際の勾配である寄与率をヒートマップとして可視化する手法である.

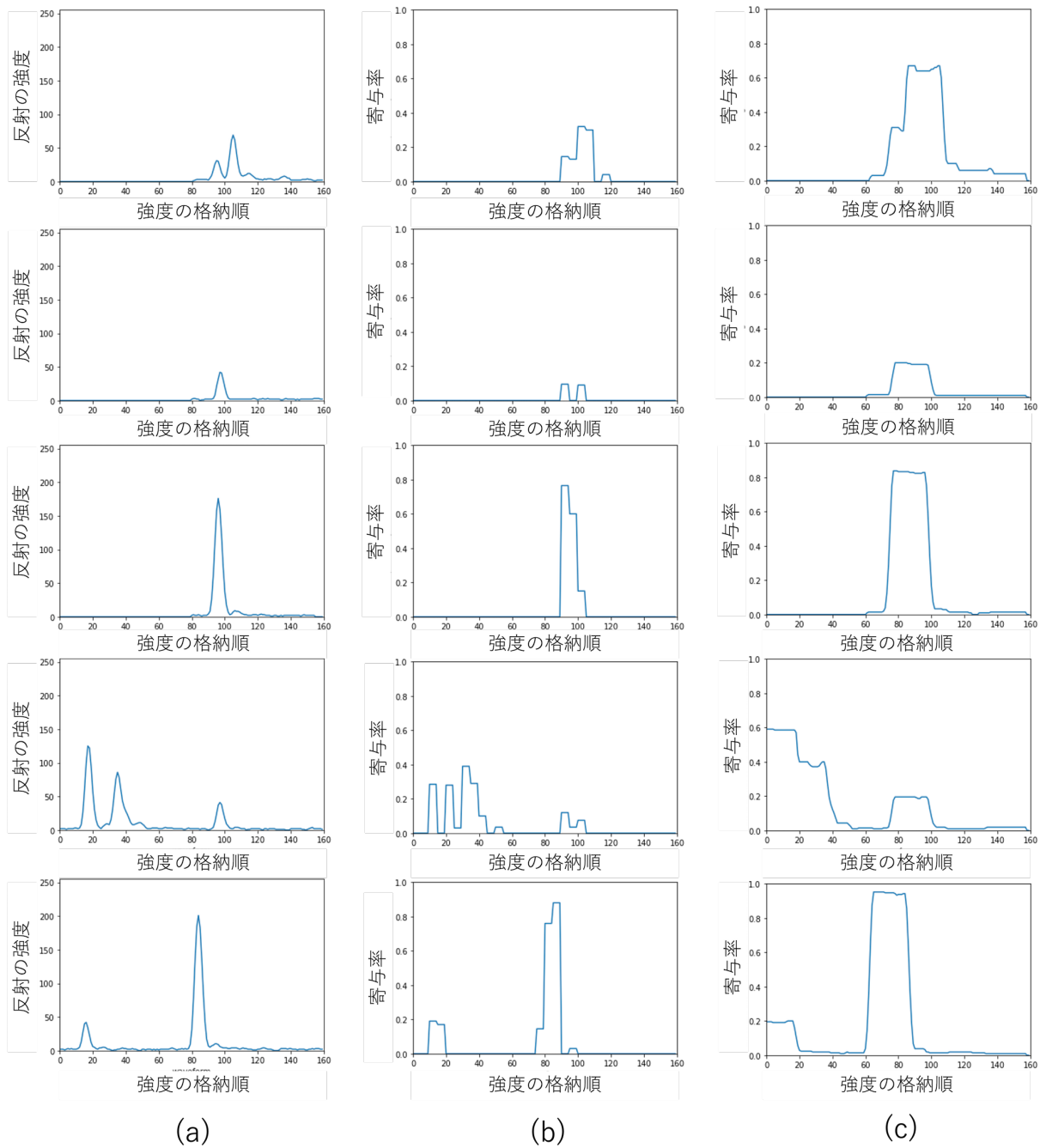


図 3.20: 波形データの判断根拠の可視化結果. (a): 入力した波形データ (b): 隠蔽感度分析による寄与率 (c): Grad-CAM を参考にした寄与率の可視化

### 3.5.8 点密度の影響

局所・大局モデルの学習データは平均で1平方メートルあたり15点の点密度で取得された点群・波形データを使用した。そこで、学習済み局所・大局モデルの多データへの適用性を確認するため、点密度を変化させた場合の分類性能へ影響を考察した。影響の調査方法は、点群・波形データの点密度を10点/m<sup>2</sup>、5点/m<sup>2</sup>、1点/m<sup>2</sup>まで等間隔に減らしたテストデータへ、学習済み局所・大局モデルをそれぞれ適用した分類結果をF1 Scoreを用いて比較した。その際、点密度を変化させるためにテストデータからランダム抽出した。表3.9にその比較結果を示した。点密度が1点/m<sup>2</sup>、5点/m<sup>2</sup>、10点/m<sup>2</sup>の場合すべてにおいて局所クラス（植生・電線・鉄塔）において大きく地物分類性能の劣化が発生しなかった。これは、局所ブロックと大局ブロックを組み合わせたことで、局所ブロックから採用する場合は個々の波形データにのみ依存するため点密度の影響は受けないためである。なお、元の15点/m<sup>2</sup>とF1 Scoreが変わる理由は点密度を下げる際に各地物の数が変わるためである。次に、大局クラスを担当する大局ブロックの比較結果を確認する。大局ブロックで採用したPointNet++[178]が点密度への依存性を軽減させる手法であるため、点密度が5点/m<sup>2</sup>、10点/m<sup>2</sup>の場合では大きく性能が劣化しなかった。PointNet++[178]を参考にした大局ブロックでは、代表点に対するその周辺の波形データで計算した特徴量の最大値を格納する集約アルゴリズムによって多少の点密度の影響を除くことが可能になる。しかし、点密度を最小の1点/m<sup>2</sup>まで減らした場合、15点/m<sup>2</sup>の場合と比較すると大局クラスである地面・建物・道路において大きく分類性能が劣化した。これは、浅い層において、グループ化に使用する半径を1mの球内に入る波形データに設定したため、その中に入る波形データの数が十分に格納されず分類に必要な特徴量が計算できなかったことに起因する。このように、局所・大局モデルの学習に使用したデータの点密度の1/10程度程度まで情報が欠損すると学習済みモデルの適用可能性が下がる。しかし、航空機で観測される点群・波形データもしくは点群データは、2012年以降の機材（ライカ社製のALS70など）であれば点密度が5点/m<sup>2</sup>以上は担保できることが多いため、学習済み局所・大局モデルによる点群・波形データ地物分類性能においては適用できる可能性が高い。

表 3.9: 点群・波形データの点密度が学習済み局所・大局モデルへ与える影響。分類結果はF1 Scoreを示す。

平均点密度 (点/m <sup>2</sup> )	地面	植生	建物	電線	鉄塔	道路
1	0.53	0.93	0.21	0.95	0.87	0.00
5	0.82	0.94	0.94	0.94	0.88	0.79
10	0.85	0.97	0.95	0.97	0.88	0.80
15	<b>0.86</b>	<b>0.97</b>	<b>0.97</b>	<b>0.99</b>	<b>0.88</b>	<b>0.82</b>

### 3.6 第3章のまとめ

本章では、点群データと波形データで構成される点群・波形データに対して各点の地物分類を行う手法を提案した。点群・波形データの分類においては既存の点群データに対する深層学習手法を参考にした。既存の点群データに対する深層学習手法を点群・波形データに適用する際において、点群・波形データを以下の2種類に大別可能であると仮説を立てた。

- ・ 局所クラス：個々の波形データだけで分類を行うことが可能な地物
- ・ 大局クラス：波形データの空間的な関係性である大局特徴で分類が可能な地物

この仮説に基づき、既存の点群データに対する深層学習手法に対して単に点群・波形データを入力するだけでなく波形データの持つ特性に着目した手法である局所・大局モデルを提案した。局所・大局モデルは、点群・波形データデータにおいて、個々の波形データで分類が可能な地物に特化した局所ブロックと、波形データの空間的な広がりである大局的な特徴を見ることが分類が可能な地物を担当する大局ブロックで構成される。なお、ベンチマークデータセットにおいては、波形データで分類可能な地物は、「植生」、「電線」、「鉄塔」である。また、波形データの空間的な広がりを見ることが分類が可能な地物としては、「地面」、「建物」、「道路」である。

具体的な局所・大局モデルの層設計は、既存の点群データに対する深層学習手法に対して波形データを扱えるように拡張した層設計を大局ブロックとし、波形データに対して独立して畳み込み演算を適用する層設計を局所ブロックとして組み合わせた。具体的には、大局ブロックはある波形データに対してその近傍の波形データをグループ化してそのグループ内において代表値を算出することで、近傍特量抽出を実行する。この近傍特徴抽出を複数回繰り返し適用することで大局的な波形データの特徴が獲得できる。局所ブロックは各波形データに対して独立して1次元の畳み込み演算を繰り返し適用する層設計である。1次元の畳み込み演算は、音声データにて使用される手法であり、音声データと同じような「時系列性」を有する波形データに適用した。これらの大局ブロック・局所ブロックを組み合わせることで、波形データで分類可能な地物とそれ以外に対して、長所短所を補うことが可能になった。

ベンチマークデータセットを用いて、提案した局所・大局モデルの有効性を確認した。比較対象として局所ブロック・大局ブロックをのみの場合を用いた。これらの比較対象よりも、提案手法の局所・大局モデルによって高い分類性能が定量的に得られた。また、定性的な評価においても、局所・大局モデルによって得た分類結果は真値ラベルと近い結果が得られた。

## 第 4 章

### 点群データからの波形推定手法の提案

## 4.1 はじめに

3次元地図の自動作成を目的とした点群データに対する地物分類を実施する際には、一般的に点群データの幾何的な情報だけでなく、観測時の波形データを用いること、つまり点群・波形データによる地物地物分類が重要になる。この観点から、本研究の第3章においても点群・波形データの地物分類を実施し、その際に深層学習手法を用いることで高い地物分類性能を達成した。しかし、点群・波形データの問題として、アーカイブデータ（自治体への納品物やオープンデータなど）では点群データのみが公開され、ストレージの削減の観点から波形データが削除されることが多い。そこで、アーカイブ点群データをより有効活用するため、及び、提案する点群・波形データの地物分類手法をアーカイブ点群データへ活用するために、点群データから波形を推定する手法を提案する。なお、波形推定とは、「点群データを深層学習モデルへ入力し、その各点に1対1に対応する波形を推定すること」と定義する。点群データから波形を推定する深層学習手法に関する先行研究が存在しないため、近い分野の似たタスクを参考にする。点群データからの波形推定は、「点群データが本来持っている別の情報を推定するタスク」と言え、そのタスクの先行研究である「点群データからその物体の色情報を推定する深層学習手法」を参考にする。点群着色手法において、条件付き敵対的生成ネットワーク（Generative Adversarial Network, GAN）（条件付きGAN）と呼ばれる深層学習技術を使用しているため、波形推定も同様に条件付きGANを使用する。

次に本章の構成を説明する。本章では、図4.1に示すような点群データから波形を推定する手法を提案する。本章の流れは、まず波形推定の問題設定を説明する。そして、次に、点群着色手法を波形推定に適用する際の仮説について説明する。仮説として、地物分類と同様に、深層学習モデルで波形データを個々に扱う局所特徴を扱う手法と、波形データの大局特徴を扱う手法が有効と考えた。そして、仮説に基づいて提案する波形推定手法（波形推定モデル）を説明する。波形推定モデルにおいては、点群データから波形を推定する生成器には点群着色手法と同様に大局的な点群データの特徴を用いて波形推定を実行する大局生成器を用いる。そして、波形推定に対する仮説に基づき、推定された波形に対して個々の波形ごとに評価する局所識別器と、波形の形状に着目する周波数回帰の損失関数を新たに提案した。まず、手法の概略を説明した後、波形推定モデルの層設計の説明を行い、最適化を行う際に用いる損失関数と、最適化アルゴリズムについて説明を行う。次に、検証実験について述べる。実験は第3章の地物分類と同様のベンチマークデータを用いて行い、学習済みモデルの波形推定性能を評価する。次に、考察を行う。考察では、提案手法における工夫点である局所識別器と大局識別器の組み合わせと新たな損失関数の効果を検証する。さらに、大局生成器を用いた妥当性と点密度の影響を検証する。最後に、本章のまとめを行う。

なお、条件付きGANは本研究で使用する屋外で取得された点群データに適用した例は無いため、付録Cに条件付きGANに関する予備実験を行った結果を記載する。

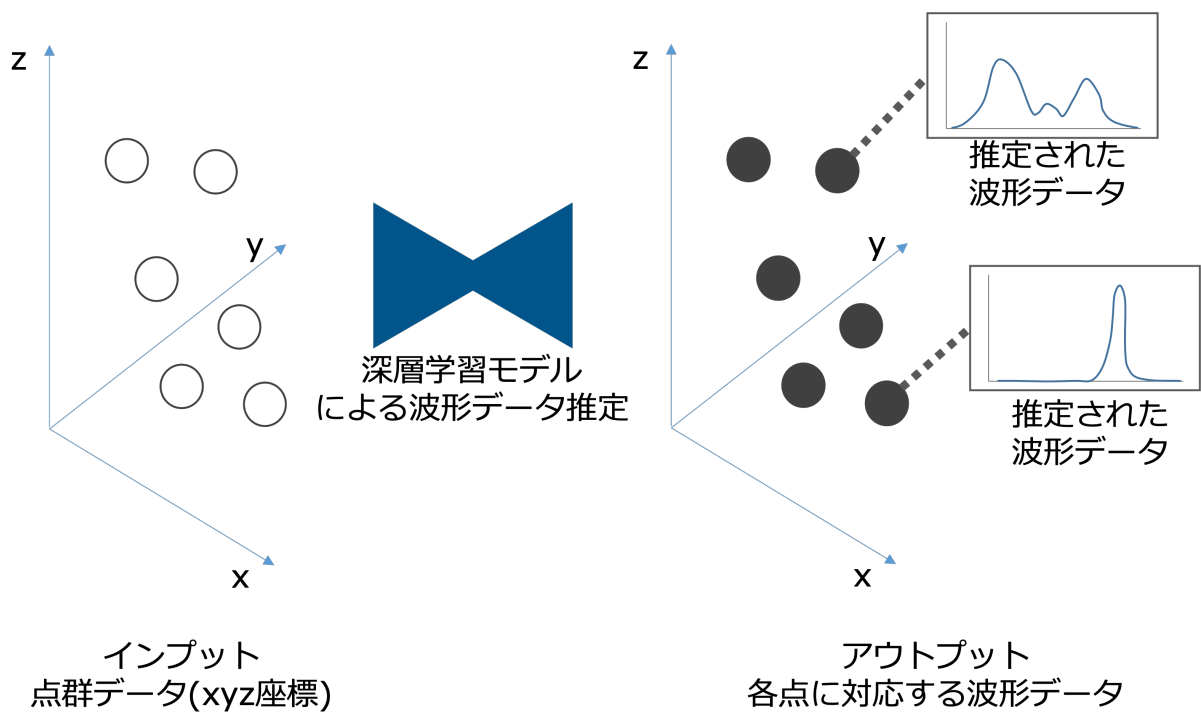


図 4.1: 点群データからの波形推定手法の概念図. 入力された点群データの各点が観測された際の波形を推定する. 条件付き GAN を参考に点群データを深層学習モデルに入力し, 波形を推定する手法を新たに提案する.

## 4.2 波形推定の問題設定

まず、深層学習モデルによる点群データからの波形推定方法について説明する。本章では、図 4.2 に示すようにファーストパルスの点群データと波形が各点において 1 対 1 に対応する学習データを使用する。この学習データによって、点群データと波形の関係性を深層学習モデルは獲得できる。具体的には学習データを用いて、点群データを深層学習モデルに入力し、その入力された点に対応する点群データを推定する。この際の仮説として、図 4.2 の赤丸で示した点に対応する波形推定を実施する際に、その近隣にある緑と青点を利用して赤丸に近い箇所で発生する弱いピークを算出する。そして、紫と水色の点を利用して赤丸から少し離れた箇所でも発生する弱いピークを算出する。最後に、遠くの地面に相当する黄色・オレンジ・グレーの点を利用してラストパルス相当のピークを算出することが可能になると考えた。真値の波形と推定した波形を最小化する最適化を実行することで、深層学習モデルが持つ学習可能な重みを獲得する。そして、学習済み深層学習モデルを用いて未知の点群データを入力した場合においても、学習時の点と波形の関係性から類推することで各点に対応した波形を統計的に推定できる。この学習済み深層学習モデルによって、単なる点群データだけではなく、学習データの点群データと波形の経験的な知識も使用可能である。

次に、点群データからの波形推定手法の問題設定を定義する。深層学習手法を用いた点群データから波形を推定する過程は、入力された点群データ  $\mathbf{P} \in \mathbb{R}^{K \times 3}$  から波形  $\mathbf{F}_{\text{real}} \in \mathbb{R}^{K \times M}$  の情報を点に 1 対 1 に対応するように推定する (図 4.3)。ここで、 $K$  は深層学習モデルに入力される点の数であり、1 点ずつモデルに入力するのではなく、ある程度の点数を一度に深層学習モデルへ入力する。また、 $M$  は波形の強度のサンプル数の次元である。つまり、ひとつの波形の中に 100 個の強度データが格納されている場合は  $M=100$  となる。点群データ  $\mathbf{P}$  は  $x, y, z$  座標情報を各点に有しており、サイズ  $K \times 3$  の実数の多次元配列として表される。 $\mathbf{F}_{\text{real}}$  は点群データ ( $\mathbf{P}$ ) に対応する波形の集合で、サイズ  $K \times M$  の実数の多次元配列として表される。ここでは点群データから波形を推定しており、深層学習モデルの学習過程では観測値の波形  $\mathbf{F}_{\text{real}}$  と推定した波形  $\mathbf{F}_{\text{pred}}$  が最小になるように深層学習モデルの重みを最適化によって調整する。

この最適化過程について詳しく説明する。波形推定における最終的な目標は、与えられた  $\mathbf{P}$  に対して、1 対 1 に対応する  $\mathbf{F}_{\text{pred}}$  を推定する大局生成器  $\mathcal{G}$  を最適化することである。これを実現するために、点群着色と同様に条件付き GAN 手法を用いて、大局生成器  $\mathcal{G}_{\theta_G}$  を学習可能な重みを有する深層学習モデル  $\theta_G$  として表現する。さらに詳しく表現すると、 $\theta_G = \{\mathbf{W}_{1:L}; \mathbf{b}_{1:L}\}$  のようになり、 $L$  層の深層学習モデルの重み  $\mathbf{W}$  とバイアス  $\mathbf{b}$  を表し、損失関数を定義して最適化することで得られる。観測値の  $\mathbf{F}_{\text{real}}$  は、最適化時にのみ利用可能であり、言い換えれば、学習済みモデルを用いた予測時には点群データ  $\mathbf{P}$  のみが使用される。本研究ではさらに識別器を  $\theta_D$  でパラメータ化された学習可能な深層学習モデル  $\mathcal{D}_{\theta_D}$  で学習し、 $\mathbf{F}_{\text{real}}$  または  $\mathbf{F}_{\text{pred}}$  を分類して、波形推定結果の評価を行い、推定される波形の質を向上させる。なお、理論的な識別器の効果として、識別器が 2 つの確率分布の距離を計算することと同義であるため、大局生成器によって推定された波形の確率分布と観測値の波形の確率分布を一致もしくはその差を最小化させる効果がある。

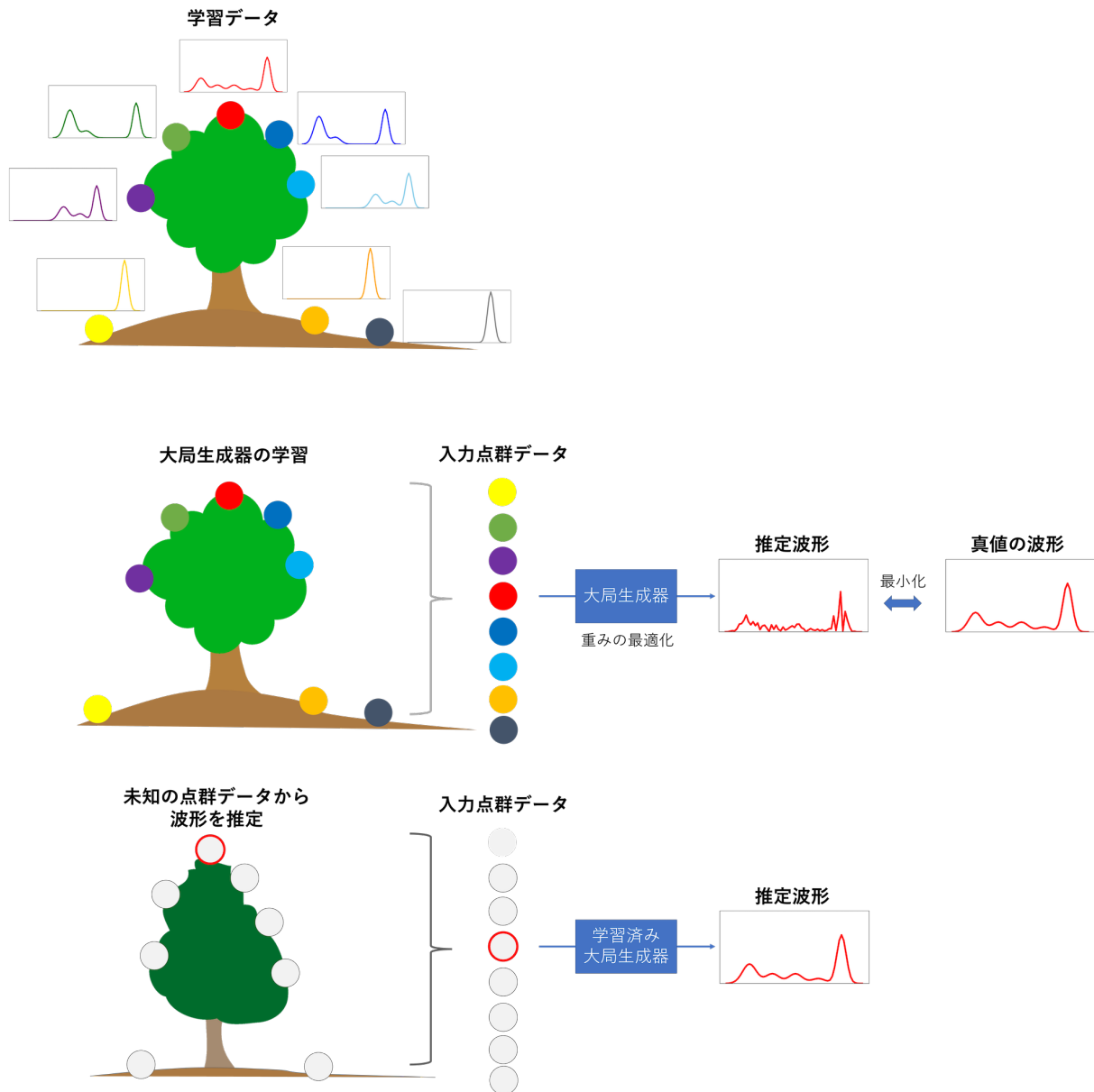


図 4.2: 点群データと波形がペアになった学習データを利用した大局生成器の学習。学習データにおいて、点群データの各点には波形の情報を保持している。この学習データを用いて点群データを深層学習モデルである大局生成器に入力して各点の波形を推定する。その際に真値の波形と推定した波形の差を最小化するように生成器の重みを最適化によって獲得する。そして、学習済み大局生成器によって未知の点群データから波形を推定する。

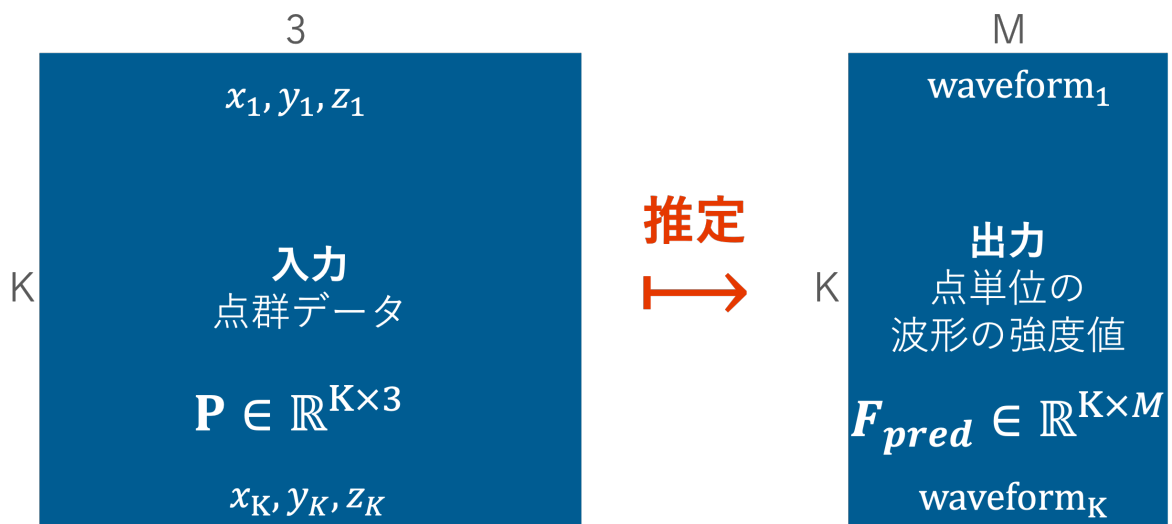


図 4.3: 点群データからの波形推定の問題設定. 入力された点群データ  $\mathbf{P} \in \mathbb{R}^{K \times 3}$  から波形  $\mathbf{F}_{pred} \in \mathbb{R}^{K \times M}$  の情報を点に 1 対 1 に対応するように推定する.

### 4.3 波形推定に点群着色手法を適用する際の仮説

着色拡張モデルにおいて、点群着色手法を波形推定に適用する際に、点群・波形データの地物分類を提案する際と同様に、波形を個々に扱うことが重要であると仮説を立てる。その際に、生成器においてはダウンサンプリング構造とアップサンプリング構造を持つ大局生成器は変更できないため、識別器と損失関数に着目する。

- 個々の波形の評価

点群データの各点の色を推定する点群着色の場合には、 $x, y, z$  座標に対応する色情報を R,G,B として推定する。この場合は、空間的な色の整合性を評価することが重要であるため、ある一定範囲（例えば、3 m の範囲内）の色の広がり方を大局識別器によって評価する。この点群着色手法の考え方は、推定された波形の評価を行う上でも同様であると考えられ、波形と波形の空間的な関係性である大局特徴を用いて評価すべきである。したがって、点群着色手法で使用された大局識別器の考え方を波形推定においても使用する。

しかし、波形を推定する場合において、第3章で個々の波形を組み合わせた地物分類手法が有効であったことから、推定された個々の波形の評価も重要であると考えられる。そのため推定された波形ごとの評価も同様に第3章の局所ブロックを参考にして、個々の波形に対する「局所識別器」を新たに提案する。

- 波形形状の評価

点群着色においては、真値の色と推定した色（RGB 値）の差分を直接最小化することで大局生成器の重みを更新する。この損失関数の設計において、色空間において青赤緑色の関係性が薄いため、簡単な損失関数の設計で十分である仮定が強く入る。しかし、波形においては、波形が持つ形状等に対する尤もらしさの評価が必要になる。そこで、波形の形状を端的に表しつつ、誤差逆伝播を実現するために微分可能な形式で表現できる損失関数が必要になる。

波形は時系列的な反射の強度が格納されていることに着目する。これは音声データ等の波形解析手法が有効であると考えられる。そこで、波形に対して高速フーリエ変換を適用して周波数特性を求め、周波数領域において推定した波形と観測値の波形を比較し、その差を最小化する損失関数を新たに提案する。この周波数領域における計算が有効な理由について説明する。この際に、既存の深層学習を用いた画像の着色手法で提示された問題点 [87] を参考にする。深層学習による画像の着色手法では推定した各画素の色と真値の色の強度差を計算する回帰の損失関数を用いただけでは推定した色がセピア色になる傾向がある。図 4.4 の左図に示すように、セピア色はどの色からも近い色、もしくは平均的な色であるため、色の強度差の損失関数を用いた場合の学習データにおける全体最適になるためである。学習データの全体最適とは、深層学習モデルの最適化に使用されるミニバッチ学習に起因する。深層学習モデルの最適化では、学習の高速化や安定化のために学習データを個々に扱わず、計算に使用する GPU のメモリサイズに比例して学習データから復習をサンプリングしたミニバッチと呼ばれる単位で扱う。したがって、ミニバッチ単位においてすべての予測結果と真値の差が最小になる場合が教師データの全体最適となる。このミニバッチ最適化における全体最適の現象を波形推定に当てはめると、図 4.4 の右図に示すように、波形の強度の差の平均絶対誤差は平坦な波形を推定すると全体最適になる可能性が高い。さらに波形特有の問題としては強度の値はゼロが多く、すべての強度の値をゼロと予測すると全体最適となり波形の予測性能が悪化する可能性が高い。図 4.5 に示すように周波数領域でも損失関数を計算することで、平坦な形状を有する波形の推定や全てゼロと予測する場合において周波数領域では観測値と推定値の差が大きくなるため、これらの問題点が解決できると考えられる。なお、この際には誤差逆伝播を用いた深層学習モデルの最適化を行う際には、全ての関数が微分可能である必要がある。そこで、高速フーリエ変換を微分可能な式へ近似する。

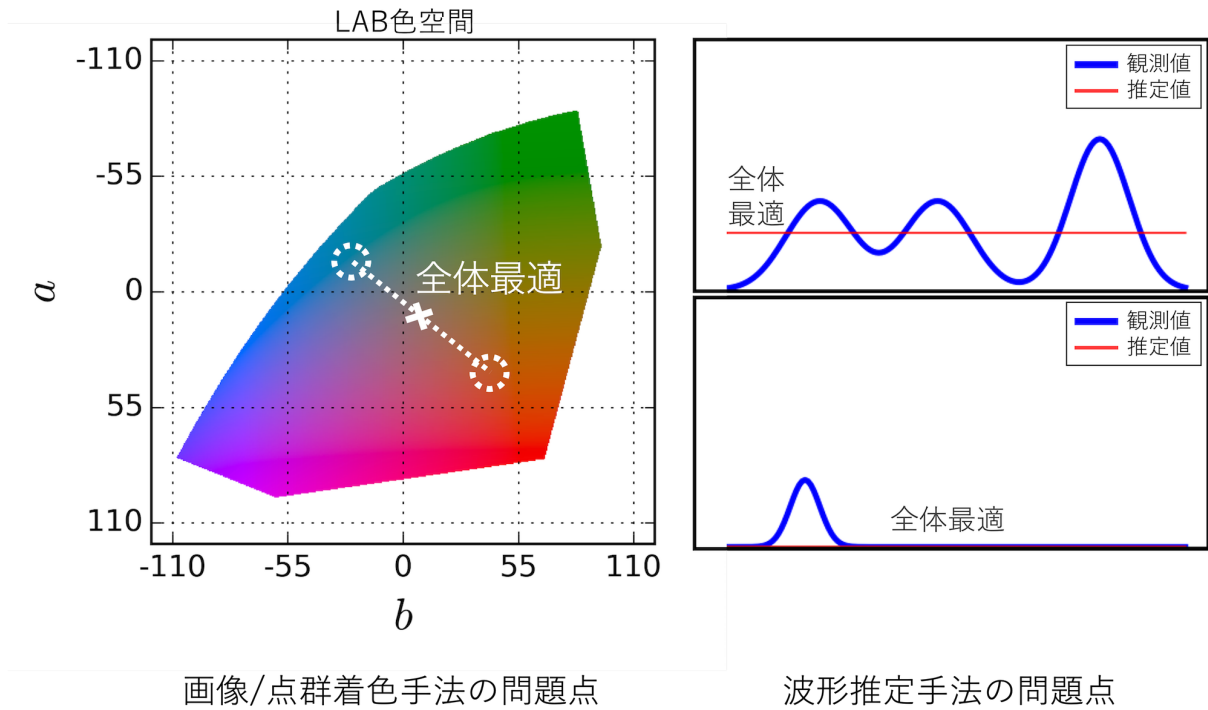


図 4.4: 強度の差を計算する損失関数の問題点. 深層学習を利用した画像の着色手法において色の強度差ではセピア色が学習データの全体最適となり, 波形の強度差を計算する場合において平坦な強度を推定すると学習データの全体最適になる可能性が高い. また, 波形はゼロが多いため, 全てゼロと推定すると学習データの全体最適になる可能性が高い.

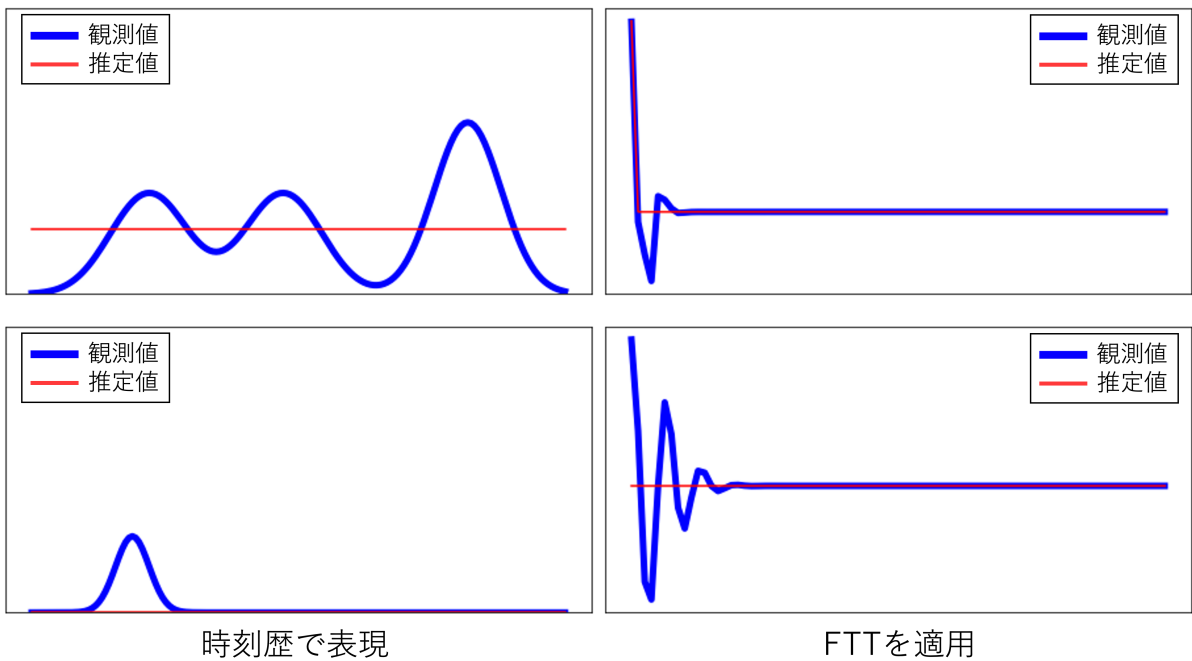


図 4.5: FFT を適用する利点. 平坦な波形や全てゼロの値を持つ波形を予測した場合, 時刻歴では回帰の損失関数の値が小さくなるが, FFT をかけた後に回帰を実行すると差が大きくなる.

## 4.4 波形推定モデルの提案

本研究ではにおける深層学習手法を用いた点群データから波形を推定する手法を説明する。波形推定では点群着色手法を参考にし、大局生成器と2つの識別器で構成される条件付きGANの考え方をを用いる。点群着色手法では大局生成器の出力がRGBという3次元の情報であったが、大局生成器の出力を連続的に記録された反射の強度情報を示す波形を推定するように変更する。識別器は点群着色手法を参考にしたものを大局識別器として、さらに個々の波形の尤もらしさを評価する識別器として「局所識別器」を新たに提案する。また、損失関数についても同様に、点群着色手法を参考にするが、波形を周波数領域で評価する損失関数である「周波数回帰」を新たに提案する。

次に、本節の構成を説明する。まず提案手法の概要を述べ、その後、大局生成器の層設計と2つの識別器の層設計を説明する。次に、最適化を行う際の損失関数と最適化アルゴリズムを説明する。

### 4.4.1 波形推定モデルの概要

ここでは提案する波形推定手法である波形推定モデルの概要を説明する。波形推定モデルは条件付きGANの枠組みを利用した点群着色手法を参考にしている。波形推定モデルでは、点群着色手法に対して「層設計の提案」と「損失関数の提案」を行ったため以降でそれぞれの概要を説明する。

まず、層設計の提案について説明する。波形推定モデルは条件付きGANの枠組みを用いるため、大局生成器と識別器の2つの深層学習モデルを用いる。この深層学習モデルの層設計において点群着色手法と同様の考えを用い、大局生成器( $G$ )と大局識別器( $D^{\text{space}}$ )を基本構造とする。大局生成器は点群データからの波形推定する深層学習モデルであり、ダウンサンプリングとアップサンプリングを有する大局的な特徴を抽出可能な構造を有する。大局識別器( $D^{\text{space}}$ )は、ある一定範囲の推定された波形に対し、空間的な整合性を評価する深層学習モデルである。しかし、大局識別器では、仮説として立てた波形ごとの評価は大局識別器だけでは達成できないため、新たに波形単位に評価する識別器を提案する。具体的には、新たに波形単位に評価する識別器として、第3章の局所ブロックを参考にしたである局所識別器( $D^{\text{wave}}$ )を提案する。局所識別器は推定した波形もしくは観測値の波形に対して1次元の畳み込み演算を繰り返し適用する。それにより、波形らしさに関する特徴量を識別器の内部で獲得でき、観測値の波形か大局生成器で推定した波形かを分類する。この波形単位の識別器によって、空間的な整合性だけでなく、波形を個別に見た際の波形らしさの評価が可能になる。なお、2つの識別器は訓練データを用いた最適化の際にのみ用い、学習済みモデルによる点群データから波形を推定する際には大局生成器のみを用いる。

次に損失関数の提案について説明する。最適化に用いる損失関数も点群着色手法を参考にしており、波形の反射の強度を評価する損失関数を用いる。しかし、反射の強度だけでは、仮説をして立てた形状の評価が行えない。そこで、波形形状を評価する損失関数(周波数回帰)を新たに提案する。周波数回帰では波形の形状を評価するために、観測値の波形と大局生成器で推定した波形に対して微分可能なFFTを適用して波形の周波数特性を獲得し、周波数に関する情報の差を計算する。推定した波形と観測値の波形の周波数特性の差を最小化することで、強度の直接的な最小化では不十分であった波形形状の評価を補助できる。また、最適化の計算時において、強度の差分では観測値の波形にゼロが多い場合において大局生成器はゼロを出力すると損失関数の値は減るため、適切な最適化が行えない可能性が高い。そこで、非ゼロの値を有しつつ等価な形式に変換して最適化を行う必要がある。

### 4.4.2 大局生成器の層設計

まず、大局生成器の層設計の概要を説明する。大局生成器( $G$ )は、点群データ( $P$ )を入力し、その点に対して1対1に対応する波形( $F_{\text{pred}}$ )を推定する。この大局生成器による波形推定は写像  $G: \mathbb{R}^{K \times 3} \rightarrow \mathbb{R}^{K \times C}$  として定義できる。なお、大局生成器は点群データに対する深層学習手法を基本的な構造として、第3章の大局ブロックと同様にダウンサンプリングと

近傍特徴抽出で構成されるエンコーダと、アップサンプリング層で構成されるデコーダを備えた構造である (図 4.8 参照)。

次に、本項の流れを説明する。まず、エンコーダで使用する点群データに対する近傍特徴抽出を説明する。近傍特徴抽出は、第 3 章の点群・波形データの地物分類を実行した大局モデルを参考にしており、まず点群データを扱う上での変更点を述べる。その後、各点に対する特徴抽出方法と、点群データに対する近傍特徴抽出の仕組みを説明する。次に、具体的な層設計について説明する。層設計については、エンコーダについて説明した後でデコーダを説明する。以降で詳しい説明を行う。

まず、エンコーダで使用する点群データに対する近傍特徴抽出について説明する。点群データに対する近傍特徴抽出は、点群データと点群データの近傍関係の特徴抽出器として使用する。点群データに対する近傍特徴抽出は第 3 章の波形に対する近傍特徴抽出と基本的には同じであるが、波形とは違い、点群データの  $x, y, z$  座標には時系列的な要素を有していないため、点群データの特徴抽出手法に変更する。

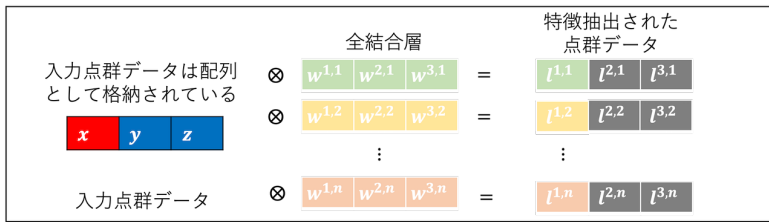
次に、全結合層を使用した個々の点群データに対する特徴を抽出する手法について説明する。各点の特徴量を計算する全結合層は図 4.6 のように定義できる。まず、入力された点群データに対してどのように演算が適用されるかを説明する。入力層では、 $1 \times 1 \times M$  の配列として点群データは扱われる。これに対して、全結合層を適用すると、ある点群データの値に対して、全ての点の値を用いて積和演算を行う。この際、積和演算を行う際の重みは学習可能な値である。このような積和演算を全ての配列の値に対して適用する。なお、積和演算は各層に 1 つだけでなく、64 個などある程度の量を設定する。この積和演算によって、入力された  $1 \times 1 \times M$  の点群データに対し、前述の 64 個の積和演算である場合、 $64 \times 1 \times M$  の特徴マップが作成される。次に、入力ではなく、隠れ層における積和演算を説明する。隠れ層  $l$  では、 $l-1$  層の積和演算によって 128 個の特徴マップが作成された場合、 $128 \times 1 \times M$  の特徴マップである。これに対し、全結合層を適用する場合を考える。入力層とは違い、128 個の特徴マップがあるため、特徴マップ全てに対して積和演算を適用する。この積和演算を例えば 256 個用意すると、 $l+1$  層に対しては、 $256 \times 1 \times M$  の特徴マップを送ることになる。なお、これらの積和演算は、学習可能な重みを持つため、最適化によって得られる。つまり、波形推定に寄与するように点群データの特徴を自動的に抽出できることを示す。

次に、図 4.7 に示す点群データの近傍特徴抽出について説明する。波形の地物分類と同様に、入力された点群データの代表点を FPS により算出する。代表点に対して半径  $r$  で定義される球の内部に入る点を近隣点として定義する。近傍点に対しては全結合層を使用した個々の点群データに対する特徴抽出を実施する。その後、グループ内の点における代表値を Max-Pooling によって算出し特徴量を代表点へ集約する。

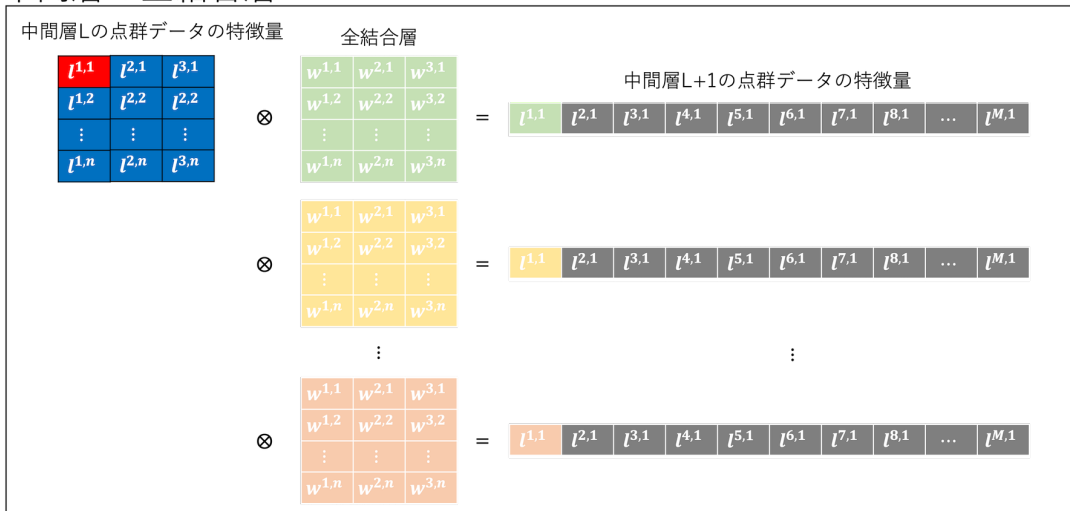
次に、大局生成器の層設計についてエンコーダ、デコーダの順で述べる。まず、エンコーダについて説明する。エンコーダは、点群データ ( $P$ ) を入力し、3 つのダウンサンプリング層と活性化関数としてパラメトリック ReLU [74] を持つ。エンコーダの各ダウンサンプリング層は、浅い層から深い層へ向けて [8, 192, 4, 096, 2, 048] 個の点数になるように徐々に疎に FPS を用いてダウンサンプリングして代表点を決定する。各ダウンサンプリングの際には、単に点数を落とすのではなく、上記で説明した近傍特徴抽出による情報の集約を各代表点に対して行う。この近傍特徴抽出によって作成される特徴マップの数は、浅い層から深い層へ向けて [256, 512, 1, 024] 個である。また近傍特徴抽出内部の特徴抽出の後にはバッチ正規化層とパラメトリック ReLU が続き、次の層への入力になる。

次に、デコーダの層設計を紹介する。デコーダは 3 つのアップサンプリング層を持ちエンコーダからの特徴マップが入力される。デコーダの構成は、点数を増やすための 3 つのアップサンプリング層である。アップサンプリング層は、3 次元内挿を行った後に MLP 層を適用する。この全結合層では、深い層から浅い層へ向けて [1024, 512, 256] の特徴マップを持ち、活性化関数はパラメトリック ReLU が使用される。全結合層では、エンコーダとデコーダの間にスキップコネクションを使用している。スキップコネクションを入れることで、エンコーダで失った高い密度の点群データの情報をデコーダ側へ直接伝えることが可能になる。デコーダの最後の層では波形推定を実施し、波形のサンプル数 (観測した反射の強度数) と同じ次元数の出力を行う。なお、その際には、非線形な活性化関数を用いずに、線形に出力する。

## 入力層の全結合層



## 中間層の全結合層



## 分類層の全結合層（6クラスの例）

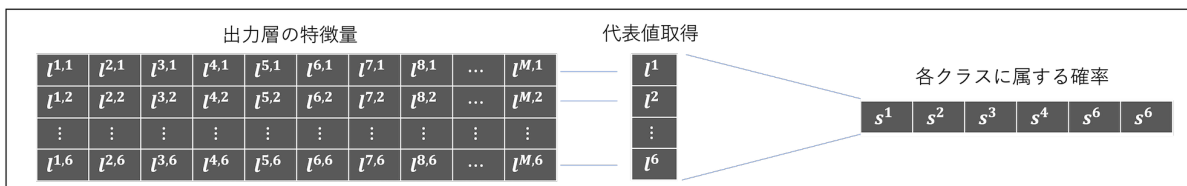


図 4.6: 個々の点群データに対する特徴抽出の仕組み。入力された点に対して全結合層を適用し点の特徴を抽出する。

以上のことから、大局生成器  $\mathcal{G}$  は、点群データを入力して波形を推定する非線形な関数である ( $\mathbf{F}_{\text{pred}} = \mathcal{G}(\mathbf{P})$ ) と言える。

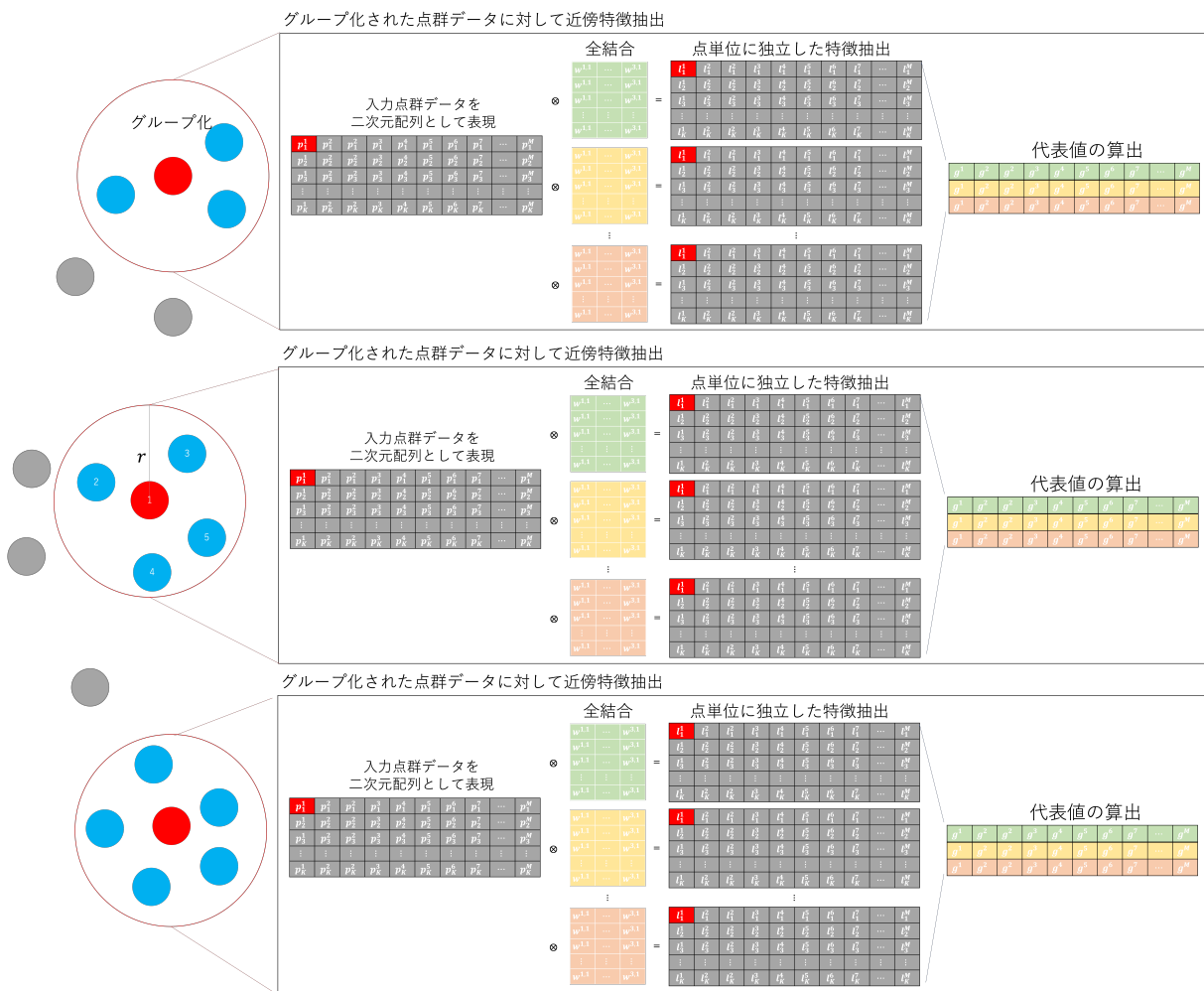


図 4.7: 点群データに対する近傍特徴抽出の仕組み. ある点に対してその周辺の点をグループ化する. このグループ化された点に対して独立した特徴抽出を適用する. そして, 各点の特徴から最大値を用いて代表値を取得する.



### 4.4.3 識別器の層設計

次に、識別器の層設計について説明する。識別器は、点群着色手法と同様に空間関係を評価する大局識別器 ( $\mathcal{D}^{\text{space}}$ ) と、本研究で新たに提案する波形に特化した局所識別器 ( $\mathcal{D}^{\text{wave}}$ ) で構成される。以降で、大局識別器 ( $\mathcal{D}^{\text{space}}$ ) と局所識別器 ( $\mathcal{D}^{\text{wave}}$ ) それぞれについて説明する。

- 大局識別器

まず、大局識別器  $\mathcal{D}^{\text{space}}$  を説明する。大局識別器 ( $\mathcal{D}^{\text{space}}$ ) は、写像 ( $\mathcal{D}^{\text{space}} : \mathbb{R}^{K \times (M+3)} \rightarrow \mathbb{R}^{2048 \times 1}$ ) として定義される。 $\mathcal{D}^{\text{space}}$  は、推定した波形と  $\mathbf{P}$  のペアもしくは、観測値の波形と  $\mathbf{P}$  のペアが入力され、出力は観測値である確率である。この際、 $\mathbf{P}$  は点群データを示し、推定した波形だけでなく座標を用いることで推定した波形の大局的な特徴抽出を実現するために使用する。

次に、図 4.9 に示す大局識別器  $\mathcal{D}^{\text{space}}$  の層設計の概略を説明する。 $\mathcal{D}^{\text{space}}$  は第 3 章の地物分類手法における大局ブロックのエンコーダ部分と同様である。ただし出力は地物分類では各地物に属する確率から観測値の波形である確率に変更する。 $\mathcal{D}^{\text{space}}$  の機能は、推定された結果が観測値かを、点群データと波形の集合として比較する。この集合を大局的に評価することで、波形の空間的な関係や配置の正しさを評価できる。

次に具体的な層設計について説明する。 $\mathcal{D}^{\text{space}}$  のダウンサンプリング層は 3 層で、点数は浅い層から順に [8,192, 4,096, 2,048] である。それぞれのダウンサンプリングの際に波形の近傍特徴抽出を実行する。波形の近傍特徴抽出で作成する特徴マップの数はそれぞれ浅いものから深いもの順で 256 から 1,024 とする。なお、活性化関数として Leaky ReLU を用いる。なお、Leaky ReLU のハイパーパラメータは多くの論文で使用されるデフォルト値である 0.2 に設定した。識別器の出力は、 $2,048 \times 1$  の実数の多次元配列で表現され、全結合層の後にシグモイド関数で観測値の波形である確率を計算する。

- 局所識別器

次に、新たに提案する波形単位の識別器である局所識別器 ( $\mathcal{D}^{\text{wave}}$ ) を説明する。図 4.10 に示す  $\mathcal{D}^{\text{wave}}$  は、大局生成器で推定した波形と観測値の波形を個々の波形単位の分類するもので、その層設計は単純な 1D CNN である。これは、第 3 章の地物分類手法における局所ブロックと同様の層設計である。この 1D CNN で表現される  $\mathcal{D}^{\text{wave}}$  は写像 ( $\mathcal{D}^{\text{wave}} : \mathbb{R}^{K \times M} \rightarrow \mathbb{R}^{K \times 1}$ ) として定義される。つまり、個々の波形に対して波形として最ももらしいか否かを判別することが可能である。

次に、具体的な局所識別器の層設計を説明する。基本的な層設計は第 3 章で提案した局所ブロックと同様であるが、出力は地物の分類から観測値の波形である確率に変更した。局所識別器は、図 4.10 の赤三角で示す  $1 \times 3$  の畳み込み層を三回適用し、特徴マップは浅い方から深い方へ 256, 512, 1,024 に設定する。畳み込み層の後には、非線形活性化関数として ReLU が続き、次の層へ渡す際にはドロップアウトを利用した。 $\mathcal{D}^{\text{wave}}$  の出力は、入力する点群データと波形の数である  $K \times 1$  の実数値の多次元配列で波形単位の観測値の波形である確率を算出する。なお、確率を求める際には畳み込み層から得られた特徴量に対し図 4.10 のオレンジ色の三角で示す全結合層を適用し、その後シグモイド関数を適用する。

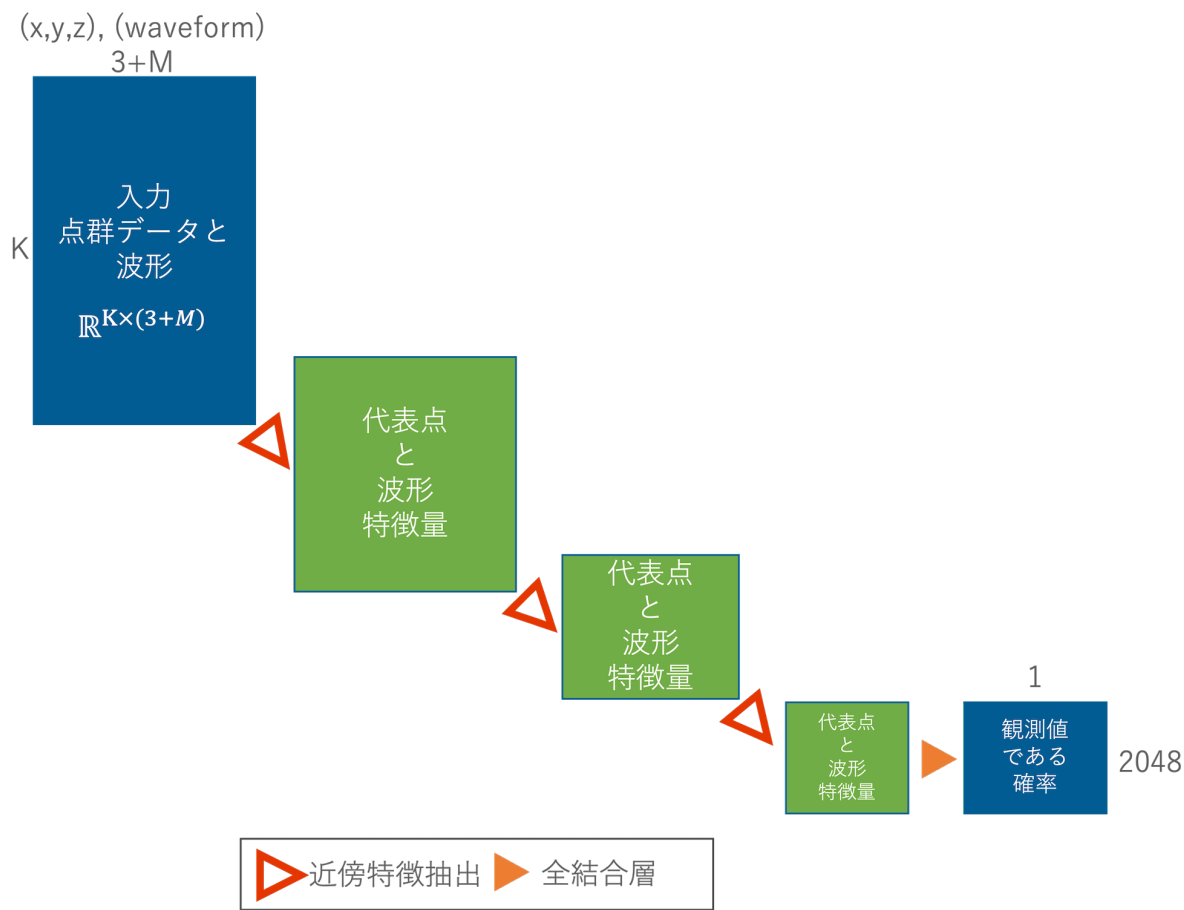


図 4.9: 大局識別器の層設計. 基本的な構造は第 3 章における大局ブロックのエンコーダ部分と同様である. 出力がダウンサンプリングされた次元数の 2048 次元になり, 観測値の波形である確率の情報を保持している.

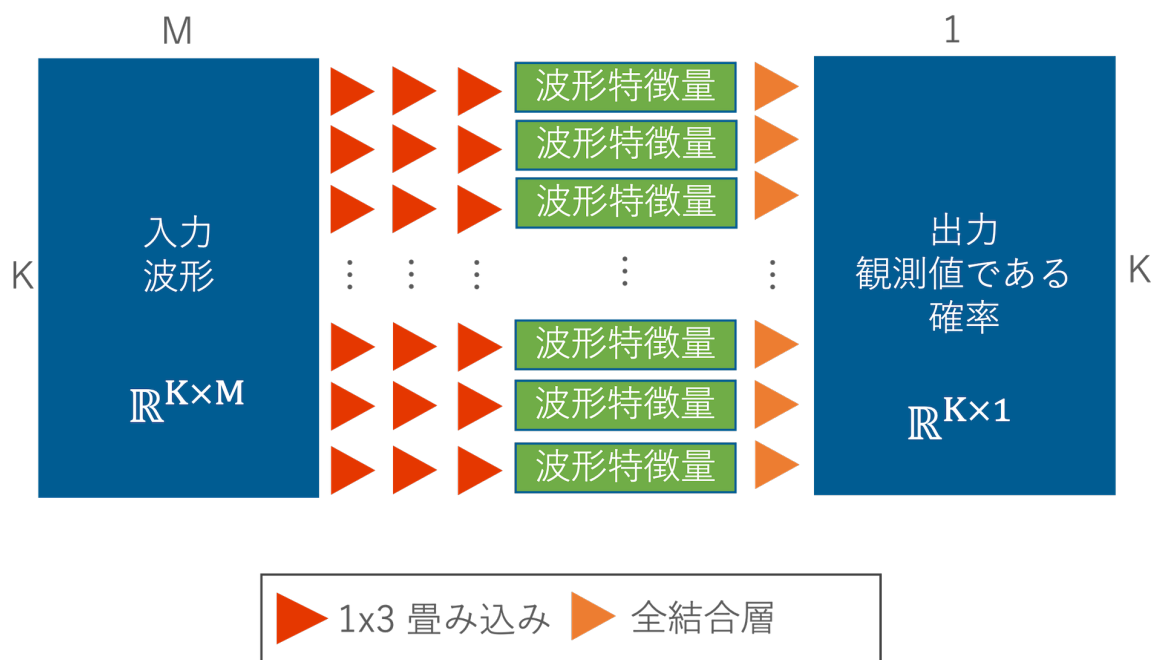


図 4.10: 局所識別器の層設計. 基本的な構造は第 3 章の局所ブロックと同様である. 入力された波形に対して独立して特徴を抽出し, それぞれが観測値の波形である確率を算出する.

#### 4.4.4 損失関数

ここでは、最適化に使用される損失関数を説明する。大きく分けると、点群着色手法を参考にした損失関数と、波形推定向けに新たに定義した損失関数の2種類の損失関数を用いる。さらに、新たに定義した損失関数は、波形の形状についての再現性を評価する損失関数(周波数回帰)と局所識別器の損失関数で構成される。以降でそれぞれについて詳しく説明する。

- 点群着色手法を参考にした損失関数

まず、点群着色手法を参考にした損失関数について説明する。点群着色手法を参考にした損失関数は、波形が持つ反射の強度の再現性を評価する損失関数(強度回帰)と、大局識別器の損失関数の2つで構成される。強度回帰では、大局生成器で推定した波形と観測値の波形における反射の強度の差を最小化することで、大局生成器の波形推定能力を向上させる。また、大局識別器の損失関数は大局識別器で使用される唯一の損失関数であり、大局生成器で推定した波形か観測値の波形かを分類する機能を有する。このような仕組みは条件付きGANで提案されたものであり、条件付きGANの仕組みを取り入れることで、大局生成器 $\mathcal{G}$ の波形推定能力を高められる。この理由は、大局識別器による分類は推定した波形と観測値の波形の確率分布を一致させる働きがある。つまり、波形の反射の強度だけの評価では成し得なかった、推定した波形が有する確率分布の評価が可能になる。

これら2つの損失関数について、以降で詳しく説明する。

- 強度回帰

まず、波形が持つ反射の強度の再現性を評価する損失関数である強度回帰について説明する。この強度回帰は大局生成器で推定した波形と観測値の波形のL1距離を計算する。L1距離を計算する際には、各波形に $M$ 個の反射の強度値があった際には、各値を個別に差の絶対値を計算した後に平均を求める。さらに、大局生成器に点群データを入力する際には $K$ 点同時に入力するため、それに応じて推定された $K$ 個の波形に対して同様に計算する。そして、計算された差の絶対値を平均することで最終的な値が算出される。この強度回帰は以下のように定義される。

$$\mathcal{L}_G^{\text{content}} = \frac{1}{K} \frac{1}{M} \sum_{i=1}^K \sum_{j=1}^M |\mathbf{F}_{\text{real},i,j} - \mathbf{F}_{\text{pred},i,j}|. \quad (4.1)$$

なお、この強度回帰は大局生成器で推定された波形と観測値の波形の強度が一致する際に最小になる。

- 大局識別器の損失関数

次に、大局識別器の損失関数を説明する。この損失関数は条件付きGANの仕組みを取り入れているため、大局生成器と大局識別器の2つの深層学習モデルに対して、別の挙動を示す。以降で、大局生成器に対する定義、大局識別器に対する定義を示す。

- \* 大局生成器

まず、大局生成器における損失関数を定義する。大局生成器においては、大局識別器に対して「観測値の波形」として分類させるように損失関数を設定する。つまり、大局識別器の出力である「観測値の波形( $\mathcal{G}(\mathbf{P}))$ )である確率」を1にするように設定する。

$$\mathcal{L}_G^{\text{space}} = \mathbb{E}[\log(1 - \mathcal{D}(\mathcal{G}(\mathbf{P})))] \quad (4.2)$$

ここで、 $1 - \mathcal{D}^{\text{space}}(\mathcal{G}(\mathbf{P}))$ は $\mathcal{G}(\mathbf{P})$ が1になる際に最小になる。つまり、大局識別器が大局生成器で推定された波形を観測値と分類すると損失関数の値が小さくなる。

\* 大局識別器

次に、大局識別器に対する損失関数を定義する。大局識別器は大局生成器の最適化と大局識別器の最適化のどちらも担うことになる。したがって、以下の式に示すようにそれぞれに対応する項が必要になる。

$$\mathcal{L}_D^{\text{space}} = \mathbb{E}[\log \mathcal{D}(\mathbf{F}_{\text{pred}})] + \mathbb{E}[\log(1 - \mathcal{D}(\mathcal{G}(\mathbf{P})))] \quad (4.3)$$

ここで、右辺の第1項は観測値の波形が大局識別器に入力された場合を表す。なお、この項は  $\mathcal{D}^{\text{space}}(\mathbf{F}_{\text{real}})$  が1になる際に最小になる。また、右辺第2項は、大局生成器で推定された波形が大局識別器に入力された場合を表す。なお、大局識別器は、大局生成器によって推定された波形が入力された際には、それを観測値と間違わないようにするため、 $\mathcal{D}^{\text{space}}(\mathcal{G}(\mathbf{P}))$  が0の時に最大になるように最適化を行う。

– 局所識別器の損失関数

局所識別器の損失関数は、層設計として新たに導入した局所識別器を最適化させるための損失関数である。この損失関数は、条件付きGANの仕組みを取り入れているため、大局生成器と局所識別器の両者に作用する。それぞれについて別の挙動を示すため、以降で個別に定義する。

\* 大局生成器

まず、大局生成器における損失関数を定義する。大局生成器においては、局所識別器に対して「観測値の波形」として分類させるように損失関数を設定する。つまり、局所識別器の出力である「観測値の波形である確率  $\mathcal{G}(\mathbf{P})$ 」を1にするように設定する。

$$\mathcal{L}_G^{\text{wave}} = \mathbb{E}[\log(1 - \mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P})))] \quad (4.4)$$

ここで、 $1 - \mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P}))$  は  $\mathcal{G}(\mathbf{P})$  が1になる際に最小になる。つまり、局所識別器が大局生成器で推定された波形を観測値であると分類する際に最小になる。

\* 局所識別器

次に、局所識別器に対する損失関数を定義する。局所識別器に対する損失関数は大局生成器の最適化と局所識別器の最適化のどちらも担うことになる。したがって、以下の式に示すようにそれぞれに対応する項が必要になる。

$$\mathcal{L}_D^{\text{wave}} = \mathbb{E}[\log \mathcal{D}^{\text{wave}}(\mathbf{F}_{\text{pred}})] + \mathbb{E}[\log(1 - \mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P})))] \quad (4.5)$$

ここで、右辺の第1項は観測値の波形が局所識別器に入力された場合を表す。なお、この項は  $\mathcal{D}^{\text{wave}}(\mathbf{F}_{\text{real}})$  が1になる際に最小になる。また、右辺第2項は、大局生成器で推定された波形が局所識別器に入力された場合を表す。なお、局所識別器は、大局生成器によって推定された波形が入力された際には、それを観測値と間違わないようにするため、 $\mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P}))$  が0の時に最大になるように最適化を行う。

• 新たに提案する損失関数

次に、波形推定に特化した新たな損失関数の説明を行う。新たに提案する損失関数は、「波形形状の再現性を評価する損失関数（周波数回帰）」である。周波数回帰、波形に対して高速フーリエ変換 (FFT) を適用した周波数特性のデータを用いる。FFT は、離散フーリエ変換を  $O(n \log n)$  時間で計算する。高速フーリエ変換は、信号の「周波数領域」での作業を信号の「空間領域」や「時間領域」での作業と同様に扱いやすくするため、さまざまな数値計算や信号処理技術の基礎となっている。FFT は音声分野等の「波形」に対して有効な前処理となり、波形も「波形」と同様の性質を持つため使用する。大局生成器で推定した波形と、観測値の波形に対してそれぞれFFTを適用し、周波数領域の情報にすることで波形の形状に関する特徴を得る。さらに、周波数回帰によって、強度

回帰では平均的な強度もしくはゼロを出力すると損失関数の値が小さくなる問題を回避できる。この周波数領域で両者の差を以下の式のように最小化する。

$$\mathcal{L}_G^{\text{spectral}} = \frac{1}{K} \frac{1}{M'} \sum_{i=1}^K \sum_{j=1}^{M'} |\mathbf{FFT}_{\text{real}_{i,j}} - \mathbf{FFT}_{\text{pred}_{i,j}}|. \quad (4.6)$$

なお、深層学習モデルの最適化時における、誤差逆伝播を実現させるために FFT を実装する際には、微分不可能な演算を用いることはできない。そこで、FFT を微分可能な形として実装するためには、深層学習のフレームワークにおける 1 次元の畳み込み演算と要素ごとの 2 乗を計算する関数の組み合わせる。

- 全体の損失関数

上記で説明した損失関数を組み合わせて波形推定モデルの最適化を行う。まず、大局生成器と大局識別器・局所識別器に関する条件付き GAN に関する損失関数は、以下のよう  
に 2 つの損失関数を足し合わせる。大局生成器の損失関数は以下のように定義される。

$$\mathcal{L}_G^{\text{GAN}} = \mathcal{L}_G^{\text{wave}} + \mathcal{L}_G^{\text{space}}. \quad (4.7)$$

また、2 つの識別器に関しては以下のように定義される。

$$\mathcal{L}_D = \mathcal{L}_D^{\text{wave}} + \mathcal{L}_D^{\text{space}}. \quad (4.8)$$

さらに、推定した波形と観測値の波形の差を最小化する大局生成器に対する損失関数においても同様に、以下のように足し合わせる。

$$\mathcal{L}_G = \mathcal{L}_G^{\text{GAN}} + \lambda \mathcal{L}_G^{\text{content}} + \lambda \mathcal{L}_G^{\text{spectral}}, \quad (4.9)$$

このように、3 つの損失関数が存在し、これらすべてを足し合わせて最終的な損失関数とする。

$$\mathcal{L}^{\text{total}}(\mathcal{D}^{\text{wave}}, \mathcal{D}^{\text{space}}, \mathcal{G}) = \mathcal{L}_G + \mathcal{L}_D. \quad (4.10)$$

#### 4.4.5 最適化アルゴリズム

以前の節にて説明した損失関数を用い波形推定モデルの最適化の計算を行う際にはアルゴリズム上の工夫が必要である。2 つの識別器と大局生成器の損失関数は、以下のようにミニマックスゲームでお互いを騙そうとする関係で非協力状態であるため、最適化が困難である。

$$\min_{\mathcal{G}} \max_{\mathcal{D}^{\text{wave}}} \max_{\mathcal{D}^{\text{space}}} \mathcal{L}^{\text{total}}(\mathcal{D}^{\text{wave}}, \mathcal{D}^{\text{space}}, \mathcal{G}) \quad (4.11)$$

これは、大局生成器は観測値の波形と大局生成器で推定した波形の差を最小化しつつ、さらに、推定された波形を 2 つの識別器 ( $\mathcal{D}^{\text{wave}}, \mathcal{D}^{\text{space}}$ ) へ入力した際に観測値である確率が 1 になるように最適化を行う。また、それと反対に、 $\mathcal{D}^{\text{wave}}, \mathcal{D}^{\text{space}}$  は、推定された波形をそれぞれの識別器へ入力した際に観測値である確率が 0 になるように、観測値の波形をそれぞれの識別器へ入力した際に観測値である確率が 1 になるように最適化を行う。このように両者は非協力的な最適化問題を解く必要がある。

このような非協力的なミニマックスゲームを実際のアルゴリズムに落とし込む際には、すべてを最小化問題に変更する必要がある。ある点群データ  $\mathbf{P}$  を大局生成器に入力し推定した波形 ( $\mathcal{G}(\mathbf{P})$ ) を 2 つの識別器 ( $\mathcal{D}^{\text{wave}}, \mathcal{D}^{\text{space}}$ ) に入力し、それぞれの識別器が観測値の波形と分類する確率を最大化する場合を考える。これは、 $\log(1.0 - (\mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P}))))$ ,  $\log(1.0 - (\mathcal{D}^{\text{space}}(\mathcal{G}(\mathbf{P}))))$  を最小化することであり、真値ラベルを観測値である確率を 1 とし  
てバイナリクロスエントロピー (BCE) を用いて損失関数を計算する。

$$\begin{aligned} \text{BCE}^{\text{wave}} &= - \sum_{k=1}^K \{1 \log\{\mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P}))\} + (1-1) \log\{1 - \mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P}))\}\}, \\ \text{BCE}^{\text{space}} &= - \sum_{k=1}^K \{1 \log\{\mathcal{D}^{\text{space}}(\mathcal{G}(\mathbf{P}))\} + (1-1) \log\{1 - \mathcal{D}^{\text{space}}(\mathcal{G}(\mathbf{P}))\}\}. \end{aligned} \quad (4.12)$$

一方、ある識別器の損失関数には、観測値の波形と推定された波形が入力された際に定義が異なる。まず、観測値の波形  $\mathbf{F}_{\text{real}}$  が入力された際の  $\log(\mathcal{D}^{\text{wave}}(\mathbf{F}_{\text{real}}))$ ,  $\log(\mathcal{D}^{\text{space}}(\mathbf{F}_{\text{real}}))$  の最大化について考える。この際には、 $\mathcal{D}^{\text{wave}}(\mathbf{F}_{\text{real}})$ ,  $\mathcal{D}^{\text{space}}(\mathbf{F}_{\text{real}})$  の値と真値ラベルの観測値である確率を 1 としてバイナリクロスエントロピー (BCE) に入力して最小化を行う。

$$\begin{aligned} \text{BCE}^{\text{wave}} &= - \sum_{k=1}^K \{1 \log\{\mathcal{D}^{\text{wave}}(\mathbf{F}_{\text{real}})\} + (1-1) \log\{1 - \mathcal{D}^{\text{wave}}(\mathbf{F}_{\text{real}})\}\}, \\ \text{BCE}^{\text{space}} &= - \sum_{k=1}^K \{1 \log\{\mathcal{D}^{\text{space}}(\mathbf{F}_{\text{real}})\} + (1-1) \log\{1 - \mathcal{D}^{\text{space}}(\mathbf{F}_{\text{real}})\}\}. \end{aligned} \quad (4.13)$$

次にある識別器に大局生成器で推定した波形が入力された際の  $\log(1.0 - (\mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P}))))$  と  $\log(1.0 - (\mathcal{D}^{\text{space}}(\mathcal{G}(\mathbf{P}))))$  の最大化について考える。これは、真値ラベルの観測値である確率を 0 としてバイナリクロスエントロピーに入力して最小化を行う。

このようにして、損失関数の定義をコードに落とし込む際には、min-max 最適化では無く、推定された波形と観測値の波形を分類する損失関数の最小化問題にする。

$$\begin{aligned} \text{BCE}^{\text{wave}} &= - \sum_{k=1}^K \{0 \log\{\mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P}))\} + (1-0) \log\{1 - \mathcal{D}^{\text{wave}}(\mathcal{G}(\mathbf{P}))\}\}, \\ \text{BCE}^{\text{space}} &= - \sum_{k=1}^K \{0 \log\{\mathcal{D}^{\text{space}}(\mathcal{G}(\mathbf{P}))\} + (1-0) \log\{1 - \mathcal{D}^{\text{space}}(\mathcal{G}(\mathbf{P}))\}\}. \end{aligned} \quad (4.14)$$

以上の内容を擬似的なアルゴリズム (Algorithm 3) で示す。最適化は、事前に設定した最大繰り返し回数 (`max_iteration`) まで実行する。まず、入力する点群データとそれに応じた真値の波形を訓練データから得る。次に、2つの識別器の学習を行う。この際、大局生成器による予測を行い波形を推定し、損失関数の計算を行う。この損失関数の値を元に2つの識別器の重み ( $\theta_{\mathcal{D}^{\text{space}}}$ ,  $\theta_{\mathcal{D}^{\text{wave}}}$ ) に対する勾配 ( $g_w$ ) を計算して、その勾配を元にそれぞれの識別器の重みを更新する。次に、大局生成器の学習を行う。この際、大局生成器によって波形を推定し、損失関数の計算を行う。そして、損失関数の値を元に大局生成器の重み ( $\theta_g$ ) に対する勾配 ( $g_w$ ) を求めて、その勾配を元に大局生成器の重みを更新する。これらの重みの更新を最大繰り返し回数まで交互に行う。

---

**Algorithm 3** 大局生成器と2つの識別器の最適化アルゴリズム

---

**Require:** training data and models  $\mathcal{G}$ ,  $\mathcal{D}^{\text{space}}$ ,  $\mathcal{D}^{\text{wave}}$ , max\_iteration,  $\alpha$ : learning rate

**Require:**  $\theta_{\mathcal{D}^{\text{space}}}$ ,  $\theta_{\mathcal{D}^{\text{wave}}}$ ,  $\theta_{\mathcal{G}}$

**Ensure:** trained model

```
1: for itr=1,...,max_iteration do
2:   Sample  $\{x\}$  a batch of point clouds.
3:   Sample  $\{\mathbf{F}_{\text{real}}\}$  a batch from the real waveforms corresponding to input point cloud.
4:    $\mathbf{F}_{\text{pred}} = \mathcal{G}(x)$ 
5:    $\mathbf{Prob}_{\text{real}} = \mathcal{D}^{\text{space}}(\mathbf{F}_{\text{pred}})$ 
6:    $g_w \leftarrow \nabla_w [\text{BCE}(0, \mathbf{Prob}_{\text{real}})]$ 
7:    $\mathbf{Prob}_{\text{real}} = \mathcal{D}^{\text{wave}}(\mathbf{F}_{\text{real}})$ 
8:    $g_w \leftarrow \nabla_w [\text{BCE}(0, \mathbf{Prob}_{\text{real}})]$ 
9:    $\theta_{\mathcal{D}^{\text{space}}} \leftarrow \theta_{\mathcal{D}^{\text{space}}} + \alpha \cdot \text{Optimizer}(\theta_{\mathcal{D}^{\text{space}}}, g_w)$ 
10:   $\theta_{\mathcal{D}^{\text{wave}}} \leftarrow \theta_{\mathcal{D}^{\text{wave}}} + \alpha \cdot \text{Optimizer}(\theta_{\mathcal{D}^{\text{wave}}}, g_w)$ 
11:   $\mathbf{F}_{\text{pred}} = \mathcal{G}(x)$ 
12:   $\mathbf{Prob}_{\text{real}} = \mathcal{D}^{\text{space}}(\mathbf{F}_{\text{pred}})$ 
13:   $g_w \leftarrow \nabla_w [\text{BCE}(1, \mathbf{Prob}_{\text{real}})]$ 
14:   $g_w \leftarrow \nabla_w [\text{L1}(\{\mathbf{F}_{\text{real}}\}, \mathcal{G}(x))]$ 
15:   $\theta_{\mathcal{G}} \leftarrow \theta_{\mathcal{G}} + \alpha \cdot \text{Optimizer}(\theta_{\mathcal{G}}, g_w)$ 
16: end for
```

---

## 4.5 ベンチマークデータによる波形推定実験

本節では、点群データから波形を推定する実験結果について述べる。まず、本実験で使用するデータセットである第3章で説明したベンチマークデータセットと、本章で新たに使用する汎化性能データセットについて述べる。次に、ベンチマークデータセットに対する波形推定の結果を述べる。この際に、提案手法である波形推定モデルと簡単な物理モデルを利用した波形推定結果を比較する。最後に、考察として、提案手法で取り入れた各工夫点の効果検証と、汎化性能を調査するために訓練済み波形推定モデルを別のデータセットに適用して汎化性能を検証する。

### 4.5.1 データセット

データは第3章と同様にベンチマークデータセットを使用する。この理由は、波形推定の最終的な目的が点群データから波形を推定したことによる地物分類性能の向上を目指すためであり、地物分類を行うデータセットと同様のデータセットに揃える必要がある。このベンチマークデータセットは波形のファーストパルス点を点群データ化しているため、各点にはひとつの波形が対応する。

波形推定における訓練データの作成方法について説明する。第4章においても第3章と同様に元の大きな領域を一度にGPUに乗らないため、小さなパッチに分割する。第3章の点群・波形データの分類モデルと比較すると、波形推定モデルでは大局生成器と2つの識別器を同時に学習するため、よりGPUのメモリ消費量が激しい。そこで、第3章では50,000点であったパッチの分割を本章では10,000点に決定した。なお、最適化過程では、5回のクロスバリデーションを行った。

### 4.5.2 実験の設定

まず、実装やハイパーパラメータについて説明する。大局生成器および2つの識別器をPyTorch [169] で実装した。全てのクロスバリデーションにおいて、Adam 最適化を使用し、初期学習率を0.001とした。最適化の繰り返し回数は、訓練データセットを100周する（100エポック）ように設定し、学習率は10エポックごとに50%ずつ低下するような設定にした。大局生成器および2つの識別器を、TSUBAME3.0の4つのNVIDIA Tesla P100 GPUを用いて最適化を実行した。バッチサイズは、大局生成器と2つの識別器を同時に学習する際のメモリ使用量を考慮して、各GPUに対して1に設定し、バッチ正規化は各層に適用した。2つの識別器における最後のFC層には0.5のドロップアウト率を使用した。また、損失関数の各項を混ぜるハイパーパラメータの設定として、 $\mathcal{L}_G$  で使用する  $\lambda$  は、10とした。学習を行う際には、事前学習モデルが存在しないため、ランダムに初期化された重みを元に大局生成器と2つの識別器の計算を途切れさせることなく最適化をした。なお、学習の最大エポック数は100としたが、損失関数が減少しない場合には早期停止を行うように設定した。

次に、訓練済みモデルによる予測方法について Algorithm 4を用いて説明する。なお、学習済みモデルは大局生成器のみを用い、最適化の際にのみ使用する識別器は予測の際には用いない。テストデータ全体の点群データに対して波形を推定する。波形推定モデルは入力点数が10,000点であるため、テストデータを固定点の10,000点に抜き出す。抜き出した点群データを波形推定モデルに入力し、擬似的な波形を推定する。このように、10,000点ごと移動しながら波形を推定する。なお、重複を許しながら、点群データを抜き出すため、重複部分では後から予測した結果を用いる。

---

**Algorithm 4** 学習済み大局生成器による波形推定のアルゴリズム

---

**Input:** test data and trained model

**Output:** predicted data

```
model = loadModel(trained model)
data = loadData(test data)
patches, indexs = loadData(data)
for  $i = 1$  to  $len(patches)$  do
    pred prob = model(patches[i])
    predWaveform = getMaxClass(pred prob)
    data[indexs[i]] = predWaveform
end for
return data
```

---

最後に、波形推定結果の定量評価方法について述べる。定量評価指標は、以下に示す学習時に用いた回帰の損失関数の定義式を用いる。

$$\mathbf{MAE}^{\text{content}} = \frac{1}{K} \frac{1}{M} \sum_{i=1}^K \sum_{j=1}^M |\mathbf{F}_{\text{real}_{i,j}} - \mathbf{F}_{\text{pred}_{i,j}}|$$
$$\mathbf{MAE}^{\text{spectral}} = \frac{1}{K} \frac{1}{M'} \sum_{i=1}^K \sum_{j=1}^{M'} |\mathbf{FFT}_{\text{real}_{i,j}} - \mathbf{FFT}_{\text{pred}_{i,j}}|.$$

これらはそれぞれ、推定した波形と、観測値の波形の強度間の平均絶対誤差と、観測値の波形のFFT後の値の平均絶対誤差を示す。なお、以降では、 $\mathbf{MAE}^{\text{content}}$  は強度の差と呼び、 $\mathbf{MAE}^{\text{spectral}}$  周波数特性の差と呼ぶ。

### 4.5.3 実験結果

波形推定モデルの学習にはTSUBAME3.0でNVIDIA P100 GPUを使用し、最大エポック数である100エポックに到達するまでに9日間要した。なお、この学習にかかる時間の要

因は大局生成器と2つの識別器を同時に学習させることに由来する。損失関数の値は100エポックまで安定して減少し続けた。なお、予測時において、テストデータを処理する時間はTSUBAME3.0にて2時間を要した。ここでは、学習済みの波形推定モデルによる波形推定の結果について、定量的・定性的な評価を行った。なお、波形推定の評価はベンチマークデータセット（図3.13の赤枠部分）のテストデータを用いて行なった。

まず、定量評価結果を述べる。ベンチマークデータを学習した波形推定モデルの波形推定結果について定量的に評価を行なった。推定した波形の評価には、強度の平均絶対誤差とFFTした周波数領域における平均絶対誤差を用いた。筆者の知る限り、点群データから波形を推定する研究はないため、先行研究との比較は行えない。そこで、提案手法である波形推定モデルの結果のみを表4.1に示している。なお、ベースライン波形推定手法を用いた結果と、物理モデルで推定した結果との比較は考察部分にて示す。

波形推定モデルはベンチマークデータセットのテストデータに対して、 $MAE^{content}$  は1.71（表4.1の”平均”の部分）という低いエラー値を達成した。また、 $MAE^{spectral}$  も同様に0.11と低いエラー値を達成した。テストデータの全体的な評価だけでなく、各クラスの平均絶対誤差も算出した。図4.11に示すような1つの強いピークを持つ道路、建物の変換波形は、 $MAE^{content}$ 、 $MAE^{spectral}$  の平均絶対誤差共に低い結果になった。その理由は、形状が単純であるため大局生成器による波形推定が容易だからである。しかし、同じく1つのピークを持つ地面（図4.11の地面に示す）は、 $MAE^{content}$  は平均絶対誤差が最も高いことが示された。これは、評価の際にクラスごとに最大値・最小値の値を揃えるような正規化を行っていないため、大きな強度がある場合に、その大きさそのものにMAEは影響を受けるためである。さらに、図4.11に示すような複雑な形状の植生も、平均絶対誤差が大きいという問題を示した。波形推定モデルでは、波形の形状と強度に対する損失関数を用いているが、波形のピーク位置に関して明示的に設定する必要があると考えられる。

表 4.1: 推定された波形と観測値の波形の平均絶対誤差

評価指標	地面	植生	建物	電線	鉄塔	地面	平均
強度の差	2.14	1.92	0.46	1.02	0.46	0.06	1.71
周波数特性の差	0.125	0.223	0.015	0.289	0.018	0.009	0.11

次に定性評価を行った。図4.11,4.12は、テストデータにおける観測値の波形と、波形推定モデルで推定した波形の例を可視化した結果である。推定した波形と観測値の波形は、それぞれのクラスからランダムに10個サンプリングされた波形の平均値を示す。図4.11に見られるように、波形推定モデルによって、両者の傾向が全体的に一致していることが確認できた。さらに詳しく見ると、全体的にピークの位置が一致していることが確認できた。

クラスごとの推定結果を確認すると、地面、建物、道路など、ピークが1つのクラスは観測値の波形に近い結果が得られた。しかし、建物の偽物の波形のピークにおける強度は観測値の波形に比べて十分ではなく、この強度の違いが定量的な評価における平均絶対誤差の大きさにつながったと考えられる。また、植生や送電線、鉄塔などの複雑な形状を有する点群データにおいてはピーク数の少ない地物と比較すると、波形の推定を十分にすることができなかった。この理由は、定量評価で述べたように、複雑な形状を評価するための損失関数が提案手法では十分に定義されていないために生じたと考えられる。そのため、複雑な地物に対する波形の推定において、より観測値に近づけるためには、波形のピークの位置情報を計算するための微分可能な演算を定義し、回帰の際に位置の値を求める損失関数を追加することで改善される可能性がある。

さらに、提案手法の有効性を示すため、以降では、提案手法の中で取り入れた工夫点についての効果検証と、物理モデルによる波形推定と提案手法の比較、汎化性能の検証を行う。

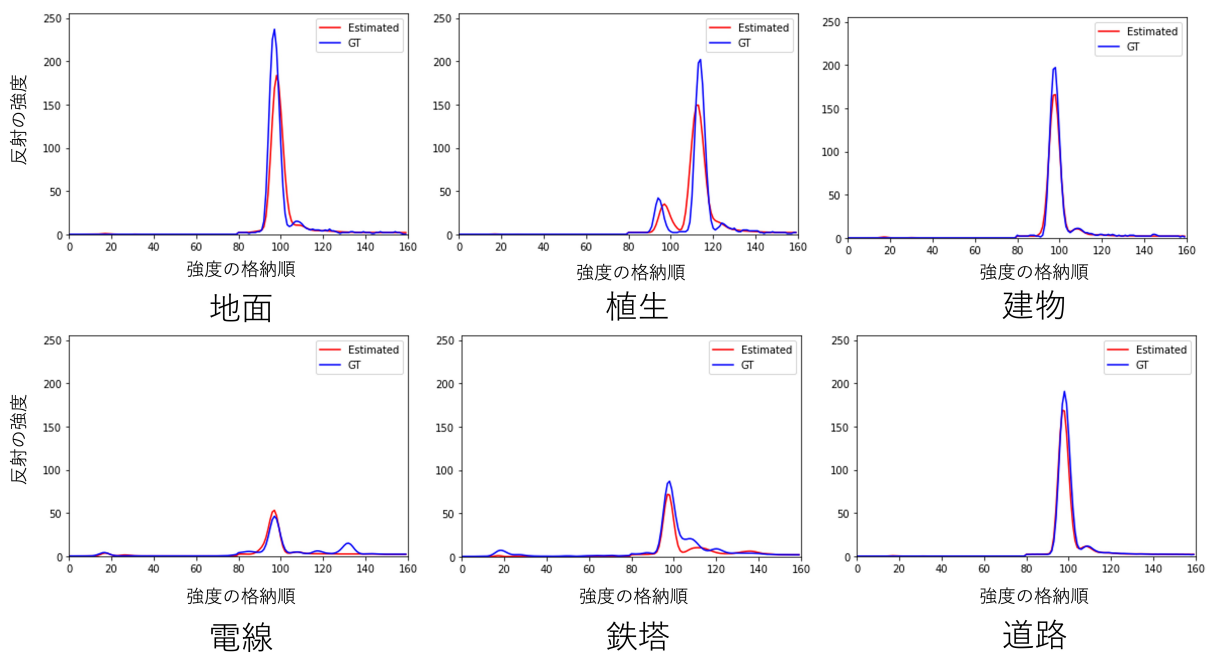


図 4.11: 波形推定モデルで推定した波形と観測値の例. 赤色が推定結果, 青色が観測値の波形を示している.

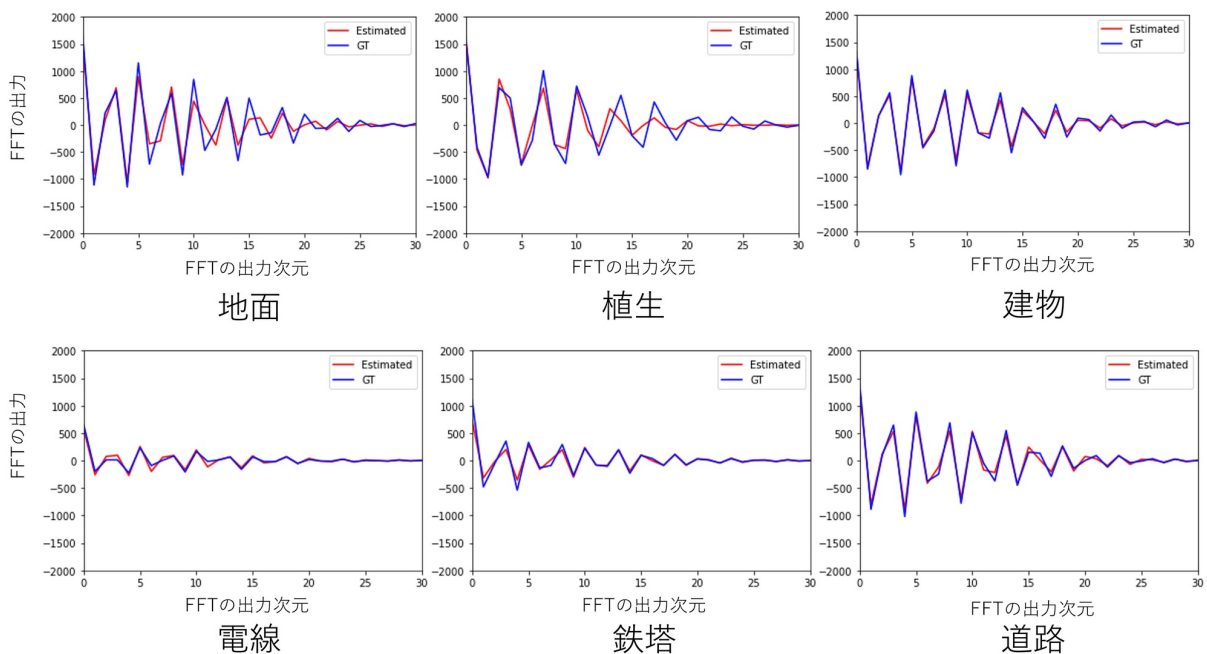


図 4.12: FFT を適用した波形推定モデルで推定した波形と観測値の例. 赤色が推定結果, 青色が観測値の波形を示している.

#### 4.5.4 大局生成器の影響

波形推定を実行する大局生成器には、大局生成器を用いた。これは、第3章の局所・大局モデルのように入力に点群・波形データではなく、点群データを入力するために点群データの深層学習手法を参考にした大局生成器を用いた。大局生成器は第3章の大局ブロックと同様にダウンサンプリング構造を持つが、点群データのみを入力する。この妥当性を実験的に評価した。比較として、第3章の局所ブロック相当の単なるMLPによる波形推定を局所生成器、第3章の局所・大局モデル相当の層設計による波形推定を局所・大局生成器と呼ぶ。全体的な結果の比較を表4.2に示し、以降で詳しく説明する。

- 局所生成器**  
 まず、局所生成器による波形推定を行なった結果を示す。これは、入力された個々の  $x, y, z$  座標の値からその点の波形を推定する層設計である。つまり、第3章の局所ブロックの層設計において、入力を点群データに変更し、さらに出力を分類ではなく波形の値に変更したものである。そのため、周辺の情報を取り込みながら波形の推定ができない。それに応じて、教師データを使用した最適化を実行する際において生成器の損失関数が発散し、波形推定ができなかった。
- 大局生成器**  
 次に、提案手法の波形推定モデルにおいて採用した大局生成器の結果を示す。テストデータの点群データを入力し、提案手法で使用した大局生成器において推定した波形と観測値の波形の平均絶対誤差は1.71である。局所生成器では波形を推定ができなかったが、大局生成器では、ある一点に対応する波形を推定する際に使用する情報を広範囲から収集できるため、波形の推定が可能になった。
- 局所・大局生成器**  
 最後に、局所・大局生成器による波形推定を行なった結果を示す。局所・大局生成器は、第3章の局所・大局モデルに点群データを入力として、出力を波形の値に変更した。地物分類の際には局所クラスと大局ブロックの出力を組み合わせた。波形推定の際には地物の情報が未知であるため組み合わせることはできない。しかし、局所・大局生成器の能力を調査する目的において地物情報が既知である条件下で実験を行った。その結果、大局生成器の波形推定結果を比べると推定能力は悪化した。これは、局所クラスに属すると仮定した地物の波形推定結果を局所生成器から得るため、推定能力の低い局所生成器の悪影響を受けたことに起因する。

表 4.2: 生成器の層設計が波形推定に与える影響

入力	局所生成器	大局生成器	平均絶対誤差 (MAE <sup>content</sup> )
点群データ	✓	-	-
点群データ	-	✓	1.71
点群データ	✓	✓	3.31

#### 4.5.5 各工夫点の効果

まず、工夫点についての効果検証に関する考察を行う。波形推定モデルの各工夫点の効果を調べるために、大局生成器と大局識別器 ( $\mathcal{D}^{\text{space}}$ ) の層設計と強度回帰で成り立つベースラインモデル (set1 と呼ぶ) に対して本研究で行なった各工夫点をそれぞれ追加し、その波形推定能力を観測値と推定した波形の平均絶対誤差の結果を比較した。具体的な工夫点は、set1 は大局生成器で推定した波形と観測値の波形の強度を回帰する強度回帰と大局識別器の損失関数で構成されており、この構成に対して提案手法では「波形の周波数特性に対する損失関数 (周波数回帰)」、「波形を個別に評価する識別器  $\mathcal{D}^{\text{wave}}$ 」という2つの工夫点を提案した。全体的な結果の比較を表 4.3 と図 4.13 に示す。図 4.13 に示した大局生成器で推定した波形はすべて、点群データを各学習済みモデルに入力した結果である。

表 4.3: 提案した各工夫点の効果検証結果

set	強度回帰	$\mathcal{D}^{\text{space}}$	周波数回帰	$\mathcal{D}^{\text{wave}}$	平均絶対誤差 ( $\text{MAE}^{\text{content}}$ )
1	✓	✓	-	-	2.82
2	✓	✓	✓	-	1.79
3	✓	✓	-	✓	2.21
Full	✓	✓	✓	✓	1.71

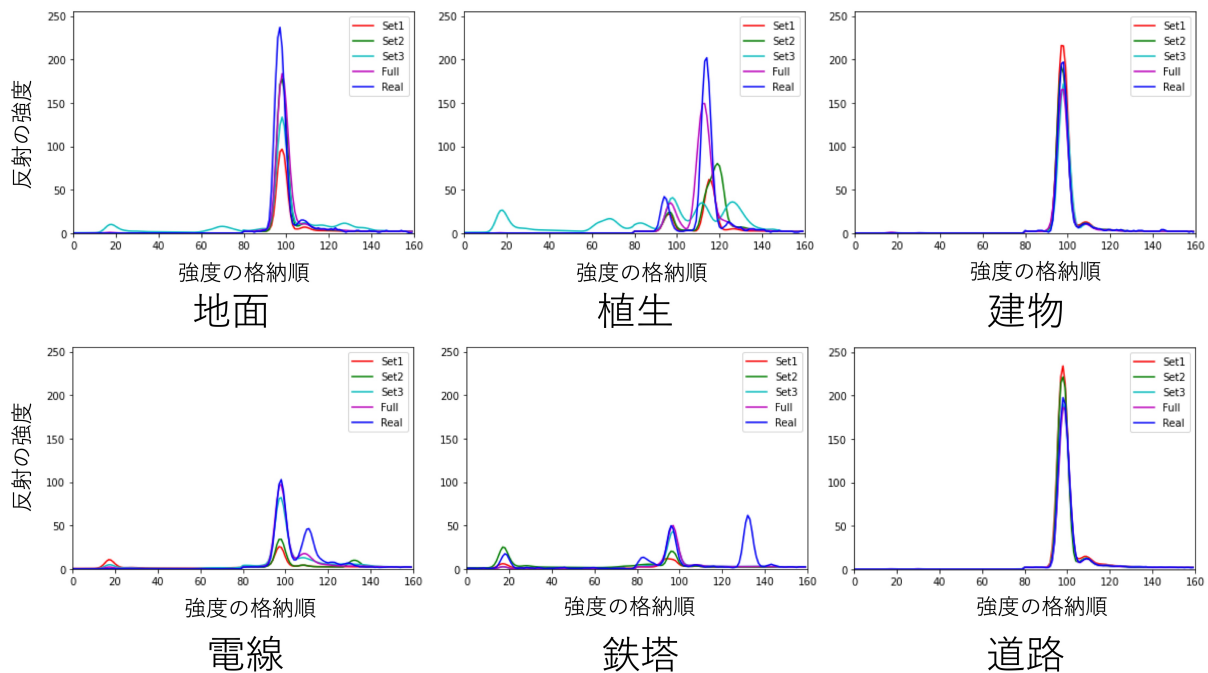


図 4.13: 各工夫点の波形推定への効果。赤色が推定結果，青色が観測値の波形を示している。set1 は工夫点を全て除去したモデル，set2 は set1 に周波数回帰を追加したモデル，set3 は set1 に局所識別器を追加したモデル，Full は提案手法の波形推定モデルを示す。

詳細な、各工夫点の効果を以下に示す。

- 周波数回帰  
まず、波形の周波数特性に対する損失関数 (周波数回帰)  $\mathcal{L}_G^{\text{spectral}}$  の効果を調べた。この比較では、ベースライン (set1) に対して、 $\mathcal{L}_G^{\text{spectral}}$  を追加して学習を行なった。この

設定を set2 とする。  $\mathcal{L}_G^{\text{spectral}}$  を用いると、ベースライン (set1) より正確な波形を推定可能であることを示した。図 4.13の緑色の線は、推定した波形と、その波形に対応する観測値の波形である。  $\mathcal{L}_G^{\text{spectral}}$  を使用することで、set1 よりも推定される波形のピーク位置をより正確に再現できた。これはゼロを多く予測する傾向にある強度回帰の欠点を周波数回帰によって補ったことに起因する。

さらに、定量評価の結果、  $\mathcal{L}_G^{\text{spectral}}$  を導入した set2 は、ベースラインモデル (set1) よりもテストデータに対する平均絶対誤差が減少した。なお、音声処理の分野では、生の音声信号だけでなく FFT を適用した後に深層学習モデルで解析する手法は広く用いられており、FFT を併用することで高性能な解析が可能であることが示されている。このような他分野における傾向と、本比較実験の結果は一致している。

- $D^{\text{wave}}$

次に、局所識別器 ( $D^{\text{wave}}$ ) の効果を確認する。set3 では、ベースラインモデル (Set1) に  $D^{\text{wave}}$  を加えたモデルである。全体的な評価として、ベースラインモデル (set1) より正確な波形の推定結果を得た。set3 では、図 4.13中のシアン色の線で、set3 の大局生成器で推定した波形と、その推定された波形に対応する観測値の波形を示している。

詳しく結果を確認すると、電線や鉄塔では、ベースラインモデルよりも、  $D^{\text{wave}}$  を使用した方が、観測値の波形に近い波形を正確に推定できることを示した。しかし、植生の波形では観測値との乖離が激しくなった。これは、一般的に GAN の学習は識別器と大局生成器の最適化を同時に行うため不安定であり、なおかつ 2 つの識別器を使用したことで、最適化が不安定になったためと考えられる。次に、定量的な評価を行ったところ、  $D^{\text{wave}}$  はベースラインよりも平均絶対誤差が小さい。これは、波形としての整合性をベースラインモデルよりも評価することができたからである。

- 全ての機能を入れた提案手法

全体的な傾向としては、周波数回帰と  $D^{\text{wave}}$  という 2 つの工夫点を入れたモデルが最も低い平均絶対誤差をもたらした。そして、その波形の推定結果は図 4.13のマゼンタ色の線で示す。上記のその設定よりも観測値に近い波形の推定が達成できたことを示した。総じて、これらの 2 つの工夫点を組み合わせることで、2 つの識別器を同時に最適化する学習の不安定さ損失関数を追加することで抑えたと考えられる。

#### 4.5.6 物理モデルとの比較

機械学習や深層学習手法で波形を推定する手法の先行研究は存在しないため、物理モデルから擬似的に推定した波形と比較する。物理モデルの計算には、ベンチマークデータのテストエリアの点群データをグリッド状に変換し、さらに、地面として考えられるグリッド内の最も低い点を元に内挿をした 3 次元モデルを用いる。物理モデルは、  $x, y, z$  座標しか公開されていない点群データのアーカイブデータと同様な問題にするため、スキャン角度等の情報は使用しない。物理モデルは第 2 章で紹介したライダー方程式は、事前に地物情報つまり地物の反射特性に関する物性値が既知であることが前提であるため、ここではより簡易的な物理モデルを用いる。本検証で使用する波形物理モデルの方程式を以下のように定義した。

$$A_{t_n}^{u_i, v_j} = A_0 \exp \left[ -4 \ln 2 \left( \frac{(t_n - \delta t_n)^2}{w^2} + \frac{(u_i^2 + v_j^2)}{\gamma^2 R^2} \right) \right]. \quad (4.15)$$

ここで、  $\delta t_n$  は  $uv$  平面から物体との交点までの距離  $g_{i,j}$  に対応する時間差で表される。

$$\delta t_n = \frac{2g_{i,j}}{V_g}. \quad (4.16)$$

$g_{i,j}$  は  $uv$  平面から  $u_i, v_j$  サブビームと物体の交点までの距離、  $V_g$  はレーザー光の速度 (シミュレーションでは  $3 + 10^9 \text{m/s}$  の値を使用)。なお  $n$  番目に取得された強度値は、取得時間  $t_n$

内のサブビームの個々の振幅のすべての取得時間  $t_n$  内のサブビームの個々の強度の合計に等しい。

$$A_n = k \sum_{u_i, v_j < -s, s >} A_{t_n}^{u_i, v_j}. \quad (4.17)$$

$k$  はスケール係数であり、表面の反射率に加えて、システムおよび大気のコэффициентを含んでいる。ここでは、事前にクラス情報は未知であるため、一様な  $k$  を用いる。事前の地物情報が点群データだけからでは得られないため、一様な素材であることや、建物屋根や木の詳細な 3 次元形状を仮定できない。なお、上記の物理モデルを点群データに適用する際において意味することは、「空間グリッド内にどれくらいの点群データがあるか」の指標から波形を推定することである。

ここからは、物理モデルで推定した波形と、提案手法の波形推定モデルで推定した波形を比較する。また、定量評価結果を表 4.4 に示す。比較すると、物理モデルは深層学習手法を用いた手法と比較すると高い誤差を示した。これは、物理モデルを適用するための事前情報としての 3 次元形状が点群データのみからでは得られないことが原因である。波形推定において、物理モデルによるシミュレーションを用いることで大きなフットプリントと単純な地物の形状においては波形をシミュレート可能である。しかし、本研究で使用したフットプリントの小さい (約 30 cm) 波形のシミュレーションは複雑な反射過程を経るため、波形のシミュレーションの計算が困難である。例えば、草やアスファルトの表面、または傾斜したタイル張りの屋根を区別するには、ランバートの仮定に基づく高度な反射モデルが必要である。さらに、反射モデルを適用するためには、事前にその点は何の地物に属するかの情報が必要となるため、シミュレートする際に地物分類された事前情報が必要となるという問題が生じる。また、物理モデルのシミュレーションに耐えられる詳細な 3 次元形状を推定するには、事前に建物や地面植生などの地物の情報が必要になるため、「点群データの地物分類を補助するために波形を推定する」という目的に対しては不適合であった。

表 4.4: 物理モデルと提案手法の定量比較結果

手法	平均絶対誤差
物理モデル	3.31
点群着色手法を参考にした手法	2.82
提案手法 (波形推定モデル)	1.71

#### 4.5.7 点密度の影響

点群データから波形の推定を実行する波形推定モデルの学習において使用した点群データと波形が1:1に対応する学習データは平均で1平方メートルあたり15点の点密度で取得されたものである。そこで、学習済み大局生成器の別のデータへの適用性を確認するため、点密度を変化させた場合の分類性能へ影響を考察した。影響の調査方法は、学習済み大局生成器を、テストデータの点密度を10点/m<sup>2</sup>、5点/m<sup>2</sup>、1点/m<sup>2</sup>まで等間隔に減らしたテストデータへそれぞれ適用した波形推定結果を比較した。その際、テストデータからランダムに点密度を減らした。表4.5にその比較結果を示した。元の15点/m<sup>2</sup>と比較して、1点/m<sup>2</sup>の場合において全ての地物において波形推定性能が劣化した。これは、第3章の地物分類と同様に、グループ化を行う際に入る波形の数が十分に確保できないことに起因する。個別の地物に着目すると、5点/m<sup>2</sup>、10点/m<sup>2</sup>の点密度において複雑な波形形状を有する植生・電線・鉄塔の波形推定性能が劣化した。一方、簡単な波形形状を有する地面・建物・道路の波形推定性能は5点/m<sup>2</sup>、10点/m<sup>2</sup>において大きく劣化しなかった。これは、大局生成器における点と点の近傍特徴抽出によって生じる問題であり、複雑な波形を推定する際には点の細かな形状が大切であるにも関わらず、点密度が少ない場合においてその形状情報を獲得できないためである。

表 4.5: 点群データの点密度が学習済み大局生成器へ与える影響。学習済みモデルによる波形推定を行う際に、入力するテストデータの点群データの点密度を減らした場合の推定結果の波形と観測値の波形強度の平均絶対誤差を各地物で計算した。

平均点密度 (点/m <sup>2</sup> )	地面	植生	建物	電線	鉄塔	道路
1	12.5	22.2	10.2	12.4	10.2	12.6
5	2.5	2.4	1.9	2.3	2.5	0.9
10	2.2	2.2	0.6	1.1	0.7	0.4
15	2.1	1.9	0.5	1.0	0.5	0.1

## 4.6 第4章のまとめ

本章では、深層学習手法による点群データからの波形推定手法である波形推定モデルを提案した。提案する波形推定モデルは、点群データに対してその点単位に色情報を推定する点群着色手法を参考にした。点群着色手法では色情報を推定する大局生成器と推定された色の空間的な整合性を評価する大局識別器によって成り立つ。しかし、本研究にて対象にする波形推定では、波形の大局的な情報だけでなく、個々の波形の局所的な情報も有効であると地物分類と同様に仮定できる。具体的には、点群データ着色では、RGB という3次元の情報を評価するため、点群データ全体に対する大局識別器のみでも有効であったが、波形を推定する際には、強度の値が160個あるため高次元な評価を行う必要がある。そこで、波形推定モデルによる波形推定手法においては、点群データ着色手法に対して、波形に対して特化した局所識別器の提案を行った。さらに、推定した波形は「波形」であるため、音声波形等の処理に使用される周波数特性を評価する損失関数の使用を提案した。

波形推定モデルは、点群データから波形推定を担う大局生成器、推定された波形の空間的な整合性を評価する大局識別器、そして、推定された波形単位に整合性を評価する局所識別器の3つの深層学習モデルで成り立つ。なお、層設計に関しては、大局生成器においては既存の点群データの深層学習手法、大局識別器は第3章の大局ブロック、局所識別器は第3章の局所ブロックと同様の層設計を用いた。これらの深層学習モデルを学習する際には、推定された波形と観測値の波形間の強度の差、そして、推定された波形と観測値の波形の周波数特性の差を最小化する。さらに、推定された波形と観測値の波形間の確率分布を一致させるために、空間識別器と局所識別器それぞれに対して、KLダイバージェンスの最小化をする。

提案手法によって、ベンチマークデータにおいて波形生成の実験を行い、定量評価と定性評価を行った。定量評価は観測値の波形と波形推定モデルで推定した波形の絶対平均誤差 (Mean Absolute Error, MAE) を計算した。波形推定モデルによって推定した波形は、MAE 指標を用いて定量的に評価すると、各工夫点を抜いたベースライン手法よりも低い MAE を示した。また、定性評価においても、提案手法を用いると推定した波形と観測値の波形は近い形状を示した。

## 第 5 章

### 波形推定手法と地物分類手法の効果検証

## 5.1 はじめに

点群・波形データに対する地物分類手法を実際に企業が所有するデータやオープンデータとして公開されたデータに適用する際に、それらのデータは点群データしか有していないという問題がある。そこで、点群データから波形データを推定することで擬似的な点群・波形データとして地物分類モデルに入力することで問題を解消する。

次に本章の構成を説明する。本章では、図 5.1 に示すように、第 4 章にて提案した波形推定手法と、第 3 章にて提案した地物分類手法を適用する。また、有効性の検証方法としては、前章までに提案した 2 つの深層学習モデルを適用して評価を実施する。本章の流れは、まず提案手法を説明する。提案手法では、点群データのみが手に入るという状況を想定し、第 4 章で提案した波形推定モデルの大局生成器と第 3 章で提案した地物分類の学習済みモデル (局所・大局モデル) を用いる。次に、これらの学習済みモデルによる推論のアルゴリズムについて説明する。そして、実際に提案手法を適用する前に、学習済み波形推定モデルに関して汎化性能を評価する。次に、ベンチマークデータにおける検証実験結果について述べる。ベンチマークデータにおける検証実験において、地物分類結果の性能評価を点群データのみを用いた場合や観測値の点群・波形データを用いた場合と比較を行う。さらに、ベンチマークデータだけでなく、実際の点群データのアーカイブデータを想定した DALES データセットと日本のアーカイブデータに対して波形推定モデルと地物分類モデルの学習済みモデルを適用して汎化性能を評価する。

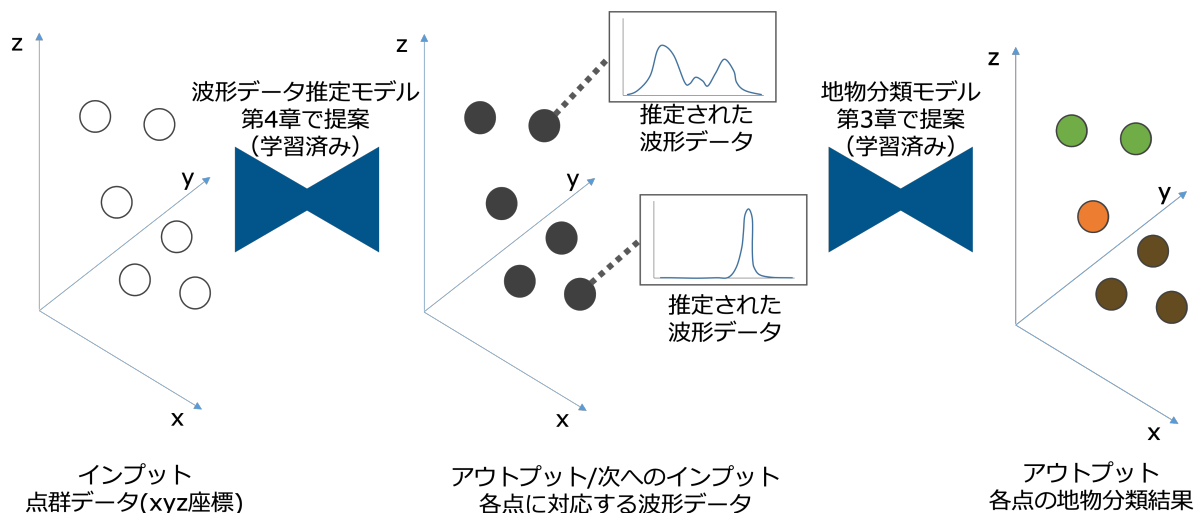


図 5.1: 本章で提案する波形推定手法と点群・波形データの地物分類手法を組み合わせる手法の概要。まず、第 4 章で提案した波形推定モデルの大局生成器によって点群データから波形データを推定し、擬似的な点群・波形データを作成する。次に、第 3 章で提案した局所・大局モデルによって擬似的な点群・波形データを地物分類する。

## 5.2 提案手法

ここでは、地物分類手法と波形推定手法を組み合わせる提案手法について説明する。まず、第4章で提案した点群データから波形を推定する手法を使用して波形データを推定する。次に、点群データと推定した波形データによる擬似的な点群・波形データを第3章で提案した地物分類モデルを使用して点群・波形データの地物分類を行う。以降では、まず本章で使用する学習済みモデルを説明し、次に推論に用いるアルゴリズムの説明を述べる。

### 5.2.1 使用する学習済みモデル

本節では、学習済みモデルの適用により、波形推定と推定した波形データによる擬似的な点群・波形データの地物分類を実施する。その際には以下の2つの学習済みモデルを使用する。

- 地物分類モデル (第3章の局所・大局モデル)  
点群データと波形データで成り立つ点群・波形データを入力して、各点を地物分類するモデル。地物分類は「地面」、「植生」、「建物」、「電線」、「鉄塔」、「道路」の6種類が対象である。なお、ベンチマークデータセットを学習したものをを用いる。
- 波形推定モデル (第4章の大局生成器)  
点群データを入力して、各点が点群および波形情報を記録するライダーで観測された際の波形を推定するモデル。出力は訓練に使用したデータと同じ次元数の波形データである。なお、ベンチマークデータセットを学習したものをを用い、大局生成器のみを波形推定に使用し、2つの識別器は使用しない。

以下、簡単のため地物分類モデルは第3章で提案した局所・大局モデルの学習済みモデルを意味し、波形推定モデルは第4章で提案した大局生成器の学習済みモデルを意味する。

### 5.2.2 アルゴリズム

本節では、第4章で提案した波形推定モデルにより点群データから波形を推定し、その後、第3章で提案した地物分類モデルに入力するため推定した波形データによる擬似的な点群・波形データを作成し地物分類を行う。なお、波形推定モデルの出力は波形と呼び、地物分類モデルに入力する際には波形データと呼ぶ。2つの処理の流れは、波形推定モデルへの入力次元数が10,000点、地物分類モデルへの入力が50,000点であるため、得た点群・波形データを直接地物分類モデルへ入力できない。そこで、2段階のアルゴリズム (Algorithm 5) を用いる。まず、波形推定を行う。入力には点群データを用いる。波形推定モデルは入力点数が10,000点であるため、解析対象の点群データを固定点の10,000点に抜き出す。抜き出した点群データを波形推定モデルに入力し、波形を推定し擬似的な点群・波形データを推定する。このように、10,000点ごとに移動しながら波形を推定する。なお、重複を許しながら、点群データを抜き出すため、重複部分では、2つの予測結果の平均を用いる。

次に擬似的な点群・波形データの地物分類を行う。地物分類モデルでは、波形データを含む50,000点ごとに地物分類する必要がある。そこで、こちらも同様に固定点で切り出しながら地物分類を実行する。なお、重複部では、後から予測した結果を採用する。

---

**Algorithm 5** Algorithm for waveform estimation and classification

---

**Input:** test data and trained models

**Output:** predicted class

```
model = loadModel(wave estimation model)
data = loadData(test data)
patches, indexes = loadData(data)
for  $i = 1$  to  $len(patches)$  do
    predProb = model(patches[i])
    predClassID = getMax(predProb)
    generatedWaveform[indexes[i]] = predClassID
end for
model = loadModel(classification trained model)
generatedFwData = loadData(generatedWaveform and pointCloud)
patches, indexes = loadData(generatedFwData)
for  $i = 1$  to  $len(patches)$  do
    predProb = model(patches[i])
    predClassID = getMax(predProb)
    classified[indexes[i]] = predClassID
end for
return classified
```

---

## 5.3 波形推定の汎化性能評価

実際に波形推定と地物分類手法を組み合わせる前に、ベンチマークデータセットで学習した波形推定モデルの汎化性能を検証する。そのために、学習済み波形推定モデルを別の点群・波形データを用いて点群データから波形を推定し、推定した波形と観測値の波形の差を計算する。

### 5.3.1 ダブリンデータセット

ベンチマークデータセットで学習した波形推定モデルの汎化性能を示すために、異なる地域で観測された別のデータセットに適用する。波形推定モデルの汎化性能を示すために使用したデータセットは、公開データの点群・波形データである Dublin city dataset[110]（以降ではダブリンデータセットと呼ぶ）である。このダブリンデータセットは、アイルランドのダブリン市中心部の 2 km<sup>2</sup> が対象になっている。なお、ダブリン市全域を対象にすると点群データ数が 6 億点以上と多いため、図 5.2 に示すような汎化性能の検証をするための推論対象としてひとつのタイルを選択した (ID : T\_315500\_234500)。また、汎化性能データセットの平均点密度が約 350 点/m<sup>2</sup> であるため、ベンチマークデータセットと同程度の点密度になるように 1/20 のダウンサンプリングを行った。また、ベンチマークデータセットと条件を揃えるため、ファーストリターンのみを用いた。

### 5.3.2 波形推定結果

ダブリンデータセットを用いて推定した結果と観測値の波形の平均絶対誤差を計算する。ダブリンデータセットには地物情報がないため、ランダムに 1,000 個抽出した推定した波形と観測値の波形の平均絶対誤差の平均を示す。その結果、推定した波形と観測値の波形の平均絶対誤差の平均は 2.52 であった。ベンチマークデータセットの場合、平均絶対誤差は 1.71 (表 4.1) であったことと比較すると、ベンチマークデータよりも高いエラーであった。そのため、汎化性能に問題あることが示された。

図 5.3 に示すように、観測値の波形と推定した波形の定性評価を行った。図 5.3 の左下図において、観測値の波形と推定した波形でピーク位置の大きなずれが生じた。これは、高い建物の点である。学習に使用したベンチマークデータセットは、ダブリン市内のような市街地のデータセットとは異なり郊外のデータセットであったため、地物の特徴の違いを吸収することができなかったと考えられる。また、ピーク位置のズレに関しては、ピーク位置を点群データから制限をかける仕組みが必要であるとも考えられる。

しかし、本研究で行うような地物分類への活用の場合においては、推定した波形の有用性に関しては、観測値との一致よりも分類性能への寄与で評価すべきである。図 5.3 の右下図は植生の点における波形推定結果を示しており、特徴的な複数ピークを持つ形状を示した。観測値との波形とは乖離しているが、地物分類への寄与は期待できる。波形の推定結果の詳細な地物分類への寄与に関しては、次節以降で行う。

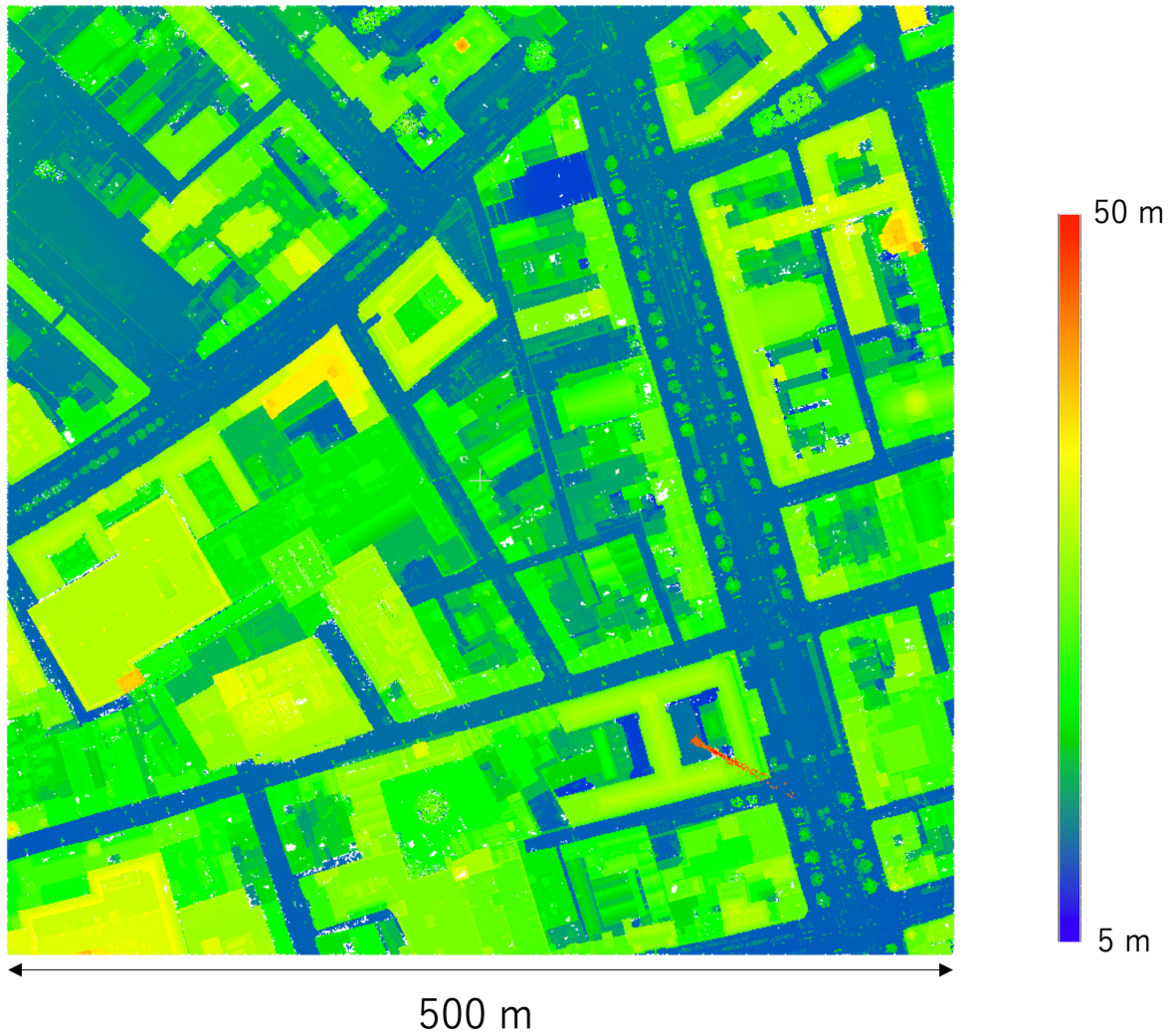


図 5.2: 波形推定モデルの汎化性能の検証実験のためのダブリンデータセット. Dublin city dataset[110] の ID : T\_315500\_234500 を汎化性能の評価に用いる.

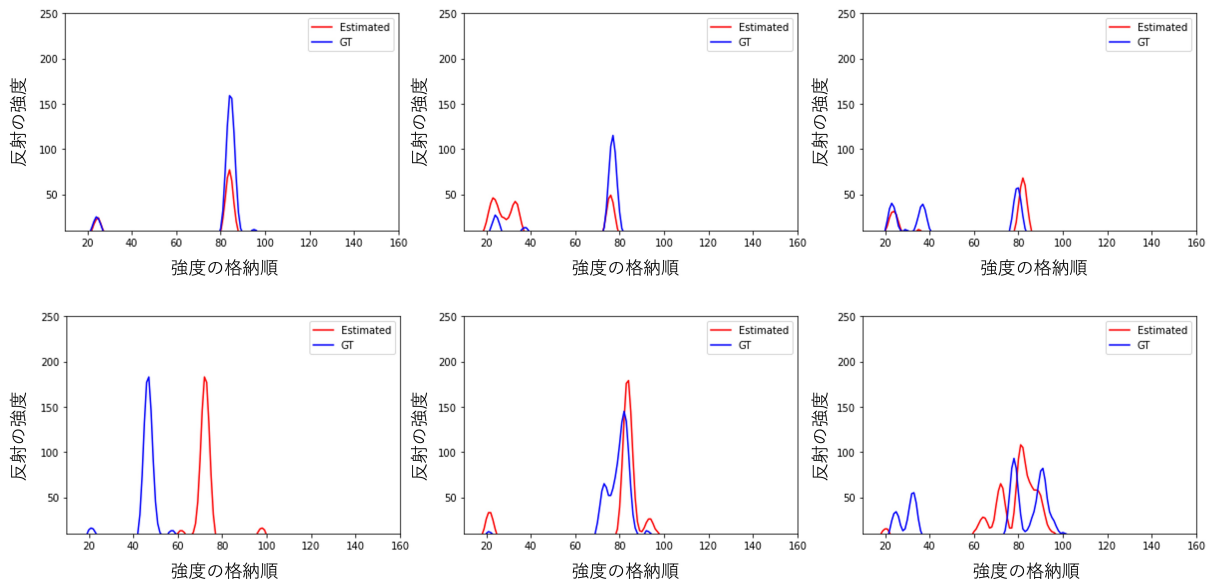


図 5.3: ダブリンデータセットに対する波形推定結果. 赤色が推定結果, 青色が観測値を示している.

## 5.4 ベンチマークデータセットへの適用

ここでは、波形推定モデルと点群・波形データの地物分類手法を組み合わせる効果を検証するために、訓練データと状況が似たベンチマークデータセットのテストデータに対して適用する。ベンチマークデータセットのテストデータを用いることで、純粋なモデルの能力を評価できる。効果検証では、観測値の点群・波形データの地物分類結果と、波形推定モデルから得た擬似的な点群・波形データを用いて地物分類を行なった結果の比較を行う。さらに、点群データのみを入力した深層学習モデル [178] と波形推定モデルから得た点群・波形データを用いて地物分類を行なった結果の比較を行い、擬似的に点群・波形データ化し、学習済み点群・波形データ地物分類モデルを活用する有効性を確認する。なお、点群データのみを入力した深層学習モデルは第3章における大局ブロックと同じ層設計を使用し、入力を点群・波形データから点群データに変更したものである。そのため、以降では点群大局モデルと呼ぶ。

### 5.4.1 データセット

本節におけるデータセットは第3章で紹介したベンチマークデータセット (3.5.1節) を用いた。なお、学習済みモデルを使用するため、訓練データは使用せず、評価用のテストデータのみを用いた。

### 5.4.2 実験結果

実験結果について説明する。まず、処理時間として、テストデータを処理する時間はTSUBAME3.0にて約2.5時間であった。この処理時間の長さの理由は、2つの学習済みモデルを適用するために長い処理時間を要したためである。

次に、ベンチマークデータセットのテストデータに対して、推定した点群・波形データを用いた地物分類結果を述べる。図5.4に地物分類結果、表5.1に定量評価結果を示す。

まず、定性的な評価結果を述べる。定性評価の比較としては、図5.4(a)は物理モデルで推定した点群・波形データを用いた結果、図5.4(b)は点群データのみを学習したモデルによる推論結果、図5.4(c)は推定した点群・波形データを用いた結果、図5.4(d)は観測値の点群・波形データを用いた結果、図5.4(e)は高さ情報、図5.4(f)は真値の地物ラベルを示す。物理モデルによる結果(図5.4(a))では、全体的に誤分類が多い傾向を確認できた。点群データのみを学習した点群大局モデル(図5.4(b))では、高さデータのみしか使用できないため、右下部分の高架道路(ピンク色の楕円で示す)になっており、高架道路を建物と誤分類した。しかし、図5.4(c)のように点群・波形データを推定し地物分類を実施した結果、高架の部分で誤分類が減少した。なお、同様の高架部分の誤分類が減少した結果は観測値の点群・波形データを使用した地物分類結果(図5.4(d))でも同様である。これは、波形データがあることで、単純に幾何的な形状が高架と建物で平坦な形状が段差になっている状況が似ていても、素材の違いが波形データに現れること反映しているためである。また、地物分類手法はマルチタスク学習と定義でき、局所ブロック・大局ブロックそれぞれが得意な地物の特徴抽出と分類に寄与するため、より詳細な分類が可能になる。

しかしながら、細い道路(白い楕円で示す)において、推定した波形データは観測値の波形データを用いた場合と比較すると地物分類が困難であった。これらの道路に着目し、深層学習モデルによって推定した波形データを図5.5に表示した。図5.5上図は失敗例を示し、細い道路における波形データ推定結果である。赤色が推定結果を示し、青色が観測値を示す。上図で示す細い道路では、観測値の波形データは細かなピークと大きなピークがあり、推定された波形データではこの複数のピークを推定できていない。一方、同じ道路であっても、下図に示すような広い道路であれば、推定した波形データが観測値に近い分類結果を示した。これは、教師データにおいて、広い道路のデータや、ひらけた場所のデータが多く、木下の細い道路のデータが少ないことに起因する。これを克服するためには、訓練データの作成時に道路の太さや植生の下にある道路などを偏りなく用意する必要がある。

次に、定量評価結果を述べる。多クラス地物分類を行う先行研究に倣い、ここではF1 Scoreで地物分類結果を定量的に評価した。点群データのみを学習した点群大局モデルによる推論

結果(表 5.1の 1 行目)と比較すると、擬似的な点群・波形データを用いた推論結果(表 5.1の 3 行目)はより高い F1 Score を示した。性能が向上した理由は、入力特徴として波形データが入力されたことで、幾何学的な情報だけでは判別が困難なクラスの地物分類が可能になったためと考えられる。また、観測値の点群・波形データを学習したモデルによる推論結果(表 5.1の 4 行目)と比較すると、擬似的な点群・波形データを用いたモデル(表 5.1の 3 行目)は近い性能を示した。第 3 章で提案した点群・波形データの地物分類モデルは植生、電線・鉄塔は波形データを用いて地物分類するので、細かな波形データの再現ができていなくても典型的な波形データもしくは波形データの特徴を再現できていれば、地物分類性能に寄与できる。さらに、物理モデルで作成した点群・波形データと第 4 章の深層学習手法で推定した手法を比較する。物理モデルで作成した点群・波形データを用いて地物分類した結果(表 5.1の 2 行目)は、深層学習手法で推定した結果(表 5.1の 3 行目)と比較すると、地物分類性能が低い結果になった。これは、物理モデルで推定した点群・波形データは地面や建物のような単純な形状を持つ地物においては平均的な結果を予測するが、複雑な形状を持つ地物においては波形データの予測が困難であることに起因する。物理モデルを有効活用するためには、計算時に仮定する 3 次元モデルの作成方法に依存するため、より現実的な 3 次元モデルを作成する必要がある。物理モデルを機能させるには、より現実的な 3 次元モデルが事前に必要になる。

手法自身の課題として、学習をする際に波形データのクラス地物分類をするための損失関数を入れていないため、画像の分野にて行われているようなタスク特化の超解像のような地物分類器に寄与する特徴抽出ができない。具体的には、画像の超解像については、セグメンテーションに特化した研究 [195] や、検出に特化した研究 [171] があり、これらの方法に準じて、それぞれのタスクに特化した手法を提案する必要がある。しかし、このような戦略は点群データに対する深層学習手法に対して適用するのは、計算に使用する GPU のメモリ容量というハードウェアの制限によって今回は検証できない。具体的には、提案した波形推定手法は、生成器と 2 つの識別器によって成立つため、さらに地物分類モデルを追加しながら学習を行うと、本研究にて使用した現状の TSUBAME3.0 では GPU メモリ不足に陥る。将来的には、GPU メモリの問題はハードウェアの技術躍進によって解決できると考える。特に、深層学習が普及した当初である 2015 年では民生機としては 12 GB の GPU メモリが最大であったが、現在では 24GB の GPU メモリを有する民生機が登場した。このようなハードウェアの進歩によって、波形データを推定しながら地物分類を行うモデルの構築も可能になると考えられる。

表 5.1: ベンチマークデータにおける定量評価結果。推定した点群・波形データによる地物分類結果と観測値の点群・波形データを用いた際の地物分類結果と物理モデルで作成した波形データによる地物分類結果と点群データのみでの地物分類結果の比較。指標は F1 Score を用いた。

点群データ	波形データ	使用モデル	建物	地面	植生	電線	鉄塔	道路	平均値
✓	-	点群大局モデル	0.67	0.89	0.64	0.87	0.33	0.60	0.67
✓	物理モデル	局所・大局モデル	0.52	0.87	0.05	0.00	0.00	0.18	0.27
✓	推定	局所・大局モデル	0.85	0.93	0.92	0.95	0.84	0.79	0.88
✓	観測値	局所・大局モデル	<b>0.86</b>	<b>0.97</b>	<b>0.97</b>	<b>0.99</b>	<b>0.88</b>	<b>0.82</b>	<b>0.92</b>

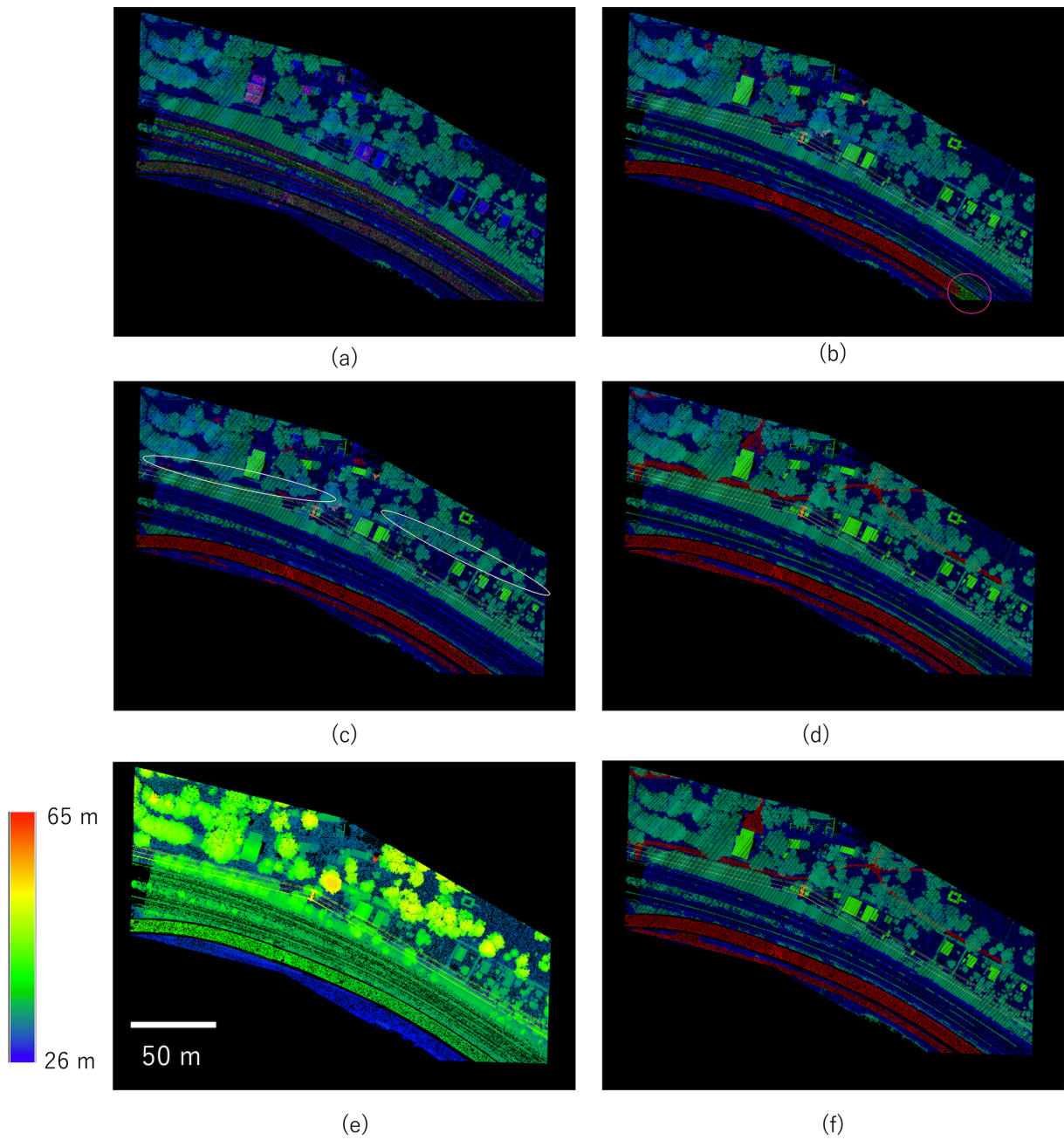


図 5.4: 推定した点群・波形データによるベンチマークデータセットに対する地物分類結果. (a): 物理モデルで推定した波形データを入力した局所・大局モデルによる結果, (b): 点群データのみを学習した点群大局モデルの結果, (c): 擬似的な点群・波形データを入力した局所・大局モデルによる結果, (d): 観測値の点群・波形データを入力した局所・大局モデルによる結果, (e): 高さ情報, (f): 真値の地物ラベルを示す.

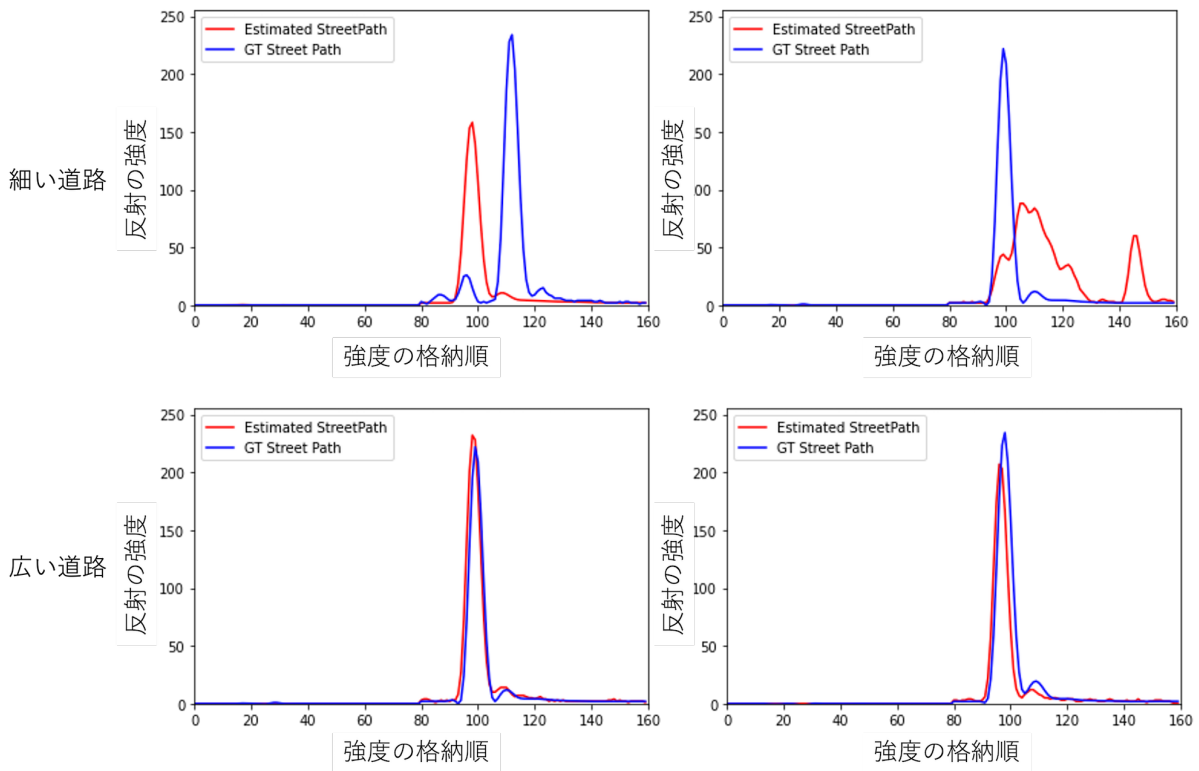


図 5.5: ベンチマークデータにおける波形推定の失敗例と成功例. 上図は細い道路, 下図は広い道路の波形推定結果. 推定結果は赤色, 観測値は青色で示している.

## 5.5 アーカイブデータへの適用実験

本節では、地物分類モデルと波形推定モデルの学習済みモデルの汎化性能を示すために、アーカイブされた点群データを用いて波形を推定し擬似的な点群・波形データを作成する。そしてその推定した点群・波形データによる地物分類性能を評価する。まず、本節で汎化性能の検証のために使用するデータセットである点群データ地物分類の真値付きデータである DALES データセット、および日本のアーカイブデータについて説明する。次に DALES データセットの実験結果、日本のアーカイブデータにおける実験結果について説明する。

### 5.5.1 アーカイブデータ

まず、本節で使用する点群データの真値付きデータセット（DALES データセット）と、日本のアーカイブデータを説明する。

#### 1. DALES データセット

図 5.6 に示す DALES データセットは点群データの地物分類の評価を行うための公開データセットである。DALES データセットのテストデータを使用して、定量評価を行った。なお、このデータセットはベンチマークデータセットよりも新しい世代の点群および波形情報を記録するライダーである Riegl Q1560 にて取得され、計測高度は 1,300 m であり、点密度はファーストパルスで 20 点/m<sup>2</sup> である。

DALES データセットは、訓練に使用したベンチマークデータと地物の状況が似ている地域かつ同じメーカーのライダーであるため、提案手法の純粋な汎化性能を評価可能である。DALES データセットは地面、植生、車、トラック、電線、フェンス、ポール、建物の真値が各点に付与されている。この DALES データセットの真値はベンチマークデータセットと対応関係を整理すると表 5.2 のようになる。DALES データセットにおいて、道路と地面の区別をしていないため、ベンチマークデータにおける道路の地物を地面として扱った。そのため、提案手法による地物分類結果は道路を地物に統合した 5 つの地物で評価する。なお、推論対象のファイルは ID: 5100\_54440 を用い、南北 500 m に区切られている。

表 5.2: DALES データセットとベンチマークデータセットとの対応

第 3 章の地物分類	DALES データセットとの対応
地面	地面
植生	植生
建物	建物
電線	電線
鉄塔	ポール
道路	地面

この DALES データセットを用いることで、ベンチマークデータセットによる地物分類と波形推定モデルの学習済みモデルの地域差やセンサ差に対する汎化性能を定量的に示す。学習済みモデルによる推論を実行する際には、ベンチマークデータセットと条件を近づけるために前処理を実行する。前処理は、ベンチマークデータセットと同様にファーストパルスのみを対象にするため、DALES データセットにおいてもファーストパルスの点のみを抽出する。また、点密度をベンチマークデータセットである 15 点/m<sup>2</sup> と同程度に点を間引いた。

#### 2. 日本のアーカイブデータ

日本のアーカイブデータは日本で実際に運用される地理空間データに関するオープンデータを公開する Web サイトである G 空間情報センター [69] で公開された点群データを

使用する。本実験では、2021年7月6日静岡県熱海市土石流災害の被災地域を観測したデータを対象にする。解析対象にする点群データを図5.7に示す。図5.7に示した範囲は、図郭に区切られたデータ(08nf2372)であり、南北は300m、東西は400mである。また、図5.7が示すように、植生や建物、電線等が入り乱れた地域であり、発災後のデータであるため、北西から南東に向かって土砂が多く確認できる。この日本のアーカイブデータは、訓練データとは撮影機材と地物の状況の条件が離れており、汎化性能を評価する上で難易度の高い状況として設定した。

このデータセットにおいて、目視判読にて作成した地表のデータを用いて定量評価を行う。この理由は今回用いた熱海のデータのような土石流の判読においては土砂の判読が重要になるためである。まず、地表の定義と学習に使用したベンチマークデータにおける地物分類の定義との関係性を述べる。訓練に使用したベンチマークデータにおける地物は地面・植生・建物・電線・鉄塔・道路である。表5.3に示すように、これらの地物を、地表(地面・道路)と非地表(建物・植生・電線・鉄塔)と定義した。

表 5.3: 日本のアーカイブセットとベンチマークデータセットとの対応。日本のアーカイブデータセットにおいては土石流に着目し地表を対象にする。

第3章の地物分類	地表との対応
地面	地表
植生	非地表
建物	非地表
電線	非地表
鉄塔	非地表
道路	地表

学習済みモデルによる推論を実行する際には、ベンチマークデータセットと条件を近づけるために前処理を実行する。前処理は、ベンチマークデータセットと同様にファーストパルスのみを対象にするため、日本のベンチマークデータセットにおいてもファーストパルスの点のみを抽出する。また、点密度をベンチマークデータセットである15点/m<sup>2</sup>と同程度に点を間引いた。

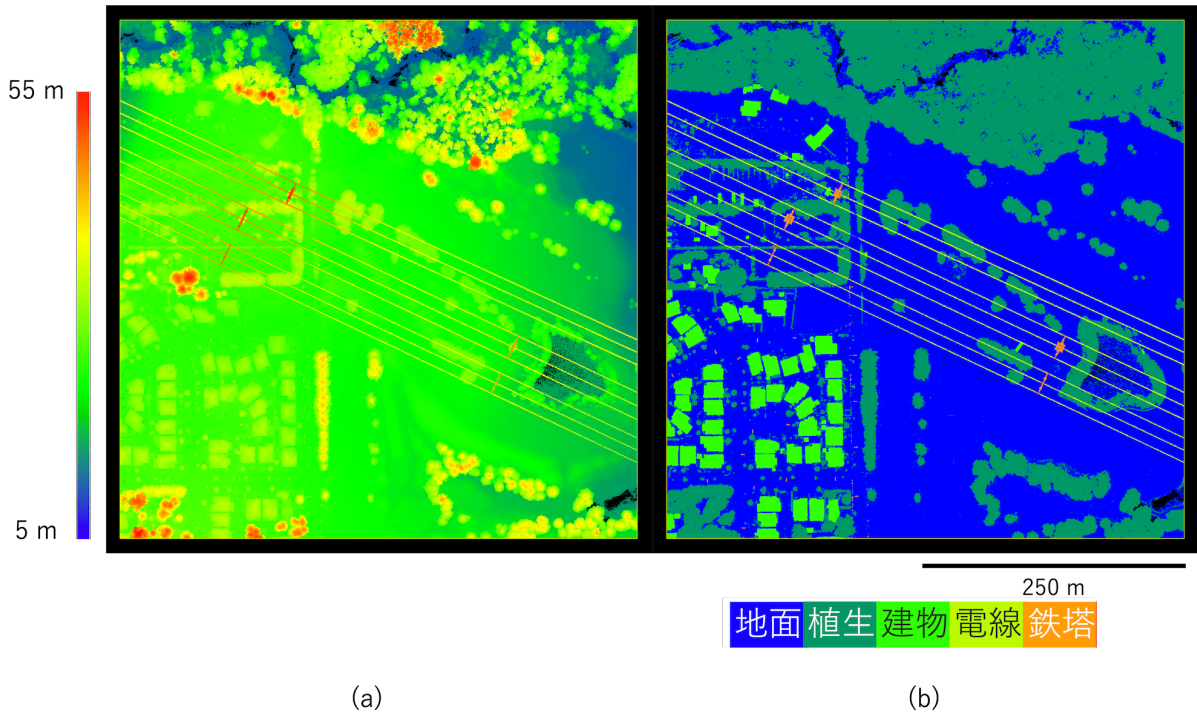


図 5.6: 真値付き点群データの DALES データセット. DALES[215] データセットでは, 地面, 植生, 車, トラック, 電線, フェンス, ポール, 建物の真値が各点に付与されているため, ベンチマークデータセットと同等にラベルを振り直した. (a): 高さの可視化, (b): 真値をベンチマークデータセットと同様に振り直した結果.

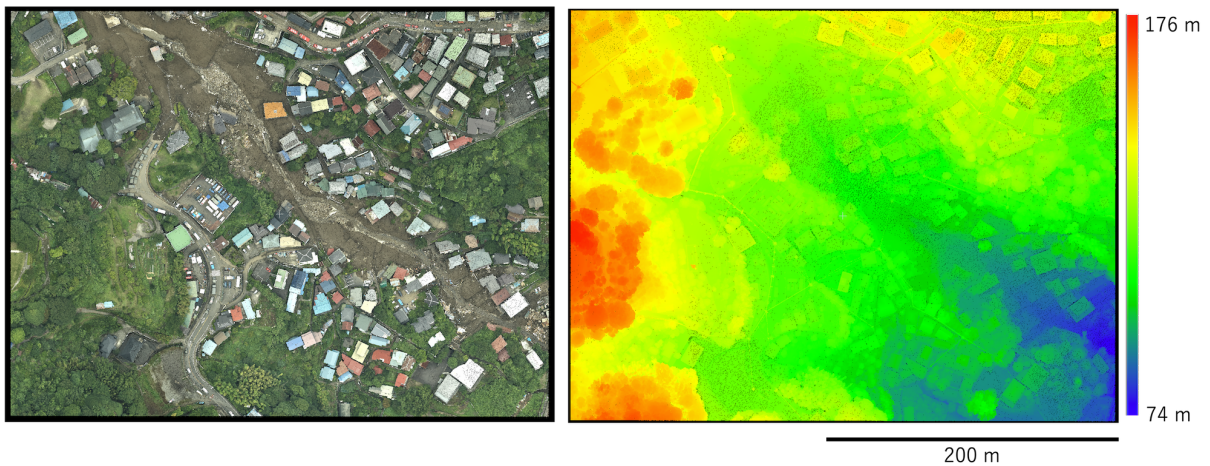


図 5.7: 日本のアーカイブデータ. G 空間情報センター [69] で公開された熱海の点群データを使用する. なお, 2021 年 7 月 6 日静岡県熱海市土石流災害の被災地域を観測したデータである. 真値は付与された状態で公開されていないため, 目視にて土石流被害に関する地表/非地表の真値を新たに付与した.

### 5.5.2 DALES データセット地物分類実験

提案手法を用いた DALES データセットに対する地物分類結果について述べる。ベンチマークデータセットを学習した地物分類モデルは地面・植生・建物・電線・鉄塔・道路の6種類の地物を分類するが、DALES データセットでは、地面と道路をひとまとめにして、地面・植生・建物・電線・鉄塔・道路の5種類の地物の真値が付与されている。そこで、地物分類モデルが出力した地物分類結果において、道路と予測された場合は、道路の地物に融合する後処理を実行した。なお、波形推定モデルにおいては地物情報の条件を受けないため、分類する対象の数に変更があった場合でも影響を受けない。提案手法である波形推定によって得た擬似的な点群・波形データの地物分類結果を図 5.8(b) に示す。図 5.8(b)において、地面（青色）、植生（緑色）、建物（黄緑色）、電線（黄色）の地物の抽出が良好であることが確認できた。点群データのみを用いた深層学習モデル（点群大局モデル）の地物分類結果（図 5.8(a)）と比較すると、植生と電線と鉄塔の抽出性能が向上したことが確認できる。これらは局所クラスとして設定した地物であり、局所・大局モデルの効果が現れたと言える。なお、点群データのみを用いた点群大局モデルは地物分類と同様のベンチマークデータを学習し、その層設計は第3章における大局ブロックと同様の構造を持つ。

詳細に地物分類結果を確認すると、図 5.9(a) に示すように、点群データのみでの点群大局モデルでは植生を建物として分類していた箇所において提案手法の効果が現れた。図 5.9(b) に示すように擬似的に点群・波形データ化することにより、植生として分類可能になった。これは、擬似的な点群・波形データ化したことで、波形データを用いる局所ブロックが地物分類結果に効果的に作用したためと考えられる。

次に、DALES データセットに対する地物分類結果を定量的に評価する。表 5.6 に DALES データセットにおける地物分類結果の F1 Score を示す。提案手法における擬似的な点群・波形データによる地物分類結果は、地面・植生・建物・電線において 0.9 を超えた。しかし、上記の定性的な評価と同じく、鉄塔の分類性能は 0.43 と他の地物と比較して低い F1 Score になった。これは、送電線を支える鉄塔が訓練データに存在しないため、波形推定、および地物分類が困難であったことに起因する。また、点群データのみを用いた点群大局モデルによる分類結果と比較すると、提案手法は波形データの形状が地物分類に寄与すると考えられる植生と鉄塔で F1 Score が向上した。擬似的な点群・波形データ化によって、点群データのみを用いた点群大局モデルよりも高い地物分類性能を持つ点群・波形データの地物分類モデルを有効的に活用できたためである。

総じて、訓練データとは地域の違う DALES データセットに対し、ベンチマークデータを学習した波形推定モデルと地物分類モデルは汎化性能を示した。一方、送電線の鉄塔のような訓練データに存在しない状況においては地物分類が困難な場合があり、継続的なデータセットの収集を行うことは必要である。

表 5.4: DALES データセットにおける定量評価結果。擬似的な点群・波形データを第 4 章で提案した地物分類モデルによる結果と点群データのみを用いる深層学習モデル（点群大局モデル）による結果の比較。評価指標は F1 Score を用いた。なお、訓練データに使用したベンチマークデータの地物は地面と道路を分けていたが、DALES データは区別していないため定量評価時においては道路を地面として扱った。

点群データ	波形データ	使用モデル	地面	植生	建物	電線	鉄塔	平均値
✓	-	点群大局モデル	0.97	0.89	0.91	0.81	0.00	0.72
✓	推定	局所・大局モデル	0.97	0.95	0.94	0.91	0.43	0.84

最後に、点群データのみを用いた点群大局モデルによる分類結果において、植生と建物の誤分類が発生したが、この誤分類を少なくするためのハイパーパラメータを探索した。点群大局モデルのハイパーパラメータは点群データの深層学習手法を用いて航空機で観測した点群データに対する地物分類の先行研究 [229] において最適とされた値を使用した。そのため、点群データのみを用いた点群大局モデルの地物分類において様々な地物を同時に扱う際にはその

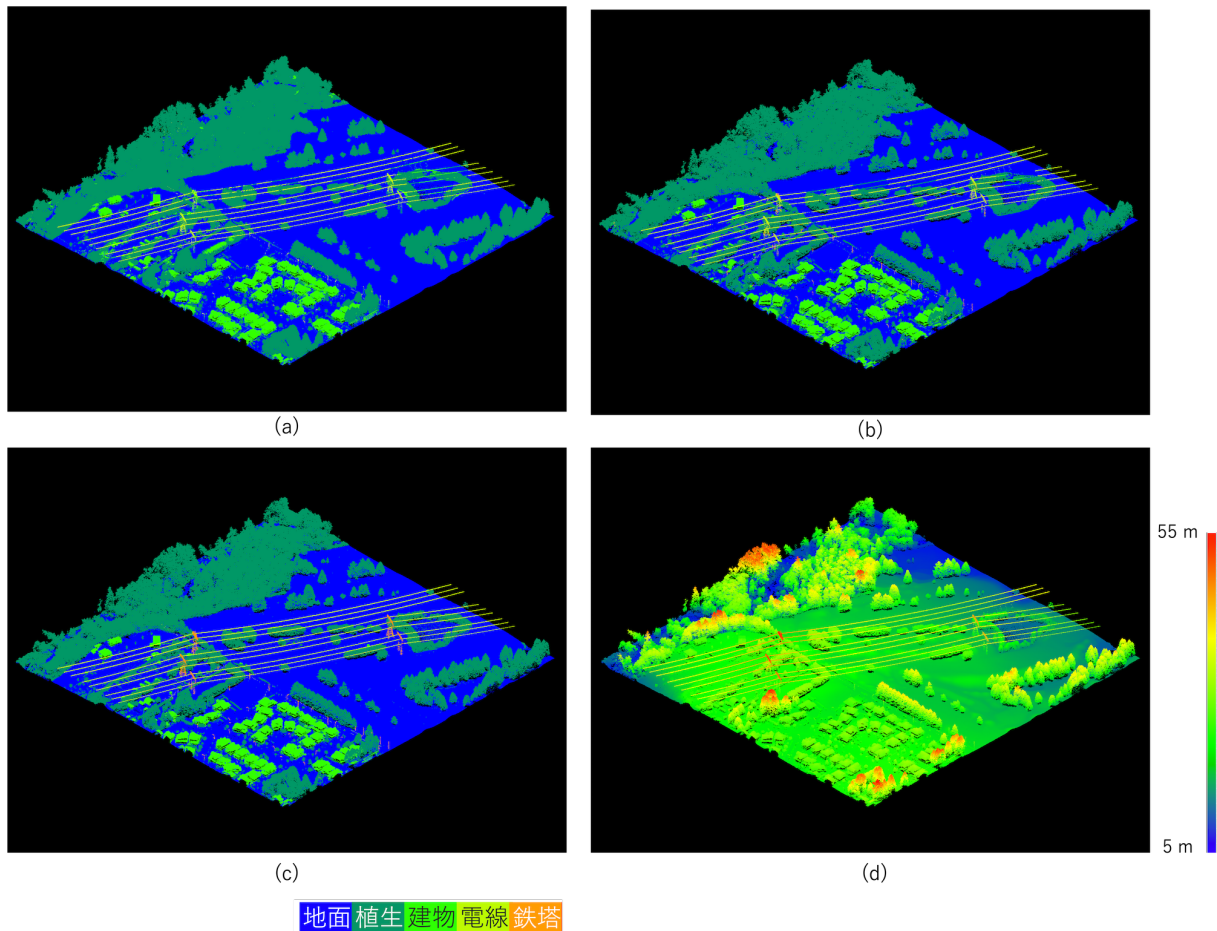
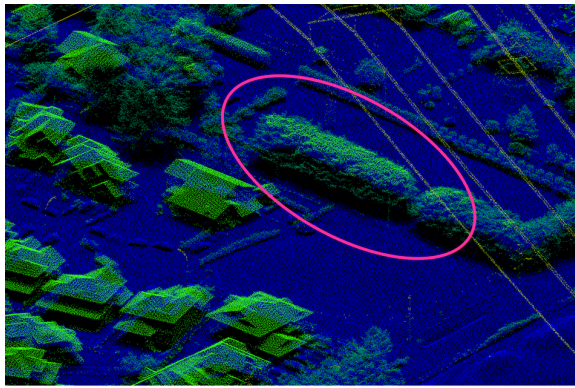
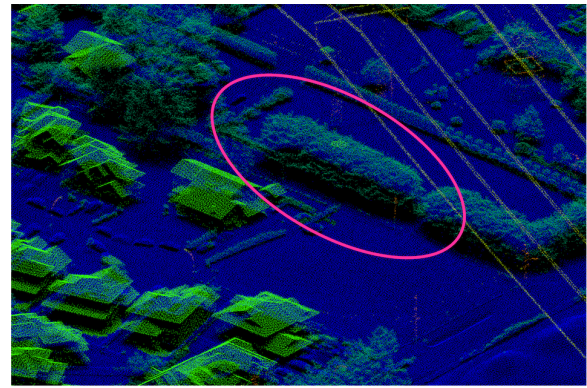


図 5.8: DALES データセットに対する地物分類結果. (a): 点群データのみを用いた場合, (b) 提案手法による擬似的な点群・波形データ化による地物分類結果, (c) 真値, (d) 高さデータ. ベンチマークデータでは地面と道路を区別していたが, DALES データセットでは地面と道路を地面としてまとめているため, 5つの地物へ分類した結果を示す.

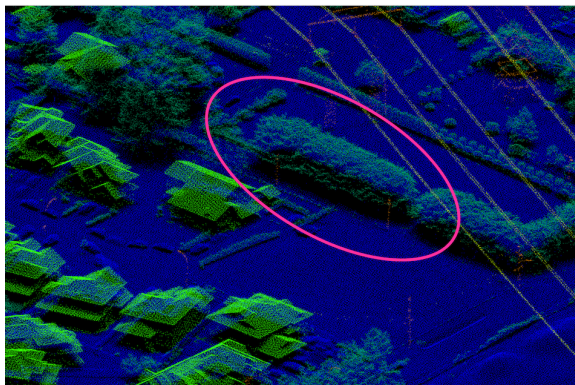
地物の大きさ等に応じて繊細なチューニングが必要になる. そこで, 点群大局モデルのハイパーパラメータをチューニングした場合の分類結果と, 提案手法の擬似的な点群・波形データを入力した局所・大局モデルの分類結果と比較した. ハイパーパラメータは点群大局モデルにおいて, 代表点に対するグループ化の際に使用する半径 (グループ化半径) のハイパーパラメータが植生と建物を分類する際に効果的であるため, ベンチマークデータの学習を実行する際にその条件を変えてモデルを複数個構築した. 点群データのみを用いた点群大局モデルではグループ化を三回実行するため, 表 5.6 で示した結果の場合ではグループ化半径を浅い層から 1m, 5m, 15m に設定している. このグループ化半径を 1/5 倍から 5 倍まで等間隔に条件を変えた点群データのみを用いた点群大局モデルをベンチマークデータを学習して新たに作成する. なお, このグループ化半径の大きさは, ある一点を分類する際に使用する周辺の点群データの範囲であり, 深層学習分野では受容野と呼ばれる範囲と関係する. グループ化半径を変えた学習済み点群データ点群大局モデルによる DALES データの分類結果を表 5.5 に示す. 半径 1m, 5m, 15m の条件を表 5.6 の点群大局モデルにおいて採用していたが, それぞれの半径を 2.5 倍にして広い範囲の点の情報を取り込む条件にした点群大局モデルが最も建物と植生の分類性能が高くなったが, 電線の分類性能結果が悪化した. 一方, それぞれの半径を 1/5 倍にして狭い範囲の点の情報を取り込む条件にした点群大局モデルが最も電線と鉄塔の分類性能が高くなったが, 地面・植生・建物の分類性能結果が悪化した. このように, 一つの地物に着目してグループ化半径をチューニングすることで, 点群大局モデルはその着目



(a)



(b)



(c)

地面 植生 建物 電線 鉄塔

図 5.9: DALES データセットにおける提案手法の効果. (a) 点群データのみを用いた場合, (b) 提案手法による擬似的な点群・波形データ化による地物分類結果, (c) 真値. 植生の誤分類の解消に着目して地物分類の結果を表示している.

した対象の地物に対する分類性能を向上させられるが、他の地物の分類性能が劣化する悪影響がある。なお、局所・大局モデルの大局ブロックでは先行研究 [229] において最適とされた値をそのまま使用したが、局所ブロックと大局ブロックでそれぞれ対象にする地物を切り分けたことで点群大局モデルのようなグループ化半径への依存を少ないため影響は少ない。これは、局所ブロックと大局ブロックでそれぞれ得意/不得意を切り分けたため、それぞれのブロックでは簡単な分類問題に帰結し、その結果、繊細なハイパーパラメータのチューニングが必要なくなったためである。本研究で提案した点群・波形データの地物分類を実行する局所・大局モデルはこの悪影響を受けずに多数の地物を一つの深層学習モデルで分類可能である。これにより、深層学習手法を用いた点群・波形データの処理や点群データの処理において計算時間が膨大であるため、一度の計算で多数の地物の情報が得られる利点もある。

表 5.5: 点群大局モデルにおけるベンチマークデータの学習時のチューニングによる効果. グループ化で使用する半径の値を元のハイパーパラメータ (1m, 5m, 15m) に対して係数をかけて条件を変えた. 分類結果は F1 Score を示す.

モデル	グループ化半径 (倍)	地面	植生	建物	電線	鉄塔
大局点群モデル	1/5	0.32	0.20	0.48	0.87	0.42
	1/2.5	0.84	0.82	0.89	0.82	0.12
	1/1	0.97	0.89	0.91	0.81	0.00
	2.5	0.98	0.91	0.92	0.51	0.00
	5	0.49	0.36	0.31	0.00	0.00
局所・大局モデル	1/1	0.97	0.95	0.94	0.91	0.43

### 5.5.3 日本のアーカイブデータにおける検証実験

提案手法を用いた日本のアーカイブデータに対する地物分類を行なった。図 5.10(b) に提案手法による擬似的な点群・波形データ化による地物分類結果を示す。また、図 5.10(a) に点群データのみを用いた深層学習モデルによる地物分類結果を示す。定性的には DALES データセットとは違い、訓練データとは特性が異なる地域性であることの影響を受けた。この結果から、地表面を推定する。その際に、ファーストパルスを用いたため全ての点に地物結果を付与するための後処理を実行した。後処理としては、ファーストパルス以外の点に対して近くの点の地物分類結果を伝播させた。そして、すべての点の地物分類結果から、地面と道路を「地表」としてラベルを振り直した。

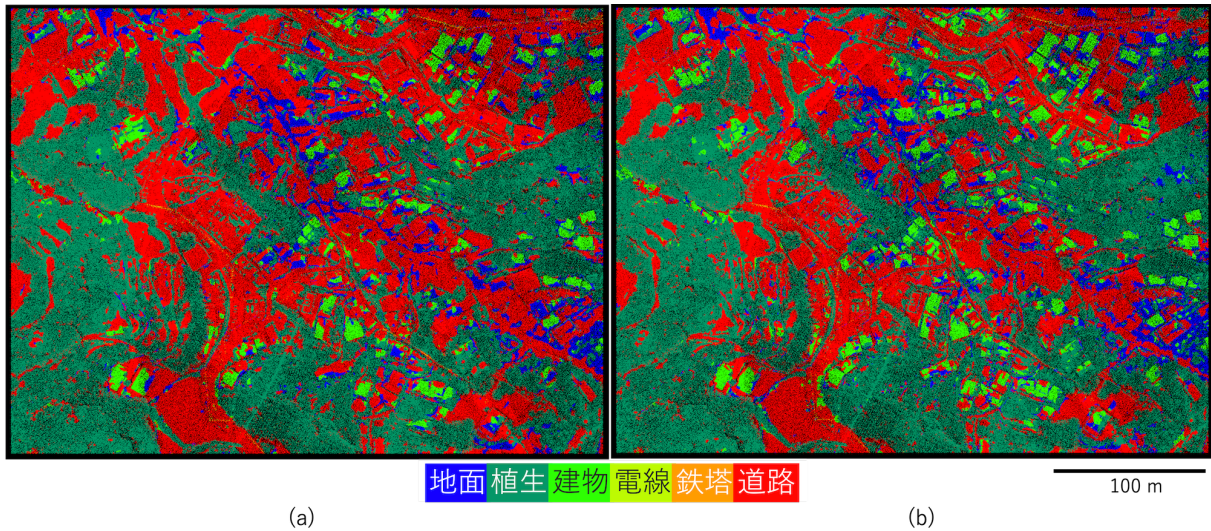


図 5.10: 日本のアーカイブデータに対する地物分類結果。(a): 点群データのみを用いた場合、(b) 提案手法による擬似的な点群・波形データ化による地物分類結果。

次に「地表」の分類結果を定量的に評価した結果を述べる。評価指標は F1 Score を用いた。提案手法の擬似的な点群・波形データを推定し地物分類を行う手法によって、表 5.6 の 2 行目に示すように地表は 0.83, 非地表は 0.97 を示した。また、点群データのみを用いた点群大局モデルの場合 (表 5.6 の 1 行目) は、地表は 0.73, 非地表は 0.97 を示した。なお、点の数としては非地表が 10 倍ほど多くなるため、定量評価上では非地表の評価は相対的に高くなる。提案手法によって、点群データのみによる点群大局モデルによる分類結果よりも高い指標を示した。

次に、詳細に「地表」の抽出結果を確認する。波形推定による具体的な効果を確認する。図 5.11 に土石流被害を受けた建物の例を示している。図 5.11(a) は入力点群データ、図 5.11(b) は点群データのみを用いた点群大局モデルの地表 (地面・道路) 分類結果、図 5.11(c) は擬似的な点群・波形データを地物分類深層学習モデル (局所・大局モデル) に入力した地面・道路の分類結果、図 5.11(d) は真値を示す。ピンク色の丸で示した建物に着目すると、点群データのみによる点群大局モデルでは屋根部分を地表 (地面・道路) と分類しているが、擬似的な点群・波形データによって、地面と連続している屋根の部分地表から除外することができた。一方、地表の分類に失敗している箇所もあり、白丸に示す平屋の建物において点群・波形データの推定結果を用いた場合、点群データのみを用いた点群大局モデルによる結果では共に誤分類が発生した。これは、訓練データ由来の問題であるため、今後、地域性を吸収する仕組みやデータセットの拡充が必要と考えられる。

表 5.6: 日本のアーカイブデータセットにおける定量評価結果. 地表と非地表の分類において, 擬似的な点群・波形データによる結果と観測値の点群・波形データを用いた結果を比較した. 指標は F1 Score を用いた.

点群データ	波形データ	使用モデル	地表	非地表
✓	-	点群大局モデル	0.73	0.97
✓	推定	局所・大局モデル	0.83	0.97

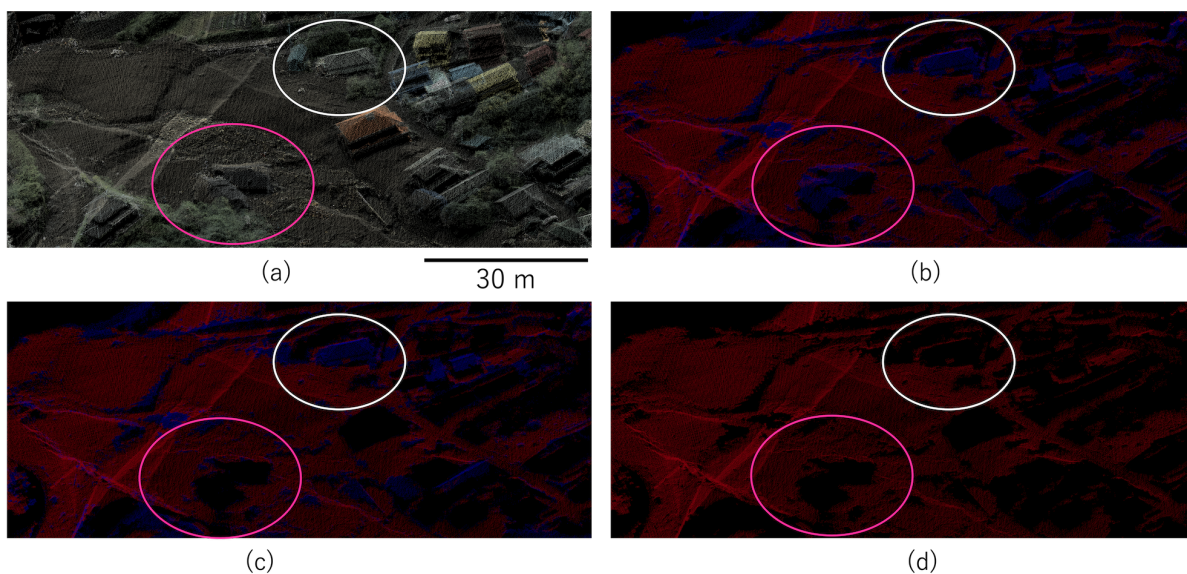


図 5.11: 日本のアーカイブデータにおける地表の抽出結果の例. (a): 入力点群データ, (b): 点群データのみを学習した点群大局モデルによる抽出結果, (c): 擬似的に点群・波形データを作成して局所・大局モデルに入力した抽出結果, (d): 作成した真値の地表の点群データを示している.

## 5.6 第5章のまとめ

本章では、第3章で提案した点群・波形データを地物分類するモデルを実際に公開されているアーカイブデータに適用する際には波形データが入手し難いという課題に対応した。課題を解決するために、点群データから波形を推定する第4章の手法を適用して、点群データから波形データを推定する手法と組み合わせてアーカイブ点群データに対して点群・波形データの地物分類モデルを適用した。

実際に提案手法の波形推定モデルと地物分類モデルを組み合わせる前に、波形推定の汎化性能を評価した。その際には、点群データと波形データがペアになっているダブリンデータセットを用いた。ダブリンデータセットに学習済みの波形推定モデルを適用した結果、ある程度の汎化性能を示した。

次に、波形推定手法と点群・波形データの地物分類を組み合わせる枠組みの有効性を確認するために、ベンチマークデータに対して検証実験を行った。ベンチマークデータを用いた検証実験の結果、実際に観測された波形データを用いた場合に迫る地物分類性能を示した。特に、点群データのみを用いて地物分類を行う点群大局モデルの学習済みモデルと比較すると、地物分類性能の向上が確認できた。

次に、学習済みモデルの汎化性能を示すために、アーカイブ点群データへの適用実験を行なった。アーカイブ点群データとして、訓練データと似た地物が存在する DALES データセットと日本のアーカイブデータを用いた。DALES データセットは点群データの精度比較用のデータセットであるため真値が付与されており、その真値と比較して学習済みモデルによる地物分類結果の定量評価を行なった。DALES データセットに対して、波形推定による擬似的な点群・波形データの地物分類を適用した結果点群データのみを用いた点群大局モデルの場合よりも高い分類性能を示した。DALES データセットは、地域が異なるが、建物の大きさや植生の繁茂の仕方などは訓練データに近い状況であり、学習済みモデルを用いて波形推定を実施し擬似的な点群・波形データの地物分類を行った結果、訓練データと離れた場所・異なるセンサにおいてもある程度は手法の有効性が確認できた。また、点群データのみを用いた深層学習モデルである点群大局モデルによる地物分類を行った結果と比較すると、植生域においては、波形推定し擬似的な点群・波形データとして地物分類する手法の有効性が特に確認できた。

次に、日本のアーカイブデータに対して適用した。日本のアーカイブデータは2021年熱海土石流を観測した点群データであり、訓練データとはかけ離れた難易度の高い汎化性能の評価が可能である。日本のアーカイブデータに対しては土石流という性質に着目して地表の点群データの分類性能を評価した。訓練に使用したベンチマークデータにおける地物は地面・植生・建物・電線・鉄塔・道路である。これら6つの地物を、地表（地面・道路）と非地面（建物・植生・電線・鉄塔）と定義した。定量的な比較を行なった結果、提案手法の波形推定による擬似的な点群・波形データ化により地表の分類性能が点群データのみによる点群大局モデルによる分類結果よりも高性能を示した。

## 第 6 章

### 結論

## 6.1 博士論文全体のまとめ

本研究では、3次元地図の自動作成を実現するために点群・波形データを自動分類する手法の提案を行なった。そのために、深層学習技術を用いた点群・波形データの地物手法の開発に取り組んだ。行うタスクとしては、点群データに地物情報を付与することであり、その際に波形データ情報の有効活用を目指した。さらに、提案した点群・波形データの地物分類手法を活用する際には、アーカイブデータにおいて波形データが格納されない問題点を克服する必要がある。そこで、点群データから波形を推定し、擬似的に点群・波形データ化する手法を提案した。以降でそれぞれについて説明する。

本研究では、点群・波形データの地物分類を新たに提案した。点群・波形データの地物分類を行う際に、「個々の波形データが有効な地物」と「波形データと波形データの空間的な関係性である大局情報が有効な地物」に大別できると仮説を立てた。それぞれの地物に対して、前者には個々の波形データの地物分類を実施する局所ブロック、後者には近傍特徴抽出を階層的に適用して地物分類を実施する大局ブロックを適用し、それぞれを組み合わせる。提案手法の効果として、深層学習を使用しており一般的な機械学習手法の波形データからのヒューリスティックな特徴量抽出を行わないため、エンドユーザーの解析能力に依存しない。また、提案手法は波形データと点群データを直接取り扱うため、深層学習を利用した点群・波形データ解析の従来手法で用いられた点群データと波形データを画像に変換する手法の2次元に投影する際の3次元構造の情報欠損の影響を受けない。ベンチマークデータセットを用いた学習済みモデルの分類性能評価によって、提案手法の大局ブロックと局所ブロックの組み合わせの効果を示した。

次に点群データからの波形推定手法について説明する。波形推定における既往研究は存在しないため、点群データから別の情報を推定する点において似ているタスクである点群着色手法を参考にした。その際に、点群・波形データの地物分類と同様に、個々の波形に着目した。さらに、推定した波形を評価する損失関数として、その形状を周波数特性にて評価する新たな損失関数を提案した。単に点群着色手法を適用するのではなく、新たな提案を用いることで、点群データからその点群データが観測された際の波形を推定する性能が向上した。

さらに、点群データからの波形データを推定する手法と、点群・波形データの地物分類手法の2つの深層学習手法を組み合わせることで、点群データから擬似的に波形データを作成し、その作成した波形データと点群データ(擬似的な点群・波形データ)を点群・波形データの地物分類手法に入力し地物分類を行った。この擬似的な点群・波形データの地物分類手法をベンチマークデータとアーカイブデータに適用した。ベンチマークデータにおいては、点群データのみを入力する深層学習モデルよりも提案手法である擬似的な点群・波形データ化は地物分類性能として、高いF1 Scoreを示した。アーカイブデータにおいては、点群データの公開データセットである DALES データセットに対して提案手法を適用した。DALES データセットにおいては真値が付与されているため、その真値を用いて地物分類性能を定量評価した。その結果、単に点群データの地物分類を実施するよりも、提案手法の擬似的な点群・波形データ化により高い地物分類性能を示した。さらに、日本のアーカイブデータに対して地物分類を実行した結果、地表の抽出性能に着目して手法の有効性を確認した。

## 6.2 詳細なまとめ

### 6.2.1 第3章: 点群・波形データを用いた地物分類手法の提案

第3章では、点群・波形データの地物分類を行う深層学習手法（局所・大局モデル）を提案した。局所・大局モデルでは、波形データと点群データを画像に投影する先行研究 [255] のように、別のデータ形式に変換することなく、生の点群・波形データの地物分類を実現した。局所・大局モデルを提案する際には、点群データに対する深層学習手法を参考にした。その際に、点群・波形データの特性に基づき、分類対象の地物を2つに大別できると仮説を立てた。それは、個々の波形データのみで分類が可能な地物と、波形データの空間的な関係性である大局情報を用いると分類が可能になる地物である。点群データに対する深層学習手法は、ある点とその近傍点の特徴を抽出する「近傍特徴抽出」を繰り返し階層的に適用することで大局的な特徴が得られる一方、個々の点の情報は有効活用できない問題がある。そこで、個々の波形データを有効的に活用する深層学習モデルを組み合わせた。提案手法である局所・大局モデルは、点群データに対する深層学習手法に対して点群・波形データを扱えるように適用した層設計と大局ブロックとして、さらに個々の波形データを用いた地物分類を実行する層設計を局所ブロックとして取り入れた。以降で、それぞれのブロックの概要を示す。

まず、大局ブロックについて概要を示す。大局ブロックは、点群データの深層学習手法を参考にした。これは点群データ入力し、その点単位にクラスの地物分類を行う手法である。この点群データに対する深層学習手法は、階層的に近傍点の情報を集約しながら点群データの特徴を抽出するエンコーダーと、抽出された特徴から元の点群データ数までアップサンプリングしながらクラス地物分類を行うデコーダーで構成される。特に、エンコーダでは、画像による地物分類等で良好な結果が得られる畳み込みニューラルネットワークのように近傍特徴抽出とプーリング構造を、ダウンサンプリングによる代表点の抽出とその代表点に対して近傍点を球内に入る点とて定義してその近傍点の情報を集約する。しかし大局ブロックは地物分類する際には地物に対して仮定に対して個々の波形データの形状を扱うことが困難である。これは、エンコーダ-デコーダ構造によって、広い範囲の大局情報を集約するため、広範囲に分布した地物には有効に働くが、局所的に分布する地物には不利になると考えられる。そこで、局所・大局モデルでは波形データのみを用いて地物分類する新たな構造を導入した。具体的には、波形データのみを用いて地物分類する構造を局所ブロックと定義し、波形データの特徴を抽出するための一次元の畳み込み演算を適用した。これは、波形データは照射したレーザーパルスの反射エコーを反射して帰ってきた時間順に格納されているため、時系列のデータと定義でき、時系列のデータの代表例である音声処理分野で使用される一次元の畳み込み演算を導入した。なお、これらの2つの局所ブロック・大局ブロックの2つは同時に学習を行う。さらに、局所・大局モデルの学習済みモデルを用いて推論を行う際は、波形データが有効であると考えられるクラスに属する確率が0.5を超えた場合を局所ブロックからの出力を用い、それ以外を大局ブロックからの出力を採用した。

提案手法の局所・大局モデルの有効性を確認するため、ベンチマークデータを用いて評価実験を行った。評価を行う際には、比較実験として、先行研究の画像に変換する点群・波形データの地物分類手法、点群データに対する深層学習手法に対して波形データを扱えるように純粋に拡張した手法を学習した。評価の結果、提案手法は先行研究よりも高い性能を示し、純粋に波形データを扱えるように拡張した点群データに対する深層学習手法よりも高い性能を示した。提案手法による分類結果は人間による目視判読において発生する6%の誤差以内にほとんどの地物が入った。また、局所ブロックの有効性を示すため、最終的な地物分類結果に対する寄与を局所ブロックからの採用率を計算した。その結果、個々の波形データで分類が可能と仮定した地物において高い採用率を示した。また、局所ブロックの判断根拠を可視化した結果、波形データの変化点に着目していることが判明した。

点群・波形データに対して、データの持つ特性を考慮し適切に層設計を提案することが重要であることが示された。

### 6.2.2 第4章: 点群データからの波形推定手法の提案

第4章では、点群データからその点群データが取得された際の波形を推定する深層学習手法(波形推定モデル)を提案した。本章で対象にするようなタスクを一般化し、点群データからその点群データが持つ別の情報を推定するタスクとすると、近いタスクとしては点群着色の手法が提案されている。

そこで、点群着色手法を参考にして、点群データからの波形推定手法の提案(波形推定モデル)を行った。点群着色手法よりも、波形推定はRGB情報という比較的低次元な情報の推定よりも高次元な強度情報の時系列情報を扱う必要がある。そこで、地物分類手法と同様に、個々の波形に着目する仕組みが必要であると仮説を立てた。この仮説を参考に、波形推定モデルでは2つの新たな手法を追加した。一つ目は、波形の形状に着目した損失関数(周波数回帰)の提案である。二つ目は、波形単位に評価を行う識別器の提案である。波形推定では、大局生成器と空間的な整合性を評価する識別器(大局識別器)と波形単位の識別器(局所識別器)の3つの深層学習モデルを使用した。点群データと波形が点単位で1:1に対応しているベンチマークデータセットを用いて、大局生成器と大局識別器・局所識別器の学習を実行した。このような提案手法を用い、検証実験を行った結果、点群データから波形を推定することは可能であった。また、新たに追加した手法に対する効果検証を行った結果、損失関数の提案と識別器の提案は有効であった。なお、単純な建物などの地物は両者ではあまり差がなかった。

深層学習を用いて点群データから波形を推定する手法として条件付きGANを採用し、波形に特化した周波数回帰と局所識別器を新たに提案することで、真値に近い性能を示した。

### 6.2.3 第5章: 波形推定手法と地物分類手法の効果検証

第5章では、第4章で提案した学習済み波形推定モデルによる点群データからの波形推定と、点群・波形データの地物分類を行う学習済み局所・大局モデルによって、波形データが無い状態においても擬似的な点群・波形データとして扱い地物分類する手法を提案した。実際に2つの学習済みモデルを組み合わせる前に、波形推定モデルの汎化性能を評価した。次に検証実験を実施し、ベンチマークデータと点群データのみが提供されるアーカイブデータとして DALES データセットと、日本のアーカイブデータを対象にした。

波形推定モデルの汎化性能評価には、点群データと波形データがペアとして存在するダブルデータセットを用い、学習に使用したベンチマークデータセットとは別の地域の点群データにおいても波形が推定できることを示した。

ベンチマークデータにおいては、点群データのみを学習した深層学習モデルに点群データを入力した地物分類結果、第3章の局所・大局モデルに観測された点群・波形データを入力した地物分類結果と、第3章の局所・大局モデルに波形推定モデルで推定した点群・波形データを入力した地物分類結果を比較した。比較の結果、擬似的な点群・波形データを用いることで、点群データのみを学習した深層学習よりも高い分類性能を示した。しかし、実際に観測された点群・波形データを入力した場合よりも地物分類性能は低い結果になった。定性的に確認すると、実際に観測された点群・波形データにおいては良好な地物分類結果を示した細い道路において、波形推定モデルでは波形の推定が十分に行えなかった。

次に、点群データのみが与えられる場合を想定したアーカイブデータに対して提案手法を適用した。地物分類性能の定量評価を行うために、波形データが存在せず、かつ点群データに真値が付与された公開データセットである DALES データセットを用いた。DALES データセットに対して提案手法である擬似的な点群・波形データの地物分類手法を適用した結果、単なる点群データのみを入力する深層学習モデルを用いる場合よりも高い F1 Score を示した。

さらに、日本のアーカイブデータを用いた検証実験を行なった。日本のアーカイブデータは、2021年に発生した熱海の土石流データを使用した。DALES データセットより困難な状況下を想定し、訓練に使用したデータとはセンサと地域特性が異なる日本のオープンデータの公開サイトである G 空間情報センターで公開されている点群データを用いた。日本のアーカイブデータでの実験は土石流という現象に着目し、地表分類を実施した。地表とは、ベンチマークデータセットにおいて6つの地物が存在したうちの、地面と道路が相当する。定量評価を実施するために、地表の点を目視にて抽出し真値を作成した。提案手法である擬似的な点群・波形データの地物分類手法を適用した結果、単なる点群データのみを入力する深層学習モデルを用いる場合よりも高い F1 Score を示した。

このように、波形データを推定は第3章にて提案した局所・大局モデルを活用する上で有効であり、点群データのみを使用する場合よりも地物分類性能への寄与が示された。

## 6.3 手法の貢献点と適用範囲

ここでは、提案手法の貢献点を確認し、また、提案手法の適用範囲を説明する。

### 6.3.1 貢献点

#### 1. 点群・波形データの地物分類手法の提案

- ワークフロー  
本研究では、点群・波形データを地物分類するためのワークフローを提案した。このワークフローは、点群データと波形データを画像等に変換することなく深層学習モデルに入力し地物分類結果の出力で終了する。これにより、標準的な点群データ地物分類のワークフローと大きく変えることなく、単に波形データを別途用意するだけで、点群・波形データの地物分類が実行可能になる。
- 完全自動処理  
提案する点群・波形データの地物分類手法を適用するには、ライダーの仕様に関する専門的な知識は必要ではない。さらに、深層学習モデルの内部で自動的に点群・波形データの特徴が計算されるため、波形データに対する専門家でなくても使用できる。これは、学習済みモデルを適用するだけで結果が得られるためである。
- 点群・波形データに特化した地物分類手法の提案  
点群データ処理のための深層学習手法を基に、点群・波形データに特化した層設計として、波形データと幾何的な情報に特化したモデルを組み合わせる新たな層設計を提案した。
- 高性能な地物分類  
先行研究と比較すると比較的高い地物分類結果が得られた。また、点群データのみを学習した深層学習モデルで行った地物分類結果よりも提案した点群・波形データの地物分類手法は高性能を示した。さらに、既存の点群データに対する深層学習手法を単純に波形データを扱えるように拡張した手法（提案手法における大局ブロック）と比較しても高性能を示した。
- 新たな解析ツール  
このようなワークフローを開発したことで、点群・波形データの処理ツールを作成することができた。これらのツールを OSS として公開した。

#### 2. 波形推定手法の提案

- 点群データからの波形推定タスクの提案  
点群・波形データの解析によって、高性能な地物分類が達成できた。しかし、波形データはデータ容量の大きさからアーカイブ時には捨てられることが多い。そこで、点群データから波形を推定して擬似的な点群・波形データとして点群・波形データの地物分類モデルへ入力する手法を提案した。
- 波形推定モデルの提案  
点群データから波形を推定するタスクは、点群データからその点群データ本来持っている情報を推定するタスクと定義できる。近いタスクとしては、点群着色の手法が提案されているが、これは CAD データで作成した単純な形状のみを対象にした。そこで、点群着色手法を参考に波形推定を実行する新たな深層学習モデルを提案した。
- アーカイブ点群データの有効活用  
点群データから推定した波形データにより擬似的に点群・波形データとして地物分類モデルにて扱えるようになった。単に点群データのみを地物分類する深層学習モデルよりも、擬似的に点群・波形データ化して点群・波形データの地物分類を行う深層学習モデルに入力することで、高性能な地物分類を実現した。

### 6.3.2 適用範囲

#### 1. 地物分類の対象とその分類性能

本研究は、3次元地図作成の基盤情報になる点群データに対する目視判読を効率化するために、点群・波形データの地物分類手法を開発した。その際に、標準的な点群データや点群・波形データの仕様定義される15種類の地物のうち、代表的な地面・植生・建物・電線・鉄塔・道路の6つに限定し、求められる地物よりも分類可能な地物の種類が少ない。そして、限定した地面・植生・建物・電線・鉄塔・道路の6つの地物分類において、人間が目視判読する際に発生する誤分類である6%未満を達成した地物は地面・植生・建物・電線の4つであった。そのため、分類性能が未達成の地物において分類性能を向上させる必要と、分類可能な地物数の向上が必要である。

#### 2. 点群・波形データが得られない

本研究では、点群・波形データの地物分類を行っているが、実運用上ではストレージの関係から波形データは捨てられる傾向にある。この欠点は点群データから波形を推定すること克服した。しかし、波形推定によって得た擬似的な点群・波形データによる地物分類への寄与は示されたが、観測値の点群・波形データを用いた地物分類結果と比較すると性能が劣る結果になった。今後、波形データの有用性を航空測量分野に訴えかけ、波形データを保持するようにする必要がある。なお、将来的に点群・波形データが整備された場合、本研究にて提案した波形推定手法は過去の点群データのアーカイブデータを有効活用するためのツールとして使用が想定される。

#### 3. 地物分類時の波形データと点群データの組み合わせ方

点群・波形データの地物分類を行う深層学習モデルは、波形データで地物分類を行う局所ブロックと、点群データと波形データを組み合わせて地物分類を行う大局ブロックをそれぞれ学習を同時に行った。そして、予測時に簡単なルールで両者の出力を組み合わせて出力する。つまり、大局ブロックと局所ブロックは入力共通して完全に独立して学習が行われていることと同義である。そのため、両者の特徴抽出が相互に作用しない問題がある。また、単純な閾値処理のルールで局所ブロックの結果を採用しているため、過剰に局所ブロックから採用する傾向もある。

#### 4. 波形推定時の制約条件

波形推定においては、各点の情報から直接深層学習モデルで推定する制約を設けていない手法を用いた。そのため、計測原理上はピーク位置と点の位置は一致するはずだが、推定されたピークの位置が真値とずれる場合が発生した。これは、深層学習のモデルに対して、計測過程の制約条件を入れていないことで生じたと考えられる。

#### 5. 学習データの依存

地物分類手法と波形データ推定手法を用いた地物分類結果は訓練データに近い条件のDALESデータセットであれば高い分類性能があることが示された。しかし、訓練データと観測値とセンサが状況である日本で取得された別のデータにおいてはベンチマークデータやDALESデータセットほどの分類性能は確認できなかった。これは、研究の対象としたアーカイブデータにて使用されることが多い2012年から2019以降の商用点群および波形情報を記録するライダー以外の新しいセンサであり、観測方式もレーザを円形に回転させながら観測する方式で取得されるため、学習データの並行スキャン方式の点群および波形情報を記録するライダーとは観測方式が異なることが理由である。地域性や観測方式の他にも、提案手法の適用範囲はフットプリントの小さく点密度の多い点群・波形データに限られ、地物分類と波形データ推定の学習済みモデルは共に点密度を疎にすると分類性能が低下した。また、本研究は航空機で取得した点群・波形データを対象にしたため、他の観測方法で取得されたデータにおいては対応が困難である。具体的には、3次元地図の作成においては航空機だけでなく自動車で取得した点群データに対しても同様に地物分類が必要であるが、本提案手法で作成した学習済みモデルは自動車で取得した点群データへ適用ができない。自動車で取得した点群データでは、航空機では観測できない信号機やポールや標識などの地物を観測可能であるため、ロボット

の自律走行などの応用を考える上では重要になる。さらに、本研究では点群・波形データに深層学習手法を適用する最初の研究であることから、都市域に存在する可能性の高い代表的な地物に絞って実験を行なった。そのため、実世界に存在するすべての地物は対象ではない。最後に、本研究では3次元地図作成の用途やアーカイブデータへの適用性に鑑み波形データのピークから取得したファーストパルスの点群データのみを使用した。そのため、波形データの複数のピークから取得した点群データと波形データの組み合わせた樹木の解析や地物分類手法については研究できていない。

上記の適用範囲の課題をまとめると、3次元地図の自動作成において、提案手法の適用可能な点群・波形データの観測方法を飛行機以外にも広げ、目視作業の分類性能をすべての地物で上回る手法の提案、他地域への汎化性能や分類対象数を向上させる手法の提案が必要である。

## 6.4 将来の展望

### 6.4.1 点群・波形データの普及

提案手法の地物分類手法は点群・波形データの普及によってその能力を發揮できる。しかし、点群および波形情報を記録するライダー観測業務の標準的な成果物は「点群データ」であり、生産ラインにおいて観測された波形データは点群データ作成時にのみ利用する。また、オープンデータ等にて公開される際にも波形データは用いられないことが多い。そのため、点群・波形データの利活用が進んでいない現状にあり、深層学習モデルの汎化性能の観点においても様々な地域やセンサで観測された点群・波形データの整備が重要になる。

本研究で示したように点群データだけでなく波形データも併用することで自動処理の高度化が可能になる。点群・波形データの波形データの利用を促進するためにまず考えられるのは、研究者コミュニティがオープンソースのライブラリやソフトウェア等を整備し配布することであり、エンドユーザーが点群・波形データの地物分類を簡単にできるようにすることである。エンドユーザーが実際に解析を行い波形データの有用性が確認できると、ストレージ容量の増加のデメリットに対し解析結果の良好さのメリットが上回るため波形データが普及すると考える。このように、実益があることを研究者コミュニティだけでなくエンドユーザーへ訴えることが重要になる。なお、本研究で提案した深層学習モデルに関しては、機械学習コミュニティの通例に従って、コードの共有サイトである GitHub に公開してオープンソースにしている。

さらに、波形データ量が大きくアーカイブサイトのストレージを圧迫する問題に対してデータ量の削減も今後着目すべき課題になる。データ量の削減としては、深層学習を利用した軽量なデータ表現手法が提案されている。例えば、点群データを符号化する FoldingNet[237] が提案されており、これを応用して波形データを符号化する手法へ拡張可能になる。さらに、3次元的な光線空間をモデリングする深層学習手法 [151] も提案されており、センサと物体間のレーザの関係性も同様に深層学習技術でモデリングできる可能性がある。これらの手法は、一度学習を行う計算機のコストはかかるが、実際にストレージに格納される際には、軽量なバイナリファイルとなり、そのバイナリファイルから深層学習モデルで符号になった状態から元の状態に戻すことが可能である。

### 6.4.2 点群・波形データの地物分類性能の向上

提案した点群・波形データの地物分類手法は、点群データと波形データを別の形式に変換せずに地物分類を行う手法の提案を解決すべき課題と設定し、それを解決した。しかし、地物分類性能の観点からは、人間の分類性能を下回る地物が存在した。そこで、深層学習の分野にて提案された技術を取り入れることで多くの改善が考えられる。課題の例として、提案した深層学習手法の組み合わせ方法が挙げられる。点群・波形データの地物分類手法は、局所ブロックと大局ブロックを独立に同時に学習して、予測時に簡単なルールで組み合わせる最終的なクラス地物分類を行った。このような手法では、局所ブロックと大局ブロックの間の相互作用が発生せず、両者の相互作用が発生しない。このような相互作用は、学習時に両者を混ぜる仕組みを入れることで達成できると考えられる。具体的には、局所ブロックで作成した特徴量と、大局ブロックで学習した特徴量を混ぜつつ予測を行うブロックを新たに追加すると性能が向上すると期待される。なお、この際には、結果に対して寄与するか否かを深層学習モデル内部に持たせる機構 [164] を持たせて後段のクラス地物分類に寄与する特徴量を渡す層設計が望ましい。

### 6.4.3 点群データからの波形推定性能の向上

本研究では、単純な条件付き GAN の考えを使用した点群データからの波形推定を行った。波形推定は点群データの着色手法を基にしており、それに対して波形を個別に評価する識別器と波形の形状の類似度を評価する損失関数を新たに提案した。この手法は、現在コンピュータビジョン分野における画像の処理の分野で行われる着色手法と比較すると、工夫点が少ない。例えば、画像地物分類等のタスクの訓練済み深層学習モデルを利用した知覚損失 [118]

である。知覚損失は訓練済みモデルで抽出した特徴量を推定した結果と真値のデータで一致させる働きがある。また、画像変換の分野において、特徴量を分離して特徴量ごとに変換を行う手法も提案されている [119]。特徴量を分離できることで、画像に写った物体単位に画像変換が可能になる。これ応用すると、地物単位に変換を陰的に点群データから波形を推定可能になると期待できる。また、位置情報が既知であるため、波形のピーク位置は任意に決定できるため、これを波形推定における制約条件として使うことも考えられる。

#### 6.4.4 より波形データを活かした地物の地物分類

また、本研究では、波形データと点群データを組み合わせることで点群データの地物分類性能の向上に着目した。そのため、波形データを詳細に分析することで判別可能になる地物は分類対象にしなかった。波形データの形状で地物分類を行うタスクに変更すると、より詳細な地物分類が実行できる可能性がある。例えば、地面からの反射エコーの存在や波形データの形状の複雑さに応じて、建物のエッジ、密な植生、疎な植生などに分類できる。これらは、点群データという幾何的な情報だけでは目視にて地物分類が困難であったため、データセットを作ることができなかった。

具体的に建物を例に波形データを利用した地物分類のメリットを考える。生の波形データを維持することで容易に建物を検出できる可能性がある。建物の端にある波形データは、建物境界と地面からの反射によって、2つのモードを有する典型的な形状を有している。このように波形データを用いると、建物境界部分の検出に関する良い情報が得られ、同じ境界線上にある複数の波形データを使用して改良することができる。さらに、建物の屋根材の識別や地面の詳細な情報の抽出の問題も扱うことができる。屋根の場合は瓦、地面の場合は低い草といった表面の微細構造が、波形データの形状に影響を与える。このように、点群データのみもしくは光学画像では地物分類が困難であった詳細な地物の地物分類が可能になるであろう。

そのためには、波形データのデータセットの構築が必要になる。大量のデータを収集し、地面の素材や屋根の素材や建物境界部分の情報を真値として作成する必要がある。このような大量のデータを扱わなければならない場合、効率的なデータ解析・処理方法を開発しなければならない。深層学習を利用した計算の場合には、年々 GPU の性能が向上しているため、計算速度の問題は解決されるであろう。

また、点群データの標準仕様である LAS の仕様 [7] では、19 クラスの地物分類が想定されている (表 6.1)。本研究では、地面・植生・建物・電線・鉄塔・道路の 6 種類を分類した。LAS の仕様の 19 クラスにおいて、ノイズや分類不能等、ユーザが任意に指定できる予約 ID (ユーザ定義) を除いた意味のある地物は 15 クラスである。将来的には本研究で行なった 6 つの地物分類だけでなく、表 6.1 中の植生のように高さによって分ける等の細かな分類が必要になる。このような細かな分類において、15 種類の地物の真値を点群・波形データへ付与して状態で再学習することによって対応が可能になる。特に、本研究で作成した学習済みモデルが持つ重みを初期値として、新たなデータセットにて再学習を実行する転移学習が有効になると考えられる。点群・波形データの地物分類手法を 15 種類の地物分類に適用する際において、局所クラスと大局クラスの振り分けを慎重に行う必要がある。将来的にはこのような細かなクラス地物分類が自動で行えると、点群データの利活用が促進すると著者は確信している。

#### 6.4.5 航空機で観測した点群・波形データ以外への拡張

本研究では、航空機で観測した点群・波形データの地物分類を行ったが提案手法は様々なライダーに広げられる。車載型のレーザスキャナのデータに対して点群データだけでなく波形データに着目したデータセット [40] が公開されているため、本研究の内容を航空機以外にも適用できる。また、提案手法をより一般化すると、点群データとその点群データが取得された際の生の計測信号に着目する手法と定義でき、生の信号に着目する手法はコンピュータビジョン分野では達成できない測量分野だからこその強みと言える。

本研究では、3次元地図の作成を目的として研究を行った。人手による3次元地図の作成コストが点群データを取得する費用の10倍近く発生するため、点群データの利活用は3次元的な測量や可視化を行うのみに留まり、3次元地図の作成はなされていない。本研究で行った、

表 6.1: 点群データの標準仕様である LAS 形式で定義される地物分類対象

ID	地物名 (原義)	地物名 (和訳)	ベンチマークデータセット
0	never classified	地物分類不能	-
1	Unclassified	未地物分類	-
2	Ground	地面	地面
3	Low Vegetation	植生 (低)	植生
4	Medium Vegetation	植生 (中)	植生
5	High Vegetation	植生 (高)	植生
6	Building	建築物	建物
7	Low Point (noise)	低点 (ノイズ)	-
8	Reserved	予約	-
9	Water	水部	-
10	Rail	鉄道	-
11	Road Surface	道路	道路
12	Reserved	予約	-
13	Wire - Guard (Shield)	電線 - 保護線	電線
14	Wire - Conductor (Phase)	電線 - 導線	電線
15	Transmission Tower	鉄塔	鉄塔
16	Wire-structure Connector	電線接続構造物	-
17	Bridge Deck	橋梁床版	-
18	High Noise	高ノイズ	-
19-63	Reserved	予約	-
64-255	User definable	ユーザー定義	-

高性能な点群データの地物分類を実現する手法があることで、この人手による 3 次元地図の作成コストを削減可能である。3 次元地図が整備されることで、今までは単に人間が目視で確認するだけにとどまっていた点群データ等の 3 次元データが、コンピュータが我々の住む実世界を解釈するための基盤的な情報になる。そして、3 次元地図の普及によってある都市をコンピュータに再現して、シミュレーションや自律移動ロボットの経路の計算等の基盤情報として使用可能になる。また、人間の普段の生活においても、位置情報と 3 次元地図が組み合わさることで、移動の経路案内も簡単になる。今後は、提案した手法を基に、3 次元地図の整備に向けた大規模な計算を行うようにしたい。

## 付録

## A 在学中の採択論文

- 査読付き雑誌
  - 第3章 : Shinohara, Takayuki, Haoyi Xiu, and Masashi Matsuoka. "FWNet: Semantic segmentation for full-waveform LiDAR data using deep learning." *Sensors* 20, no. 12 (2020): 3568.
  - 第4,5章 : Shinohara, Takayuki, Haoyi Xiu, and Masashi Matsuoka. "Point2Wave: 3-D Point Cloud to Waveform Translation Using a Conditional Generative Adversarial Network With Dual Discriminators." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021): 11630-11642.
- 査読付き国際学会
  - 第3章 : Shinohara, Takayuki, Haoyi Xiu, and Masashi Matsuoka. "Semantic segmentation for full-waveform LiDAR data using local and hierarchical global feature extraction." In *Proceedings of the 28th International Conference on Advances in Geographic Information Systems*, pp. 640-650. 2020. (acceptance rate: 22.1%)
  - 第4章 : Shinohara, Takayuki, Haoyi Xiu, and Masashi Matsuoka. "Point2color: 3d point cloud colorization using a conditional generative network and differentiable rendering for airborne lidar." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1062-1071. 2021. (EarthVision: acceptance rate: 30%)
- 査読無し国際学会
  - Shinohara, T., H. Xiu, and M. Matsuoka. "Image to Point Cloud Translation Using Conditional Generative Adversarial Network for Airborne LIDAR Data." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2 (2021): 169-174.
  - Shinohara, Takayuki, Haoyi Xiu, and Masashi Matsuoka. "3D Point Cloud Generation Using Adversarial Training for Large-Scale Outdoor Scene." In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 2935-2938. IEEE, 2021.
  - Shinohara, Takayuki, Haoyi Xiu, and Masashi Matsuoka. "FWNetAE: Spatial representation learning for full waveform data using deep learning." In *2019 IEEE International Symposium on Multimedia (ISM)*, pp. 259-2597. IEEE, 2019.

## B 地物分類へ深層学習モデルの層設計が与える影響

この付録では第3章の考察にて行なった深層学習モデルの影響で使用した層設計について述べる。本研究の地物分類において大局ブロックに PointNet++[177] を用いた。PointNet++[177] は点群データに対する深層学習手法であり、点群データに対する深層学習手法は他にも存在する。そこで、PointNet++[177] を大局ブロックに使用した妥当性を検証するため、他の深層学習を用いた地物分類結果と比較する。比較対象として、点群・波形データを扱えるように既存の点群データに対する深層学習手法である PointNet を拡張した全体横断モデルを紹介する。図 B.1 に全体横断モデルの概要を示す。全体横断モデルは、PointNet アーキテクチャを採用しており、波形データを入力データに対応する地物分類結果  $S_{seg}$  に変換する。入力されるデータ ( $\mathbf{P}$ ) は、2次元 ( $K \times (M + 3)$ ) の座標と波形データの行列である。PointNet ベースの特徴抽出器 (図 B.1 の左側) は、3つのブロックで構成されている。

### 1. 座標と波形データごとの演算

最初のブロックは、入力された点群データと波形データを受け取り、各点とその波形データの波形データ単位に独立した局所的な波形データ特徴量を計算する (図 B.1 の赤い三角形)。この波形データ特徴量は、局所波形モデルと同様に  $1 \times 3$  のフィルタサイズを持つ1次元の畳み込み層を介して追加される。さらに、畳み込み演算を適用された後に、活性化関数の ReLU を適用することで、入力に対する非線形変換とする。この非線形変換は、ある層への入力  $p_k$  に対する畳み込み演算を  $h$ 、活性化関数を ReLU とすると、

$$\text{ReLU}(h(p_0)), \text{ReLU}(h(p_1)), \dots, \text{ReLU}(h(p_{K-1})), \quad (\text{B.1})$$

と示される。

### 2. 全点の代表値の取得

点と波形データごとの演算  $h$  を行い抽出した特徴に対して全点から対象関数によって代表値を取得する。対称関数とは、入力点の数を  $K$  としたときに、 $K!$  の入力順番に変動があった際に、入力順序に依存せずに同じ出力が得られる関数のことである。この場合、 $h$  は単純な1次元の畳み込み演算、活性化関数 ReLU で成り立ち、 $g$  は max-pooling 演算で計算される。max-pooling によって、入力順に不変的な全点を横断した特徴量の代表値を与える。

これらの畳み込みと代表値の取得を定式化すると以下ようになる。

$$f_\theta = g(\text{ReLU}(h(xp_0)), \text{ReLU}(h(p_1)), \dots, \text{ReLU}(h(p_{K-1}))), \quad (\text{B.2})$$

ここで、 $p_k$  は  $K$  個の波形データが入力された時の個々のデータ、 $f_\theta$  は重み  $\theta$  を持つ関数、 $h$  は個々の入力データに応じた畳み込み、ReLU は活性化関数、 $g$  は対称関数を表す。

### 3. 情報の結合

点ごとの特徴に対して全点で計算した独立した特徴量に対して大局的な情報を加えるために、max-pooling 層の出力の後に、各点の特徴と大局的な特徴を結合する (図 B.1 の  $\oplus$ )。

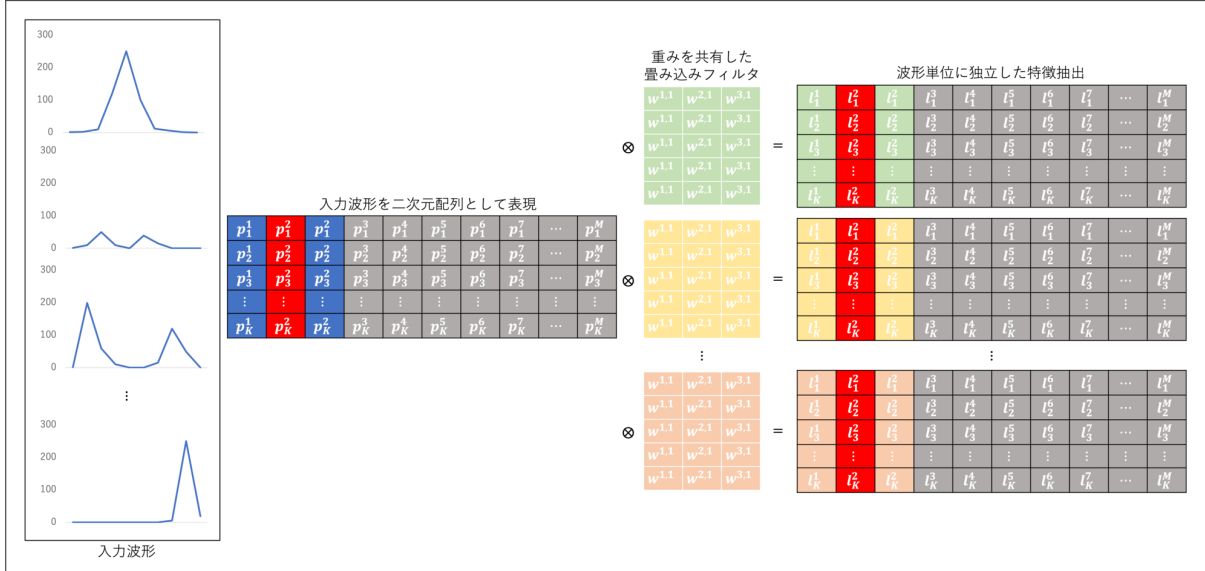
これら3つの要素を用いることで、入力点ごとの座標と波形データに対し独立した1次元の畳み込みによる特徴抽出だけでなく、入力データ全体を横断した特徴抽出も可能になる。このようにして得られた特徴量に対して地物分類を担う全結合層を適用して、点群・波形データの地物分類を行う (図 B.1) の右側)。なお、地物分類層として、単純な1次元畳み込み演算を使用する。3つの1次元畳み込み演算 (図 B.1) のオレンジ色の三角形) によって非線形特徴変換と、地物数と同等の次元数にされた後、ソフトマックス層によってそれぞれの地物に属する確率が推定される。

$$\text{Softmax} = \frac{\exp(a_i)}{\sum_j^n \exp(a_j)}, \quad i = 1, \dots, n \quad (\text{B.3})$$

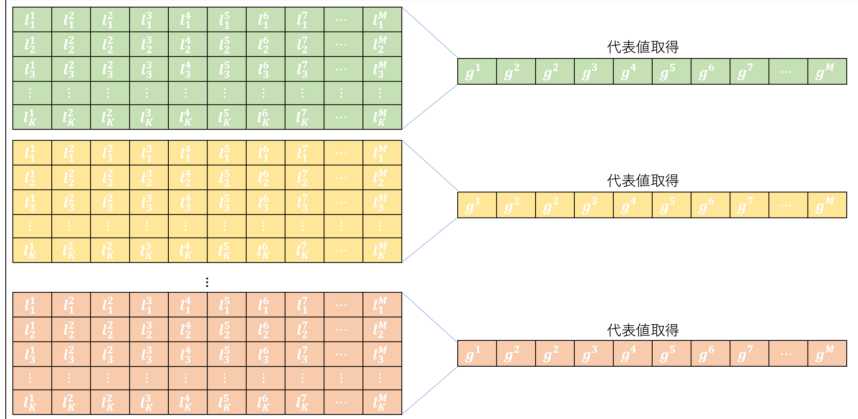
ここで、 $a_i$  はクラス  $i$  のクラス確率、 $\exp(a_j)$  は各クラスの地物分類層からの最終出力値、 $n$  はクラスの総数。

この層設計を用いた結果は、表 B.1 に示す通り、提案手法である局所・大局モデルよりも分類性能が低い結果となり、提案手法で採用した層設計の妥当性が示された。

波形単位の特徴抽出



全体の代表値の取得



分類層の演算 (6クラスの例)

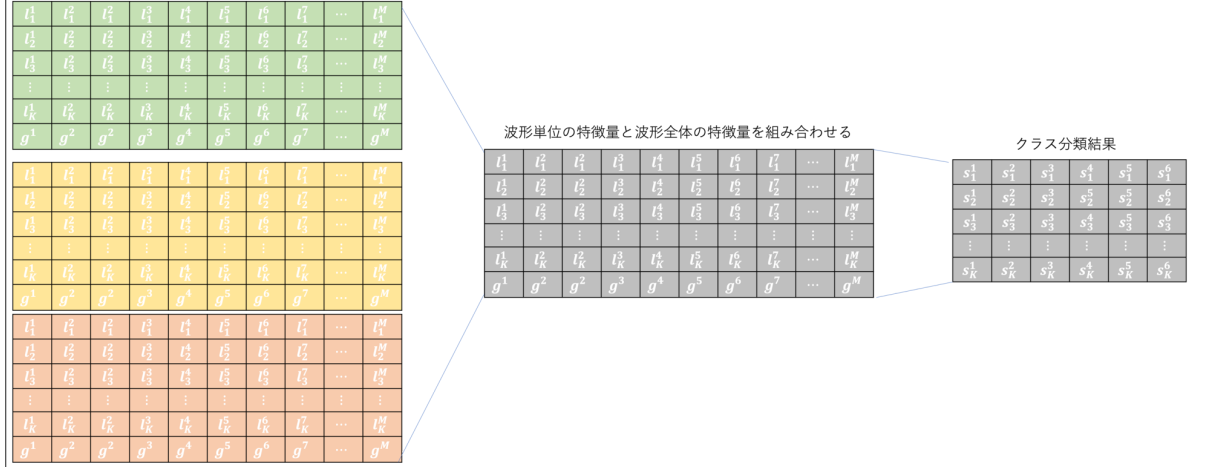


図 B.1: 全体横断モデルで使用する演算

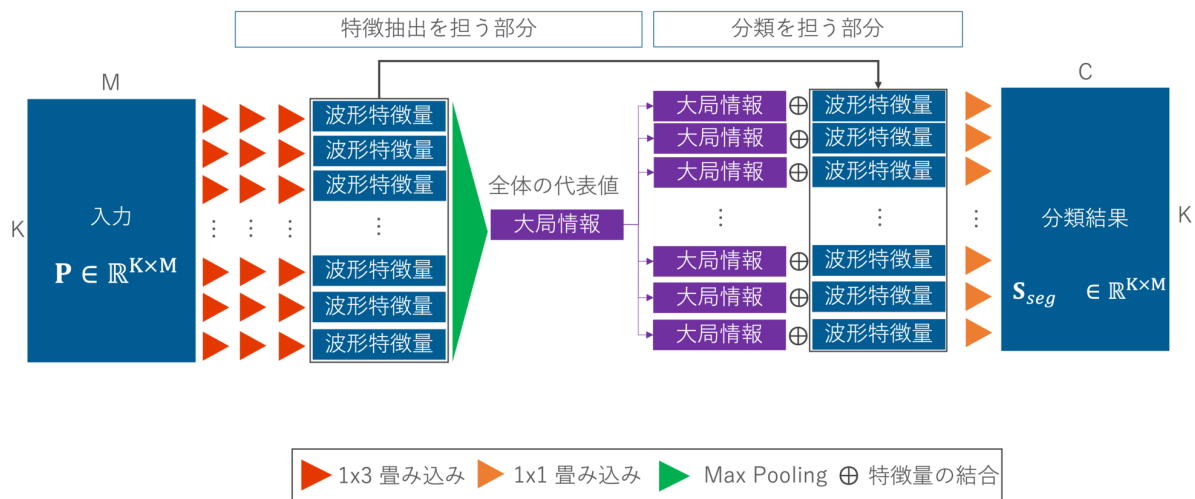


図 B.2: 全体横断モデルの層設計

表 B.1: ベンチマークデータにおける深層学習モデル構造の比較結果。太字は最高性能を示す。また、示している結果は5回のクロスバリデーションで得た学習済みモデルによる分類結果の平均値である。

手法	評価指標	地面	植生	建物	電線	鉄塔	道路	平均
全体横断モデル	Precision	0.56	0.97	0.95	0.92	0.61	0.94	0.81
	Recall	0.91	0.85	0.83	0.84	0.48	0.62	0.73
	F1 Score	0.69	0.91	0.88	0.88	0.53	0.75	0.76
局所・大局モデル	Precision	<b>0.80</b>	<b>0.98</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.97</b>	<b>0.95</b>
	Recall	<b>0.94</b>	<b>0.96</b>	<b>0.95</b>	<b>0.99</b>	<b>0.79</b>	<b>0.71</b>	<b>0.89</b>
	F1 Score	<b>0.86</b>	<b>0.97</b>	<b>0.97</b>	<b>0.99</b>	<b>0.88</b>	<b>0.82</b>	<b>0.92</b>

## C 点群データに対する条件付き GAN の予備実験

本節では、条件付き GAN を航空機に搭載されたライダーで観測された点群データへ適用する手法を検討する。条件付き GAN を単一物体かつ単純な形状を持つ点群データに適用した例はあるが、点群および波形情報を記録するライダーをはじめとした航空機に搭載されたライダーで取得した点群データに適用した例は無い。そのため、条件付き GAN を本研究で対象にするような屋外で計測された点群データに適用する際の特徴を把握する必要がある。

条件付き GAN を点群データへ適用して、別の情報を推定する先行研究の例は、点群着色 [131, 21] がある。しかし、これらの点群着色手法はカップやテーブルや車などの CAD データから作成した理想条件の点群データに対してのみ適用されている。この理想条件とは、「ひとつの物体のみが深層学習モデルに入力される」、「形状が単純」であることである。しかし、博士論文で扱うような航空機に搭載されたライダーで観測される点群データは様々な地物が入り乱れており、CAD データよりも複雑であることが多い。

そこで、CAD データに提案された手法を実際に屋外で観測された航空機に搭載されたライダーで観測された点群データに対する点群着色に対応できるように手法を改良する。この改良した点群着色手法を点群データ着色モデルと呼ぶ。そして、ここで提案する点群データ着色モデルを点群データから波形データを推定する手法への基本構造とする。

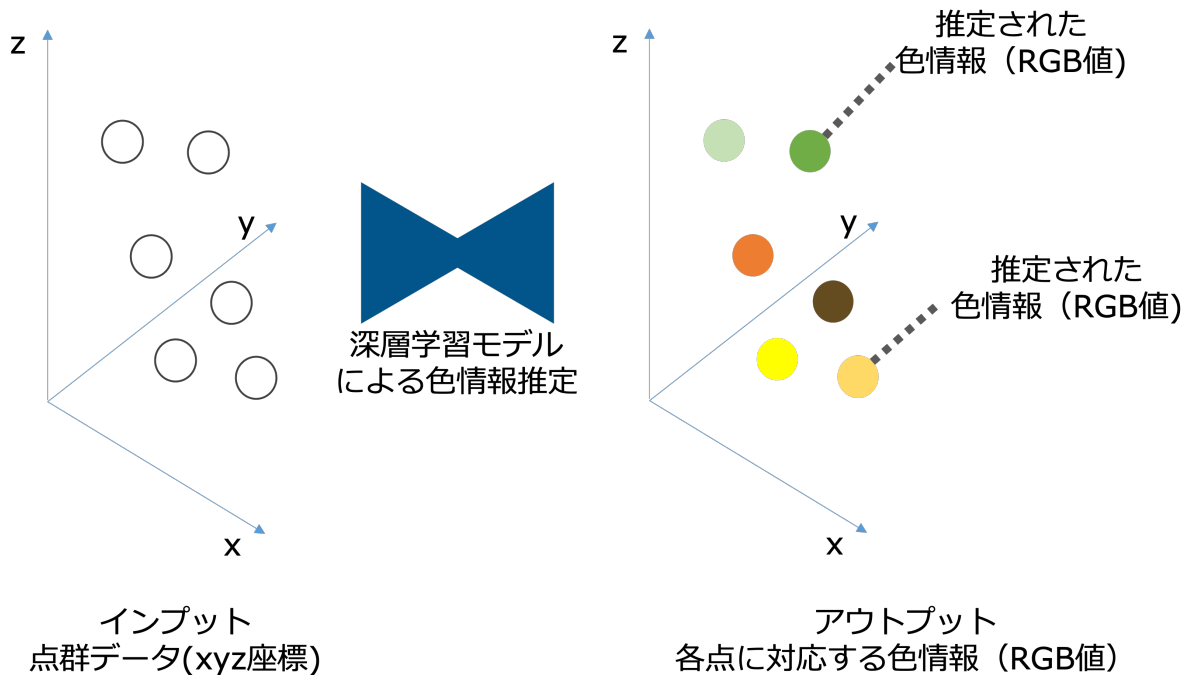


図 C.3: 点群着色手法の概念図 点群データにおける各点に対応する色情報を推定する深層学習モデルを用いる。

### C.1 点群着色手法の問題設定

まず、点群着色の問題設定について説明する。図 C.4に示すように、点群着色を行う深層学習モデルは、点群データ  $\mathbf{P} \in \mathbb{R}^{N \times 3}$  を入力とし、その点群データの点単位に推定された色情報  $\mathbf{V}_{\text{pred}} \in \mathbb{R}^{N \times 3}$  を出力する。具体的には、入力された点群データ  $\mathbf{P}$  に対応する推定した色  $\mathbf{V}_{\text{pred}}$  を、生成器 ( $g$ ) を用いて獲得する。ここで、各点  $\mathbf{P}$  は幾何学的情報 ( $x, y, z$  座標) を持ち、 $\mathbf{V}_{\text{pred}}$  は色情報 (RGB 値) を持つ。

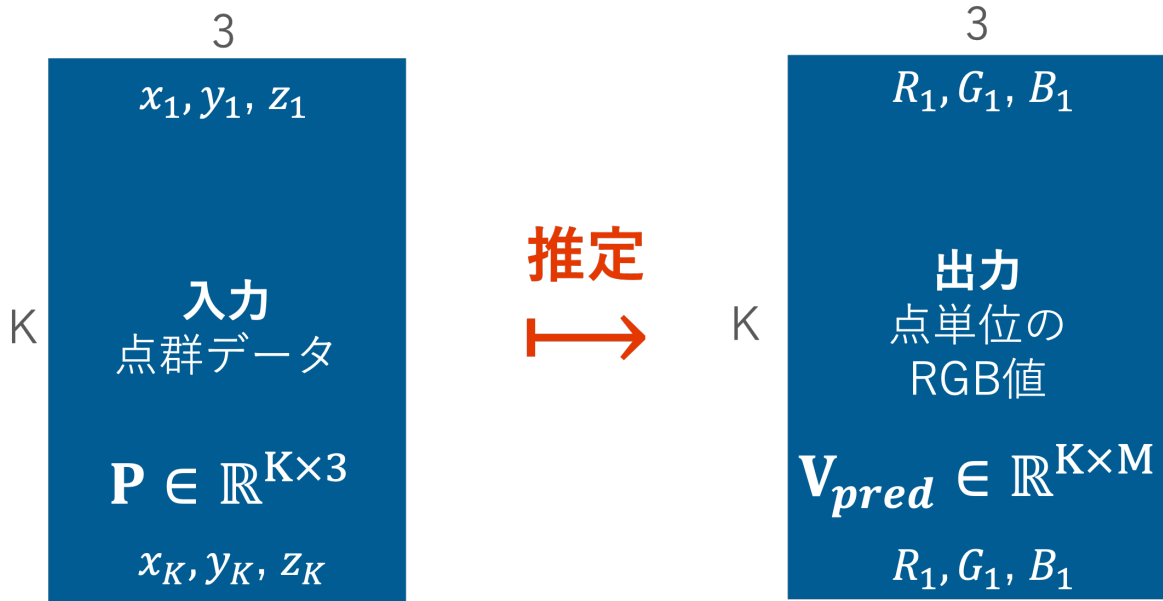


図 C.4: 点群着色手法の問題設定 点群データ  $P \in \mathbb{R}^{K \times 3}$  を入力とし、その点群データの点単位に推定された色情報  $V_{pred} \in \mathbb{R}^{K \times 3}$  を出力する。

## C.2 改良した点群着色手法

本節では、航空機に搭載されたライダーで観測された点群データに対する着色手法 (点群データ着色モデル) を説明する。基本的には、CAD データに対する条件付き GAN による点群着色手法 [131, 21] を参考にする。これらの点群着色手法では、生成器で入力された点群データの色を推定し、識別器で推定された色の評価を行う。生成器には点群データを扱うことが可能な深層学習モデルである PointNet の出力層である分類部分を色推定に変更した層設計を使用している。また、生成器によって推定された色なのか、真値の色なのかを分類する識別器は PointNet の分類機能を使用している。これらの 2 つのネットワークを用いることで、点群着色の結果が直接的な色情報の回帰による最小化だけでなく、識別器によって全体を通した整合性のような「人間の知覚的な評価」が可能になる。

しかし、先行研究 [131, 21] は簡単な CAD データを対象にしているため、それに応じて深層学習モデルも単純な PointNet を参考にしている。PointNet では、本研究で使用する航空機に搭載されたライダーで観測した点群データに対しては不向きである。この理由は、PointNet では、入力された点群データに対して点単位に独立して特徴量を抽出し、その後、入力点群データ全体の特徴量を計算するため、点と点の近傍関係を考慮できていない。また、画像で使用される畳み込みニューラルネットワークのような近傍のピクセルの情報を階層的に特徴抽出する処理が行えないことが挙げられる。そこで、点群着色を担う生成器の層設計を PointNet から、近傍の情報を階層的に特徴抽出可能な PointNet++ 構造に変更した。また、識別器も同様に PointNet++ 構造に変更した。このような手法を点群データ着色モデルと呼ぶ。以降で、提案する点群データ着色モデルの具体的な層設計を説明するが、基本的な構造は、前章の大局ブロックと同じである。違いは、入力が点群・波形データではなく点群データであること、出力が各点に対する地物情報ではなく色情報であること、そして、入力に波形データが無いため波形データに対する畳み込み機能を削除し単なる点群データの特徴抽出を全結合層で構成したことである。以降で詳しく点群データ着色モデルについて説明する。

まず、層設計について説明する。提案する点群データ着色モデルでは、CAD データに対する点群着色手法を参考に条件付き GAN の考え方を用いる。条件付き GAN を用いた点群着色では、着色を担う生成器と、着色された点群データか真値の色を持つ点群データかを分類する識別器で成り立つ。以降で、これらの生成器と識別器の説明を行う。

- 生成器の層設計

まず，生成器の層設計について説明する．生成器 ( $\mathcal{G}$ ) は，点群データ ( $\mathbf{P}$ ) から RGB の 3 次元な色情報 ( $\mathbf{V}_{\text{pred}}$ ) を推定するもので，この生成器による色の推定を写像  $\mathcal{G} : \mathbb{R}^{K \times 3} \rightarrow \mathbb{R}^{K \times 3}$  として定義できる．なお， $K$  は生成器に入力する点群データ数であり，これは計算機のメモリ量に依存する．

生成器は PointNet++ [178] を基本的な構造として，ダウンサンプリングと近傍特徴抽出で構成されるエンコーダと，アップサンプリング層で構成されるデコーダを備えた構造である (図 C.5 参照)．なお，この PointNet++ 基本的な構造は，前章の大局ブロックと同じであるため，代表点の算出やグルーピング等の基本的な演算の説明は省略する．以降でそれぞれエンコーダとデコーダについて説明する．

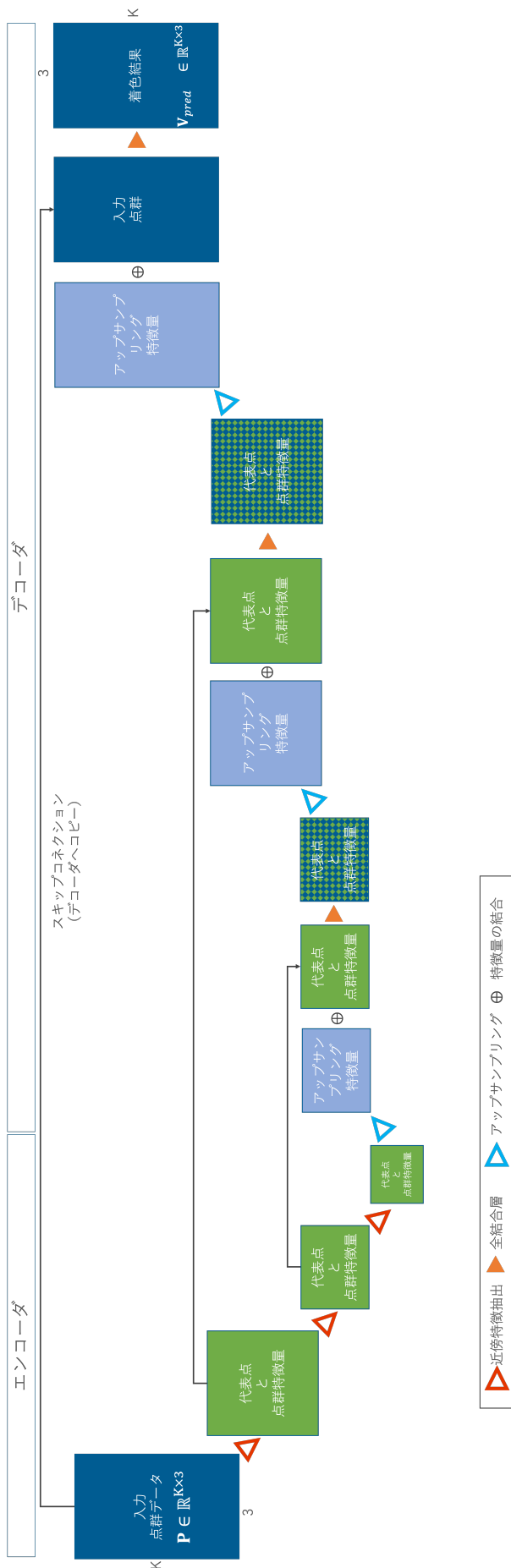


図 C.5: 点群着色手法における生成器の層設計. 第3章における大局ブロックと同じ考え方を用い, エンコーダによって入力された点群データから階層的に各点の特徴を抽出, デコーダによって抽出された特徴を元の点数に復元し点の色情報を推定する. なお, 波形データではないため, 近傍学習では単純な全結合層を用いる.

まず、エンコーダについて説明する。エンコーダは、点群データ ( $\mathbf{P}$ ) を入力し、3つのダウンサンプリング層と活性化関数としてパラメトリック ReLU [74] を持つ。エンコーダの各ダウンサンプリング層は、浅い層から深い層へ向けて [8,192, 4,096, 2,048] 個の点数になるように徐々に疎にサブサンプリングする。各ダウンサンプリングの際には、単に点数を落とすのではなく、近隣点の情報集約も同時に行う。情報の集約は、ダウンサンプリングによって得られる点を代表点もしくはクエリ点として、元の点数においてその近隣の点の情報を集約する。これを点群データの近傍特徴抽出と呼ぶ。この点群データの近傍特徴抽出を行う回数はダウンサンプリング回数に依存し、3回である。この畳み込み演算によって作成される特徴量の数は、浅い層から深い層へ向けて [256, 512, 1,024] 個である。また畳み込みの後にはバッチ正規化層とパラメトリック ReLU が続き、次の層への入力になる。

次に、デコーダを説明する。エンコーダからの特徴量を、3つのアップサンプリング層を持つデコーダに入力する。デコーダの構成要素は、ダウンサンプリングで少なくなった点群データ数を増やすための3つのアップサンプリング層である。アップサンプリング層は、3次元内挿と全結合層で構成されている。FC層では、深い層から浅い層へ向けて [1024, 512, 256] の特徴マップを持ち、活性化関数はパラメトリック ReLU を用いた。FC層では、エンコーダとデコーダの間にスキップコネクションを使用している。スキップコネクションを入れることで、エンコーダで失った高い密度の点群データの情報をデコーダ側へ直接伝えることが可能になる。デコーダの最後の層では、点群データの点数と同じ次元数の出力を行う。その際には、非線形な活性化関数を用いずに、線形に出力する。以上のことから、生成器  $\mathcal{G}$  は、点群データを入力して色情報を推定する非線形な関数 ( $\mathbf{V}_{\text{pred}} = \mathcal{G}(\mathbf{P})$ ) と言える。

- 識別器の層設計

次に、識別器の説明を行う。図 C.6に識別器  $\mathcal{D}$  の層設計を示す。 $\mathcal{D}$  は、写像 ( $\mathcal{D} : \mathbb{R}^{K \times (3+3)} \rightarrow \mathbb{R}^{2048 \times 1}$ ) として定義できる。つまり、 $\mathcal{D}$  は、推定した点群データとその色情報を  $\mathbf{V}_{\text{pred}}$  with  $\mathbf{P}$ 、真値の色を  $\mathbf{V}_{\text{real}}$  with  $\mathbf{P}$  を分類する。なお、 $\mathbf{P}$  は点群データを示し、座標を用いることで空間的な学習を実現するために使用する。そして、層設計としては、第3章で説明した PointNet++ [178] を参考にした層設計 (大局ブロック) のエンコーダ部分と同様である。識別器の機能は、推定した色否かを、点群データと色情報を用いて分類する。識別器のダウンサンプリング層は3層で、特徴マップは256から1,024へと2倍ずつ増加させる。それぞれのダウンサンプリングの際に畳み込み層を設置して、畳み込み層の後に挿入される活性化関数として leaky ReLU を用いる。なお、leaky ReLU のハイパーパラメータは多くの論文で使用されるデフォルト値である0.2に設定した。点数はそれぞれ浅いものから深いもの順で [8,192, 4,096, 2,048] とした。識別器の出力は、 $2,048 \times 1$  の実数の多次元配列で表現され、FC層の後にシグモイド関数で真値の色である確率を計算する。

層設計の説明の次に、点群データ着色モデルの最適化を行う際に使用する損失関数について説明する。点群データ着色モデルは条件付き GAN を用いているため、生成器と識別器の2つのニューラルネットワークが存在するため、それぞれに対応する損失関数を定義する必要がある。

- 生成器の損失関数

まずは、生成器の損失関数を説明する。生成器の損失関数では、生成器が推定した色と真値の色の差を直接回帰するための項と、識別器が真値の色と見間違えるように仕掛ける項の2種類存在する。一つ目の回帰の項は以下のようにになっている。

$$\mathcal{L}_G^{\text{content}} = \frac{1}{K} \frac{1}{3} \sum_{i=1}^K \sum_{j=1}^3 |\mathbf{V}_{\text{real},i,j} - \mathbf{V}_{\text{pred},i,j}|. \quad (\text{C.4})$$

この式において、また、推定した色と真値の色の強度が完全に合う場合において、この項は0になる。

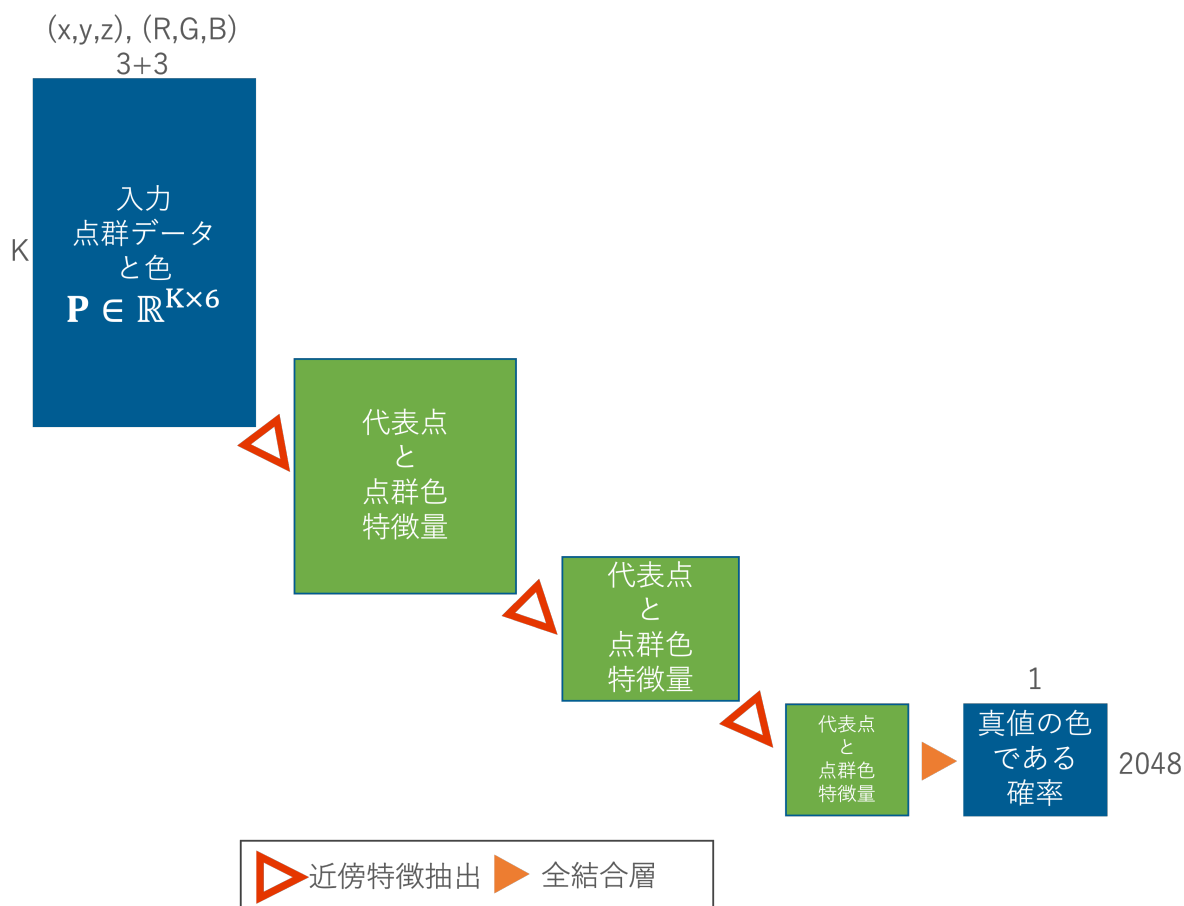


図 C.6: 点群着色手法における識別器の層設計生成器と同様に、第 3 章の大局モデルと同様の考え方をを用いる。層設計としては生成器のエンコーダ部分のみと同じであるが、出力がダウンサンプリングされた 2048 次元になり、この次元数の分だけ真値の色である確率を出力する。

また、後者の項では、識別器が算出した真値の色である確率を用いて、以下のような式で表現する。

$$\mathcal{L}_G^{\text{adv}} = \mathbb{E}[\log(1 - \mathcal{D}(\mathcal{G}(\mathbf{P})))] \quad (\text{C.5})$$

これは、真値の色である確率が 1 の時に最小になるような設定である。

- 識別器の損失関数

次に、識別器の損失関数を説明する。識別器の学習には、識別器に真値の色が入力された際の項と、生成器によって推定した色が入力された際の項の 2 つ存在する。これらをまとめて表現すると以下のような式になる。

$$\mathcal{L}_D^{\text{adv}} = \mathbb{E}[\log \mathcal{D}(\mathbf{V}_{\text{pred}})] + \mathbb{E}[\log(1 - \mathcal{D}(\mathcal{G}(\mathbf{P})))] \quad (\text{C.6})$$

まず、右辺の 1 項では、真値の色が入力された場合の項を示している。この場合は、識別器が算出した真値の色である確率が 1 になる場合に最小化される。右辺の 2 項は推定した色が入力された際の項である。この際には、識別器が算出した真値の色である確率が 1 になる場合に最小化される。逆に確率が 0 つまり偽物と判断した場合には最大になる。

- 全体の損失関数

実際の学習時には、上記の生成器と識別器の 2 つの損失関数を同時に最適化する。

$$\mathcal{L}^{\text{total}}(\mathcal{D}^{\text{adv}}, \mathcal{G}) = \mathcal{L}_G^{\text{adv}} + \lambda \mathcal{L}_G^{\text{content}} + \mathcal{L}_D^{\text{Adv}}. \quad (\text{C.7})$$

この損失関数を用いた最適化を行う際には、生成器と識別器において非協力な Min-Max 最適化戦略が用いられる。

$$\min_G \max_D \mathcal{L}^{\text{total}}(\mathcal{D}, \mathcal{G}). \quad (\text{C.8})$$

この際には、生成器は損失関数を最小になるように学習を行い、その反対に識別器は損失関数を最大化するように学習を行う。

### C.3 実験結果

提案する点群着色手法が実際に航空機に搭載されたライダー点群データで取得された点群データに適用可能かを調査するために、真値の色情報を持つ点群データを用いて検証実験を行った。まず、実験に使用するデータセットを説明し、その後、実験結果について説明する。

点群着色の実験を行うために、航空機に搭載されたライダーで観測された点群データを用いて提案するモデルの学習と、学習済みモデルの評価を行う。学習と評価に使用する着色データセットとしては、2018 IEEE GRSS Data Fusion Contest (DFC2018) [72] と呼ばれる屋外で取得された点群データと航空写真のペアを用いた (図 C.7)。DFC2018 には、家屋、地面、樹木、道路、アメリカンフットボール場など、多種多様な地物が含まれている (図 C.7 (a))。DFC2018 (図 C.7 (b)) は、米国の国立航空宇宙センター (NASA) がカメラ一体型の Optech Titan MW を用いて取得したもので、密度は約 5 点/m<sup>2</sup> である。DFC2018 には点群データと同じ地域を撮影された航空写真が収録されているため、航空写真を用いて点群データに色を付け、実際の真値の色を作成した。この際に、PDAL [32] を使用して、航空写真から各点の実際の色を作成した。さらに、真値の点群データの色を 0-1 に正規化した。DFC2018 の元データは、数百メートル単位で区切られている。そのため、計算機のメモリの上限を超えるため、そのままでは深層学習モデルに入力することができない。そこで本研究では、色をつけた点群データを 30 m<sup>2</sup> に分割して、小領域のパッチを作成して現状の GPU のスペックで計算可能にした。なお、中心部の 25 m<sup>2</sup> を損失関数の計算の対象とし、その外側は単に深層学習が着色を行うためのコンテキスト情報として使用した。また、点群データには地中の点や空中に浮かぶ独立点などのノイズが含まれているため、孤立した点をノイズとして除去し、Open3D [250] を用いて各パッチの点を 2 万個にサブサンプリングした。また、パッチの  $xyz$  座標を [-0.5, 0.5] に正規化した。モデルの学習と評価のために、すべてのパッチデータを 2 つのデータセットに分けた。すなわち、モデルの学習に用いた訓練データと、訓練済みモデルの評価に用いたテストデータである。なお、図 C.7 の青い四角で示したのがテスト領域である。また、深層学習モデルの学習時には訓練データをさらに分割し、5 倍のクロスバリデーションを行った。

ここでは、学習済みモデルを用いた点群着色手法の検証実験を行った結果を述べる。比較対象として CAD データに対する点群着色手法 [131, 21] を用いた。

まず、点群着色結果を定量的に評価する。性能を測定するために、0-1 に正規化された RGB 空間での平均絶対誤差 (Mean Absolute Error, MAE) を採用することにした。着色データセットにおけるテストデータ全体の推定した色と真値の色の MAE は、各色チャンネルの点単位で色の絶対誤差の平均をとることで算出した。点群データ着色モデルの学習結果では、テストデータの MAE は 0.10 であった。次に、PointNet ベースの条件付き GAN を用いた従来の点群着色手法 [131, 21] と比較した。PointNet を用いた Dense-Point [21] では、テストデータにおいて MAE が 0.25 であった。また、同じような PointNet を用いた PCCN [131] では、テストデータにおける MAE が 0.22 であった。これらの PointNet をベースにした先行研究と比較して、点群データ着色モデルは低い MAE を達成した。これは、単純な 3 次元物体を対象とした先行研究の手法では、航空機に搭載されたライダーで観測された複雑な物体に対応できないためである。特に、最大の問題は、これらの手法では PointNet が色付けのための生成器として使用されているため、近傍の特徴と大局的な特徴を効果的に組み合わせられないことである。



図 C.7: 航空写真と航空機に搭載されたライダーで観測された点群データによる着色データセット 2018 IEEE GRSS Data Fusion Contest (DFC2018) [72] の点群データと航空写真が同時に観測されたデータを用いる。青い四角で示した箇所がテスト領域である。

次に、定性評価結果を述べる。点群着色データセットを用いた代表的な結果の一部を図 C.8 に示す。点群データ着色モデルで推定した着色結果（図 C.8 (d)）とそのレンダリング画像（図 C.8 (f)）には、鮮やかな色が含まれている。一般的に、着色された点群データは、真値の色をほぼ再現する傾向がある（図 C.8 (e)）。また、点群データでは違いがわかりにくいので、上空から見た画像相当にレンダリングして比較する。レンダリング画像も、真値のレンダリング画像をほぼ再現している（図 C.8 (g)）。図 C.8 の下図に示すように、空白の入力でも着色が可能である。また、点群データ着色モデルは、同じような低い場所かつ平坦性を持つ草と道路を区別して着色している。このように、点群データ着色モデルは、目視判読性を向上させ、対象物の理解を容易にすることが期待される。点群データ着色モデルは、全体的には入力された点群データの色を比較的妥当に推定したが、図 C.8 に丸で示したように、いくつかの失敗例があった。図 C.8 の丸印のように、車の色が推定できなかった。これは、点群データの深層学習手法が学習データの量に敏感であるため、点数が少ない自動車のような小さな物体は無視されてしまう傾向があるからである。また、点群データ着色モデルでは、データセットに頻出する色を着色する傾向があるのも欠点であった。例えば、図 C.8 の青丸で示した裸地は、草のような緑に着色されていた。これは、裸地の点数が草地の点群データ数よりも圧倒的に多かったためと考えられる。今後は、データセットにおいて色のバリエーションを増やしたり、小さな物体に対応したりしながら、より尤もらしい色を推定するための新しい手法を開発する必要がある。

総じて、屋外で取得された航空機に搭載されたライダー点群データに対しては、点群データ着色モデルのように PointNet++ をベースとする方が PointNet を元にするよりも、3次元的な構造を把握することができ、より真値の色に近い着色結果が得られることがわかった。また条件付き GAN の仕組みも点群データに対しても有効であることが判明した。

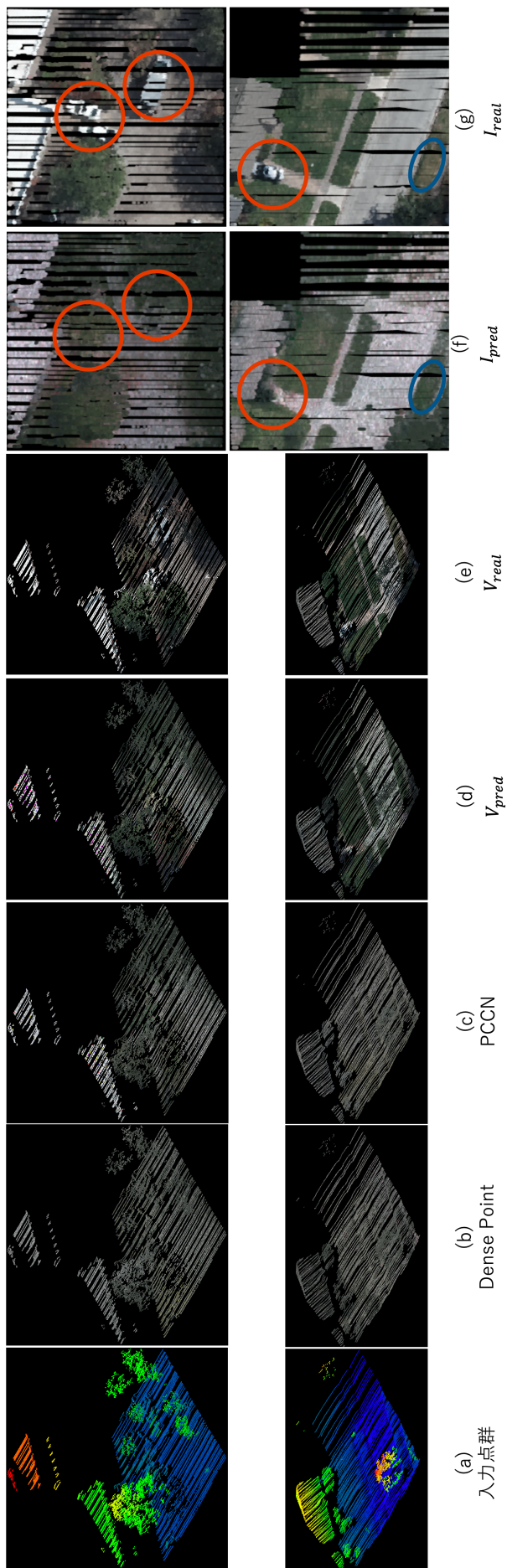


図 C.8: 点群着色結果 (a) 入力点群データ, (b) Dense-Point [21], (c) PCCN [131], (d) 点群データ着色モデル, (e) 真値の色, (f) 点群データ着色モデルの結果を画像にレンダリングした結果, (g) 真値の色を画像にレンダリングした結果

## 参考文献

- [1] *3DDB Model Viewer*. [https://gsrt.digiarc.aist.go.jp/3ddb\\_demo/tdv/index.html](https://gsrt.digiarc.aist.go.jp/3ddb_demo/tdv/index.html). (Accessed on 11/03/2021).
- [2] 3Dビジュアライゼーション実証プロジェクト / 東京都デジタルサービス局. <https://www.digitalservice.metro.tokyo.lg.jp/society5.0/digitaltwin.html>. (Accessed on 11/01/2021).
- [3] Eman Ahmed et al. “Deep Learning Advances on Different 3D Data Representations: A Survey”. In: *arXiv preprint arXiv:1808.01462* (2018).
- [4] Cici Alexander et al. “Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 65.5 (2010), pp. 423–432.
- [5] Hans-Erik Andersen, Robert J McGaughey, and Stephen E Reutebuch. “Estimating forest canopy fuel parameters using LIDAR data”. In: *Remote sensing of Environment* 94.4 (2005), pp. 441–449.
- [6] Iro Armeni et al. “3D semantic parsing of large-scale indoor spaces”. In: *CVPR*. 2016.
- [7] *ASPRSorg/LAS: LAS Specification*. <https://github.com/ASPRSorg/LAS>. (Accessed on 11/10/2021).
- [8] Andreas Aßmann, Brian Stewart, and Andrew M. Wallace. “Deep Learning for LiDAR Waveforms with Multiple Returns”. In: *2020 28th European Signal Processing Conference (EUSIPCO)*. 2021, pp. 1571–1575. DOI: 10.23919/Eusipco47968.2020.9287545.
- [9] Peter Axelsson. “DEM generation from laser scanner data using adaptive TIN models”. In: *International archives of photogrammetry and remote sensing* 33.4 (2000), pp. 110–117.
- [10] Mohsen Azadbakht, Clive Fraser, and Kouros Khoshelham. “The role of full-waveform LiDAR features in improving urban scene classification”. In: *Proceedings of the 9th International Symposium on Mobile Mapping Technology (MMT 2015), Sydney, Australia*. 2015, pp. 9–11.
- [11] Mohsen Azadbakht, Clive S Fraser, and Kouros Khoshelham. “Synergy of sampling techniques and ensemble classifiers for classification of urban environments using full-waveform LiDAR data”. In: *International journal of applied earth observation and geoinformation* 73 (2018), pp. 277–291.
- [12] Emanuel P Baltsavias. “Airborne laser scanning: existing systems and firms and other resources”. In: *ISPRS Journal of Photogrammetry and Remote sensing* 54.2-3 (1999), pp. 164–198.
- [13] Emmanuel P Baltsavias. “Airborne laser scanning: basic relations and formulas”. In: *ISPRS Journal of photogrammetry and remote sensing* 54.2-3 (1999), pp. 199–214.
- [14] Jens Behley et al. “SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences”. In: *ICCV*. 2019.

- [15] M Brédif. “3D Building Modeling: Topology-Aware Kinetic Fitting of Polyhedral Roofs and Automatic Roof Superstructure Reconstruction”. In: *Télécom ParisTech, France* (2010).
- [16] A Brenner. “Derivation of Range and Range Distributions From Laser Pulse Waveform Analysis for Surface Elevations, Roughness, Slope, and Vegetation Heights. Algorithm Theoretical Basis Document V4. 1”. In: [http://www.csr.utexas.edu/glas/pdf/Atbd\\_20031224.pdf](http://www.csr.utexas.edu/glas/pdf/Atbd_20031224.pdf) (2003).
- [17] Frédéric Bretar and Nesrine Chehata. “Terrain modeling from lidar range data in natural landscapes: A predictive and bayesian framework”. In: *IEEE Transactions on Geoscience and Remote Sensing* 48.3 (2009), pp. 1568–1578.
- [18] S Briechle, P Krzystek, and G Vosselman. “SEMANTIC LABELING OF ALS POINT CLOUDS FOR TREE SPECIES MAPPING USING THE DEEP NEURAL NETWORK POINTNET++.” In: *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* (2019).
- [19] Eoin Brophy, Zhengwei Wang, and Tomas E Ward. “Quick and Easy Time Series Generation with Established Image-based GANs”. In: *arXiv preprint arXiv:1902.05624* (2019).
- [20] Holger Caesar et al. “nuScenes: A Multimodal Dataset for Autonomous Driving”. In: *CVPR*. 2020.
- [21] Xu Cao and Katashi Nagao. “Point cloud colorization based on densely annotated 3d shape dataset”. In: *International Conference on Multimedia Modeling*. Springer. 2019, pp. 436–446.
- [22] Tomas Carlsson, Ove Steinvall, and Dietmar Letalick. “Signature simulation and signal analysis for 3-D laser radar”. In: *Sweden: Swedish Defence Research Agency* (2001), pp. 7–8.
- [23] Angel X Chang et al. “ShapeNet: An Information-Rich 3D Model Repository”. In: *arXiv preprint arXiv:1512.03012* (2015).
- [24] Ming-Fang Chang et al. “Argoverse: 3D tracking and forecasting with rich maps”. In: *CVPR*. 2019.
- [25] Amin P Charaniya, Roberto Manduchi, and Suresh K Lodha. “Supervised parametric classification of aerial lidar data”. In: *2004 Conference on Computer Vision and Pattern Recognition Workshop*. IEEE. 2004, pp. 30–30.
- [26] Adrien Chauve et al. “FullAnalyze: A Research tool for handling, processing and analyzing full-waveform lidar data”. In: *2009 IEEE International Geoscience and Remote Sensing Symposium*. Vol. 4. IEEE. 2009, pp. IV–841.
- [27] Adrien Chauve et al. “Processing full-waveform lidar data in an alpine coniferous forest: Assessing terrain and tree height quality”. In: *International journal of remote sensing* 30.19 (2009), pp. 5211–5228.
- [28] Yang Chen et al. “PointNet++ Network Architecture with Individual Point Level and Global Features on Centroid for ALS Point Cloud Classification”. In: *Remote Sensing* 13.3 (2021), p. 472.
- [29] Zeyuan Chen et al. “High resolution face completion with multiple controllable attributes via fully end-to-end progressive generative adversarial networks”. In: *arXiv preprint arXiv:1801.07632* (2018).
- [30] Junsuk Choe et al. “Face generation for low-shot learning using generative adversarial networks”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, pp. 1940–1948.

- [31] Yunjey Choi et al. “StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 8789–8797.
- [32] PDAL Contributors. *PDAL Point Data Abstraction Library*. Nov. 2018. DOI: 10.5281/zenodo.2556738. URL: <https://doi.org/10.5281/zenodo.2556738>.
- [33] Marius Cordts et al. “The cityscapes dataset for semantic urban scene understanding”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 3213–3223.
- [34] Angela Dai et al. “ScanNet: Richly-annotated 3D reconstructions of indoor scenes”. In: *CVPR*. 2017.
- [35] Quanyu Dai et al. “Adversarial network embedding”. In: *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.
- [36] Zihang Dai et al. “Good semi-supervised learning that requires a bad GAN”. In: *Advances in neural information processing systems*. 2017, pp. 6510–6520.
- [37] Mark De Deuge et al. “Unsupervised Feature Learning for Classification of Outdoor 3D Scans”. In: *ACRA*. 2013.
- [38] Arthur P Dempster, Nan M Laird, and Donald B Rubin. “Maximum likelihood from incomplete data via the EM algorithm”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 39.1 (1977), pp. 1–22.
- [39] Sandor Der, Brian Redman, and Rama Chellappa. “Simulation of error in optical radar range measurements”. In: *Applied Optics* 36.27 (1997), pp. 6869–6874.
- [40] Jean-Luc Déziel et al. “PixSet: An Opportunity for 3D Computer Vision to Go Beyond Point Clouds With a Full-Waveform LiDAR Dataset”. In: *arXiv preprint arXiv:2102.12010* (2021).
- [41] Carl Doersch. “Tutorial on variational autoencoders”. In: *arXiv preprint arXiv:1606.05908* (2016).
- [42] Brian Dolhansky and Cristian Canton Ferrer. “Eye in-painting with exemplar generative adversarial networks”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 7902–7911.
- [43] Chris Donahue, Julian McAuley, and Miller Puckette. “Synthesizing Audio with Generative Adversarial Networks”. In: *arXiv preprint arXiv:1802.04208* (2018).
- [44] Michael Doneus et al. “Archaeological prospection of forested areas using full-waveform airborne laser scanning”. In: *Journal of Archaeological Science* 35.4 (2008), pp. 882–893.
- [45] Hao Dong et al. “Semantic image synthesis via adversarial learning”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, pp. 5706–5714.
- [46] Ralph O Dubayah and Jason B Drake. “Lidar remote sensing for forestry”. In: *Journal of forestry* 98.6 (2000), pp. 44–46.
- [47] Vesna Ducic et al. *3D vegetation mapping and classification using full-waveform laser scanning*. na, 2006.
- [48] Hieu Duong, Norbert Pfeifer, and Roderik Lindenbergh. “Full waveform analysis: ICESat laser data for land cover classification”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.Part 7 (2006), pp. 30–35.
- [49] VH Duong et al. “Single and two epoch analysis of ICESat full waveform data over forested areas”. In: *International Journal of Remote Sensing* 29.5 (2008), pp. 1453–1473.

- [50] Gintare Karolina Dziugaite, Daniel M Roy, and Zoubin Ghahramani. “Training generative neural networks via maximum mean discrepancy optimization”. In: *arXiv preprint arXiv:1505.03906* (2015).
- [51] Gil Elbaz, Tamar Avraham, and Anath Fischer. “3D point cloud registration for localization using a deep neural network auto-encoder”. In: *CVPR*. 2017, pp. 4631–4640.
- [52] Francis Engelmann et al. “Know what your neighbors do: 3D semantic segmentation of point clouds”. In: *ECCV*. 2018.
- [53] Cristóbal Esteban, Stephanie L Hyland, and Gunnar Rätsch. “Real-valued (medical) time series generation with recurrent conditional GANs”. In: *arXiv preprint arXiv:1706.02633* (2017).
- [54] Mathieu Fauvel. “Spectral and spatial methods for the classification of urban remote sensing data”. PhD thesis. Institut National Polytechnique de Grenoble-INPG; Université d’Islande, 2007.
- [55] William Fedus, Ian Goodfellow, and Andrew M Dai. “MaskGAN: Better Text Generation via Filling in the \_.” In: *arXiv preprint arXiv:1801.07736* (2018).
- [56] Artur Fidera, Michael A Chapman, and Jingook Hong. “Terrestrial lidar for industrial metrology applications: modelling, enhancement and reconstruction”. In: *XXth ISPRS congress*. Vol. 5. Citeseer. 2004, pp. 880–886.
- [57] Karolina D Fieber et al. “Analysis of full-waveform LiDAR data for classification of an orange orchard scene”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 82 (2013), pp. 63–82.
- [58] Mário AT Figueiredo and Robert D Nowak. “Wavelet-based image estimation: An empirical Bayes approach using Jeffrey’s noninformative prior”. In: *IEEE Transactions on Image Processing* 10.9 (2001), pp. 1322–1331.
- [59] Michele Fumarola and Ronald Poelman. “Generating virtual environments of real world facilities: Discussing four different approaches”. In: *Automation in Construction* 20.3 (2011). Augmented and Virtual Reality in Architecture, Engineering and Construction (CONVR2009), pp. 263–269. ISSN: 0926-5805. DOI: <https://doi.org/10.1016/j.autcon.2010.08.004>. URL: <https://www.sciencedirect.com/science/article/pii/S0926580510002074>.
- [60] Andreas Geiger, Philip Lenz, and Raquel Urtasun. “Are we ready for autonomous driving”. In: *CVPR*. 2012.
- [61] Ian Goodfellow. “NIPS 2016 tutorial: Generative adversarial networks”. In: *arXiv preprint arXiv:1701.00160* (2016).
- [62] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.
- [63] Ian Goodfellow et al. “Generative adversarial nets”. In: *Advances in Neural Information Processing Systems*. 2014, pp. 2672–2680.
- [64] H Gross, B Jutzi, and U Thoennessen. “Segmentation of tree regions using data of a full-waveform laser”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.part 3 (2007), W49A.
- [65] G Guenther and H Mesick. “Analysis of airborne lidar bathymetric waveforms”. In: *Proc. of the 9th Ocean Optics. Orlando, FA, USA, SPIE* (1988), pp. 4–6.
- [66] Gary C Guenther et al. *Meeting the accuracy challenge in airborne bathymetry*. Tech. rep. NATIONAL OCEANIC ATMOSPHERIC ADMINISTRATION/NESDIS SILVER SPRING MD, 2000.

- [67] Li Guo et al. “Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 66.1 (2011), pp. 56–66.
- [68] Yunhui Guo. “A survey on methods and theories of quantized neural networks”. In: *arXiv preprint arXiv:1808.04752* (2018).
- [69] G空間情報センター. [https://www.geospatial.jp/gp\\_front/](https://www.geospatial.jp/gp_front/). (Accessed on 10/10/2021).
- [70] T Hackel et al. “Semantic3D.net: A new large-scale point cloud classification benchmark”. In: *ISPRS* (2017).
- [71] Xianfeng Han, Hamid Laga, and Mohammed Bennamoun. “Image-based 3D Object Reconstruction: State-of-the-Art and Trends in the Deep Learning Era”. In: *IEEE TPAMI* (2019).
- [72] Saurabh Prasad; Bertrand Le Saux; Naoto Yokoya; Ronny Hansch. “2018 IEEE GRSS Data Fusion Challenge -Fusion of Multispectral LiDAR and Hyperspectral Data”. In: (2020). DOI: 10.21227/jnh9-nz89. URL: <https://dx.doi.org/10.21227/jnh9-nz89>.
- [73] Kay Gregor Hartmann, Robin Tibor Schirrmeyer, and Tonio Ball. “EEG-GAN: Generative adversarial networks for electroencephalographic brain signals”. In: *arXiv preprint arXiv:1806.01875* (2018).
- [74] Kaiming He et al. “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 1026–1034.
- [75] Christian Heipke, Karsten Jacobsen, and Helge Wegmann. “Analysis of the results of the OEEPE test “Integrated Sensor Orientation”. In: *OEEPE Integrated Sensor Orientation Test Report and Workshop Proceedings, Editors*. Citeseer. 2002.
- [76] Bernhard Höfle and Markus Hollaus. *Urban vegetation detection using high density full-waveform airborne lidar data-combination of object-based image and point cloud analysis*. na, 2010.
- [77] Bernhard Höfle, Markus Hollaus, and Julian Hagenauer. “Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 67 (2012), pp. 134–147.
- [78] Bernhard Höfle and Norbert Pfeifer. “Correction of laser scanning intensity data: Data and model-driven approaches”. In: *ISPRS journal of photogrammetry and remote sensing* 62.6 (2007), pp. 415–433.
- [79] Michelle A Hofton, Jean-Bernard Minster, and J Bryan Blair. “Decomposition of laser altimeter waveforms”. In: *IEEE Transactions on geoscience and remote sensing* 38.4 (2000), pp. 1989–1996.
- [80] Markus Hollaus and Bernhard Höfle. *Terrain roughness parameters from full-waveform airborne LIDAR data*. na, 2010.
- [81] Qingyong Hu et al. “RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds”. In: *CVPR* (2020).
- [82] C Hug, A Ullrich, and A Grimm. “Litmapper-5600-a waveform-digitizing LiDAR terrain and vegetation mapping system”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.Part 8 (2004), W2.
- [83] Jean-François Hullo et al. “Multi-Sensor As-Built Models of Complex Industrial Architectures”. In: *Remote Sensing* 7.12 (2015), pp. 16339–16362. ISSN: 2072-4292. DOI: 10.3390/rs71215827. URL: <https://www.mdpi.com/2072-4292/7/12/15827>.
- [84] Anastasia Ioannidou et al. “Deep Learning Advances in Computer Vision with 3D Data: A Survey”. In: *ACM Computing Surveys* (2017).

- [85] Jennifer L Irish and W Jeff Lillycrop. “Scanning laser mapping of the coastal zone: the SHOALS system”. In: *ISPRS Journal of Photogrammetry and remote sensing* 54.2-3 (1999), pp. 123–129.
- [86] ISO - ISO 23629-7:2021 - UAS traffic management (UTM) —Part 7: Data model for spatial data. <https://www.iso.org/standard/76973.html>. (Accessed on 10/24/2021).
- [87] Phillip Isola et al. “Image-to-image translation with conditional adversarial networks”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1125–1134.
- [88] Zwally H Jay. “ICESat’s Laser Measurements of Polar Ice, Atmosphere, Ocean, and Land”. In: (2003).
- [89] Albert V Jelalian. “Laser radar systems”. In: *EASCON’80; Electronics and Aerospace Systems Conference*. 1980, pp. 546–554.
- [90] Nikolay Jetchev, Urs Bergmann, and Roland Vollgraf. “Texture synthesis with spatial generative adversarial networks”. In: *arXiv preprint arXiv:1611.08207* (2016).
- [91] Mingyang Jiang, Yiran Wu, and Cewu Lu. “PointSIFT: A sift-like network module for 3D point cloud semantic segmentation”. In: *arXiv preprint arXiv:1807.00652* (2018).
- [92] B Jutzi, J Neulist, and U Stilla. “Sub-pixel edge localization based on laser waveform analysis”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.3/W19 (2005), pp. 109–114.
- [93] B Jutzi and U Stilla. “Laser pulse analysis for reconstruction and classification of urban objects”. In: *International archives of photogrammetry remote sensing and spatial information sciences* 34.3/W8 (2003), pp. 151–156.
- [94] Boris Jutzi, Bernd Eberle, and Uwe Stilla. “Estimation and measurement of backscattered signals from pulsed laser radar”. In: *Image and Signal Processing for Remote Sensing VIII*. Vol. 4885. International Society for Optics and Photonics. 2003, pp. 256–267.
- [95] Boris Jutzi and Uwe Stilla. “Range determination with waveform recording laser systems using a Wiener Filter”. In: *ISPRS Journal of Photogrammetry and Remote sensing* 61.2 (2006), pp. 95–107.
- [96] Yoonseok Jwa and Gunho Sohn. “A piecewise catenary curve model growing for 3D power line reconstruction”. In: *Photogrammetric Engineering & Remote Sensing* 78.12 (2012), pp. 1227–1240.
- [97] Tero Karras, Samuli Laine, and Timo Aila. “A Style-Based Generator Architecture for Generative Adversarial Networks”. In: *arXiv preprint arXiv:1812.04948* (2018).
- [98] R Kesten et al. *Lyft level 5 av dataset 2019*. 2019.
- [99] Asharul Islam Khan and Salim Al-Habsi. “Machine Learning in Computer Vision”. In: *Procedia Computer Science* 167 (2020). International Conference on Computational Intelligence and Data Science, pp. 1444–1451. ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2020.03.355>. URL: <https://www.sciencedirect.com/science/article/pii/S1877050920308218>.
- [100] Diederik P Kingma and Max Welling. “Auto-encoding variational Bayes”. In: *arXiv preprint arXiv:1312.6114* (2013).
- [101] Serkan Kiranyaz et al. “1D convolutional neural networks and applications: A survey”. In: *Mechanical Systems and Signal Processing* 151 (2021), p. 107398. ISSN: 0888-3270. DOI: <https://doi.org/10.1016/j.ymsp.2020.107398>. URL: <https://www.sciencedirect.com/science/article/pii/S0888327020307846>.

- [102] Michael Kirchhof, Boris Jutzi, and Uwe Stilla. “Iterative processing of laser scanning data by full waveform analysis”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 63.1 (2008), pp. 99–114.
- [103] Naveen Kodali et al. “On convergence and stability of GANs”. In: *arXiv preprint arXiv:1705.07215* (2017).
- [104] Jean Kossaifi et al. “GANGAN: Geometry-aware generative adversarial networks”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 878–887.
- [105] Karl Kraus and Norbert Pfeifer. “Determination of terrain models in wooded areas with airborne laser scanner data”. In: *ISPRS Journal of Photogrammetry and remote Sensing* 53.4 (1998), pp. 193–203.
- [106] Alex Krizhevsky and Geoffrey Hinton. *Learning multiple layers of features from tiny images*. Tech. rep. Citeseer, 2009.
- [107] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*. NIPS’12. Lake Tahoe, Nevada: Curran Associates Inc., 2012, pp. 1097–1105.
- [108] Antero Kukko, Sanna Kaasalainen, and Paula Litkey. “Effect of incidence angle on laser scanner intensity and surface data”. In: *Applied optics* 47.7 (2008), pp. 986–992.
- [109] S. Lacherade et al. “Spectral variability and bidirectional reflectance behaviour of urban materials at a 20 cm spatial resolution in the visible and near-infrared wavelengths. A case study over Toulouse (France)”. In: *International Journal of Remote Sensing* 26.17 (2005), pp. 3859–3866. DOI: 10.1080/01431160500177414. eprint: <https://doi.org/10.1080/01431160500177414>. URL: <https://doi.org/10.1080/01431160500177414>.
- [110] Debra F Laefer et al. “2015 aerial laser and photogrammetry survey of Dublin city collection record”. In: (2017). DOI: 10.1109/TGRS.2019.2919472. URL: <https://geo.nyu.edu/catalog/nyu-2451-38684/>.
- [111] Xudong Lai et al. “Full-waveform LiDAR point clouds classification based on wavelet support vector machine and ensemble learning”. In: *Sensors* 19.14 (2019), p. 3191.
- [112] S Laky, P Zaletnyik, and C Toth. “Land classification of wavelet-compressed full-waveform LiDAR data”. In: *International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences (ISSN: 1682-1750)* 38 (2010), 3A.
- [113] Gustav Larsson, Michael Maire, and Gregory Shakhnarovich. “Colorization as a proxy task for visual understanding”. In: *Proc. CVPR*. 2017, pp. 6874–6883.
- [114] Rosa Lasaponara and Nicola Masini. “Full-waveform Airborne Laser Scanning for the detection of medieval archaeological microtopographic relief”. In: *Journal of Cultural Heritage* 10 (2009), e78–e82.
- [115] Christoph Lassner, Gerard Pons-Moll, and Peter V Gehler. “A generative model of people in clothing”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, pp. 853–862.
- [116] Yann Lecun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. In: *Nature* 521.7553 (2015), pp. 436–444. ISSN: 14764687. DOI: 10.1038/nature14539.
- [117] Yann LeCun et al. “Gradient-based learning applied to document recognition”. In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324.
- [118] Christian Ledig et al. “Photo-realistic single image super-resolution using a generative adversarial network”. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 2017, pp. 105–114.

- [119] Hsin-Ying Lee et al. “Diverse image-to-image translation via disentangled representations”. In: *Proceedings of the European conference on computer vision (ECCV)*. 2018, pp. 35–51.
- [120] Dan Li et al. “MAD-GAN: Multivariate Anomaly Detection for Time Series Data with Generative Adversarial Networks”. In: *arXiv preprint arXiv:1901.04997* (2019).
- [121] Nan Li, Olaf Kähler, and Norbert Pfeifer. “A comparison of deep learning methods for airborne lidar point clouds classification”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021), pp. 6467–6486.
- [122] Yijun Li et al. “Generative face completion”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 3911–3919.
- [123] Ying Li et al. *Deep Learning for LiDAR Point Clouds in Autonomous Driving: A Review*. 2020. arXiv: 2005.09830 [cs.CV].
- [124] Yujia Li, Kevin Swersky, and Rich Zemel. “Generative moment matching networks”. In: *International Conference on Machine Learning*. 2015, pp. 1718–1727.
- [125] Yujia Li et al. “Dualing GANs”. In: *Advances in Neural Information Processing Systems*. 2017, pp. 5606–5616.
- [126] Yu-Ching Lin, JP Mills, and S Smith-Voysey. “Rigorous pulse detection from full-waveform airborne laser scanning data”. In: *International Journal of Remote Sensing* 31.5 (2010), pp. 1303–1324.
- [127] Tsung-Yi Lin et al. *Microsoft COCO: Common Objects in Context*. cite arxiv:1405.0312Comment: 1) updated annotation pipeline description and figures; 2) added new section describing datasets splits; 3) updated author list. 2014. URL: <http://arxiv.org/abs/1405.0312>.
- [128] YC Lin and JP Mills. “Integration of full-waveform information into the airborne laser scanning data filtering process”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.Part 3/W8 (2009), pp. 224–229.
- [129] Gangping Liu and Jun Ke. “Deep-learning for super-resolution full-waveform lidar”. In: *Optoelectronic Imaging and Multimedia Technology VI*. Ed. by Qionghai Dai, Tsutomu Shimura, and Zhenrong Zheng. Vol. 11187. International Society for Optics and Photonics. SPIE, 2019, pp. 245–252. URL: <https://doi.org/10.1117/12.2536719>.
- [130] Gangping Liu, Jun Ke, and Edmund Y Lam. “CNN-based Super-resolution Full-waveform LiDAR”. In: *Imaging and Applied Optics Congress*. Optical Society of America, 2020, JW2A.29. DOI: 10.1364/3D.2020.JW2A.29. URL: <http://www.osapublishing.org/abstract.cfm?URI=3D-2020-JW2A.29>.
- [131] Jitao Liu, Songmin Dai, and Xiaoqiang Li. “PCCN: POINT Cloud Colorization Network”. In: *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2019, pp. 3716–3720.
- [132] Ming-Yu Liu, Thomas Breuel, and Jan Kautz. “Unsupervised image-to-image translation networks”. In: *Advances in Neural Information Processing Systems*. 2017, pp. 700–708.
- [133] Ming-Yu Liu and Oncel Tuzel. “Coupled generative adversarial networks”. In: *Advances in neural information processing systems*. 2016, pp. 469–477.
- [134] Weiping Liu et al. “Deep learning on point clouds and its application: A survey”. In: *Sensors* 19.19 (2019), p. 4188.
- [135] Jonathan Long, Evan Shelhamer, and Trevor Darrell. “Fully convolutional networks for semantic segmentation”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 3431–3440.
- [136] Haihao Lu and Kenji Kawaguchi. “Depth creates no bad local minima”. In: *arXiv preprint arXiv:1702.08580* (2017).

- [137] Ruodan Lu and Ioannis Brilakis. *Recursive segmentation for as-is bridge information modelling*. Jan. 2017. URL: <https://hdl.handle.net/2134/37755>.
- [138] Pauline Luc et al. “Semantic segmentation using adversarial networks”. In: *arXiv preprint arXiv:1611.08408* (2016).
- [139] Shezhou Luo et al. “Fusion of airborne discrete-return LiDAR and hyperspectral data for land cover classification”. In: *Remote Sensing* 8.1 (2016), p. 3.
- [140] Lian Ma, Mei Zhou, and Chuanrong Li. “Land covers classification based on Random Forest method using features from full-waveform LiDAR data”. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42.2/W7 (2017).
- [141] Liqian Ma et al. “Pose guided person image generation”. In: *Advances in Neural Information Processing Systems*. 2017, pp. 406–416.
- [142] Shuang Ma et al. “DA-GAN: Instance-level image translation by deep attention generative adversarial networks”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 5657–5666.
- [143] Dwarikanath Mahapatra, Behzad Bozorgtabar, and Rahil Garnavi. “Image super-resolution using progressive generative adversarial networks for medical image analysis”. In: *Computerized Medical Imaging and Graphics* 71 (2019), pp. 30–39.
- [144] Dwarikanath Mahapatra et al. “Image super resolution using generative adversarial networks and local saliency maps for retinal image analysis”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2017, pp. 382–390.
- [145] Clément Mallet. “Analyse des données lidar aéroportées à Retour d’Onde Complète pour la cartographie des milieux urbains. (Analysis of Full-Waveform lidar data for urban area mapping)”. In: 2010.
- [146] Clément Mallet, Uwe Soergel, and Frédéric Bretar. “Analysis of full-waveform lidar data for classification of urban areas”. In: *ISPRS Congress 2008*. 2008.
- [147] Clément Mallet et al. “Relevance assessment of full-waveform lidar data for urban area classification”. In: *ISPRS journal of photogrammetry and remote sensing* 66.6 (2011), S71–S84.
- [148] G Mandlbürger et al. “Orientation and processing of airborne laser scanning data (OPALS) —Concept and first results of a comprehensive ALS software”. In: *ISPRS Workshop Laser-scanning*. Vol. 2009. 2009.
- [149] Donald W Marquardt. “An algorithm for least-squares estimation of nonlinear parameters”. In: *Journal of the society for Industrial and Applied Mathematics* 11.2 (1963), pp. 431–441.
- [150] Eleonora Maset, Roberto Carniel, and Fabio Crosilla. “Unsupervised classification of raw full-waveform airborne lidar data by self organizing maps”. In: *International Conference on Image Analysis and Processing*. Springer. 2015, pp. 62–72.
- [151] Ben Mildenhall et al. “Nerf: Representing scenes as neural radiance fields for view synthesis”. In: *European conference on computer vision*. Springer. 2020, pp. 405–421.
- [152] Shervin Minaee et al. *Image Segmentation Using Deep Learning: A Survey*. 2020. arXiv: 2001.05566 [cs.CV].
- [153] Mehdi Mirza and Simon Osindero. “Conditional generative adversarial nets”. In: *arXiv preprint arXiv:1411.1784* (2014).
- [154] Kaichun Mo et al. “PartNet: A Large-scale Benchmark for Fine-grained and Hierarchical Part-level 3D Object Understanding”. In: *CVPR*. 2019.

- [155] Werner Muecke, Christian Briese, and Markus Hollaus. “Terrain echo probability assignment based on full-waveform airborne laser scanning observables”. In: (2010).
- [156] Daniel Munoz et al. “Contextual classification with functional max-margin markov networks”. In: *CVPR*. 2009, pp. 975–982.
- [157] NEDO. 同一空域・複数事業者のドローン運航管理システムとの相互接続試験の環境を構築 / NEDO. <https://www.nedo.go.jp/news/press/AA5\101208.html>. (Accessed on 10/24/2021).
- [158] M Netzband, CL Redman, and WL Stefanov. “Challenges for applied remote sensing science in the urban environment”. In: *The international archives of the photogrammetry, remote sensing, and spatial information sciences* 36.8/W27 (2005).
- [159] Amy L Neuenschwander, Lori A Magruder, and Marcus Tyler. “Landcover classification of small-footprint, full-waveform lidar data”. In: *Journal of applied remote sensing* 3.1 (2009), p. 033544.
- [160] Linda Nordin. *Analysis of waveform data from airborne laser scanner systems*. 2006.
- [161] Curtis G. Northcutt, Anish Athalye, and Jonas Mueller. *Pervasive Label Errors in Test Sets Destabilize Machine Learning Benchmarks*. 2021. arXiv: 2103.14749 [stat.ML].
- [162] NYC Street Tree Map. <https://tree-map.nycgovparks.org/>. (Accessed on 10/16/2021).
- [163] Augustus Odena, Christopher Olah, and Jonathon Shlens. “Conditional image synthesis with auxiliary classifier gans”. In: *Proceedings of the 34th International Conference on Machine Learning*. Vol. 70. JMLR. 2017, pp. 2642–2651.
- [164] Ozan Oktay et al. “Attention u-net: Learning where to look for the pancreas”. In: *arXiv preprint arXiv:1804.03999* (2018).
- [165] Haruki Oshio, Tomoki Kiyono, and Takashi Asawa. “Numerical simulation of the nocturnal cooling effect of urban trees considering the leaf area density distribution”. In: *Urban Forestry & Urban Greening* 66 (2021), p. 127391.
- [166] Junting Pan et al. “SalGAN: Visual saliency prediction with generative adversarial networks”. In: *arXiv preprint arXiv:1701.01081* (2017).
- [167] Christopher E Parrish. *Vertical object extraction from full-waveform lidar data using a 3D wavelet based approach*. University of Wisconsin–Madison, 2007.
- [168] Christopher E Parrish and Robert D Nowak. “Improved approach to lidar airport obstruction surveying using full-waveform data”. In: *Journal of Surveying Engineering* 135.2 (2009), pp. 72–82.
- [169] Adam Paszke et al. “Pytorch: An imperative style, high-performance deep learning library”. In: *Advances in neural information processing systems* 32 (2019), pp. 8026–8037.
- [170] Abhishek Patil et al. “The H3D dataset for full-surround 3D multi-object detection and tracking in crowded urban scenes”. In: *ICRA*. 2019.
- [171] Matheus Barros Pereira and Jefersson Alex dos Santos. “An End-to-end Framework For Low-Resolution Remote Sensing Semantic Segmentation”. In: *arXiv preprint arXiv:2003.07955* (2020).
- [172] Åsa Persson et al. “Visualization and analysis of full-waveform airborne laser scanner data”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.3/W19 (2005), pp. 103–108.
- [173] Quang-Hieu Pham et al. “A\*3D Dataset: Towards Autonomous Driving in Challenging Environments”. In: *ICRA* (2020).
- [174] PLATEAU [プレート]. <https://www.mlit.go.jp/plateau/>. (Accessed on 11/01/2021).

- [175] *PointCNN: replacing 50,000 man hours with AI* | by Dmitry Kudinov | GeoAI | Medium. <https://medium.com/geoai/pointcnn-replacing-50-000-man-hours-with-ai-d7397c1e7ffe>. (Accessed on 12/19/2021).
- [176] Ben Poole et al. “Improved generator objectives for GANs”. In: *arXiv preprint arXiv:1612.02780* (2016).
- [177] Charles R Qi et al. “Pointnet: Deep learning on point sets for 3d classification and segmentation”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 652–660.
- [178] Charles R Qi et al. “PointNet++ deep hierarchical feature learning on point sets in a metric space”. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017, pp. 5105–5114.
- [179] Zhaofan Qiu et al. “Deep semantic hashing with generative adversarial networks”. In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. 2017, pp. 225–234.
- [180] Mohammad Muntasir Rahman et al. “Recent Advances in 3D Object Detection in the Era of Deep Neural Networks: A Survey”. In: *IEEE TIP* (2019).
- [181] J Reitberger, Pl Krzystek, and U Stilla. “Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees”. In: *International journal of remote sensing* 29.5 (2008), pp. 1407–1431.
- [182] Josef Reitberger, Peter Krzystek, and Uwe Stilla. “Benefit of airborne full waveform LiDAR for 3D segmentation and classification of single trees”. In: *ASPRS 2009 Annual Conference*. 2009, pp. 1–9.
- [183] A Roncat, G Bergauer, and N Pfeifer. “Retrieval of the backscatter cross-section in full-waveform lidar data using b-splines”. In: *Proc. Int. Archives Photogramm. Remote Sens. Spatial Inf. Sci* (2010), pp. 137–142.
- [184] Andreas Roncat et al. *Echo detection and localization in full-waveform airborne laser scanner data using the averaged square difference function estimator*. na, 2008.
- [185] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9351 (2015), pp. 234–241. ISSN: 16113349. DOI: 10.1007/978-3-319-24574-4\_28. arXiv: 1505.04597.
- [186] RB Roth and J Thompson. “Practical application of multiple pulse in air (MPiA) Lidar in large-area surveys”. In: *Proceedings of international archives of the photogrammetry, remote sensing and spatial information sciences*. Vol. 37. 2008, pp. 183–188.
- [187] Franz Rottensteiner et al. “The ISPRS benchmark on urban object classification and 3D building reconstruction”. In: *ISPRS* (2012).
- [188] Xavier Roynard, Jean-Emmanuel Deschaud, and François Goulette. “Paris-Lille-3D: A large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification”. In: *IJRR* (2018).
- [189] Martin Rutzinger et al. “Object-based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification”. In: *Sensors* 8.8 (2008), pp. 4505–4528.
- [190] Tim Salimans et al. “Improved techniques for training GANs”. In: *Advances in Neural Information Processing Systems*. 2016, pp. 2234–2242.
- [191] Farhad Samadzadegan, Behnaz Bigdeli, and Pouria Ramzi. “A multiple classifier system for classification of LIDAR remote sensing data using multi-class SVM”. In: *International Workshop on Multiple Classifier Systems*. Springer. 2010, pp. 254–263.

- [192] Ramprasaath R Selvaraju et al. “Grad-cam: Visual explanations from deep networks via gradient-based localization”. In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 618–626.
- [193] Andrés Serna et al. “Paris-rue-Madame database: a 3D mobile laser scanner dataset for benchmarking urban detection, segmentation and classification methods”. In: *ICRA*. 2014.
- [194] Jie Shan and Charles K Toth. *Topographic laser ranging and scanning: principles and processing*. CRC press, 2018.
- [195] Jacob Shermeyer and Adam Van Etten. “The effects of super-resolution on object detection performance in satellite imagery”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2019, pp. 0–0.
- [196] *Shizuoka Point Cloud DB*. Dec. 2020. URL: <https://pointcloud.pref.shizuoka.jp/>.
- [197] Kaleem Siddiqi et al. “Retrieving articulated 3-D models using medial surfaces”. In: *Machine Vision and Applications* 19.4 (2008), pp. 261–275.
- [198] George Sithole and George Vosselman. “Bridge detection in airborne laser scanner data”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 61.1 (2006), pp. 33–46.
- [199] Uwe Soergel. “Review of radar remote sensing on urban areas”. In: *Radar remote sensing of urban areas*. Springer, 2010, pp. 1–47.
- [200] Shuran Song, Samuel P Lichtenberg, and Jianxiong Xiao. “Sun RGB-D: A RGB-D scene understanding benchmark suite”. In: *CVPR*. 2015.
- [201] Xibin Song et al. “Apollocar3D: A large 3D car instance understanding benchmark for autonomous driving”. In: *CVPR*. 2019.
- [202] Nasim Souly, Concetto Spampinato, and Mubarak Shah. “Semi supervised semantic segmentation using generative adversarial network”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, pp. 5688–5696.
- [203] Ove Steinvall. “Waveform simulation for 3-D sensing laser radar”. In: *2000*. (2000).
- [204] U Stilla, Wei Yao, and B Jutzi. “Detection of weak laser pulses by full waveform stacking”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.Part 3 (2007), W49A.
- [205] Chujin Sun et al. “Automated Simulation Framework for Urban Wind Environments Based on Aerial Point Clouds and Deep Learning”. In: *Remote Sensing* 13.12 (2021). ISSN: 2072-4292. DOI: 10.3390/rs13122383. URL: <https://www.mdpi.com/2072-4292/13/12/2383>.
- [206] Pei Sun et al. “Scalability in Perception for Autonomous Driving: Waymo Open Dataset”. In: *CVPR*. 2020.
- [207] Weikai Tan et al. “Toronto-3D: A Large-scale Mobile LiDAR Dataset for Semantic Segmentation of Urban Roadways”. In: *arXiv preprint arXiv:2003.08284* (2020).
- [208] Karl-Heinz Thiel, Aloysius Wehr, and Christoph Hug. “A new algorithm for processing fullwave laser scanner data”. In: *EARSeL 3D-Remote Sensing Workshop, CDROM*. 2005.
- [209] KH Thiel and A Wehr. “Performance capabilities of laser-scanners-an overview and measurement principle analysis”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36.Part 8 (2004), W2.
- [210] Matteo Tomei et al. “Art2Real: Unfolding the Reality of Artworks via Semantically-Aware Image-to-Image Translation”. In: *arXiv preprint arXiv:1811.10666* (2018).

- [211] Mehmet Ozgur Turkoglu et al. “A Layer-Based Sequential Framework for Scene Generation with GANs”. In: *Thirty-Third AAAI Conference on Artificial Intelligence*. Honolulu, Hawaii, United States, 2019.
- [212] Mikaela Angelina Uy et al. “Revisiting Point Cloud Classification: A New Benchmark Dataset and Classification Model on Real-World Data”. In: *ICCV*. 2019.
- [213] Bruno Vallet et al. “TerraMobilita/iQmulus urban point cloud analysis benchmark”. In: *Computers & Graphics* 49 (2015), pp. 126–133.
- [214] Nicolas Vandapel, Omead Amidi, and J Rayan Miller. “Toward laser pulse waveform analysis for scene interpretation”. In: *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*. Vol. 1. IEEE. 2004, pp. 950–955.
- [215] N. Varney, V. K. Asari, and Q. Graehling. “DALES: A Large-scale Aerial LiDAR Data Set for Semantic Segmentation”. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Los Alamitos, CA, USA: IEEE Computer Society, June 2020, pp. 717–726. DOI: 10.1109/CVPRW50498.2020.00101. URL: <https://doi.ieeecomputersociety.org/10.1109/CVPRW50498.2020.00101>.
- [216] Nina Varney, Vijayan K Asari, and Quinn Graehling. “DALES: A Large-scale Aerial LiDAR Data Set for Semantic Segmentation”. In: *arXiv preprint arXiv:2004.11985* (2020).
- [217] Carl Vondrick, Hamed Pirsiavash, and Antonio Torralba. “Generating videos with scene dynamics”. In: *Advances In Neural Information Processing Systems*. 2016, pp. 613–621.
- [218] George Vosselman and Hans-Gerd Maas. *Airborne and terrestrial laser scanning*. CRC press, 2010.
- [219] W Wagner et al. “Radiometric calibration of full-waveform small-footprint airborne laser scanners”. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 37.B1 (2008), pp. 163–168.
- [220] Wolfgang Wagner et al. “3D vegetation mapping using small-footprint full-waveform airborne laser scanners”. In: *International Journal of Remote Sensing* 29.5 (2008), pp. 1433–1452.
- [221] Wolfgang Wagner et al. *From single-pulse to full-waveform airborne laser scanners: potential and practical challenges*. Vol. 35. na, 2004.
- [222] Wolfgang Wagner et al. “Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner”. In: *ISPRS journal of Photogrammetry and Remote Sensing* 60.2 (2006), pp. 100–112.
- [223] Chisheng Wang et al. “A random forest classifier based on pixel comparison features for urban LiDAR data”. In: *ISPRS journal of photogrammetry and remote sensing* 148 (2019), pp. 75–86.
- [224] Hongzhou Wang and Craig Glennie. “Fusion of waveform LiDAR data and hyperspectral imagery for land cover classification”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 108 (2015), pp. 1–11.
- [225] Ting-Chun Wang et al. “High-resolution image synthesis and semantic manipulation with conditional GANs”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 8798–8807.
- [226] Xintao Wang et al. “ESRGAN: Enhanced super-resolution generative adversarial networks”. In: *European Conference on Computer Vision Workshop*. 2018.
- [227] Yanfei Wang et al. “Regularizing method for the determination of the backscatter cross section in lidar data”. In: *JOSA A* 26.5 (2009), pp. 1071–1079.
- [228] Yue Wang et al. “Dynamic Graph CNN for Learning on Point Clouds”. In: *ACM TOG* (2019).

- [229] Lukas Winiwarter et al. “Classification of ALS point clouds using end-to-end deep learning”. In: *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 87.3 (2019), pp. 75–90.
- [230] H Wong and Andreas Antoniou. “Characterization and decomposition of waveforms for LARSEN 500 airborne system”. In: *IEEE transactions on geoscience and remote sensing* 29.6 (1991), pp. 912–921.
- [231] Huikai Wu et al. “GP-GAN: Towards realistic high-resolution image blending”. In: *arXiv preprint arXiv:1703.07195* (2017).
- [232] Zhirong Wu et al. “3D ShapeNets: A Deep Representation for Volumetric Shapes”. In: *CVPR*. 2015.
- [233] Yuxing Xie, Jiaojiao Tian, and Xiao Xiang Zhu. “Linking Points With Labels in 3D: A Review of Point Cloud Semantic Segmentation”. In: *IEEE Geoscience and Remote Sensing Magazine* 8.4 (2020), pp. 38–59. DOI: 10.1109/MGRS.2019.2937630.
- [234] Yuxing Xie, Jiaojiao Tian, and Xiaoxiang Zhu. “A Review of Point Cloud Semantic Segmentation”. In: *IEEE GRSM* (2020).
- [235] Chao Yang et al. “High-resolution image inpainting using multi-scale neural patch synthesis”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 6721–6729.
- [236] Yao-Yuan Yang et al. “TorchAudio: Building Blocks for Audio and Speech Processing”. In: *arXiv preprint arXiv:2110.15018* (2021).
- [237] Yaoqing Yang et al. “Foldingnet: Point cloud auto-encoder via deep grid deformation”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 206–215.
- [238] Zhilin Yang et al. “Semi-Supervised QA with Generative Domain-Adaptive Nets”. In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Vancouver, Canada, 2017, pp. 1040–1050.
- [239] Raymond A Yeh et al. “Semantic image inpainting with deep generative models”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 5485–5493.
- [240] Jiahui Yu et al. “Generative image inpainting with contextual attention”. In: *arXiv preprint arXiv:1801.07892* (2018).
- [241] Xiaohui Yuan, Jianfang Shi, and Lichuan Gu. “A review of deep learning methods for semantic segmentation of remote sensing imagery”. In: *Expert Systems with Applications* 169 (2021), p. 114417. ISSN: 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2020.114417>. URL: <https://www.sciencedirect.com/science/article/pii/S0957417420310836>.
- [242] Manzil Zaheer et al. “Deep Sets”. In: *NeurIPS*. 2017.
- [243] Andy Zeng et al. “Multi-view self-supervised deep learning for 6D pose estimation in the amazon picking challenge”. In: *ICRA*. 2017, pp. 1386–1383.
- [244] Han Zhang et al. “Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, pp. 5907–5915.
- [245] Jinming Zhang et al. “DEM extraction from ALS point clouds in forest areas via graph convolution network”. In: *Remote Sensing* 12.1 (2020), p. 178.
- [246] Zhiyuan Zhang, Binh-Son Hua, and Sai-Kit Yeung. “ShellNet: Efficient Point Cloud Convolutional Neural Networks using Concentric Shells Statistics”. In: *ICCV*. 2019.

- [247] Hengshuang Zhao et al. “PointWeb: Enhancing Local Neighborhood Features for Point Cloud Processing”. In: *CVPR*. 2019.
- [248] Lichen Zhou, Chuang Zhang, and Ming Wu. “D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2018, pp. 182–186.
- [249] Mei Zhou et al. “LAND COVER CLASSIFICATION FROM FULL-WAVEFORM LIDAR DATA BASED ON SUPPORT VECTOR MACHINES.” In: *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 41 (2016).
- [250] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. “Open3D: A modern library for 3D data processing”. In: *arXiv preprint arXiv:1801.09847* (2018).
- [251] Jun-Yan Zhu et al. “Generative visual manipulation on the natural image manifold”. In: *European Conference on Computer Vision*. Springer. 2016, pp. 597–613.
- [252] Jun-Yan Zhu et al. “Toward multimodal image-to-image translation”. In: *Advances in Neural Information Processing Systems*. 2017, pp. 465–476.
- [253] Jun-Yan Zhu et al. “Unpaired image-to-image translation using cycle-consistent adversarial networks”. In: *arXiv preprint arXiv:1703.10593v6* (2017).
- [254] Wentao Zhu et al. “Adversarial deep structural networks for mammographic mass segmentation”. In: *arXiv preprint arXiv:1612.05970* (2016).
- [255] Stefano Zorzi et al. “Full-Waveform Airborne LiDAR Data Classification Using Convolutional Neural Networks”. In: *IEEE Transactions on Geoscience and Remote Sensing* PP (2019), pp. 1–7. ISSN: 0196-2892. DOI: 10.1109/TGRS.2019.2919472. URL: <https://ieeexplore.ieee.org/document/8737008/>.
- [256] 今井優. “深層学習による空中写真を対象とした土地被覆分類の試み”. In: 国際 ICT 利用研究会研究会研究論文誌 1.1 (2018), pp. 16–21.
- [257] 佐藤俊明 et al. “様々なプラットフォームによる 3 次元情報取得技術の活用と取り組み”. In: *IEICE Conferences Archives*. The Institute of Electronics, Information and Communication Engineers. 2019.
- [258] 前田佳子, 本田謙一, and 今井靖晃. “波形記録式航空機 LiDAR データを用いた森林モニタリング技術の紹介”. In: 日本リモートセンシング学会誌 36.5 (2016), pp. 501–506.
- [259] 浩之 吉田, 健一 石橋, and 直子 栗崎. “航空機レーザ測量データによる複合日影シミュレーション: CBD を対象として”. In: 地理情報システム学会講演論文集 = *Papers and proceedings of the Geographic Information Systems Association* 10 (Sept. 2001), pp. 315–319. URL: <https://ci.nii.ac.jp/naid/10009667379/>.
- [260] 国土交通省. 報道発表資料: 産学官連携によるイノベーション創出を目指して～国土交通データプラットフォームの公開イベントを開催～- 国土交通省. [https://www.mlit.go.jp/report/press/kanbo08\\_hh\\_000633.html](https://www.mlit.go.jp/report/press/kanbo08_hh_000633.html). (Accessed on 10/24/2021).
- [261] 国土地理院. 3次元地図検討部会報告書～3次元地図の適切な整備と活用促進のために～. <https://www.gsi.go.jp/common/000228237.pdf>. (Accessed on 10/24/2021).
- [262] 小山浩 and 柴田泰秀. “自動走行におけるダイナミックマップ整備”. In: システム/制御/情報 60.11 (2016), pp. 463–468.
- [263] 本田謙一, 前田佳子, and 今井靖晃. “波形記録式航空レーザスキャナによる地盤面抽出に対する効果の検証”. In: 応用測量論文集 25 (2014), pp. 105–113.
- [264] 朱林, 橘菊生, 島村秀樹, et al. “フルウェーブフォーム航空レーザの解析手法の研究”. In: 写真測量とリモートセンシング 54.1 (2015), pp. 4–19.

- [265] 李勇鶴. “日本写真測量学会学術講演会の発表論文からみた写真測量・リモートセンシング分野における深層学習・機械学習の活用状況”. In: 写真測量とリモートセンシング 59.1 (2020), pp. 49–53.
- [266] 濱口竜平 and 彦坂修平. “畳み込みニューラルネットを用いた衛星画像からの建物検出”. In: 写真測量とリモートセンシング 57.6 (2018), pp. 246–247.
- [267] 織田和夫. “ISO/TC204/WG3 ITS における空間情報の国際標準化”. In: 写真測量とリモートセンシング 58.3 (2019), pp. 98–103.
- [268] 自律航行ドローン、安全飛行のために検知すべきデータや技術は？(深掘り！自動運転 × データ第 20 回) / 自動運転ラボ. [https://jidounten-lab.com/u\\\_autonomous-drone-20\#\\\_-7](https://jidounten-lab.com/u\_autonomous-drone-20\#\_-7). (Accessed on 10/24/2021).
- [269] 萩原晟也 et al. “波形記録式航空機 LiDAR による林相区分への応用可能性の検討”. In: 日本森林学会大会発表データベース第 126 回日本森林学会大会. 日本森林学会. 2015, p. 41.
- [270] 長谷川裕之 and 都竹正志. “LIDAR データを用いた道路領域抽出における反射強度の有用性”. In: 写真測量とリモートセンシング 45.4 (2006), pp. 26–34.
- [271] 関本義秀. “私のブックマーク: 都市空間の情報処理—データセットの世界動向”. In: 人工知能 36.4 (2021), pp. 512–520.
- [272] 高精度 3次元地図データ / SERVICE / ダイナミックマップ基盤株式会社. <https://www.dynamic-maps.co.jp/service/hdmap/index.html>. (Accessed on 10/24/2021).