

卒業論文

トポロジカルマップを用いた シナリオによるナビゲーション

- 全天球カメラ画像に基づく通路認識手法の
提案 -

Proposal of Navigation by Scenarios Using
Topological Map

- Proposal of passage recognition method
based on spherical camera images -

指導教員 林原 靖男 教授

2022年1月11日

千葉工業大学 先進工学部 未来ロボティクス学科

18C1095 原桃子

目次

第 1 章 序論	8
1.1 背景	8
1.2 目的	10
1.3 関連研究	11
1.4 本論文の構成	11
第 2 章 要素技術	12
2.1 トポロジカルマップ	12
2.2 シナリオ	13
2.3 Neural Network	15
2.4 Convolutional Neural Network(CNN)	16
2.5 You Only Look Once(YOLO)	17
第 3 章 トポロジカルを用いたシナリオによるナビゲーション	18
第 4 章 提案手法	22
4.1 提案手法の概要	22
4.2 データセットの作成	24
4.3 学習の結果	26
4.4 通路分類の手法	27
第 5 章 実験	29

5.1	実験目的	29
5.2	実験の概要	29
5.3	実験装置	30
5.4	全天球カメラを用いた通路分類の検証	31
5.5	25 places to place robots in aisle classification experiments	32
第 6 章 まとめ		33
謝辞		33

図目次

1.1	Detecting an object using YOLO	10
2.1	Example of metric and topological map	13
2.2	An example of human directions	14
2.3	Neural network architecture	15
2.4	Architecture of CNN (出典 : [11])	16
2.5	The YOLO Detection System.(出典:[4])	17
3.1	Topological map of previous research	20
3.2	Type of passage	21
4.1	Flow of passage recognition method	23
4.2	Preprocessing of spherical camera images	24
4.3	An example of a dataset	25
5.1	Experiment environment	30
5.2	ORNE-	30

表目次

4.1	Class name to be labeled	25
4.2	Structure of data used for learning	26
5.1	Specification of PC	31
5.2	Specification of ORNE-	31

要約

トポロジカルマップを用いた シナリオによるナビゲーション

- 全天球カメラ画像に基づく通路分類手法の提案 -

人の移動する能力をロボットの自律移動に応用する手法として、先行研究によりトポロジカルマップとシナリオを用いたナビゲーション手法が提案された。人は道案内により移動する際、三叉路や突き当たりなどの通路の特徴を重視していることから、この手法ではナビゲーションを行なながら通路の特徴を分類し、ナビゲーションに利用する。しかし、この手法の問題点としてナビゲーションで必要となる通路の分類に失敗することにより、目的地までたどり着くことができずにナビゲーションに失敗するということがあった。通路の分類は、Chen らが提案する LiDAR を用いた通路検出手法(Toe-Finding Algolithm)[?]を参考にしており、LiDAR の周囲に壁などの遮蔽物がなく、開けている方向があればその方向に通路があると検出する。そのため、開いているドアや隙間などを通路と誤検出してしまった。そこで、本研究では通路を分類する手法として、全天球カメラ画像を用いた手法を検討する。具体的には、全天球カメラで水平 360 度の画像を取得し、取得した画像から YOLO の学習器を用いて通路やドアなどの物体を検出する。また、検出した通路とその方向からどの通路の特徴に相当するかを分類する手法である。本研究では、先行研究と同様の 9 種類の通路の特徴を分類する。本手法で全天球カメラを用いた理由は、全天球カメラはカメラの周囲 360 度の画像データを一度の撮影により取得することができるため、複数台のカメラを使わずに通路分類に必要となるデータを取得することができるからである。初めに、通路の画像を集めたデータセットを用いてトレーニングを行い、学習モデルを作成する。次に、YOLO により検出した画像中の通路の位置座標に基づき、通路の特徴を分類するシステムを作成した。最後に、提案した手法により実環境において通路の分類が

できるかどうかを検証する。

キーワード：トポロジカルマップ，機械学習，道案内

Abstract

第1章

序論

1.1 背景

近年，商業施設での清掃作業や屋外での宅配など，屋内外を問わず多くの場所で自律移動ロボットが活用されている。それらのロボットを安全かつ正確に運用するためには，自己位置推定をはじめとする様々な機能を持たせる必要がある。現在，商業施設などで用いられている自律移動ロボットは，あらかじめセンサにより取得した環境の詳細なデータを持つメトリックマップと呼ばれる地図を用いて，自己位置推定を行なっている。しかし，詳細な地図を用いた自己位置推定は，事前に取得した地図に存在しないものが環境中に置かれると自己位置推定に致命的な支障をきたすという問題がある。

一方，人はそのような詳細な地図がなくても，「次の交差点で右」のような言葉の情報に基づき目的地まで移動することができる。そこで，そのような人の移動する能力をロボットの自律移動に応用する手法が研究されている。例えば，島田らは人の道案内に注目し，人の道案内のようにロボットを目的地まで移動させるナビゲーションの手法を提案した[1][2]。この研究では，人が道案内により移動する際，歩行者の向きや三叉路や突き当たりなどの通路の特徴（形状）を重視しているということをアンケートにより収集し，ナビゲーションに用いるトポロジカルマップと呼ばれる地図と，シナリオと呼ばれる道案内を文章で表現したものの形式を決め，実ロボットでの実験により提案したナビゲーション手法の有効性を検証した。検証の結果，通路の特徴の分類が正しく行われた場合は提案したナビゲーション手法により目的地に到達できるが，通路の特徴の分類に失敗した場合はロボットが経路から外れ，ナビゲーションに失敗してしまうということが報告されている。提案手法の通路の分類はChenらが提案するLiDARを用いた通路検出手法（Toe-Finding Algolithm）[?]を参考にしており，LiDARの周囲に壁などの遮蔽物がなく，開けている方向があればその方向に通路があると検出する。先行研究の通路分類に失敗した原因は，開いているドアや隙間がある方向に通路があv誤検出した

からであると述べられている。そこで、通路の特徴分類にカメラ画像を用いることで通路の誤検出を解消できるのではないかと考えた。

1.2 目的

本研究は、全天球カメラ画像に基づく通路の分類手法を提案する。具体的には、全天球カメラで水平360度の画像を取得し、取得した画像からYOLOの学習器を用いて Fig. 1.1 に示すように通路やドアなどの物体を検出する。また、検出した通路とその方向からどの通路の特徴に相当するかを分類する手法である。全天球カメラを取り付けたロボットを実験環境の24箇所の通路に移動させ、提案した手法により通路の分類が行えるかどうかを検証する。

また、本研究では全天球カメラの標準的なフォーマットである正距円筒図法という形式で画像を扱う。



Fig. 1.1: Detecting an object using YOLO

1.3 関連研究

1.4 本論文の構成

本論文ではまず、第1章で研究背景、目的、関連研究について述べた。第2章では、本研究で用いる要素技術について述べる。また、第3章では提案した手法について述べ、第4章では提案した手法の有効性の検証を行う。また、第5章では4章で行なった実験の結果をまとめ、考察を行う。最後に、第6章で本研究のまとめを行う。

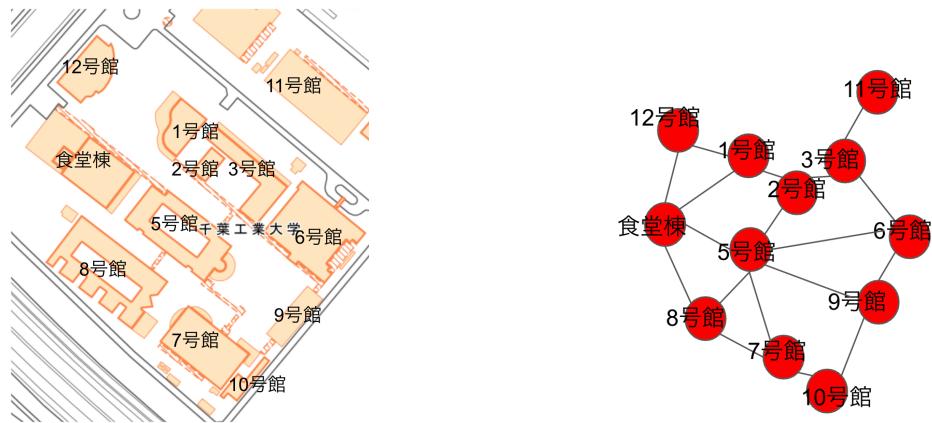
第2章

要素技術

2.1 トポロジカルマップ

私たちの身の回りには様々な種類の地図があり、活用されている。例えば、Fig. 2.1a に表すメトリックマップと呼ばれる地図は、普段人が目的地まで移動する際に用いられる。しかし、本研究で用いているトポロジカルマップは Fig. 2.1b のような形をしている。メトリックマップがやや複雑な形をしているのに対し、トポロジカルマップはより簡潔に、環境を抽象的に表現することができる。

トポロジカルマップは、大きく分けてノードとエッジの 2 つの要素により構成されている。Fig. 2.1b では、赤い丸の図形で表現されているのがノードである。ノードには、地図の作成者が好きな情報を入れることができる。もう 1 つの要素であるエッジは、それぞれのノード同士を接続するのに用いられる。ノード同士に関係性がある場合、ノードとノードはエッジにより接続される。



(a) Map representing Chiba Institute of Technology Shin-Narashino Campus with metric map (出典：国土地理院より一部を加工して作成 [10])

(b) Map representing Chiba Institute of Technology Shin-Narashino Campus with topological map

Fig. 2.1: Example of metric and topological map

2.2 シナリオ

本研究におけるシナリオとは、人間の道案内を文章で表現したものである。例えば、Fig. 2.2 に表すような道案内の場合は、シナリオは「次の角まで直進。右折。突き当たりまで直進。停止。」のようなになる。また、シナリオは「次の角」や「突き当たり」のような条件と「直進」、「右折」のような行動の組み合わせにより作成される。

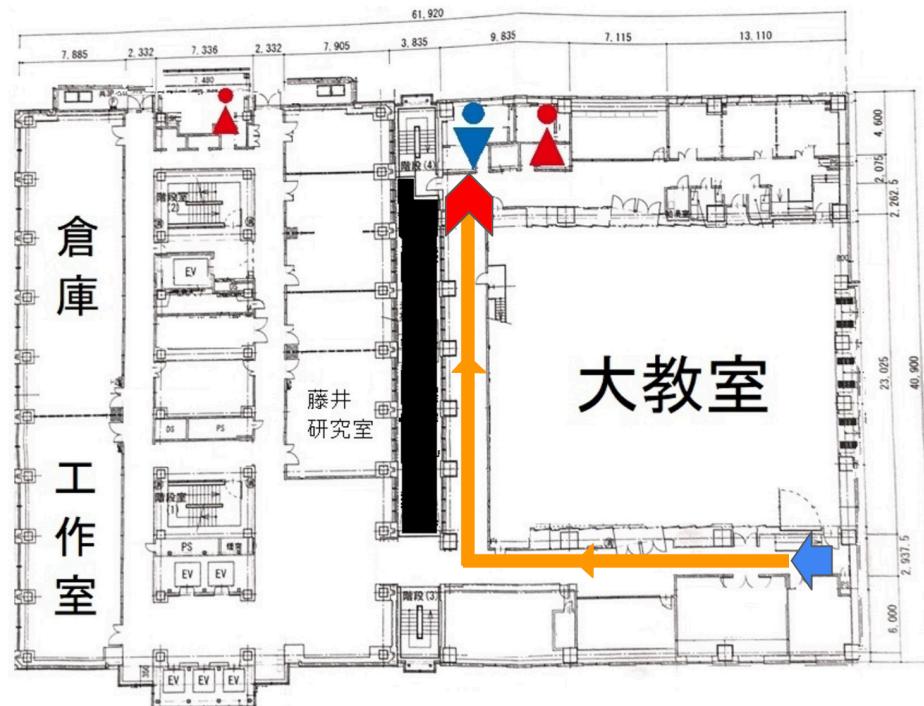


Fig. 2.2: An example of human directions

2.3 Neural Network

ニューラルネットワークとは、人間の脳内の神経細胞（ニューロン）のネットワーク構造を模して作られた数式的なモデルである。ネットワークは、Fig. 2.3 に示すように入力層、出力層、1つ以上の隠れ層（中間層）により構成されており、円で表されているものはニューロンと呼ぶ。また、線により結ばれているニューロンとニューロンの間には重みを持っており、重みは接続されているニューロン間のつながりの強さを表現している。ニューラルネットワークを用いることにより、複雑な回帰問題や分類問題などを解くことができる。

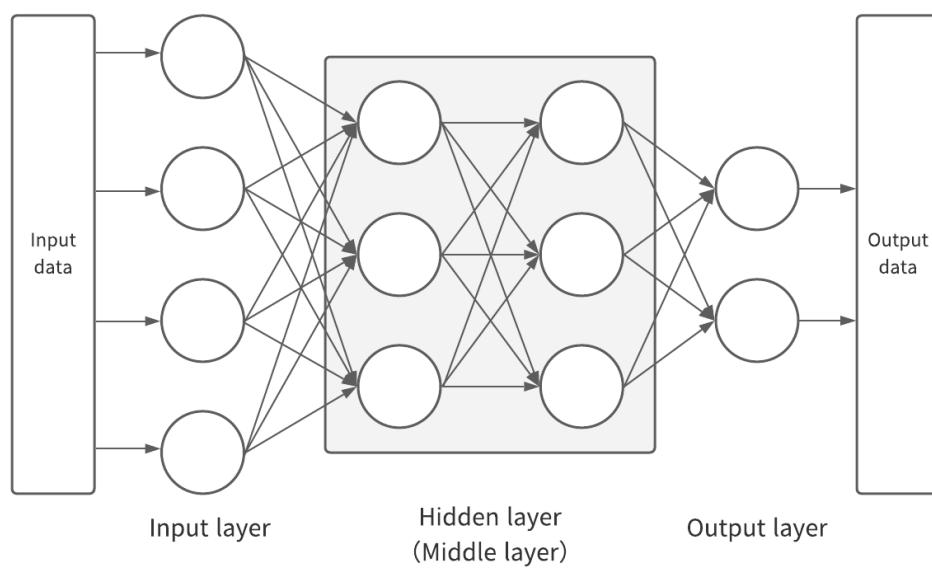


Fig. 2.3: Neural network architecture

2.4 Convolutional Neural Network(CNN)

CNNはディープラーニングの手法の1つであり、ニューラルネットワークを元に作られている。ニューラルネットワークは入力層、中間層、出力層により構成されているが、CNNはニューラルネットワークの中間層に畳み込み層、プーリング層を組み込んだネットワークである。CNNにより画像認識をする場合の各層の役割を以下に記す。

- ・畳み込み層：入力した画像データにフィルタをかけ、特徴マップを出力することで特徴抽出を行う
- ・プーリング層：特徴マップをいくつかの領域に分割し、各領域で最大値または平均化を行うことで重要な情報を残しながら画像を縮小する。

CNNは機械学習分野において主に画像・動画認識の分野で高い性能を発揮する。LeCunら[11]はCNNを用いて手書きの画像データから数字を識別できることを示した。また、使用されたネットワーク構造を Fig. 2.4 に表す。

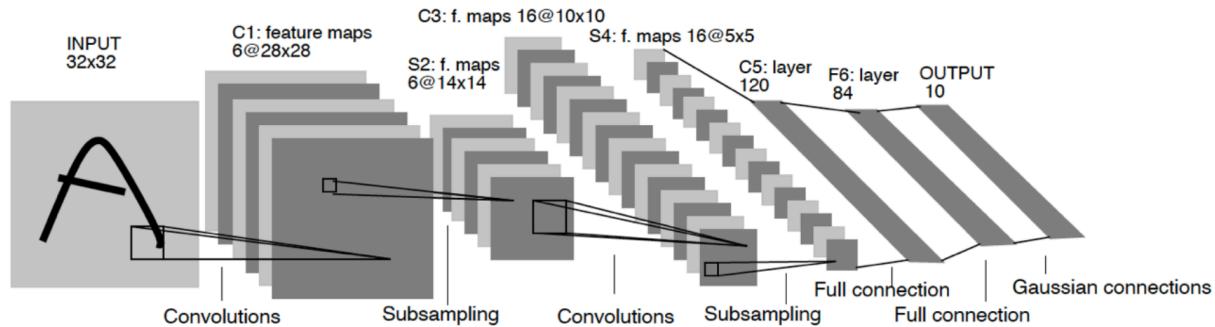


Fig. 2.4: Architecture of CNN (出典 : [11])

2.5 You Only Look Once(YOLO)

本研究で用いる YOLO[4][5][6][7] は、リアルタイム物体検出アルゴリズムである。YOLO は、画像の RGB データの配列を CNN に入力し、画像中のどの範囲に物体が存在しているのかを表すバウンディングボックスの情報と、ボックス内の物体がどのクラスに属しているのかを確率とともに表すクラス確率の情報を出力する。Fig. 2.5 は、YOLO を用いて画像中の物体を検出している様子である。

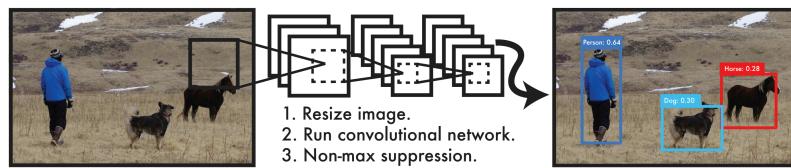


Fig. 2.5: The YOLO Detection System.(出典:[4])

左の画像データを入力した結果、画像からは 3 つの物体が検出されている。また、それぞれの物体が Person 0.64 , Dog 0.30 , Hose 0.28 の確率で予測されていることが確認できる。

第3章

トポロジカルを用いたシナリオによるナビゲーション

トポロジカルマップとシナリオを用いたナビゲーションは，人の道案内を自律移動ロボットのナビゲーションに応用した手法である。この手法では，Fig. 3.1 に示すようなトポロジカルマップと人の道案内を文章で表現したシナリオを用いてナビゲーションを行う。この研究では，初めに，人が道案内により移動する際にどのような情報を必要としているのかを調査するためにアンケートを実施した。アンケートにより，道案内による移動の際に，歩行者の向きと三叉路や突き当たりのような通路の特徴を重視しているということがわかった。そこで，アンケートの結果を基にナビゲーションに用いるトポロジカルマップとシナリオの形式をそれぞれ以下のように定めた。

- ・トポロジカルマップ：ノードは通路の特徴的な箇所に配置する。隣り合うノード間に通路がある場合はエッジにより接続する。また，各ノードは接続しているエッジの相対角度，その地点の通路の特徴情報を保持する。
- ・シナリオ：道案内を文章により表現したもので，「突き当たり」のような条件と「直進」，「右折」のような行動の組み合わせにより生成する。

本ナビゲーション手法では，まず作成したシナリオをロボットに読み込ませる。そして，スタート地点からロボットを発進させ，シナリオの条件と行動に基づき経路に沿ってロボットを走行させることで目的地までのナビゲーションを行う。また，この手法では経路を走行中にシナリオに含ま

れる通路の特徴を検出する。通路の特徴検出は Chen らが提案する LiDAR を用いた通路検出手法 (Toe-Finding Algolithm)[?] を参考にしており、次頁に示す 8 種類に通路を分類する。通路の図は Fig. 3.2 に示す。

- ①一本道 (straight)
- ②行き止まり (dead end)
- ③右のみ曲がれる角 (right)
- ④左のみ曲がれる角 (left)
- ⑤十字路 (cross)
- ⑥右に曲がれる三叉路 (3-way junction_right)
- ⑦左に曲がれる三叉路 (3-way junction_left)
- ⑧突き当たりの三叉路 (3-way junction_center)



Fig. 3.1: Topological map of previous research

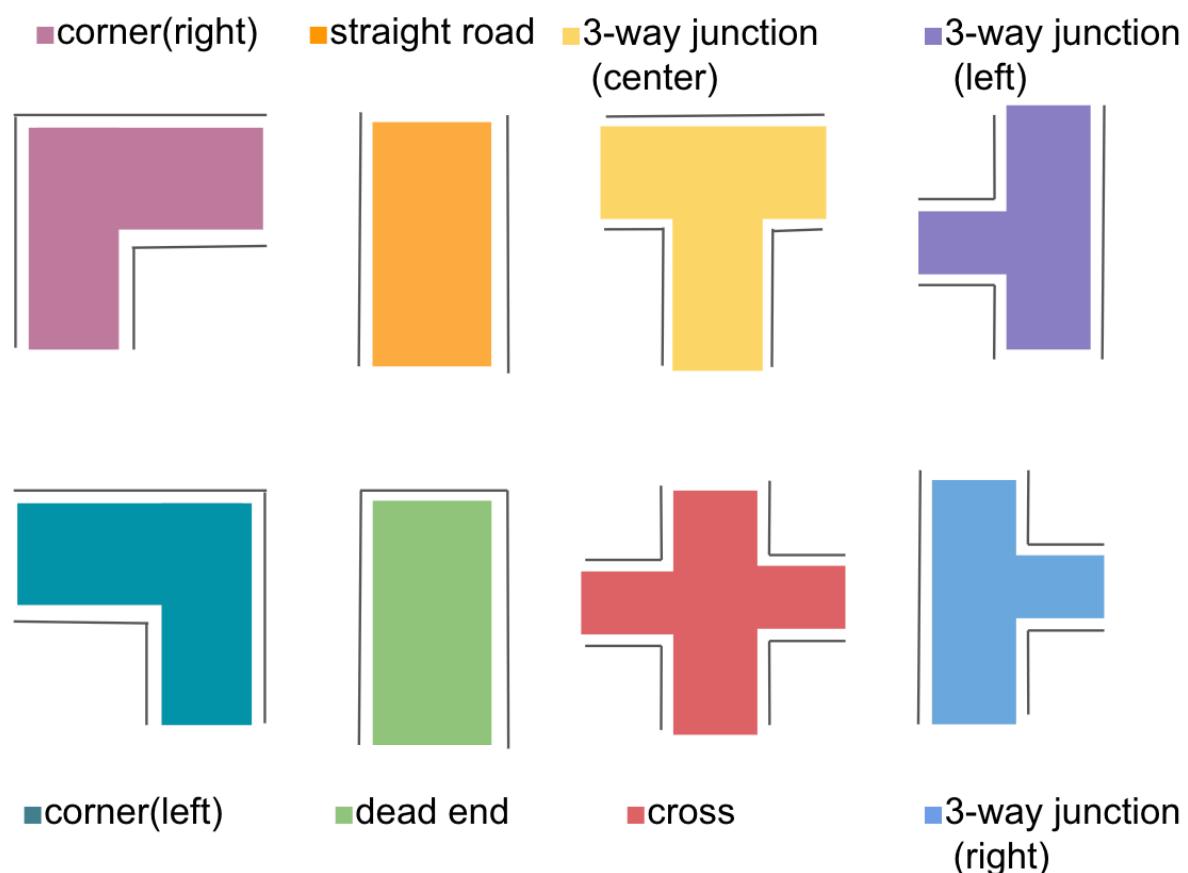


Fig. 3.2: Type of passage

第4章

提案手法

4.1 提案手法の概要

本研究は、全天球カメラから取得した画像に基づき通路の特徴を分類する手法を提案する。また、本研究では先行研究と同様の8種類の通路の特徴を分類する。

本手法の通路分類の流れを Fig. 4.1 に示す。まず、全天球カメラで水平 360 度の画像データを収集する。次に、YOLO の学習器を用いて画像中の通路やドアなどの物体を検出する。通路が検出された場合、通路が検出された方向からどの通路の特徴に相当するかを分類する。Fig. 4.1 の例では、学習器の出力結果から通路が 3 つ検出されている。また、通路はカメラを取り付けたロボットに対して前、後ろ、右方向に検出されている。この通路の特徴から、通路は右に曲がれる三叉路に分類される。

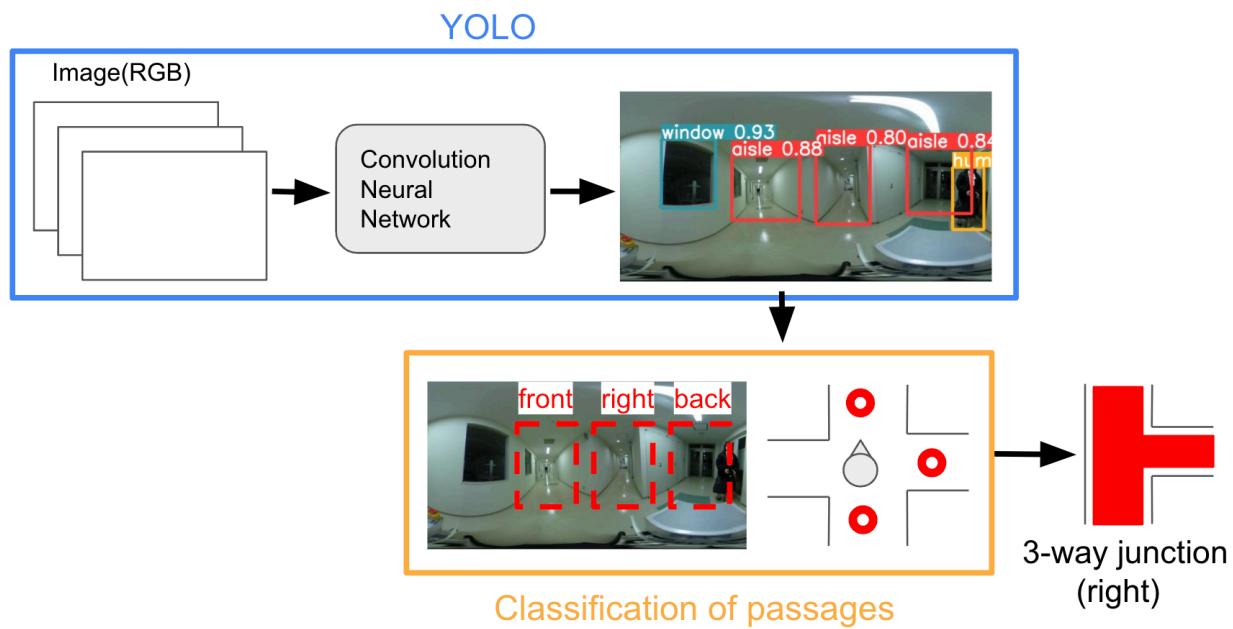
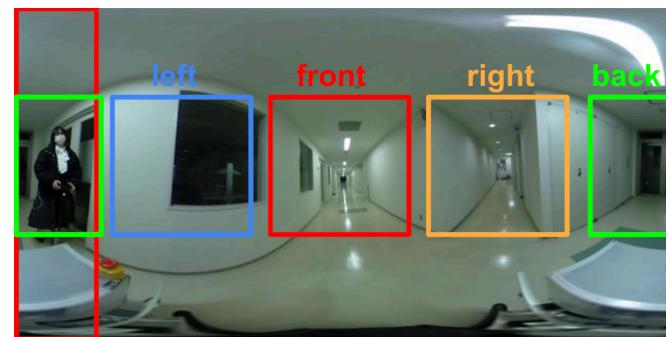


Fig. 4.1: Flow of passage recognition method

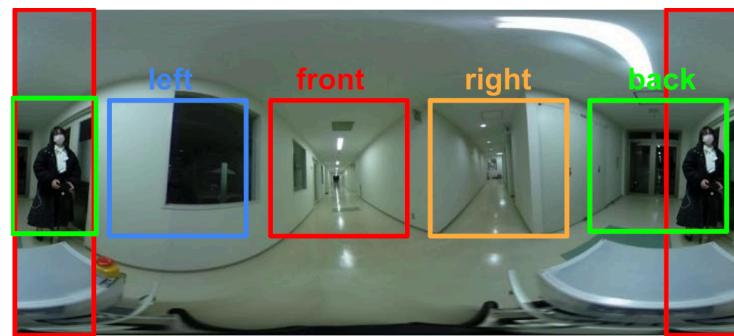
4.2 データセットの作成

学習モデル作成のため、独自に作成した通路の画像を集めたデータセットを用いて学習を行う。学習に用いる画像は千葉工業大学津田沼キャンパス2号館3階の廊下で収集した。全天球カメラで取得した画像は通常、Fig. ??に示すようにカメラの正面にくる物体が画像の中心に写り、後方の物体は画像の左端と右端で分割されてしまう。本手法ではカメラの前後左右の各方向にある物体を正しく検出する必要があるため、後方の物体が分割されないようにデータセットの画像に対して Fig. ??に示すような前処理を行う。この処理では、画像の左端1/8を切り取り、右端にスライドさせるという内容の処理を行う。

データセットは合計3554枚の画像により構成されており、その一部を Fig. 4.3に示す。また、データセットのクラスは Fig. 4.1に示す11クラスとした。作成したデータセットで人をラベル付した理由は、実験中にロボットの後ろについて歩く際、距離を取っている場合でも細い通路などで人が通路と被ることがあり、検出結果に支障をきたさないようにするためである。



(a) no processing image



(b) Preprocessing images

Fig. 4.2: Preprocessing of spherical camera images

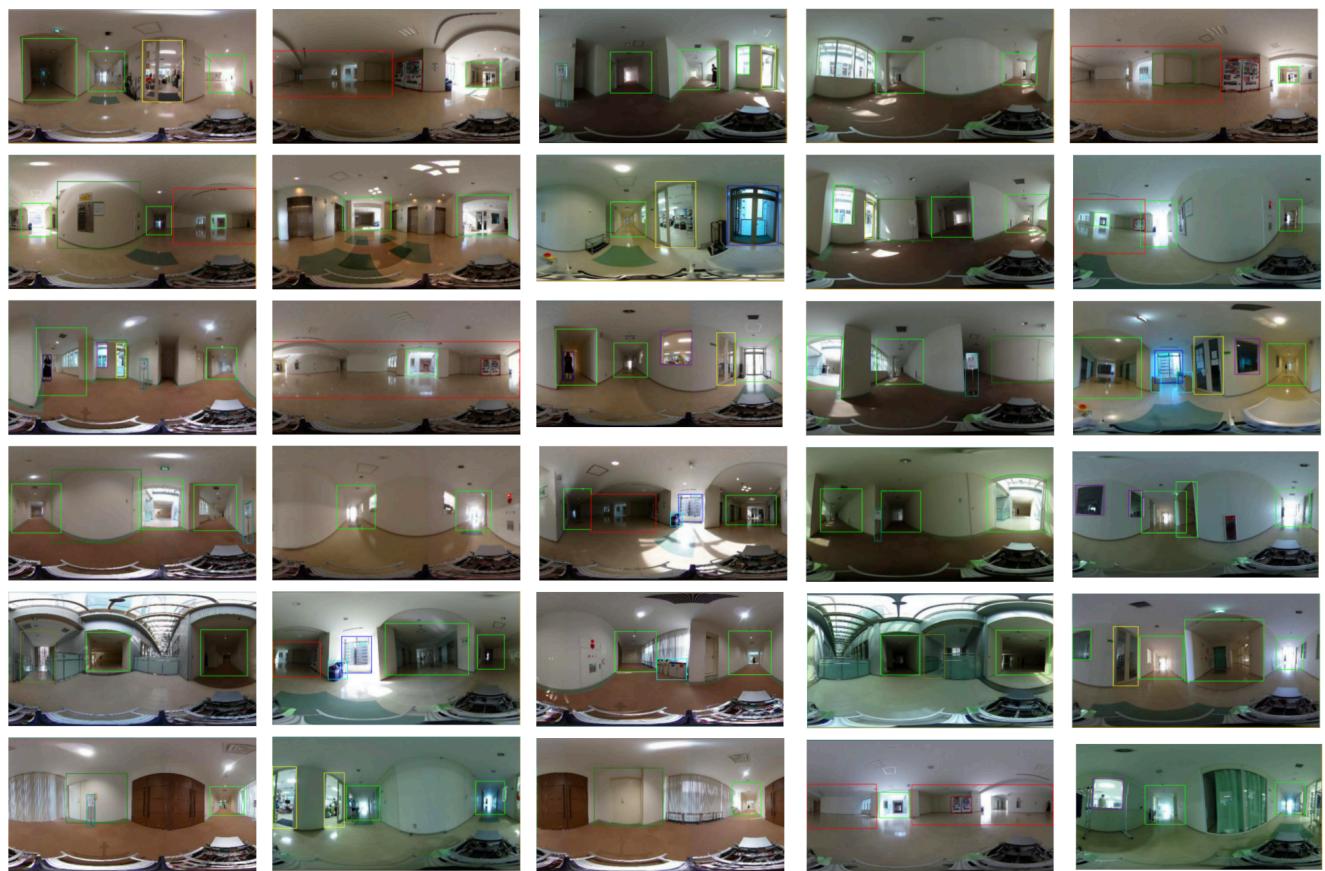


Fig. 4.3: An example of a dataset

Table 4.1: Class name to be labeled

name of the class
aisle
end
door_end
human
door
step
square
vending_machine
trash_can
signboard
window

4.3 学習の結果

作成したデータセットを用いて YOLOv5 により学習を行う。トレーニングに用いたネットワークを Fig. ?? に示す。また、学習に用いるデータの構成とハイパーパラメータの設定を Fig. 4.2 に示す。

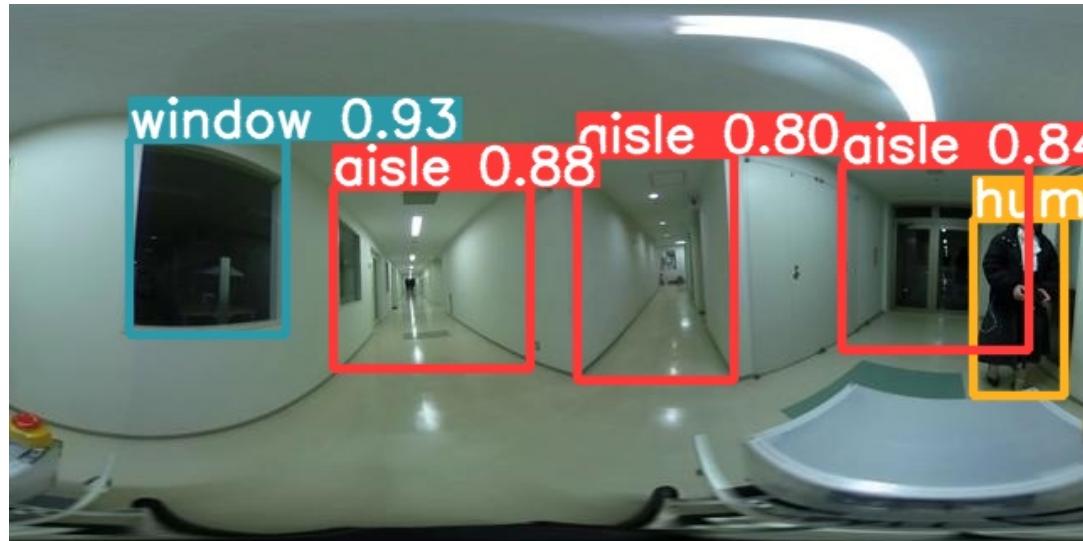
学習の結果、作成したモデルの精度は 0.95 であった。また、学習中の Loss の推移は Fig. ?? のようになり、過学習が起きていないことが確認できる。

Table 4.2: Structure of data used for learning

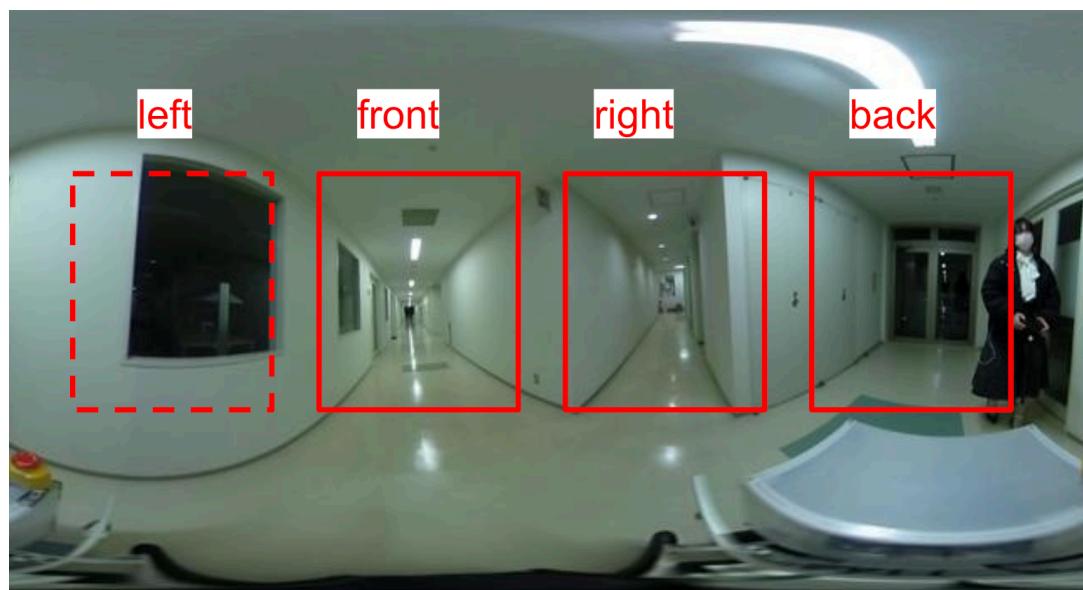
training	3540
validation	14
epoch	300
batch	8

4.4 通路分類の手法

学習器により出力された物体検出の結果から、通路の特徴を分類する手法について述べる。作成した学習器による物体検出の結果を Fig. ?? に示す。また、カメラに対する前後左右の各方向と、それに対応する画像の領域を Fig. 4.4b の赤枠により表す。本手法では、学習器により通路が検出された場合、検出された各通路のバウンディングボックスの中心座標を計算する。そして、中心座標が Fig. 4.4b の各赤枠の領域内にある場合、対応する方向に通路があるとして扱う。最終的に、Table ?? に示すように、通路がある方向から画像の通路がどの通路の特徴に相当するのかを分類する。



(a) Output result of object detection by learner



(b) Positional relationship between the area in the image and each direction with respect to the camera

第5章

実験

5.1 実験目的

本研究で提案した全天球カメラ画像に基づく通路分類手法の有効性をロボットを用いた実環境での実験により検証する。

5.2 実験の概要

提案した通路分類手法の有効性をロボットを用いた実環境での実験により検証する。実験環境は Fig. 5.1 に示す千葉工業大学津田沼キャンパス 2 号館 3 階の廊下とした。

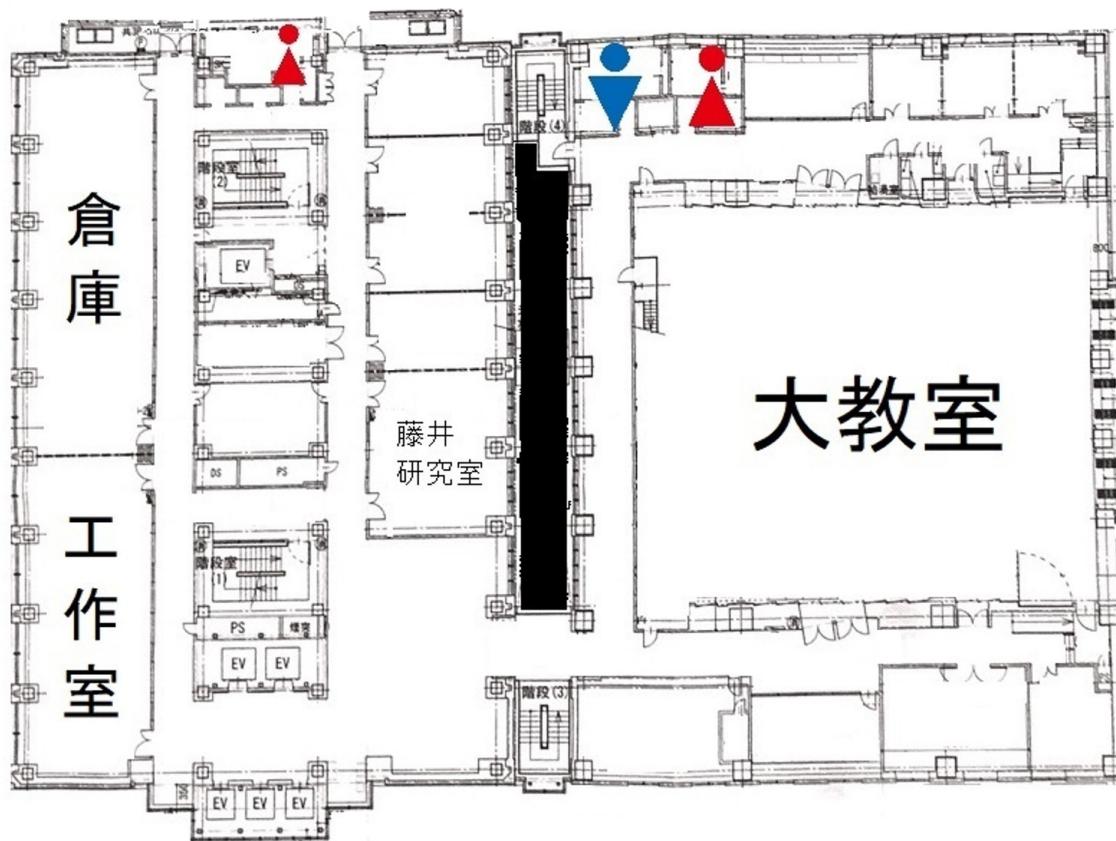


Fig. 5.1: Experiment environment

5.3 実験装置

実験には本研究室で開発をしている ORNE- を用いた。また、機体とその構成、使用した PC のスペックをそれぞれ Fig. 5.2 , Table ?? , Table 5.1 に示す。

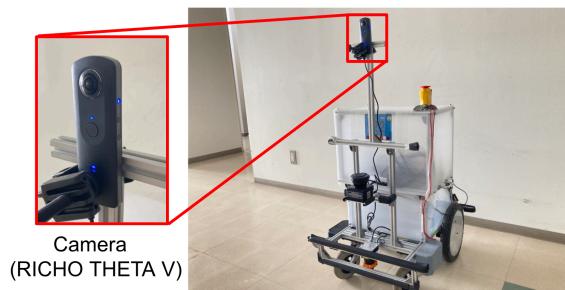


Fig. 5.2: ORNE-

Table 5.1: Specification of PC

CPU	Core i7-9750H(Intel)
RAM	16GB
GPU	RTX 2070 Max-Q

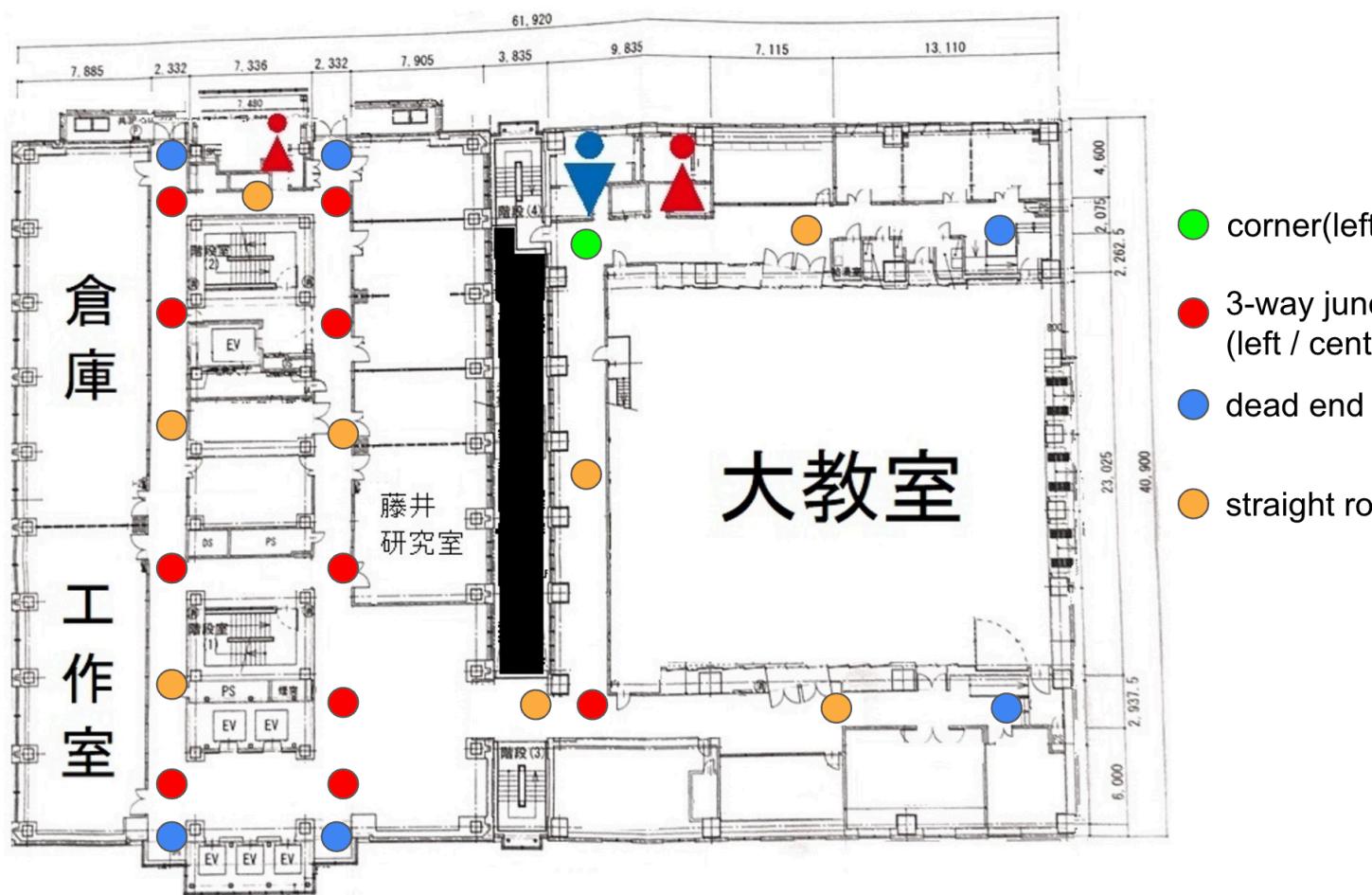
Table 5.2: Specification of ORNE-

item	ORNE-
Depth[mm]	710
Wide[mm]	560
Height[mm]	810
Weight[kg]	20
Wheel diameter [mm]	304
Battery	LONG WP12-12 × 2
Motor	Oriental motor TF-M30-24-3500-G15L/R
Spherical camera	RICOH THETA V

5.4 全天球カメラを用いた通路分類の検証

5.4.1 実験方法

全天球カメラを取り付けた実ロボットを実験環境の 25 地点に移動させ，各地点で全天球カメラにより取得した画像データを用いて，提案手法により通路の特徴分類ができるかどうかを検証する。Fig. 5.5 の丸い図形により示す実験環境の 25 箇所にロボットを移動させ，各地点で取得した全天球カメラ画像から提案した手法により通路の分類を行う。また，先行研究で述べられた通り，人が道案内により移動する際には，向いている方向の情報も必要としているため，同一箇所においてもロボットの向いている方向により分類される通路の特徴が変化する場合，ロボットの向きを変え，各向きを向いた状態を 1 例とし，正しく通路分類が行えるかどうか検証する。本実験の環境では，同一箇所においてもロボットの向きで通路の特徴が異なる箇所が 箇所存在したため，全 例の通路分類ができるかを検証する。ロボットの移動にはジョイスティックコントローラーを使用する。実験の条件として，突き当たりと通路の判定する境界は，5 m 先が行き止まりかどうかにより判断を行うこととする。



5.5 25 places to place robots in aisle classification experiments

5.5.1 結果と考察

実験の結果。例中 例で正しく通路の特徴を分類することができた。正しく分類ができた例の一部を Fig に示す。また、分類に失敗してしまった例の一部を Fig ~ に示す。今回、分類に失敗した原因として、 ということが考えられる。

第6章

まとめ

謝辞

本研究を進めるにあたり、熱心にご指導を頂いた林原靖男教授に深く感謝いたします。また、島田先輩には研究を引き継がせていただき、多くの知識や経験をもとに研究のサポートをしていただきました。また、高橋先輩にも多くのサポートをいただきました。日常の議論を通じて多くの知識や示唆を頂いたロボット設計制御研究室の皆様に謝意を表します。

参考文献

- [1] 島田滉己 , 上田隆一 , 林原靖男 , " トポロジカルマップを用いたシナリオによるナビゲーションの提案 ー人が道案内に用いる情報の取得と評価ー ", 日本機械学会ロボティクス・メカトロニクス講演会'20 予稿集 , 2P1-K02 (2020)
- [2] 島田滉己 , 上田隆一 , 林原靖男 , " トポロジカルマップを用いたシナリオによるナビゲーションの提案 ーシナリオに基づく実ロボットのナビゲーションー ", 1H2-04 , SI2020 (2020)
- [3] Chen,Tongtong,etal. " Lidar-based long range road intersection detection. " 2011 Sixth International Conferenceon Imageand Graphics.IEEE,(2011)
- [4] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [5] Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- [6] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).
- [7] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection." arXiv preprint arXiv:2004.10934 (2020).
- [8] 澤橋遼太 , 細田佑樹 , 町中希彰 , 山崎亮太 , 定國裕大 , 草刈亮輔 , 黒田洋司 (2018)Edge-Node-Graph 及び分岐点検出に基づく道なり走行ナビゲーションシステムの開発 第 23 回ロボティクスシンポジア No.18-3 , pp.67-72
- [9] YOLOv5: <https://github.com/ultralytics/yolov5>
- [10] 国土地理院: 基盤地図情報ダウンロードサービス, <https://maps.gsi.go.jp/development/ichiran.html>
- [11] LeCun , Yann , et al . "Gradient-based learning applied to document recognition ." Proceedings of the IEEE 86 . 11 (1998): 2278-2324 .