

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-

指導教員 林原 靖男 教授

2021年12月31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	7
1.1 背景	7
1.2 目的	8
1.3 関連研究	9
1.4 本論文の構成	9
第 2 章 要素技術	10
2.1 トポロジカルマップ	10
2.2 シナリオ	11
2.3 Neural Network	13
2.4 Convolutional Neural Network(CNN)	14
2.5 You Only Look Once(YOLO)	15
第 3 章 提案手法	16
3.1 提案手法の概要	16
3.2 画像の前処理	18
3.3 学習用データの作成	19
3.4 学習の結果	20
第 4 章 実験	21
4.1 実験の手順	21

4.2	実験装置	22
4.3	実験 1 事前取得のデータを用いた通路認識の検証	24
4.4	実験 2 ナビゲーション適用時の通路認識の検証	25
第 5 章 まとめ		26
謝辞		26

図目次

1.1	Example of spherical camera image of equirectangular projection	8
2.1	Example of metric and topological map	11
2.2	An example of human directions	12
2.3	Neural network architecture	13
2.4	Architecture of CNN from[9]	14
2.5	The YOLO Detection System.(出典:[3])	15
3.1	Type of passage	17
3.2	Flow of passage recognition method	17
3.3	Preprocessing of spherical camera images	18
3.4	An example of a dataset.	19
4.1	Experiment environment	22
4.2	ORNE-	23

表目次

3.1	Class name to be labeled	19
3.2	20
4.1	Specification of PC	23
4.2	Specification of ORNE-	24

概要

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 ー全天球カメラ画像に基づく通路認識手法の提案ー

人の道案内のようにロボットを目的地まで移動させる手法として、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。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特徴抽出に失敗することにより、ナビゲーションにも失敗するということがあった。通路の特徴抽出には LiDAR を用いており、センサが開いているドアなどに反応してしまうことで誤認識が発生した。そこで、本研究では通路の特徴情報を抽出する手法として、全天球カメラ画像を用いた手法を検討する。具体的には、YOLO を用いて全天球カメラ画像から通路を検出し、画像中の通路の位置座標に基づくことにより通路の特徴を認識する。初めに、通路のデータセットを用いて学習モデルを作成し、この手法により通路の特徴を検出できるか検証を行う。その後、先行研究の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に提案手法を適用し、ナビゲーションの成功回数を先行研究の結果と比較することにより、提案した手法の有効性の検証を行う。

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中になかった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，人が道案内で移動するには通路の特徴（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言葉に表現したもの的形式を決め，実ロボット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，通路の特徴の認識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，誤認識が起き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。また，この手法では通路の特徴の認識にLiDARを使用しており，誤認識は開いているドアや隙間にLiDARが反応したことが原因であると述べられている。そこで，通路の特徴認識にカメラ画像を用いることで誤認識を解消し，ナビゲーション途中に経路から外れるという問題を解決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特徴認識の手法を提案する。そして、先行研究により提案された実ロボットを用いたトポロジカルマップとシナリオに基づくナビゲーションに対し、認識した通路の特徴情報を適用することで本手法の有効性を検証する。検証の際は、本研究と先行研究のナビゲーション結果に着目し、その成功回数を比較することとする。

また、本研究では Fig. 1.1 に示すような、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: Example of spherical camera image of equirectangular projection

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。

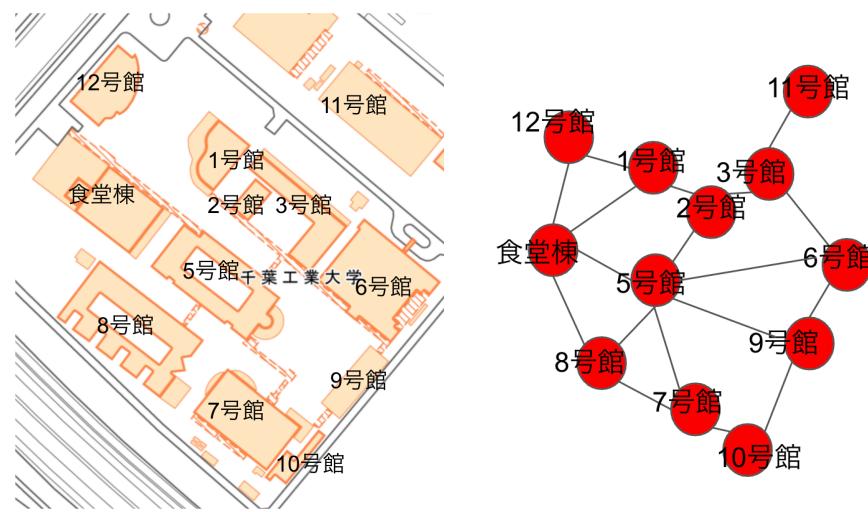
第2章

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. 2.1(a) 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. 2.1(b) 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. 2.1(b) 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 [8]

Fig. 2.1: Example of metric and topological map

2.2 シナリオ

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。」のようになる。また、シナリオは「次の角まで」や「突き当たりまで」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、行動の後には「.」をつけて区切る必要がある。

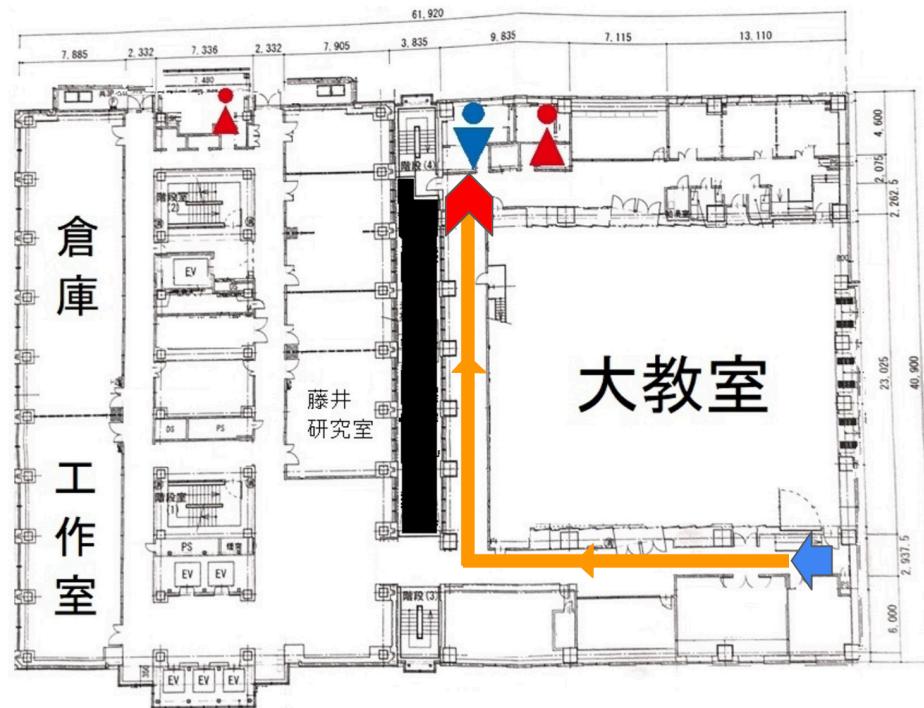


Fig. 2.2: An example of human directions

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。

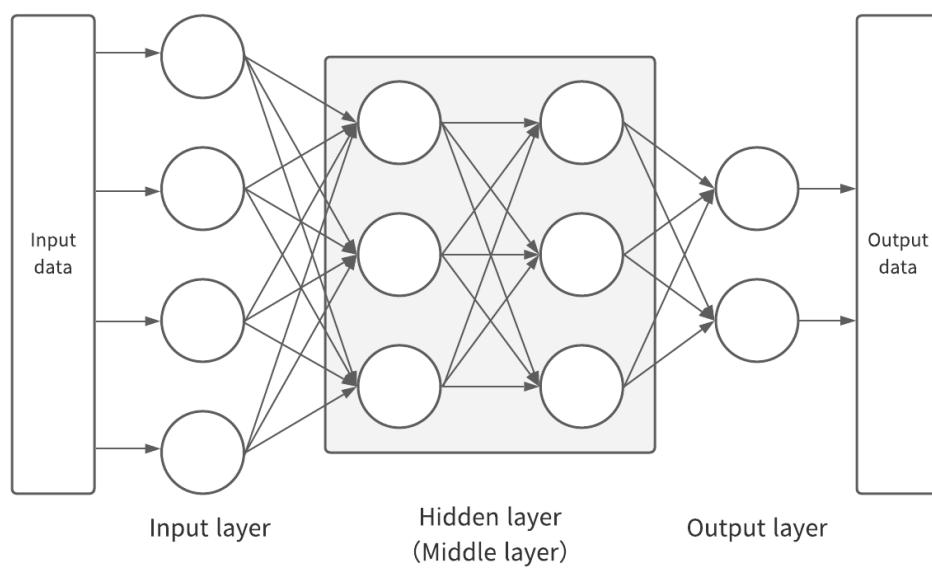


Fig. 2.3: Neural network architecture

2.4 Convolutional Neural Network(CNN)

CNNはディープラーニングに手法の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[9]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、その時のネットワークを Fig. 2.4 に表す。

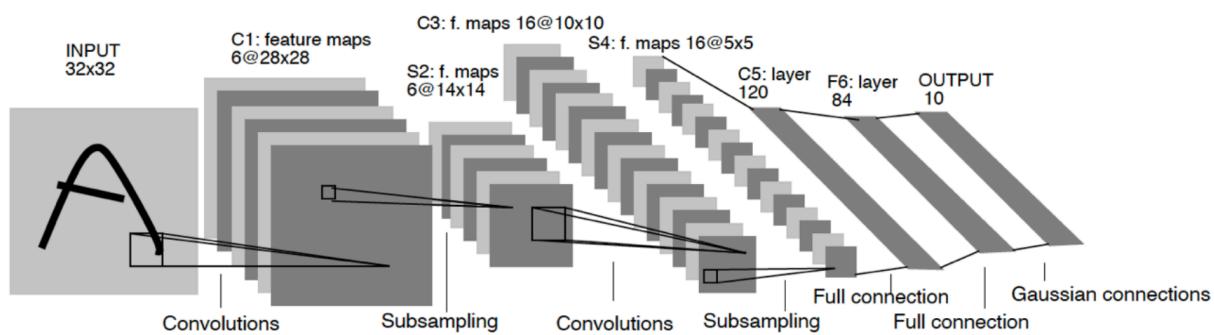


Fig. 2.4: Architecture of CNN from[9]

2.5 You Only Look Once(YOLO)

本研究で用いる YOLO[3][4][5][6] 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。

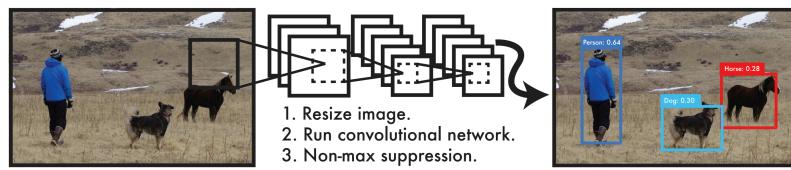


Fig. 2.5: The YOLO Detection System.(出典:[3])

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

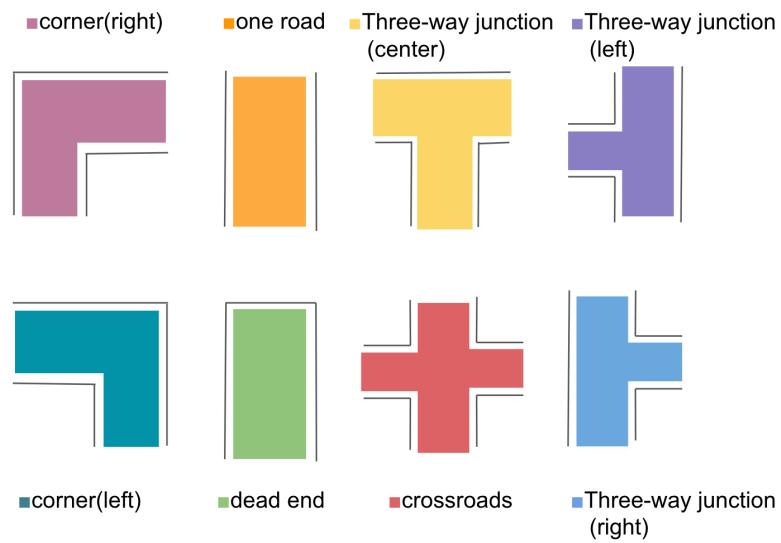
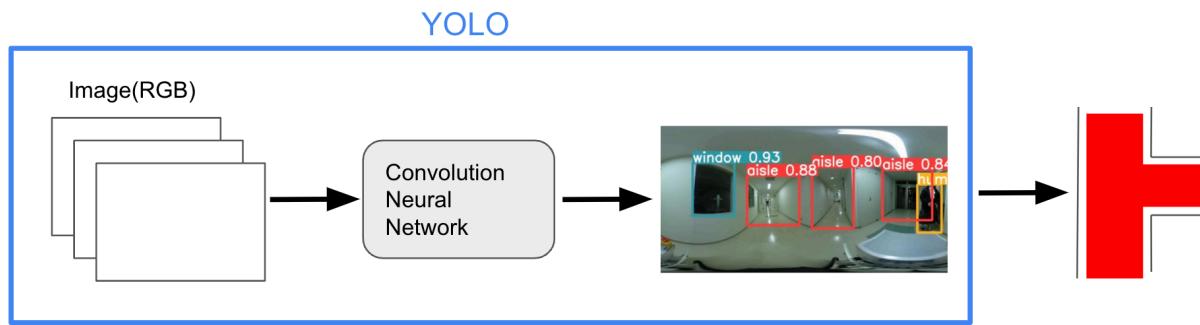
提案手法

3.1 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，通路認識を行う手法の検証を行う．ここでいう通路認識とは，画像中に通路があるかどうかを認識するだけでなく，画像中に通路がある場合，その通路がどのタイプに属するのかといった，通路の特徴を抽出することである．また，本研究で扱う通路のタイプは以下の8種類である．それぞれの図を Fig. 3.1 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)

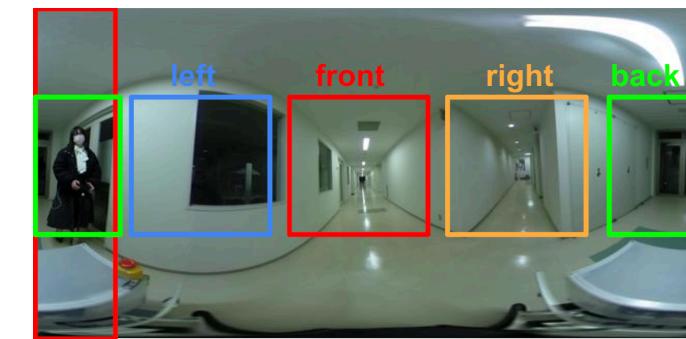
本手法の通路認識の流れを Fig. 3.2 に示す．まず，全天球カメラにより取得したRGBの画像配列データを YOLO の入力とし，出力されたバウンディングボックスとクラス確率の情報をもとに画像中に通路があるかどうかを検出する．画像中に通路が検出された場合，通路の座標位置に基づき，通路のタイプを決定する．

**Fig. 3.1:** Type of passage**Fig. 3.2:** Flow of passage recognition method

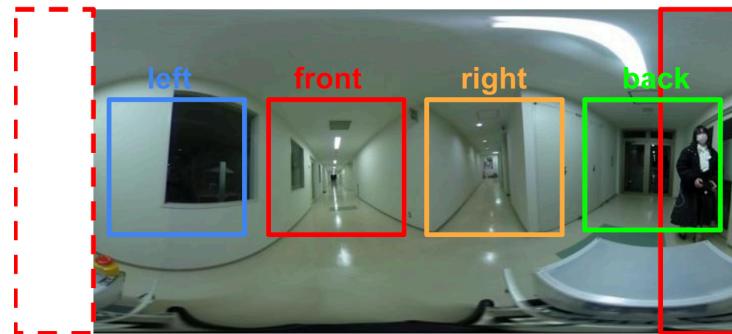
この例では、YOLO の出力した画像から通路が 3 つ検出されている。
 また、通路はロボットに対して前方、後方、右手側に検出されているため、
 通路のタイプは⑥右に曲がれる三叉路であるということがわかる。

3.2 画像の前処理

全天球カメラで取得した画像は通常、Fig. 3.3(a) に示すようにカメラの正面が画像の中心となり、後方の物体は画像の左端と右端で見切れてしまう。本手法では前後左右の通路を扱うため、見切れている状態は通路の認識精度に影響を与える恐れがある。そこで、YOLO に入力する前の段階で Fig. 3.3(b) に示すような画像の前処理を行う。画像の前処理としては以下のことを行う。
・画像のピクセル数を減らすため、 960×1920 から 320×640 の大きさにリサイズする
・カメラの前後左右全ての通路が画像内に収まるように、画像の左端 $1/8$ を切り取り右端にスライドさせる



(a) no processing image



(b) preprocessing image

Fig. 3.3: Preprocessing of spherical camera images

3.3 学習用データの作成

学習モデル作成のため，自作のデータセットにより学習を行う．学習に用いるデータは津田沼キャンパス2号館3階により収集し，作成したデータセットの一部を Fig. 3.4 に示す．データセットのラベル数は11クラスで設定し，クラス名を Table 3.1 に示す．

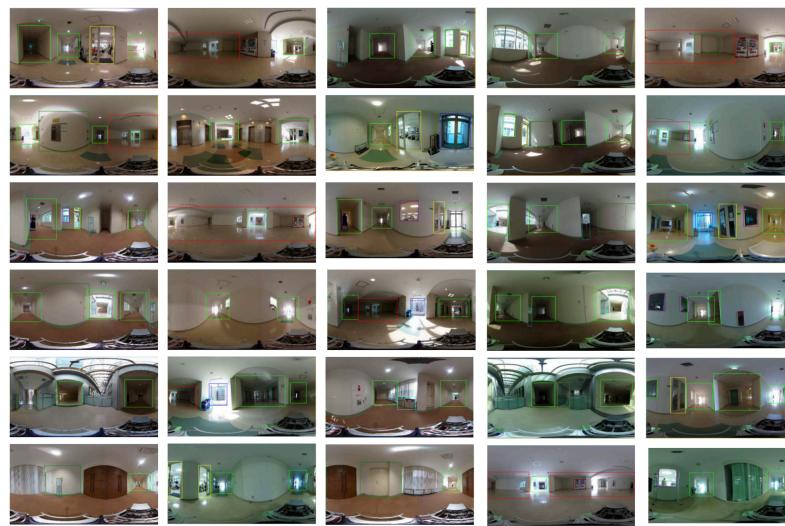


Fig. 3.4: An example of a dataset.

Table 3.1: Class name to be labeled

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

3.4 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セットの構成を Fig. 3.2 に示す。

Table 3.2

training	3540
test	126
validation	14
epoch	300
batch	8

第4章

実験

4.1 実験の手順

提案した通路認識の手法を 2 つのフェーズに分けて検証する。1 つ目のフェーズでは事前に実験環境で取得した画像データを用いて通路の認識ができるか検証する。また、2 つ目のフェーズでは機体に搭載したカメラからリアルタイムに取得する画像データを用いて通路認識を行い、トポロジカルマップとシナリオによるナビゲーションに適用し、先行研究の LiDAR を用いた手法と比較して有効性を検証する。実験環境は Fig. 4.1 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。

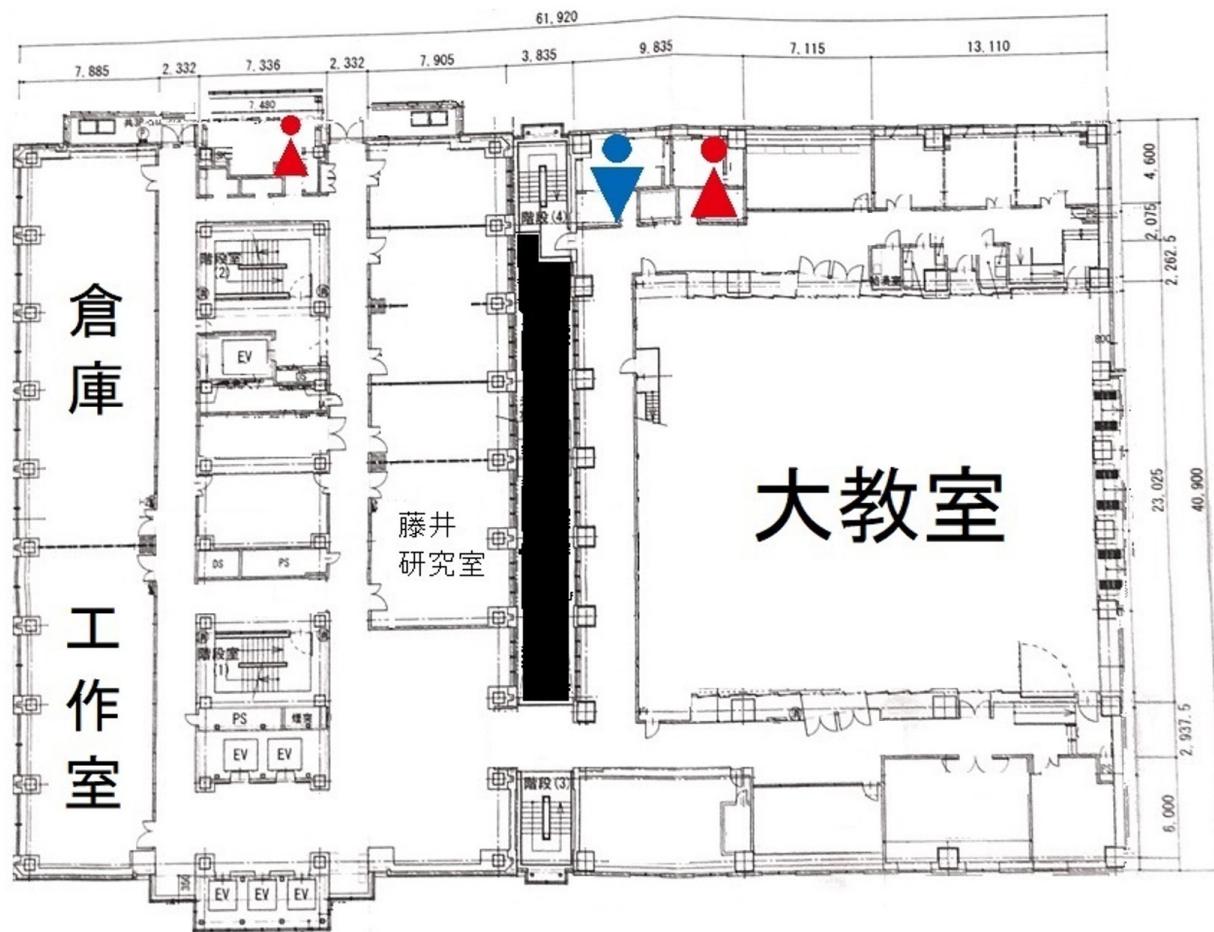


Fig. 4.1: Experiment environment

4.2 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. 4.2 , Table ?? , Table 4.1 に示す。



Fig. 4.2: ORNE-

Table 4.1: Specification of PC

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

Table 4.2: Specification of ORNE-

item	ORNE-
Depth[mm]	710
Wide[mm]	560
Height[mm]	810
Weight[kg]	20
Wheel diameter [mm]	304
2D LiDAR(front)	Hokuyo URM-40LC-EW
2D LiDAR(back)	
IMU	Analog devices ADIS16465-2BMLZ
Battery	LONG WP12-12 × 2
Motor	Oriental motor TF-M30-24-3500-G15L/R

4.3 実験 1 事前取得のデータを用いた通路認識の検証

4.3.1 実験目的

実験 1 では、あらかじめ実験環境で取得していた画像データを用いて通路の認識ができるか検証する。

4.3.2 実験方法

実験環境で取得した Fig. ?? に示すような画像を YOLO に入力し、出力されたバウンディングボックスの情報とクラスの属性情報をテキストファイルに保存する。次に、保存したテキストファイルを読み込み、通路のタイプを判定する。用意した画像 12 枚を上記の方法で 1 枚ずつ読み込み、そこに写っている通路の認識を行う。

4.3.3 結果

4.3.4 考察

4.4 実験2 ナビゲーション適用時の通路認識の検証

4.4.1 実験目的

4.4.2 実験方法

4.4.3 結果

4.4.4 考察

第5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 一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] 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.
- [4] Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- [5] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).
- [6] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection." arXiv preprint arXiv:2004.10934 (2020).
- [7] 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
- [8] 国土地理院 : 基盤地図情報ダウンロードサービス,
- [9] LeCun , Yann , et al . "Gradient-based learning applied to document recognition ." Proceedings of the IEEE 86 . 11 (1998): 2278-2324 .