

摘要

近年の建築では BIM (Building Information Modeling) と呼ばれるコンピュータ上に現実と同じ建物の立体モデルを再現し，可視化するワークフローが注目されている．従来の配管 BIM は高精度な Lidar センサを用いて配管モデルの推定を行なわれていたが，振動に弱く高価である．そのため，Lidar センサより安価である RGBD カメラを採用する．本研究は従来の点群データのみを用いた 3D 再構築を行わず，取得画像と関連する点群データに基づき配管の 6D 姿勢推定を行い 3D 再構築の高精度化を目的とする．また，姿勢推定の従来の方法では手作業で行われていたが，深層学習を取り入れることにより高精度かつ高速化を図る．

謝辞

本稿の内容は，著者が京都大学大学院工学研究科機械工学専攻メカトロニクス研究室（吉川研）での博士後期課程において，そして 2007 年度現在，講師として在籍している立命館大学理工学部ロボティクス学科において学んだ内容をまとめたものです．あまり明文化されない，しかし非常に重要なこのような知識を与えて頂いた，吉川研と立命館大学ロボティクス学科のすべての方々に，深く感謝の意を表します．

なお，本稿の表紙には，指導教員として川村貞夫教授の名前がありますが，これは修士論文，卒業論文のテンプレートとして出力されたものです．したがって，本稿の文責は川村教授ではなく，すべて著者（金岡）にあることをここに明記しておきます．

目次

第 1 章	序論	1
1.1	研究背景	1
1.2	既存研究	2
1.3	研究目的	3
1.4	本手引の構成	4
第 2 章	深層学習による配管 6D 姿勢推定	5
2.1	配管 6D 姿勢推定方法	5
2.2	配管の特徴を活かしたアイソメ図作成方法	6
2.3	ネットワーク構造	6
2.3.1	全体構成	6
2.3.2	RXD ネットワーク	6
第 3 章	配管データセット	9
3.1	使用機材	9
3.2	物体検出のデータセット収集	10
3.3	6D 姿勢推定のデータセット収集	11
第 4 章	実験	13
4.1	評価指標	13
4.2	結果と考察	14
4.2.1	物体検出	14
第 5 章	結言	15
	参考文献	17
付録 A	数式の記述	19
A.1	記述例	19

目 次

1.1	Gen6D ネットワーク構造	3
1.2	YOLO モデルの検出の流れ	3
2.1	RGB-D カメラを用いた深層学習によるアイソメ図作成方法	5
2.2	配管の検出例	6
2.3	RXD ネットワーク構造図	7
2.4	Sigmoid 関数のグラフ	8
2.5	RXD ネットワーク構造図	8
3.1	暗闇での RGB-D カメラの撮影	9
3.2	Intel Realsense L515	10
3.3	Colmap を用いた曲管の点群データ	11
4.1	Intersection over Union(IOUS)	13
4.2	適合率と再現率	14

表 目 次

4.1	物体検出ネットワークの実行結果	14
4.2	テスト画像を用いた検出結果	14

第 1 章 序論

本章では、本研究の背景と目的、そして構成について述べる。

1.1 研究背景

配管は気体、液体、粉粒対などの流体を輸送や配線の保護などを目的とする管のことである。例を挙げると、電気配線やケーブルを保護する電気配管や生活に必要な水を家庭、学校などに輸送する水道管などに使用されている。この配管は運用と保護を確保するために、耐久性と安全性を常に保ち続ける必要がある。

BIM とは、Building Information Modeling の略称で、建築物や土木構造物などの情報をコンピュータ上に現実と同じ建物の立体モデルを形成し、設計から維持管理までのプロセスをデジタル化する新しいワークフローの一環である。この BIM モデリングはこれまでの 3 D モデリングとは大きく異なる。3 次元モデルは 2 次元上で図面を作ってから 3 次元の形状を組み立てシミュレーションするという流れが主流であった。そのため、3 D モデルに修正点があった場合に、2 D の図面を全て修正してから構築する必要があり効率的ではなかった。しかし、この BIM モデルは一つのデータを修正すると全てのデータが連動し、関係する図面の該当箇所が自動修正されるようになるため、従来の 3 次元モデルよりも高効率で作業を行うことができる。

本研究においては建築物の中でも配管に焦点をあて、配管のアイソメ図を作成することが最終目標である。アイソメ図を取得するためにはこれまでに Light Detection and Ranging(LIDAR) センサーと呼ばれるレーザー光を使用して離れた場所にある物体の形状や距離を測定できるセンサーを使用していた。このセンサーは精度が非常に良いというメリットはあるが、他のセンサーと比較すると高価であるというデメリットを抱えている。そのため、一般的に使用するためには安価なセンサーでデータ収集ができることが望まれる。また、従来の方法ではセンサーから取得されたデータを手作業で配管を選択することでアイソメ図を作成していた。手作業による手法は効率が悪いことから、自動で配管を認識できるシステムが求められる。このような背景から本研究では LIDAR センサーの代わりに RGB-D カメラを用いた深層学習によるアイソメ図の作成を目指す。RGB-D カメラは Depth 画像を使用できることから LIDAR センサー同様に距離情報を取得でき、配管の距離を算出することが可能である。また、LIDAR センサーと比較すると約 40 倍ほど安価であるため、たくさんの方が利用しやすいメリットがある。また、深層学習を取り入れることで自動で配管の認識を行い高効率化を目指す。

1.2 既存研究

アイソメ図を作成するにあたって曲管とT字管の6 D 姿勢を認識する必要がある。6 D 姿勢は座標 X、Y、Z に加え Yaw、Pitch、Roll の回転情報を加えたものである。その6 D 姿勢推定を行うネットワークである Generalizable Model-Free 6-DoF Object Pose Estimation from RGB Images(Gen6D) を紹介する。姿勢推定に必要な主なデータセットは3次元データやカラー画像、深度データなどが代表的である。しかし、3次元データを生成するためにはオブジェクトの3Dモデルを作る必要があるため、実用化するのは困難である。Gen6D ネットワークはデータセットを3 D データを必要とせずカラー画像のみで物体の姿勢推定を行うことができる。Gen6D のネットワーク構成について図3に示す。まず、Detector と呼ばれる工程では参照画像の情報をもとに認識したいオブジェクトの領域を検出する。これはCNN と呼ばれる畳み込みニューラルネットワークを用いて画像の特徴を抽出し、得られた特徴をもとに領域を予測する。次の工程である、Selector では Detector で得られた領域の画像と最も近い視点を持つ参照画像を複数枚ある中から1つ抽出する。これは選択された参照画像の視点をテスト画像の視点とほぼ同様とみなし、誤差は生じますがオブジェクトのポーズの初期姿勢を形成する。最後の工程では先程得られた姿勢の改良を試みる。まず、参照画像から近い視点の画像をさらに6枚選択し、全参照画像間の平均と分散を算出し、初期に求められた姿勢の情報を改善して最終的な結果を予測する。この研究のメリットとしてRGB画像のみを用意することで物体の姿勢を推定できるため、データセットの作成は非常に容易である。しかし、この Gen 6 D をしようするにあたって問題点が2つある。まず、一つ目にRGB画像は距離情報を持たないため、スケール情報を示さなければならないアイソメ図を作成することが不可能である。そのため、Depth画像を用いることでカラー画像に加え、震度情報が追加されるため物体間のスケールを算出することが可能である。2つ目は複数の物体を姿勢推定することができない点である。一般的に配管が設置されている現場では、単体ではなく複数の配管が張り巡らされていることがほとんどであるため、画像内の全ての配管を網羅し認識することが要求される。これを解決するためには Gen 6 D の Detector を複数物体検出可能なネットワークに変更する必要がある。

その Detector を複数物体検出可能にするにあたって、2次元上で複数の物体検出を可能にする YOLOV3 を紹介する。このモデルはほぼ同時期に発表された Fast R-CNN と同様に、物体検出に大きな影響を与えた。これ以降、End-to-End モデルとリアルタイム検出が物体検出の中で主流となった。YOLO の特徴は従来までは境界設定と物体検出を2段階に行っていた作業を一度に行うことで推定速度の高速化を行うことができた。図1.2にYOLOの物体検出までの流れを示す。

まず、入力画像を $S \times S$ のグリッドセルに分割する。物体の中心がグリッドセルに存在していた場合に、そのセルが物体を検出するように学習する。次に、バウンディングボックスの推定では各グリッドセルに B 個のバウンディングボックスを持ち、それらのボックスの信頼スコアを予測する。信頼スコアとは背景ではなく物体が含まれている確率のことである。次に、各グリッドセルは複数のクラスに対する条件付き確率を予測する算出された条件付きクラス確率と一つ前の個々のバウンディングボックスの信頼スコアを掛け合わせることで、バウンディングボックス毎のクラ

スに対する信頼スコアを得ることができる。このスコアを使用しどのバウンディングボックスが正解の物体を推定しているのかを判断する。

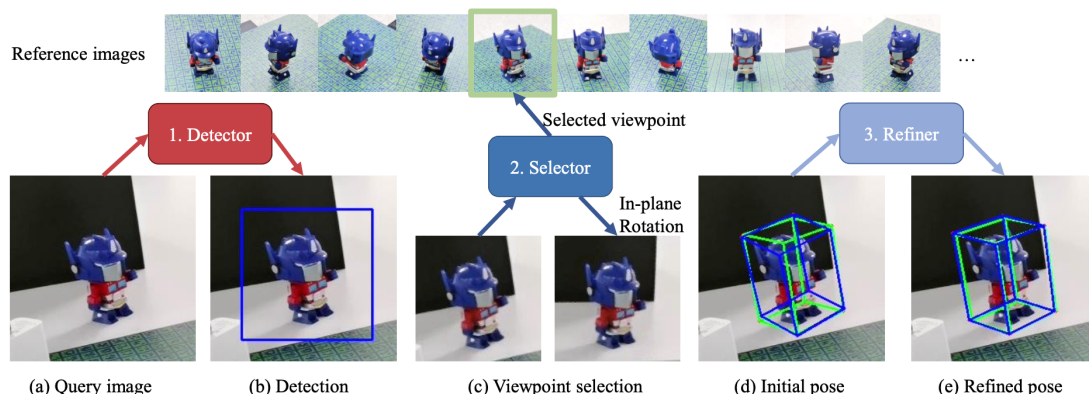


図 1.1: Gen6D ネットワーク構造

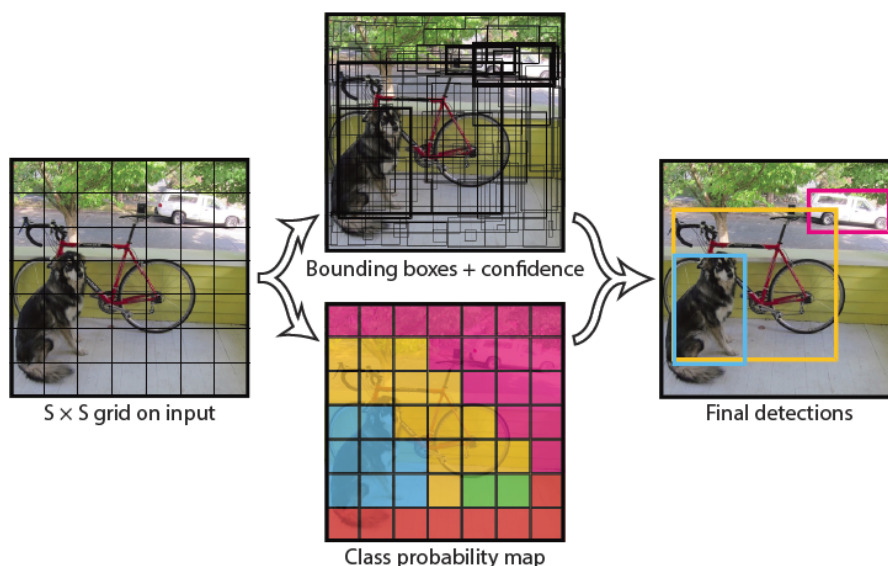


図 1.2: YOLO モデルの検出の流れ

1.3 研究目的

本研究ではRGBD データを使用した深層学習による配管 6D 姿勢推定を行い、RGBD カメラを用いることによる安価な機器での姿勢推定の実現を試みる。また、既存の RGB 画像のネットワークに Depth 画像を組み込んだモデルを提案し、認識精度向上と推定速度の高速化を目標とする。本研究の貢献は以下になる。まず一つ目は深層学習による RGB 画像と Depth 画像より物体検出ネットワークの提案である。特に、RGB 画像と Depth 画像からそれぞれ抽出された特徴を結合する RxDLayer を導入し、ネットワーク性能の向上させた。他のネットワークと精度を比較することで提案ネットワークの有効性を検証した。

2 つ目は既存の 6 D 姿勢推定ネットワークに複数物体検出を可能にさせたことである。配管は単体ではなく複数の管が張り巡らされているため、複数の認識を可能にする必要がある。提案ネットワークでは画像内部にある配管全てを網羅し、それぞれの物体の中心ピクセル座標とスケールを推定することができる。3 つ目は本研究の最終目的であるアイソメ図を作成するにあたっての必要不可欠な配管距離測定である。アイソメ図は配管の向きだけでなく、距離情報を図面に示す必要がある。そのため、Depth 画像を用いることでネットワークによって認識された情報をもとに、配管の距離情報を算出することを可能にした。

1.4 本手引の構成

本論文の構成は以下のようになる。第一章では研究背景、既存研究、研究目的について述べる。研究背景では、Building Information Modeling (BIM) についてや RGB-D カメラを使用するメリットについて述べる。既存研究では、配管 6 D 姿勢推定と物体検出のそれぞれのネットワークを紹介する。それぞれのネットワークのメリットやデメリットをもとにアイソメ図に必要となるネットワーク設計を述べる。研究目的では、本研究の目的及び貢献について述べる。第 2 章では、論文の論理構成をどのようにすべきかを述べる。第 3 章では、論文の体裁について述べる。第 4 章では、論文作成および研究一般におけるその他の重要な注意点について述べる。第 5 章は結言である。

第 2 章 深層学習による配管 6D 姿勢推定

従来のアイソメ図取得には LIDAR センサーにより 3 次元点群を取得し図面を作成していたが、センサーが高価であるというデメリットを抱えていた。そのため、本研究では LIDAR センサーよりも比較的安価な RGB-D カメラを用いてデータセット収集から深層学習やアイソメ図作成までの流れを紹介する。第 2.1 節では RGB-D カメラを用いた深層学習による配管 6D 姿勢推定の手順を述べる。第 2.2 節では物体認識のネットワーク設計を詳しく説明する。

2.1 配管 6D 姿勢推定方法

RGB-D カメラを用いた深層学習による配管のアイソメ図作成の手順はデータ収集、物体検出、6D 姿勢推定、アイソメ図作成の 4 つに分けられる。図 2.1 にそのプロセスの流れを示す。

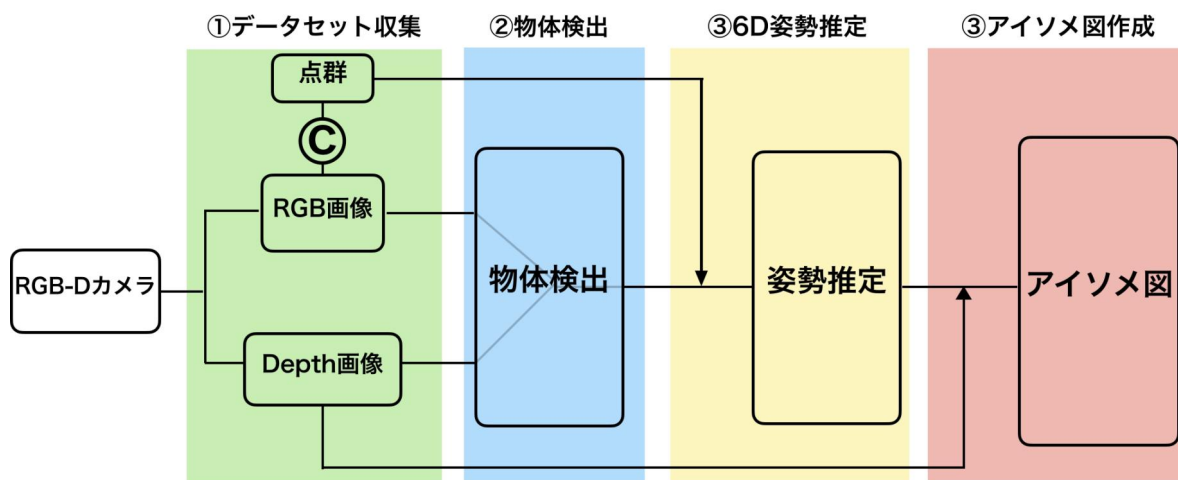


図 2.1: RGB-D カメラを用いた深層学習によるアイソメ図作成方法

まず、RGB-D カメラを用いて RGB 画像と Depth 画像を取得する。これらの画像を使用して物体検出ネットワークより複数の配管検出を行う。次に、物体検出で認識された配管の姿勢を推定する。その際に、データセットとして RGB 画像から 3 次元復元ツールである Colmap を使用して点群データを取得する。この Colmap は Structure from Motion(SfM) という技術で異なる視点からの写真を使用して 3 次元形状を復元する写真解析ソフトウェアである。Colmap から得られた点群データをもとに姿勢推定問題に取り組む。最後に姿勢推定された結果を用いてアイソメ図を

作成するが、図面には配管の距離情報を示す必要がある。そのため、Depth 画像を用いることで認識された配管のスケールを算出することができる。以上のステップを踏むことで RGB-D カメラからアイソメ図を作成することができる。

2.2 配管の特徴を活かしたアイソメ図作成方法

配管の形状には配管固有の特徴が存在する。図 2.2 に一部配管の例を示す。配管は両端部分を除くと直線であるという特徴があるため、両端の曲管又は T 字管がどの方向を向いているのかを認識できればその間を直線で結ぶことでアイソメ図を作成することができる。そのため、本研究においては配管全体を認識するのではなく、画像内の曲管及び T 字管を検出し姿勢を推定する。それに加え、両端の曲管又は T 字管間の距離を Depth 画像を用いて算出することで図面を作成することができる。

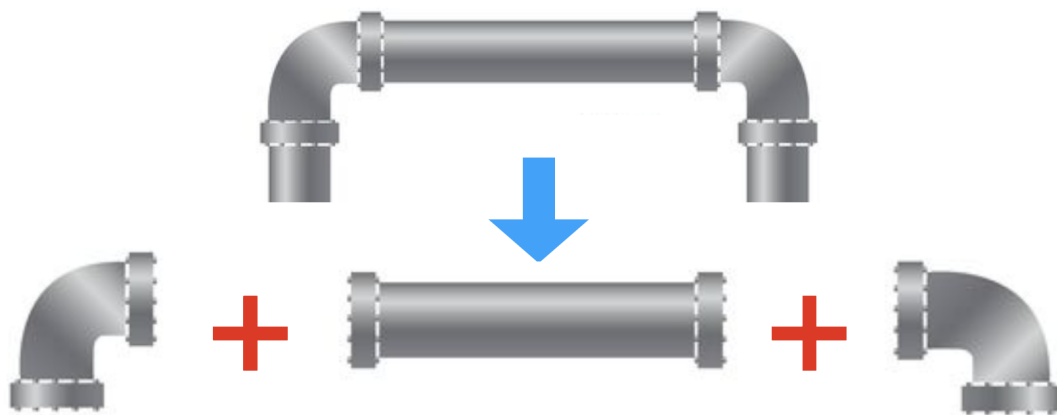


図 2.2: 配管の検出例

2.3 ネットワーク構造

2.3.1 全体構成

本研究では曲管及び、T 字管の物体検出と 6D 姿勢推定を深層学習を用いて推定する。その全体構成を図 2.3 に示す。物体検出ネットワークでは RXD Layer 用いた RXD ネットワークを提案する。RGB 画像と Depth 画像を RXD Layer に挿入することでそれぞれの画像から抽出された特徴を結合し認識精度向上を図ることができる。また、姿勢推定ネットワークでは Gen6D モデルを使用する。Gen6D は複数物体の検出ができないが、RXD ネットワークから認識された複数の配管のピクセル座標とスケールを Gen6D の Selector にそれぞれ渡すことで Refiner を通して最終的に画像内の全ての曲管及び T 字管の姿勢を求めることが可能になる。

2.3.2 RXD ネットワーク

物体検出に使用する RXD ネットワークについて紹介する。RXD ネットワークの構成を図 2.4 に示す。RGB カメラから取得された RGB 画像と Depth 画像をそれぞ

れ Convolutional set に挿入することで特徴を抽出することができる。深層学習は基本的に層を深くすることで認識精度が向上するため、RXD ネットワークでは複数畳み込み層を使用している。また、畳み込みの際には Batch Normalization(BN)、ReLU、Max Pooling(MP) を各層に取り入れている。Batch Normalization は各バッチのデータを使用し正規化を行う。その結果、出力が適度に分散され、勾配消失などの問題が起こりにくくなり、学習が適切に進む。特に深めのネットワークを使用したときに、数カ所に挟むことで効果を得る。次に Max Pooling とは CNN で用いられる基本的なプーリング層である。最大値プーリングではカーネル内の最大値のみを残すプーリング処理である。これらの層を複数利用することで特徴をより濃くすることができる。

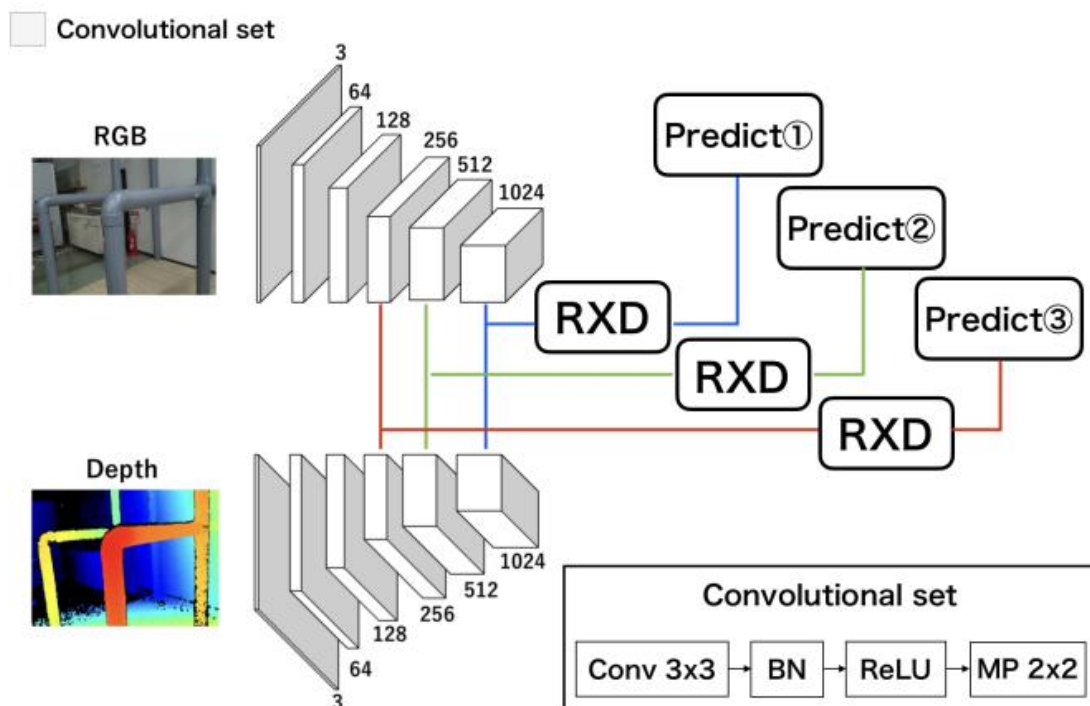


図 2.3: RXD ネットワーク構造図

RXD 層の内部構成を図 2.5 に示す。RXD 層の中身では RGB 画像と Depth 画像が RXD ネットワークで畳み込まれたデータを結合する役割を担っている。まず、RGB 画像と Depth 画像をそれぞれ畳み込み特徴マップを取り出す。それらのデータを Concatenate 関数を用いて連結させる。次に、結合されたデータを畳み込んだあとは Sigmoid 関数という活性化関数を使用する。通常の活性化関数には ReLU 関数を使用され、入力が 0 以下の時は 0 を、0 より大きい時はその値を出力する関数である。Sigmoid 関数は ReLU 関数とは違い、入力値 x の値に依らず、0 ~ 1 の数値に変換して出力する。次に、Sigmoid 関数によって出力された値をそれぞれもとの RXD ネットワークから入力されたデータと乗算する。このステップにより RGB 画像と Depth 画像の相関を利用し特徴マップの表現力を強化することができる。これによって得られたそれぞれの値を Concatenate 関数を用いることで結合し、2 度の畳み込み層を経ることで出力結果が求められる。

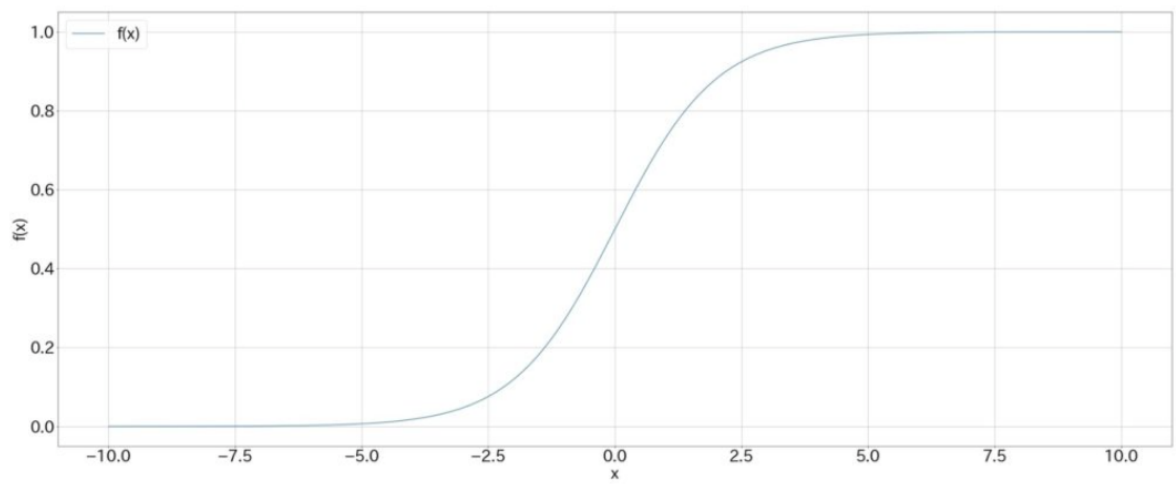


図 2.4: Sigmoid 関数のグラフ

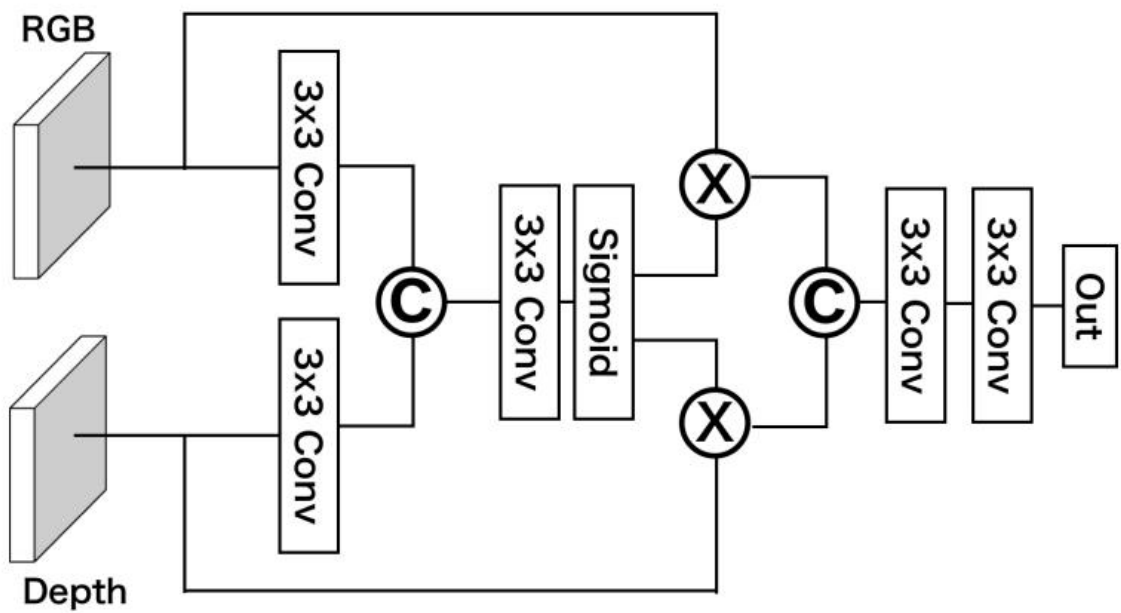


図 2.5: RXD ネットワーク構造図

第 3 章 配管データセット

前章では、アイソメ図を作成するためのネットワーク構造を提案した。本章では、深層学習に用いるデータセットの収集及び、作成方法について論じる。

3.1 使用機材

データセットの取得には RGB-D カメラを使用する。従来の方法では LIDAR センサーを用いて配管の 3 D データやアイソメ図を作成していた。しかし、LIDAR センサーは高価であり、一般的に使用することが困難であるという欠点を抱えていた。そのため、RGB-D カメラは LIDAR センサーよりも比較的安価であるため本研究のデータセット収集に使用する。次に、RGB カメラではなく RGB-D カメラを使用する利点を紹介する。RGB-D カメラの Depth 画像にはたくさんのメリットが存在する。まず、1 つ目に Depth 画像には距離情報を取得できるという点である。配管のアイソメ図には配管のそれぞれの部位の長さを正確に示す必要がある。そのため、RGB 画像には距離情報を含まれていないことからスケールを求めるには Depth 画像が重要になるのだ。2 つ目に光の明暗に影響されない点である。RGB 画像は撮影する環境が暗闇の場合、画像には何も映らない。これは RGB 画像が光に反射された物体の度合いを数値化しているため、極端に明るすぎたり暗すぎると RGB 画像が活用できなくなる。特に配管が設置されている地盤地下や天井裏などの照明を当てることが困難な環境では Depth 画像が必要になる。3 つ目に配管が背景色と同様の色を示していた場合に、区別が容易に可能であるという点である。RGB 画像では色の違いが判断できないが、Depth 画像は距離情報の違いを示すことができるため、背景と異なる物体として認識可能になる。以上の点より本研究には RGB-D カメラを採用した。

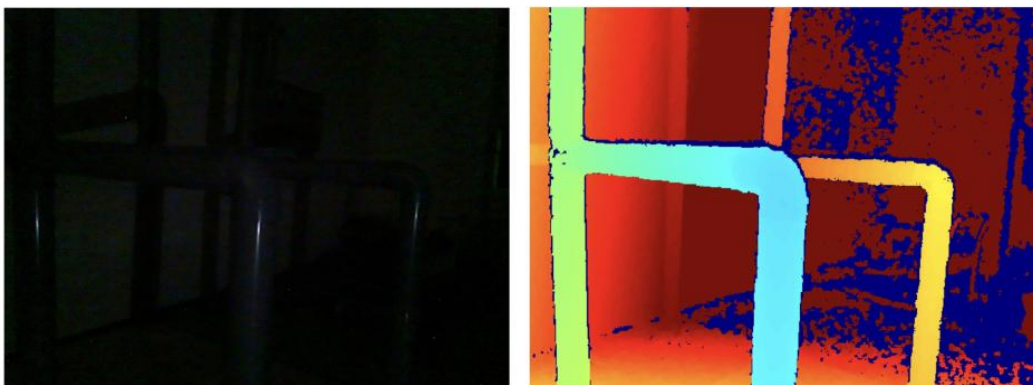


図 3.1: 暗闇での RGB-D カメラの撮影

カメラはインテル社製の Intel Realsense L515 を使用した。Realsense L515 を使用した理由は Realsense カメラの中でも屋内に適した RGB-D カメラであるからだ。このカメラは外光の影響を受けやすいが、屋内の環境であればその影響を受けないため、配管などの室内で多く使用される環境では非常に適していると判断した。



図 3.2: Intel Realsense L515

しかし、Realsense L515 は仕様上、RGB カメラと Depth カメラの位置が異なるため、撮影した際に両方の画像を比較すると画角に差異が生じてしまう。これは、データセットのラベリングを行う際に配管のピクセル座標にそれぞれの画像で異なると認識の精度に大きな誤差が生じてしまう。そのため、Realsense の alignment ライブラリを使用する。これによって両方のカメラの画角をソフトウェア上で位置合わせが可能になる。

3.2 物体検出のデータセット収集

深層学習による認識ネットワークにはデータセットの数量が多いほど精度とロバスト性が向上する。それは様々な場面での配管の写真を学習することによってどの環境においても対応できる汎用性が高まることを意味している。本研究使用するデータセットの一部を図 3.2 に示す。配管には曲管や T 字管や直管が含まれており、この画像内の中から曲管と T 字管を全て認識できることを目標とする。また、Depth 画像の有効性を示すためにテスト画像では暗闇の中に配管を設置したデータセットを用意した。Depth 画像は光の影響を受けにくいことから、暗闇の中でも配管を認識できるかを検証する。収集したデータはラベリング作業を行う。これは深層学習するにおける正解データとして、予め画像内のどの部分が曲管又は T 字管であることをアノテーションする必要がある。本研究では配管画像に対して曲管、T 字管の 2 クラスに分けてラベリング作業を行った。

3.3 6D 姿勢推定のデータセット収集

6D 姿勢推定のデータセットには Colmap を使用して点群データを取得する。Colmap は 2 D 画像から 3 D 点群を再構築するために使用されるソフトウェアである。この 2 D 画像は異なる視点から撮影された同じオブジェクトの画像を複数枚利用することで 3 次元情報を復元することができる。そのため、本研究では曲管と T 字管の周囲をそれぞれ撮影し、Colmap を使用することで点群データを取得した。

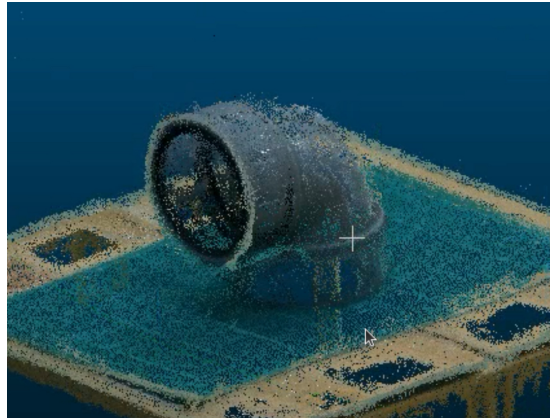


図 3.3: Colmap を用いた曲管の点群データ

第4章 実験

RGB-D カメラから取得したデータセットを RXD ネットワークを使用し曲管又は T 字管を認識できるか検証する。物体認識においては他のネットワークでも実験し、RXD ネットワークの有用性を確かめる。

4.1 評価指標

物体認識の評価指標ではパラメータ数 (Params), Intersection over Union(IoU), mean Average Precision(mAP) を用い認識ネットワークの性能評価を行う。まず、パラメータ数は認識ネットワークの学習可能なパラメータの合計数を示す。これにより、認識ネットワークの複雑度を示すことができる。次に、IoU は正解と予測のバウンディングボックスの共通の重なり部分を 2 つのバウンディングボックスを重ねたときの総面積で除算したものである。IoU は 0 1.0 の値の範囲で示され、値が大きければ大きいほどラベル付されたボックスと予測されたボックスの重なりが正しいことになり、正確に認識していると判断できる。次に、mAP は 1 つ 1 つのクラスに対して平均適合率である AP(Average Precision) を計算する。まず、モデルの予測結果を、出力する信頼度スコア順に並べる。ラベルごとに信頼度スコアがそのラベルの値以上の予測結果について、適合率と再現率を求める。適合率と再現率は図 4.1 のように True Positive(TP) と False Negative(FN) を用いて表される。その適合率と再現率のグラフから適合率の下側の面積を求める。ここで、予測されたラベルが正解なのかの判断は IoU が決められたしきい値以上で、最も信頼度スコアが高い予測ラベルが正解とするように判断される。そして最後に、クラスごとに計算された AP の平均を算出したものが mAP になる。

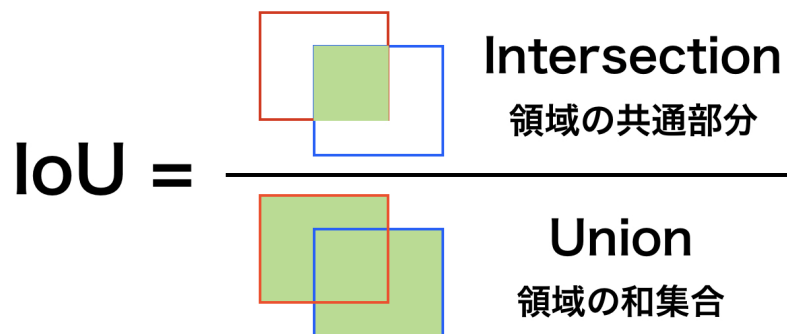


図 4.1: Intersection over Union(IoU)

		予測結果	
		曲菅	曲菅以外
正解	曲菅	TP	FN
	曲菅以外	FP	TN

$$\text{適合率} = \frac{TP}{TP + FP}$$

$$\text{再現率} = \frac{TP}{TP + FN}$$

図 4.2: 適合率と再現率

4.2 結果と考察

表 4.1: 物体検出ネットワークの実行結果

<i>Network</i>	<i>AP</i>			<i>AP50</i>			<i>Parameters millions</i>
	<i>bent</i>	<i>junction</i>	<i>all</i>	<i>bent</i>	<i>junction</i>	<i>all</i>	
YOLOV3	33.9	68.6	51.3	9.95	20.1	15.0	61.5
YOLOV3-Depth	1.3	0.0	0.7	0.4	0.0	0.2	86.3
RXD	70.9	37.2	54.1	20.8	10.9	15.8	32.4

表 4.2: テスト画像を用いた検出結果

<i>AP</i>	<i>AP50</i>	<i>Parameters</i>
-----------	-------------	-------------------

4.2.1 物体検出

第 5 章 結言

本稿では，立命館大学 理工学部 ロボティクス学科 川村・金岡研の修士論文，卒業論文執筆の共通化と効率化のために，論文の論理構造，体裁，その他一般の注意点を明文化した．

途中，読み苦しい部分もあったと思われるが，より体系的で読みやすい手引きとなるよう，改訂を重ねて行く予定である．

参考文献

- [1] 著者：書名．出版社，year.
- [2] 著者：“題名”，掲載誌名，vol. W, no. X, pp. YYY–ZZZ, year.
- [3] Author(s): Book Title. *Publisher*, City of Publisher, year.
- [4] Author(s): “Paper Title,” *Name of Journal*, vol. W, no. X, pp. YYY–ZZZ, year.
- [5] 中島 利勝，塚本 真也：知的な科学・技術文章の書き方．コロナ社，1996.
- [6] 中島 利勝，塚本 真也： \LaTeX ユーザのためのレポート・論文作成入門．共立出版，2002.
- [7] 川村 貞夫：効果的な表現戦略．森北出版，1998.
- [8] 酒井 聡樹：これから論文を書く若者のために．共立出版，pp. 138–140, 2002.
- [9] 有本 卓：ロボットの力学と制御．システム制御情報学会編，朝倉書店，1990.

付録 A 数式の記述

本付録では，数式の記述例を示す．本手引のクラスファイル `kzthesis.cls` 独自のコマンドの使用法の説明も兼ねているので，参考にとすること．

A.1 記述例

まず，受動性の定義を与えておく [9]．システムの入力 u と出力 y が同じ次元であるとする．このとき，ある有限な正の定数 γ_0^2 に対して次式が成り立つならば，システムは受動的であるという．

$$\int_0^t \mathbf{y}^T(\tau) \mathbf{u}(\tau) d\tau \geq -\gamma_0^2, \quad \forall t > 0 \quad (\text{A.1})$$

一方， n 自由度ロボットの動力学は一般に次のような微分方程式に従う．

$$\mathbf{u} = \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \frac{1}{2}\dot{\mathbf{M}}(\mathbf{q})\dot{\mathbf{q}} + \mathbf{S}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) + \mathbf{d}(\dot{\mathbf{q}}) \quad (\text{A.2})$$

ただし， $\mathbf{q} \in \mathbb{R}^n$ は関節変位ベクトル， $\mathbf{u} \in \mathbb{R}^n$ は関節駆動力ベクトルである．右辺第一項は慣性項であり， $\mathbf{M}(\mathbf{q}) \in \mathbb{R}^{n \times n}$ は実対称正定な慣性行列である．第二，三項は非線形項であり， $\mathbf{S}(\mathbf{q}, \dot{\mathbf{q}}) \in \mathbb{R}^{n \times n}$ は歪対称行列となる． $\mathbf{g}(\mathbf{q}) \in \mathbb{R}^n$ はポテンシャル項である． $\mathbf{d}(\dot{\mathbf{q}}) \in \mathbb{R}^n$ は摩擦等による散逸項であり，その各成分は $\dot{\mathbf{q}}$ の対応する成分と常に同符号である． $K(\mathbf{q}, \dot{\mathbf{q}})$ ， $P(\mathbf{q})$ をそれぞれ運動エネルギー，ポテンシャルエネルギーとすると，以下のような関係が成り立つ．

$$K(\mathbf{q}, \dot{\mathbf{q}}) = (1/2) \dot{\mathbf{q}}^T \mathbf{M}(\mathbf{q}) \dot{\mathbf{q}} \quad (\text{A.3})$$

$$\mathbf{g}(\mathbf{q}) = (\partial P(\mathbf{q}) / \partial \mathbf{q}^T)^T \quad (\text{A.4})$$

このとき，ロボットの全内部エネルギーは $E(\mathbf{q}, \dot{\mathbf{q}}) = K(\mathbf{q}, \dot{\mathbf{q}}) + P(\mathbf{q})$ となる．

ロボットシステムは，速度出力 $\mathbf{y} = \dot{\mathbf{q}}$ に関して受動的であることが知られている．すなわち，式 (A.2) から，

$$\begin{aligned} \int_0^t \dot{\mathbf{q}}^T \mathbf{u} d\tau &= \int_0^t \dot{\mathbf{q}}^T \left\{ \mathbf{M}\ddot{\mathbf{q}} + \frac{1}{2}\dot{\mathbf{M}}\dot{\mathbf{q}} + \mathbf{S}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) + \mathbf{d}(\dot{\mathbf{q}}) \right\} d\tau \\ &= \int_0^t \frac{d}{d\tau} \{K(\mathbf{q}, \dot{\mathbf{q}}) + P(\mathbf{q})\} d\tau + \int_0^t \dot{\mathbf{q}}^T \mathbf{d}(\dot{\mathbf{q}}) d\tau \\ &= E(\mathbf{q}(t), \dot{\mathbf{q}}(t)) - E(\mathbf{q}(0), \dot{\mathbf{q}}(0)) + \int_0^t \dot{\mathbf{q}}^T \mathbf{d}(\dot{\mathbf{q}}) d\tau \\ &\geq -E(\mathbf{q}(0), \dot{\mathbf{q}}(0)) = -\gamma_0^2 \end{aligned} \quad (\text{A.5})$$

となり，式 (A.1) が成立する．ただし， \mathbf{S} の歪対称性と $\dot{\mathbf{q}}^T \mathbf{d}(\dot{\mathbf{q}}) \geq 0$ を用いた．上式から， γ_0^2 はロボットの初期内部エネルギーと解釈できる．