

YOLO による物体検出を用いた全壊した住家の検出

木村 裕貴[†] 木内 一隆[†] 河合 紀彦[†] 鈴木 基之[†] 伊勢 正[‡]大阪工業大学 情報科学部 情報メディア学科[†] 国立研究開発法人防災科学技術研究所[‡]

1. はじめに

地震や水害といった災害発生後に、基礎自治体は災害の被災者から申請があった際に遅滞なく被災者の住家の被害状況を調査し、住家の被害の程度を証明する罹災証明書を交付する必要がある。現状、被害状況の調査および罹災証明書の交付は手作業で行われている^[1]。そのため、大規模災害などで多くの住家が被災し、全ての被災者から申請があると多くの時間と人員、労力が必要になる。

従来手法として、日立製作所が制作した空撮映像から災害状況を瞬時に把握できるシステムが存在する^[2]。この技術は、ドローンやヘリコプタから撮影した映像から地滑りや家屋倒壊といった災害箇所および災害状況を検出するシステムである。このシステムの問題点として、ドローンやヘリコプタといった特殊な機材や操縦技術が求められることである。

本研究では、住家の被害状況の調査において、事前スクリーニングとして明らかに全壊している住家を調査対象から省くことを目的として、被災した住家の写真をタブレットなどのモバイル端末で撮影し、その写真内から明らかに全壊している住家を検出する手法を提案する。従来手法と異なり、スマートフォンやタブレットのようなモバイル端末で、道路上や高層建造物内、見晴らしのいい場所からの撮影を想定している。

2. 提案手法

提案手法では、被災した住家が写っている写真で学習したニューラルネットワークに被災地の写真を入力することでその写真の中から全壊している住家とそうでない住家の物体検出を行う。物体検出にはYOLOv7^[3]を使用する。また、学習データが少ないため、学習済みニューラルネットワークを用いて転移学習する。バウンディングボックスによる物体検出を行うため、学習データには学習させる画像だけでなくバウンディングボックスとクラスのデータを用意する。

3. 実験と考察

3.1 実験概要

本実験では、学習データと検証データに防災科学技術研究所所有の熊本地震の被災地を撮影した写真を用いた。全3549枚の画像を学習データ2487枚、検証データ250枚、住家が写っていない未使用デー

表1 学習に使用したPCのスペック

OS	Windows11
CPU	IntelCorei9-10850K
メモリ	32.0GB
GPU	NVIDIA GeForce RTX 3080

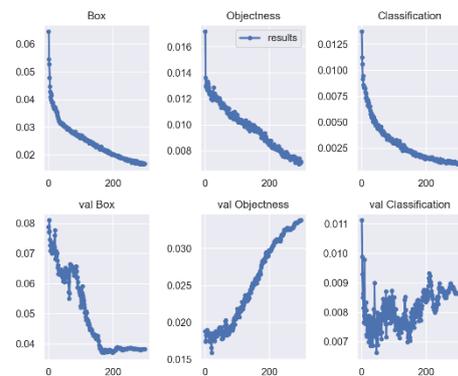


図1 学習時の各パラメータのloss

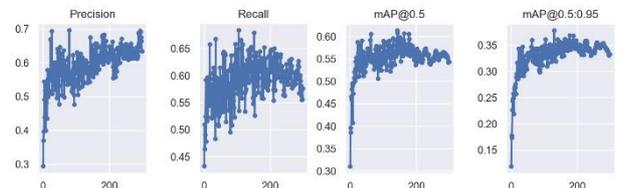


図2 学習時のPrecision, Recall, mAP

タ812枚に分類した。学習データと検証データの2737枚の画像全てにlabelImg^[4]を使用して著者の主観によりGUIでバウンディングボックスの領域とクラスを指定した。学習したニューラルネットワークを評価するため学習データ・検証データに含まれない3枚の画像で実際に物体検出を行った。

3.2 転移学習

学習データ2487枚を用いてYOLOv7の転移学習を行った。学習させる回数エポック数は300、1回に学習させる枚数バッチサイズは8で学習させた。表1に示すスペックのPCで学習したところ学習時間は6時間10分程度となった。図1に学習した際の各loss、図2にmAPなどのパラメータを示す。

図1からBox, Objectness, Classificationの各lossは、過学習もしておらず学習を重ねるごとに小さくなっていき、順調に学習出来ていることがわか

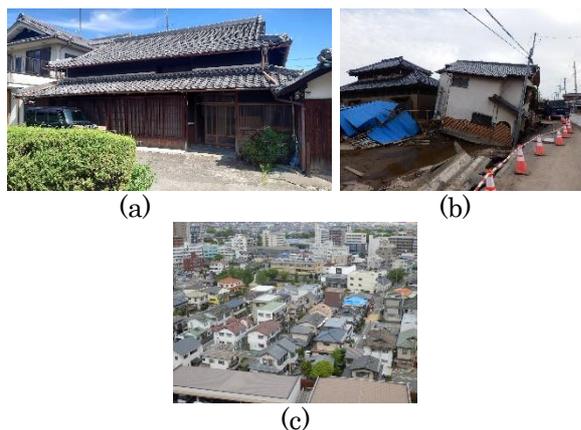


図3 評価で使った画像

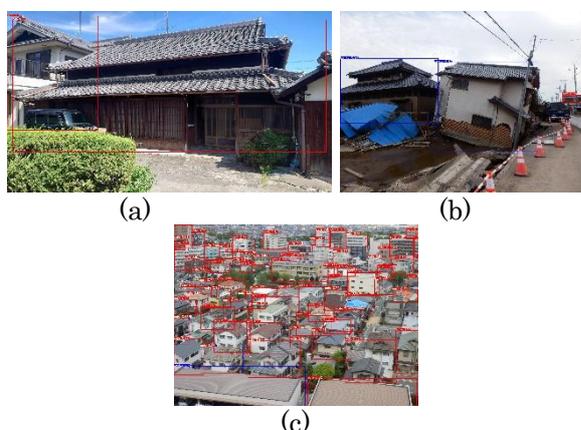


図4 物体検出した画像

る。しかし、検証データでは Box の数値は、学習を重ねるごとに小さくなっているが、Box 以外の loss の数値は過学習している結果となっている。

図2から Precision の数値には大きな停滞なく学習出来たが、Recall は学習中盤から数値が小さくなった。従って、mAP@0.5 と mAP@0.5:0.95 の数値が小さく学習終盤あたりから停滞する結果となった。mAP の数値から、本実験での学習ではバウンディングボックスの検出精度と正解率が悪い結果になっている。

3.3 評価結果例

道路上や高層建造物内から撮影した代表的な画像3枚に対して学習したニューラルネットワークを使用して物体検出を行い、結果を示し考察を行う。評価画像を図3に、物体検出した画像を図4に示す。赤いバウンディングボックスが全壊していない住家、青いバウンディングボックスが全壊した住家を示している。

図4の画像から、画像に対して住家が大きく写っている(a)、(b)はバウンディングボックスの領域やクラスが正確に検出出来ている。しかし、(b)の左の住家は前方に倒壊した建造物があるため全壊した

住家として検出している。(c)のような画像に対して住家が小さく写っている場合は、バウンディングボックスの領域、検出したクラスに正確性が欠ける結果になっている。正確に物体検出ができなかったのは、学習データの偏りが原因だと考える。本実験の学習データでは、全壊している住家と全壊していない住家の数が均等にならず全壊している住家の方が少なくなっている。また、道路上から撮影した写真が多く、見晴らしのいい場所から撮影された写真も少ないため正確に物体検出ができていないと考える。

4. まとめ

本研究では、事前スクリーニングとして、被災した住家の写真をタブレットなどのモバイル端末で撮影し、その写真内から明らかに全壊している住家を検出する手法を提案した。実験の結果、住家が大きく写っている写真での物体検出は、正確に住家そのものの領域をバウンディングボックスで捉えることが出来たが、一部の住家で正確にクラスを判別できないものもあった。また、住家が小さく写っている写真での物体検出では、バウンディングボックスの領域、検出したクラスに正確性が欠ける結果となった。

今後の展望として、学習データの種類を増やすことで住家が小さく写っている写真からでも正確に物体検出が行えるようにすることとモバイル端末のみで物体検出が行えるシステムを開発することが挙げられる。

謝辞

本研究では、JSPS 科研費 JP22H01749 の助成を受けて実施した。

<参考文献>

- [1] 内閣府：“災害に係る住家の被害認定 – 内閣府防災情報”，<https://www.bousai.go.jp/taisaku/unyou.html>，(2022年7月19日閲覧)
- [2] 日立：“AI で災害状況を瞬時に把握する「映像解析システム」を日立が開発：社会イノベーション：日立”，<https://social-innovation.hitachi/ja-jp/article/dsdi/>，(2022年10月14日閲覧)
- [3] Wang, Chien-Yao and Bochkovskiy, Alexey and Liao, Hong-Yuan Mark, “YOLOv7: Trainable bag-of-freebies sets new state-of-art for real-time object detectors”, arXiv preprint arXiv:2207.02696, 2022
- [4] Tzutalin : “LabelImg. Git code (2015)”, <https://github.com/tzutalin/labelImg>，(2023年1月11日閲覧)