

空撮映像を用いた Multi Object Tracking による倒壊建物検出 Detection of Collapsed building using Multi Object Tracking from Aerial Videos

○藤田翔乃・畑山満則

○Shono FUJITA, Michinori HATAYAMA

In the past earthquake disaster, it was too difficult to grasp overall damage of buildings, which hindered initial response such as rescue. In recent years, many fire headquarters use drones to grasp damage of disaster from aerial images or videos. In this paper, we developed detection system of collapsed buildings from aerial movies using MOT (Multiple Object Tracking) to provide effective information for rescue activity. Our model using MOT can extract collapsed buildings by circumscribed quadrangle and identify the same buildings from some frame images with various angle.

1. 背景

過去の地震災害では、建物の被害状況の把握に困難を極め、救助などの初動対応に大きな支障をもたらした。東日本大震災では被災地消防本部のうち、約 25%の消防本部で通信網が不通となり 119 番通報することができなかった。熊本地震や大阪北部地震では、緊急性の要しない 119 番通報が多く、効率的な対応を行うことができなかった。そしてこれまでは、119 番通報のピークが治ると全ての建物を調査するローラー調査を行っていたが、今後予想される大規模地震では人手不足によりこのような対応を行うことができないと考えられる。

また、現在 726 の消防本部のうち 116 の消防本部が、有毒ガス・火災・土砂災害・地震災害などの被害状況の把握のためにドローンを保有している。このことから対応現場は、ドローンで撮影された航空写真、空撮映像から被害状況を把握することを推進していることがわかる。

以上の研究背景から、本研究はドローンによる空撮映像から倒壊建物を検出し、救助・救命活動に有効な情報を提供することを目的とする。

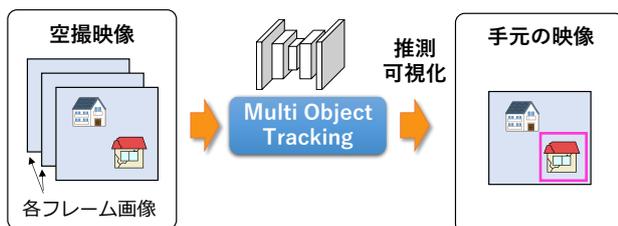


図 1 提案手法

2. 提案手法

上記の目的のため、空撮映像から MOT(Multiple Object Tracking)を用いて倒壊建物を自動検出する。空撮映像を用いることで広範囲の被害状況を認識することができ、また自動検出することで迅速に情報を得ることができる。

3. 先行研究

現状の対応現場で利用できる、リアルタイム被害推定システム[1]などは 250m メッシュ毎の被害建物数を推測しており、個々の建物を推測できない。本研究では空撮映像と MOT を利用することで個々の倒壊建物を検出することができる。また、真上からの航空写真により建物被害を検出する先行研究では、屋根被害はないが倒壊などの建物被害がある建物を検出することができない。本研究のように、斜めから撮影される空撮映像を用いることで、このような垂直方向の変化を持った倒壊建物を検出することができ、救助活動に有効な情報を提供できると考えられる。

4. Multiple Object Tracking

Multiple Object Tracking とは深層学習を用いた映像解析手法であり、映像中の物体の種類と位置(外接四角形)を推測することができる。映像中の推測された物体をフレーム間で紐づけることにより、同一物体を識別し、追跡することができる。

5. 倒壊建物検出

本研究では、熊本地震後の益城町でドローンによ



図3 推測結果の可視化動画

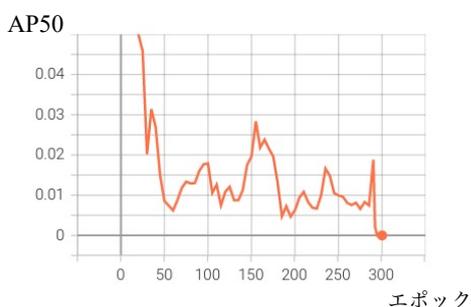


図2 検証データの AP50 の推移

り撮影された空撮映像と、MOT モデルとして Byte Track[2]を利用した。元の映像では、30FPS のフレームレートであったが、学習・検証データでは 3FPS でアノテーションを行った。アノテーションでは救助が必要であろう 1 階部分が崩壊している建物を倒壊建物として、外接四角形を描いた。

学習データとして合計 7 分 38 秒、1375 フレーム、14238 アノテーション、227 トラックを、検証データとして、合計 1 分 55 秒、345 フレーム、832 個のアノテーション、23 トラックを利用した。テストデータでは、元の映像と同じ 30FPS のフレームレートの合計 1 分 16 秒、2290 フレーム、4530 アノテーション、23 トラックで評価した。

検証データによりモデルの汎化性を確認し、過学習が生じていると考えられるエポック 20 で学習を止めたモデルを利用した。(図 2)

結果は、MOTA が 0.8%, MOTP が 0.276 であり、倒壊があった建物 11 棟のうち 5 棟を検出した。また、2 棟は誤った倒壊と推測したが、傾きがある建物で倒壊に近い状態であるものであった。図 3 に推測結果を可視化させた動画の一部を示す。

6. 考察と今後の課題

上記の結果から、トラッキングとしての精度は未だ低いが、約半分の倒壊建物を検出することができた。

検証データによるモデル精度を見ると、エポック 20 あたりで汎化性能が低下しているため、学習データを増加させる、もしくはより少ないパラメータ数を持つ MOT モデルを利用する必要があることがわかった。また、学習・検証データにおいて同一の建物が映るデータが存在し、適切にモデルの汎化性能を評価できていない恐れがある。今後は映像中に映る建物を地域ごとに、学習・検証・テストデータに振り分け、分析を行う。

今後、MOT モデルにより倒壊建物を可視化させた動画を用いて、災害対応機関にヒアリングなどを行い、よりニーズに沿ったシステムの要件を分析しなければならない。

また過去の地震災害では、被害を地図上に可視化させたことが有効であったことから、空撮映像から MOT により検出された被害を地理空間情報と紐付けるシステムを今後開発する予定である。

参考文献

- [1] 戦略的イノベーション創造プログラム (SIP) . 対応：課題⑤ リアルタイム被害推定、https://www.jst.go.jp/sip/k08_team5.html, (最終閲覧 2022/01/05)
- [2] Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Zehuan Yuan, Ping Luo, Wenyu Liu, Xinggang Wang. ByteTrack: Multi-Object Tracking by Associating Every Detection Box, arXiv:2110.06864