# 建物損傷度評価に向けた熊本地震と能登半島地震の現地写真を用いた損傷検出モデルの構築

Development of Building Damage Detection Model Using On-site Photos after the Kumamoto Earthquake and the Noto Peninsula Earthquake

> 23WM1313 秋葉 涼太 Ryota Akiba

# **SYNOPSIS**

In this study, we aim to assess building damage by detecting damage such as cracks on the walls of buildings using deep learning using photos in areas damaged by the earthquakes. There is a lack of speed and objectivity in damage assessment surveys and it is thought that using a model based on deep learning to determine building damage will contribute to improving efficiency, speed and objectivity rather than humans visit the site to carry out surveys. This study employs Mask R-CNN and YOLOv8 to achieve the objective. The images used for training are the photos of walls taken after the 2016 Kumamoto Earthquake and the 2024 Noto Peninsula Earthquake. Comparing the default model and the model with attention module, which is a mechanism to identify where to focus on in input data, the model with attention module is able to detect more flacking.

1. 研究背景

災害が発生した際には、被災者に対して義援金の支給や 税の減免など様々な支援が講じられる.これらの支援の多 くは住宅の被害の程度を基準にしているため、内閣府の被 害認定基準に基づき実施される被害認定調査<sup>11</sup>は迅速かつ 正確に行われる必要がある.この調査では、自治体職員等 が研修を受けた後に、外観目視を基本とした損傷の種類や 程度を評価する第一次調査や、第一次調査に納得のいかな かった被災者からの申請があった場合に実施される第二次 調査が行われ、建物の被害程度を5段階で評価する.得ら れた調査結果は罹災証明書として被災者に交付され、様々 な被災者支援措置に活用される.

大規模な災害が発生した際には、調査対象となる建物が 多いため人員の確保が困難になり、調査に遅れが出ること が問題視されている.2016年熊本地震の際には、地震発生 から1か月の時点の罹災証明書の交付申請受付件数に対す る交付件数は3割未満であり証明書発行の遅延が問題とな った<sup>2)</sup>.また、結果に納得がいかない場合に再調査を行う ことや、被害程度の評価が境界近くの場合に評価を一段階 あげることがあることも指摘されている.

そこで本研究では、建物被害認定調査の効率化と迅速化、 客観性を保つことを目的として、物体検出モデルである Mask R-CNN<sup>3)</sup>と You Only Look Once (YOLO)<sup>4)</sup>を適用した 深層学習により建物の損傷検出に向けた基礎検討を行う. 将来的には、被害認定調査の基準に即して建物被害を判別 することが可能となる手法の開発を目指している.

#### 2. 物体検出手法の概要と学習条件

本研究では、物体検出手法としてインスタンスセグメン テーションが可能な Mask R-CNN と YOLOv8 を使用する. インスタンスセグメンテーションとは入力画像をピクセル 単位でクラス分類し、物体がある箇所をバウンディングボ ックスと呼ばれる矩形で囲むアルゴリズムである.

Mask R-CNN のネットワークは Backbone 部, RPN 部, Head 部の 3 つから構成されている. Backbone 部は畳み込 みニューラルネットワーク (CNN) <sup>5</sup>により,画像の線,角 などの特徴を抽出する役割を担う. RPN 部では物体らしき 領域を選定し, Head 部では RPN 部で提案された領域に対 して物体の有無の判断とクラス識別を行い,物体がある領 域にマスクをかける.



図1 Mask R-CNN のネットワーク構造

本研究では、Mask R-CNN が使用できるパッケージとし て、Detecron2 と MMDetection を使用した. 両パッケージで は、主に Backbone 部の CNN に適用される初期重みパラメ ータが異なる. また、事前学習済みの重みパラメータの転 移学習を行う. Detecron2 のモデルを model1, model2, MMDetection のモデルを model3, model4 とする. Backbone 部の CNN には model1, model3 において Resnet101 を採用 し, model2, model4 ではその派生である ResneXt101 を使 用した.

YOLOv8 のネットワークは Backbone 部, Neck 部, Head 部の 3 つから構成されている. Backbone 部では Mask R-CNN と同様に CNN で入力画像の特徴量を抽出する. Neck 部は Backbone 部の CNN の様々な深さの層から出力された 複数の特徴量マップを処理し,異なるスケールの物体の検 出に対応させる役割を持つ. Head 部ではバウンディングボ ックスとクラスの予測を行う. YOLOv8 を使用したモデル を model5 とし、各モデルの学習条件を表1に示す.

	model1	model2	model3	model4	model5	
パッケージ	Deteo	tron2	MMDe	MMDetection		
Bashbara BasNat101 BasNaVt101 BasNa	RegNet101	DooNoV+101	YOLOv8-			
Dackbolle	Residector	ResNexIIII ResNetIUI	101 ResNeXt101 0.0025	Backbone		
optimizer		SGD				
学習率	0.00	00025 0.0025		0.001		
momentum		0.9				
weight decay		0.0	001		0.0005	
学習回数		300	000		300000	
batch size			1			

表1 本研究のモデルの学習条件

#### 3. 本研究で使用したデータ

本研究では、2016 年熊本地震後の被害認定調査時に調査 棟数が多かった熊本市と益城町において撮影された建物の 外壁画像と2024 年能登半島地震にて被害を受けた輪島市, 珠洲市,能登町, 穴水町にて車載カメラおよびデジタルカ メラで撮影されたものを使用する.

画像データを訓練データと検証データとテストデータに 分割した後にアノテーションを行った.アノテーションと は、物体検出における正解データとして画像上の物体の種 類と形をラベル付けする作業である.本研究では建物の損 傷度判定に用いるひび割れ、剥落、下地材の3つのクラス 分けとした.アノテーションツールには、COCO形式でア ノテーションできる COCO Annotator を利用した.本研究 で使用したデータセットの画像数とインスタンス数の内訳 を表2に示す.

	枚数	イ	インスタンス数		
		ひび割れ	剥落	下地材	
訓練データ	327	742	215	164	
検証データ	17	60	18	30	
テストデータ	12	29	7	26	

表2 画像データセットの内訳

# 4. 本研究で用いる評価指標

精度評価指標として Optimal Correction Cost (OC-Cost) <sup>(0)</sup> を用いて検討を行った. OC-Cost とは,検出結果を正解デ ータと同じような大きさ,位置,クラスに修正するための コストに基づき精度を評価する指標である. 各検出結果を それぞれの正解データに修正するためのコストを表すコス ト行列において,修正コストが最小となるような組み合わ せから OC-Cost を算出する. OC-Cost の概略図を図 2 に示 す. 精度評価指標としてよく使用される mean Average



Precision (mAP) と比較すると, OC-Cost では予測クラスが 誤りであっても, 予測位置の精度は評価される.また, 誤 検出に対して mAP よりも厳しい評価を行う点に特徴があ る.本研究で扱う被害認定調査は公的な調査であり, 誤検 出を少なくすることがより重要と考えるため,本研究では OC-Cost を採用した. OC-Cost は0から 0.5 の値をとり, 0 に近いほど精度が良いことを示す.

また、物体検出では、混同行列を用いて精度評価が行われる. 混同行列は、各クラスの予測と各クラスの正解の対応を示した表であり、定められた Intersection over Union (IoU) より大きい場合に検出が成功したものとして評価する. ここで、IoU は、検出したい物体が存在する正解の領域とモデルが予測した領域の重なる箇所の面積を、正解と予測の面積の和集合で除した値である.

さらに, 混同行列から式(1) ~(3)のようにして算出され る F 値を算出した.ここで, TP は対象のクラスにおいて正 解している予測結果の個数, FP は対象のクラスと予測した ものの不正解だった予測の個数, FN は対象のクラスだっ たものの対象のクラスと予測できなかった予測の個数を示 す.

適合率(Precision) = 
$$\frac{TP}{TP + FP}$$
(1)再現率(Recall) =  $\frac{TP}{TP + FN}$ (2)F 値 =  $\frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$ (3)

また, OC-Cost と F 値のクラス平均を同時に評価するために式(4)のような評価式を使用し, 各モデルの比較検討を行う.

$$Score = F \operatorname{\underline{\textit{i}}} - 2 \cdot OC \cdot Cost$$
 (4)

5. 学習結果の精度評価

各モデルの OC-cost を表 3, 各モデルの F 値を表 4, 各モ デルの Score を表 5 に示す. 学習した各モデルのテストデ ータの推論結果を図 3 に示す.

表 3 各モデルの OC-Cost						
	model1	model2	model3	model4	model5	
OC-Cost	0.307	0.317	0.396	0.334	0.321	

表4 各モデルのF値						
F値	model1	model2	model3	model4	model5	
ひび割れ	0.456	0.439	0.392	0.370	0.275	
剥落	0.200	0.182	0.216	0.286	0.348	
下地材	0.638	0.727	0.585	0.737	0.759	
平均	0.431	0.449	0.398	0.464	0.461	

表 5 各モデルの Score					
	model1	model2	model3	model4	model5
Score	-0.183	-0.185	-0.394	-0.204	-0.181

OC-Cost は model1, model2, model4, model5 では 0.34 未 満となり, model3 のみ約 0.4 となった. パッケージごとに 検討すると Detectron2 を用いている model1, model2 の OC- Cost の平均値は 0.312 であり, MMDetection を用いている model3, model4 の平均値 (0.365) や model5 の値 (0.321) よりも小さい.

F値のクラス平均も同様に比較すると, model5 は 0.461 となり, model1, model2 の平均値 (0.440) や model3, model4 の平均値 (0.431) よりも大きい. クラスごとの F 値に着目 すると, model1, model2, model3, model4 で同様の傾向が 見られ, 下地材の F 値は 0.58 を超え, 比較的高めであるが, 剥落の F 値はどのモデルでも 0.28 以下の値であり, 剥落の 検出精度が低く, クラスごとに検出性能に偏りがあること が分かる. これは, 下地材は比較的大きい損傷であり検出 しやすいことと, 訓練データにおいて剥落のデータが少な いことが原因と考えられる.

Score については, model1, model2, model5 では-0.183 付 近の値となり, Score の値が-0.181 の model5 が最良であり, model3 のみ-0.394 と劣る.

テストデータの推論結果(図3)においては画像2の左 側の下地材に着目すると model1 のみが検出に成功してお り, 土壁の建物の損傷の検出が難しいことが分かる.これ は訓練データにおいて土壁の建物の画像が少ないことが原 因だと考えられる.

評価指標や推論画像から誤検出や検出の見逃しが発生していることが分かり、さらなる精度改善が必要である.



6. アテンションモジュールを使用したモデルの評価

モデルの精度向上のためにアテンションモジュールの Convolutional Block Attention Module (CBAM) <sup>7)</sup> を導入し た. CBAM は画像のどのチャネルに着目し、画像のどこに 着目すべきかを強調する機構であり, model3, model4, model5 に適用した. model3, model4 では CBAM を提案し た論文と同様に Backbone 部の ResNet101 と ResNeXt101 の スキップ接続の直前に CBAM を計 33 個配置する. model5 においては、CBAM を Backbone 部に1つ追加したものを model5-1, Backbone 部に 4 つ追加したものを model5-2 とす る.また,損傷としてひび割れを検出する既往研究<sup>®</sup>にて 最良の結果となったモデルと同様に Neck 部に 3 つ追加し たものを model5-3 とする. また, 検出結果の尤もらしさを 示す信頼度スコアの閾値を 0.25, 0.5, 0.75 の 3 通りで検証 し、学習率を 0.001, 0.002, 0.003, 0.004, 0.005 に変化さ せて、4 章で示した評価式(4)で算出した Score が最大のも のを各モデルの代表値として比較検討を行う. 学習率以外 の学習条件は表1と同様である.各モデルの OC-Cost を表 6, 各モデルの F 値を表 7, 各モデルの Score を表 8, 各モ デルのテストデータの推論結果を図4に示す.

表6 CBAM を追加した各モデルの OC-Cost

	model3	model4	model5-1	model5-2	model5-3
学習率	0.003	0.002	0.003	0.005	0.002
信頼度閾値	0.5	0.5	0.25	0.25	0.25
OC-Cost	0.324	0.358	0.361	0.380	0.352

表 7 CBAM を追加した各モデルの F 値

F値	model3	model4	model5-1	model5-2	model5-3
ひび割れ	0.305	0.340	0.377	0.462	0.373
剥落	0.533	0.500	0.500	0.333	0.600
下地材	0.681	0.560	0.591	0.622	0.600
平均	0.506	0.467	0.489	0.472	0.524

表 8 CBAM を追加した各モデルの Score

	model3	model4	model5-1	model5-2	model5-3
Score	-0.142	-0.249	-0.233	-0.288	-0.180

推論条件として、パッケージに MMDetection を用いた model3, model4 では信頼度スコアの閾値は 0.5 が最適であ り、YOLOv8 を用いた model5-1, model5-2, model5-3 では 閾値 0.25 が最適であることが分かる.

OC-Cost は、CBAM を追加することで model3 のみ 0.072 低下し、精度の向上が見られるがその他のモデルでは値が 増加しており、CBAM を追加することで精度が悪化した.

F値のクラス平均はいずれのモデルでも CBAM を追加後 のモデルで精度が向上した.特に model3 では 27.1%増加 し, model5-3 では 13.7%増加している.クラスごとの F値 では,剥落の F値が model5-2 を除いたすべてのモデルで 0.5~0.6 の値をとり, CBAM により剥落の検出性能が向上す る傾向があることが分かる.5章で示した5つのモデルの 各クラスの F値を母集団としたときの標準偏差は 0.202 で あるのに対して本章で示した5つのモデルの各クラスの F 値を母集団としたときの標準偏差は 0.121 となり, CBAM を導入することで検出性能に比較的偏りのないモデルにな ることが分かる.一方で, model3 ではひび割れの F 値は 22.2%低下し, model4 では 8.1%低下していることから,パ



図4 CBAM を追加した各モデルのケストケータの推贏 結果

ッケージとして MMDetection を用いるモデルではひび割れの検出性能が悪化した.

Score は CBAM 追加後の model3 において値が 0.252 増加 し, model5-3 では 0.001 増加したものの, model4, model5-1, model5-2 では値が低下した.5 章の最良の Score (model5: -0.181) と比較すると,本章で最も優れるものは model3 (-0.142) であり, CBAM により Score が 0.039 増加した.

推論画像(図4)では、画像1においては model5-2で画 像右側の剥落の検出に初めて成功し、画像2では model5-1 で画像右側の剥落を検出できるようになり, model5-3 では 画像右側の5箇所の損傷をすべて検出できるようになった. model5-3による画像3の推論画像の画像上部のように住宅 の屋根の瓦やブロック塀を誤検出する傾向が見られ、これ が OC-Cost の値の増加に影響したと考えられる.

評価指標を鑑みると CBAM 追加前のモデルと CBAM 追 加後のモデルでは CBAM を追加した model3 が最良のモデ ルであると言えるものの, OC-Cost は model1 の値 (0.307) を下回らず, ひび割れの F 値が低下し, ひび割れの検出性 能が悪化していることから,本研究で示したモデルは改善 の余地があると考える.

# 7. まとめと今後の展望

本研究では、本研究では、国内で発生した地震の際に撮影された画像を用いて、Mask R-CNN と YOLOv8 により外壁の損傷を検出する試みを行った.2016 年熊本地震と 2024 年能登半島地震の際に撮影された画像を使用して深層学習を行ったところ、精度向上の余地があったことからアテンションモジュールである CBAM を導入したモデルによる検討を行った. 主に剥落の検出性能が向上したものの、ひび割れの検出精度と瓦やブロック塀の誤検出において改善の余地がある.

今後は各クラスの検出精度向上と誤検出を減らしていく ために Grad-CAM<sup>9)</sup>を活用し、モデルの判断の根拠となる点 を含めた考察を進めていくことを検討している. Grad-CAM とはモデルの特定の層における特徴マップと重みパ ラメータを用いて入力画像のどこに着目して結果を出力し たのかを示すアルゴリズムである.また、瓦やブロック塀 の誤検出に対応するために、瓦やブロック塀のクラスを新 に追加したデータセットで深層学習を行うことでこれらを 検出できるようにして、誤検出を防ぐことも有効だと考え ている.

最終的には建物の外壁も検出できるモデルを作成し,外 壁の大きさと損傷の大きさから建物の損傷度を算出するこ とができる機構の作成を目指している.

### 参考文献

- 1) 内閣府:災害に係る住家の被害認定, https://www.bousai.go.jp/taisaku/pdf/r303shishin all.pdf
- 総務省: 大規模災害時における罹災証明書の交付等に 関する実態調査-平成 28 年熊本地震を中心として—, https://www.soumu.go.jp/main content/000528758.pdf
- He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask r-cnn, Proceedings of the IEEE international conference on computer vision, pp. 2961-2969, 2017.
- Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi,: You Only Look Once: Unified, Real-Time Object Detection, arXiv:1506.02640 [cs.CV], 2015.
- LeCun, Y., Haffner, P., Bottou, L. and Bengio, Y.: Object Recognition with Gradient-Based Learning, Shape, Contour and Grouping in Computer Vision, Lecture Notes in Computer Science, Vol. 1681, pp. 319–345, 1999.
- Mayu Otani, Riku Togashi, Yuta Nakashima, Esa Rahtu, Janne Heikkilä, Shin'ichi Satoh:Optimal Correction Cost for Object Detection Evaluation, arXiv : 2203.14438 [cs.CV], 2022.
- Sanghyun Woo, Jongchan Park, Joon-Young Lee, In So Kweon: CBAM: Convolutional Block Attention Module, arXiv:1807.06521 [cs.CV], 2018.
- 8) Hongyu Wang, Xiao Han, Xifa Song, Jie Su, Yang Li, Wenyan Zheng, Xuejing Wu.: Research on automatic pavement crack identification Based on improved YOLOv8 International Journal on Interactive Design and Manufacturing (IJIDeM), 24 February 2024.
- 9) Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra : Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization : arXiv:1610.02391 [cs.CV], 2016.