

木造建築物の地震被害を対象とした 応急危険度判定の自動化に関する研究 「壁の被害」を対象とした画像判定

A Research on Automatization of Emergency Risk Assessment for Earthquake Damage of Wooden Buildings Image Judgment for 'wall damage'

○横井 秀平*¹, 谷 明勲*², 山邊 友一郎*³
Shuhei Yokoi*¹, Akinori Tani*² and Yuichiro Yamabe*³

*1 神戸大学大学院工学研究科建築学専攻 大学院生

Graduate Student, Department of Architecture, Graduate School of Engineering, Kobe University

*2 神戸大学大学院工学研究科建築学専攻 教授 博士(工学)

Professor, Department of Architecture, Graduate School of Engineering, Kobe University, Dr. Eng.

*3 神戸大学大学院工学研究科建築学専攻 准教授 博士(工学)

Associate Professor, Department of Architecture, Graduate School of Engineering, Kobe University, Dr. Eng.

Summary: In recent years, huge earthquakes have frequently occurred in Japan. Therefore, it is important to carry out an emergency risk assessment quickly to prevent secondary damage. However, the shortage of judges and the ambiguity of judgment criteria become problems.

Therefore, in this study, we use CNN to classify images into 3 classes based on the criteria of emergency risk assessment and verify their accuracy. 115 image data for each of the three classes were prepared, 100 image data were used as training data, and 15 image data were used as verification data. The training data was expanded 10 times by image processing. The input value is RGB image data compressed to 64 x 64 pixels, and the output value is the confidence ratio for 3 classes. The structure has three convolutional layers and two fully connected layers.

As a result, the judgment accuracy of 90% or more was obtained for the training data, and the good accuracy was obtained for the one class also for the verification data. However, the accuracy of the other two classes declined. The accuracy of this part may be improved by using multiple CNNs and increasing the learning data.

キーワード: 応急危険度判定; 建物被害; 画像判定; 深層学習; CNN; 多クラス分類

Keywords: Emergency risk assessment; Damage of buildings; Image judgment; Deep learning; CNN; Multi-class classification

1. はじめに

地震はいつ、どこで、どのような規模で発生するか詳細に予測することが不可能である。1995年の兵庫県南部地震以降、2011年の東北地方太平洋沖地震や近年では2016年の熊本地震など、大規模地震の発生が相次いでいる。また、南海トラフ地震などの大規模な被害が想定される地震の発生も既に予測されている^[1]。このような地震が発生した後、被害を受けた建造物の事後の安全性や被災度を素早く把握し、二次災害を防止することが重要である。現在、応急危険度判定^[2]が二次災害防止を目的に行われているが、地震発生後に必要な判定士の確保に要する時間や、判定士の判定基準の相違による判定結果のばらつきや曖昧さが問題点として指摘されている^[3]。このような問題の解決手法として、近年人工知能の分野

で注目されている畳み込みニューラルネットワーク (Convolutional Neural Network, 以降CNNと記す) を用いた画像判定が挙げられる。この手法を用いた研究として、建物の構造種別を問わずに、外観から「被害なし」、「被害あり」の2クラスについて、ある程度の判定精度が得られるシステムが報告されている^[4]。また、内閣府「災害に係る住家の被害認定基準運用方針 (損傷程度の例示)」^[5]を元にした3クラス (「被害小」、「被害中」、「被害大」) の分類についても、田中らによって報告されており^[6]、一定の推定精度が得られている。しかしながら、応急危険度判定は単純に損傷程度の割合に依存するのではなく、二次被害に焦点を当てた判定基準となっているため、同様に画像判定を行った場合の判定精度は未知数である。

以上を踏まえ、本研究では現行の応急危険度判定制度に焦点を当て、対象を木造建築物の「壁の被害」の項目に限定して、表1に示す被災建築物応急危険度判定マニュアル⁷⁾に定められている3段階の被害判定基準を用いる。この判定基準の元、3クラスの画像判定を行い、その判定精度の検証を行うことを目的とする。

表1 「壁の被害」判定基準

A ランク	軽微なひび割れ
B ランク	大きな亀裂、剥落
C ランク	落下の危険有り

2. 画像データ

本研究で用いる建物被害画像は、国土技術政策総合研究所から公表されている、2011年の東北地方太平洋沖地震などを含む、2003年から2019年にかけて発生した災害に関する災害調査報告書⁸⁾に掲載されている被害画像である。これに阪神大震災記録写真集⁹⁾に掲載されている被害画像を加えて用いる。図1~3に、A~Cランクの被害画像群を例示する。Aランクの被害画像については、画像が不足したため、Web画像から無被害の健全な建物画像を無作為に抽出し用いた。表1に示す判定基準の元、



図1 Aランク被害画像群



図2 Bランク被害画像群



図3 Cランク被害画像群

応急危険度判定士の講習を受けた筆者が被害画像の被害ランクを判定し、分類したものをランク毎に115枚ずつ用意し、これらの学習データをCNNの学習に用いる教師データ100枚、学習結果の検証を行うための検証用データ15枚に分けて用いる。また、教師データについては、過学習を防ぐために画像の平行移動、左右の鏡像反転、 $\pm 5^\circ$ 以内の回転、部分的な抽出をランダムに組み合わせた処理を行う事で、1枚の画像データから9枚の画像を複製する。元の画像と複製画像を合わせた3000枚の画像データをCNNの学習に用いる。

3. CNNの概要

本報で用いたCNNの基本構造を図4に示す。このCNNの構造については、先行研究¹⁴⁾で用いられているCNN構造を参考にハイパーパラメータの異なるCNNを複数作成し、それぞれの精度検証を経て構造を決定した。入力値には教師データとなる画像データを 64×64 ピクセルに圧縮し、各ピクセルの濃度値を0~1に基準化したものをCNNに与える。なお、入力値はRGBの3チャンネルである。畳み込み層(convolution layer)は3層とし、各層のkernelの大きさは 9×9 , 5×5 , 3×3 , Filter数は16, 32, 64, 活性化関数にはReLU(Rectified Linear Unit)

を用いる。各畳み込み層の後には Pooling 層を設け、 2×2 の Max Pooling を行う。出力層の前には、活性化関数に ReLU を用いたユニット数 1600 と 800 の全結合層を設け、出力層には softmax 関数を用いる。また、全結合層の前は平坦化を行い、各 Pooling 層と全結合層の後にはそれぞれ 25%、50% の Dropout 層を設けている。出力値は A, B, C ランクそれぞれに対応する信頼度である 0~1 の 3 出力である。なお、本報で用いる CNN は Python の深層学習ライブラリである Keras を用いて作成し、ハイパーパラメータ等に関する名称等は Keras の公式ホームページ^[10]に掲載されているものを引用している。

4. CNN の学習

教師データを用いて、CNN の学習を行った学習終了時の損失関数の値 (loss) と正答率 (accuracy) の結果を表

2 に示し、その学習経過を図 5, 6 に示す。なお、正答率はデータ数に対する分類の正答数の割合を示す。損失関数の値 (loss) は学習開始直後に大きく下がった後、20epoch 付近から緩やかに下降する。正答率 (accuracy) も同様に、学習開始時に大きく上昇した後はほぼ一定の勾配で上昇していく。なお、CNN の学習について、mini-batch サイズは 128, epoch 数を 1000 として、学習は誤差逆伝搬法で行い、最適化手法には SGD (Stochastic Gradient Descent : 確率的勾配降下法)^[11]、損失関数には交差エントロピー (categorical-crossentropy) を用いた。学習回数等については異なる値での正答率を検証した上で決定した。

5. 推定結果

学習後の CNN を用いて検証データに対する分類予測

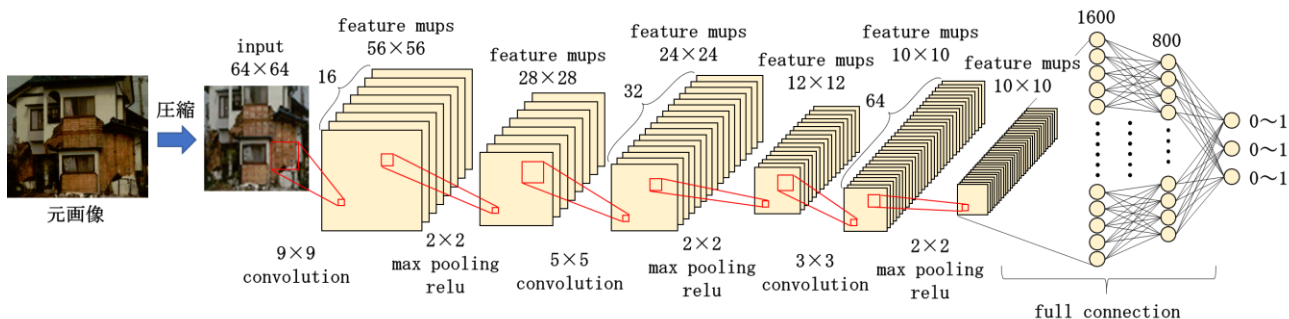


図 4 CNN の基本構造

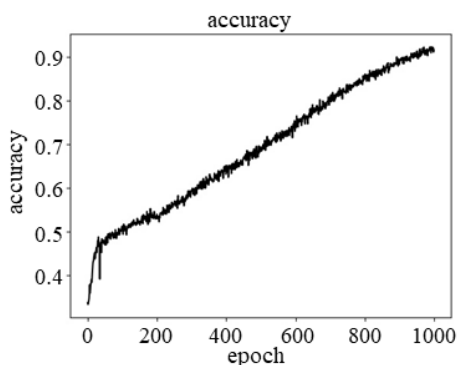


図 5 教師データに対する正答率 (accuracy) の推移

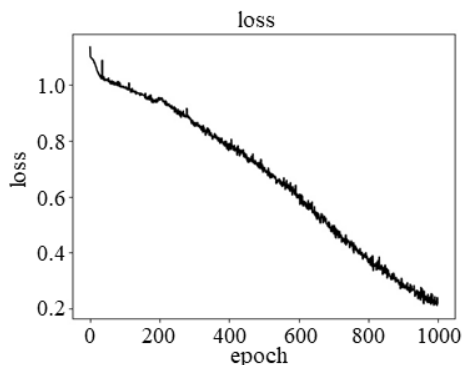


図 6 損失関数の値 (loss) の推移

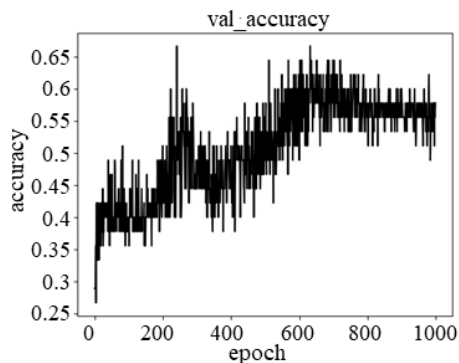


図 7 分類予測の正答率 (val_accuracy) の推移

表 2 学習結果 (学習終了時)

accuracy	loss
0.9170	0.2186

表 3 分類予測結果

true \ predict	A	B	C
A	73.3% (11)	13.3% (2)	13.3% (2)
B	26.7% (4)	40% (6)	33.3% (5)
C	20% (3)	20% (3)	60% (9)

							
True	A ランク	True	A ランク	True	A ランク	True	A ランク
分類予測結果	×	分類予測結果	×	分類予測結果	○	分類予測結果	×
A ランク	0.1848	A ランク	0.0042	A ランク	0.8542	A ランク	0.0051
B ランク	0.6708	B ランク	0.0575	B ランク	0.1350	B ランク	0.2944
C ランク	0.1444	C ランク	0.9383	C ランク	0.0107	C ランク	0.7004

図 8 検証データ A-1

図 9 検証データ A-2

図 10 検証データ A-3

図 11 検証データ A-4

							
True	A ランク	True	A ランク	True	A ランク	True	A ランク
分類予測結果	○	分類予測結果	○	分類予測結果	○	分類予測結果	○
A ランク	0.9473	A ランク	0.7161	A ランク	0.9774	A ランク	0.5452
B ランク	0.0175	B ランク	0.2067	B ランク	0.0215	B ランク	0.4368
C ランク	0.0352	C ランク	0.0771	C ランク	0.0011	C ランク	0.0180

図 12 検証データ A-5

図 13 検証データ A-6

図 14 検証データ A-7

図 15 検証データ A-8

							
True	B ランク	True	B ランク	True	B ランク	True	B ランク
分類予測結果	×	分類予測結果	×	分類予測結果	×	分類予測結果	○
A ランク	0.0736	A ランク	0.8766	A ランク	0.9608	A ランク	0.1769
B ランク	0.0211	B ランク	0.0994	B ランク	0.0388	B ランク	0.7673
C ランク	0.9052	C ランク	0.0240	C ランク	0.0004	C ランク	0.0558

図 16 検証データ B-1

図 17 検証データ B-2

図 18 検証データ B-3

図 19 検証データ B-4

を行った結果を表 3 に示す。表 3 の True は正しい判定ランク、Predict は予測による判定ランクを表している。次に CNN の学習経過に伴う検証データに対する分類予測の正答率 (val_accuracy) の変化を図 7 に示す。教師データに対する正答率と同様に、学習開始時に大きく上昇した後、300epoch、700epoch 付近でピークを迎える。正答率の最大値は 0.6778 となり、1000epoch の学習終了時には 0.5778 となった。また、各ランクの検証データから 8 枚ずつ抜粋したものと、その分類予測の信頼度を図 8～

31 に示す。なお、A ランクの画像データについては Web 画像から抽出したデータは省いて抜粋しており、各検証データに対する信頼度については学習終了時の値、画像は 64×64 ピクセルに圧縮後のものである。

6. 考察

図 5、表 2 より学習データについては 90%以上の正答率が得られ、学習が上手く行っていた。また、ランクによる精度の偏りがなくも確認している。

							
True	B ランク	True	B ランク	True	B ランク	True	B ランク
分類予測結果	×	分類予測結果	○	分類予測結果	×	分類予測結果	○
A ランク	0.6320	A ランク	0.0000	A ランク	0.9493	A ランク	0.1468
B ランク	0.2962	B ランク	0.9995	B ランク	0.0109	B ランク	0.4484
C ランク	0.0718	C ランク	0.0005	C ランク	0.0398	C ランク	0.4048

図 20 検証データ B-5

図 21 検証データ B-6

図 22 検証データ B-7

図 23 検証データ B-8

							
True	C ランク	True	C ランク	True	C ランク	True	C ランク
分類予測結果	○	分類予測結果	○	分類予測結果	×	分類予測結果	×
A ランク	0.0002	A ランク	0.0564	A ランク	0.6637	A ランク	0.2250
B ランク	0.0069	B ランク	0.0016	B ランク	0.2841	B ランク	0.6199
C ランク	0.9929	C ランク	0.9420	C ランク	0.0522	C ランク	0.1551

図 24 検証データ C-1

図 25 検証データ C-2

図 26 検証データ C-3

図 27 検証データ C-4

							
True	C ランク	True	C ランク	True	C ランク	True	C ランク
分類予測結果	×	分類予測結果	○	分類予測結果	×	分類予測結果	×
A ランク	0.2084	A ランク	0.0169	A ランク	0.5305	A ランク	0.8538
B ランク	0.4012	B ランク	0.3591	B ランク	0.3753	B ランク	0.0249
C ランク	0.3903	C ランク	0.6239	C ランク	0.0942	C ランク	0.1212

図 28 検証データ C-5

図 29 検証データ C-6

図 30 検証データ C-7

図 31 検証データ C-8

検証データについては、図 7 の分類予測経過から、1epoch 学習が進む毎に正答率が細かく変動していた。これは、図 9, 18, 24 の信頼度がいずれかのランクで 90% 以上となっているように、学習経過中にある特定のランクの信頼度が高くなったデータについては、学習が進んでも分類予測が変化することが少なく、図 15, 23 のように信頼度が各クラスに分散しているデータは学習の進行に伴い頻繁に分類が変わるためだと考えられる。また、図 7 より、分類予測の正答率は 200epoch 付近で一度高く

なるが、異なるパラメータ設定や教師データ・検証データの組み合わせを用いた場合の分類予測においても同様の傾向が多く見られた。このことから、教師データへの正答率は低くなるが、学習回数を少なくすることで検証データに対する正答率を向上させることができる見込みがある。

次にランク別の正答率に関して、表 3 の結果より、A ランクについては 73% と比較的高い正答率が得られた。これは、教師データと検証データをランダムに入れ替え

た場合でも同じで、A ランクについては学習後の CNN で常に 70%以上の正答率を得ることが可能である。一方で、B、C ランクの正答率は 40%、60%と低下し、特に B ランクの正答率が低いことが確認できる。この理由としては、応急危険度判定では、判定ランクが損傷程度の割合に単純に依存するのではなく、二次被害防止に焦点を当てた判定基準であることが考えられる。ここで、図 32、33 に示すのは図 23、28 の圧縮前の画像であり、壁の被害として破壊度が大きい場合でも、図 32 の様に外装材が完全に剥落し、落下の危険性が小さい場合は B ランクになることがある。反対に図 33 の様に破壊度が小さい場合でも、落下の危険性が高い場合は C ランクとなる。破壊度に依存する判定であれば、外装材の剥落の割合や木材の露出割合などから予測が容易であると考えられるが、応急危険度判定の判定基準では落下の危険性という部分に着目しており、亀裂の大小など細かな部材の状況によって判定が変わるため、今回用いた教師データ量では学習が不足したと考えられる。その結果として、A ランクに比べ、B、C ランクの各検証データに対する信頼度は複数のクラスに分散していることが多く、1 クラスに高い信頼度を持って分類することが出来なかったことが分かる。このことから、教師データを増やすことによって、1 クラスへの高い信頼度を持った判定を行う事ができ、正答率が向上する可能性が考えられる。また、今回用いた画像データには、建物全体が含まれている場合や周辺状況まで含んでいるデータがあったため、外壁のみを抽出することで精度の向上に繋がる可能性がある。

実際に応急危険度判定を判定士が行った場合でも、基準にばらつきがあるのは B、C ランクの判定であると考えられるため、正答率を向上させることが今後の重要課題となる。

また、図 23 は完全に外壁が剥落しきって下地の木材が大部分で露出しており、分類予測に関しては B ランクだが、信頼度に関しては B、C ランクではほぼ同等の値となっている。同様に B、C ランクの教師データには下地の木材が露出しているものが多くある。このことから、CNN が教師データから下地の木材の色データを学習し、木材の色が占めている部分が多い場合に B、C ランクの信頼度が高くなると推定することができる。



図 32 図 23 の元データ



図 33 図 28 の元データ

7. まとめ

本報を通して以下のことが明らかになった。

- CNN の学習により正答率の精度を 90%以上まで高めることができる。
- 検証データに対しては、A ランクの被害画像について正答率は 73%で判定を行う事が可能である。一方で、B、C ランクの被害画像については 40~60%程度と正答率は低下する。

以上より、CNN を用いて応急危険度判定の「壁の被害」の判定基準に基づいたクラス判定を行った場合、部分的に良い精度で判定を行う事が可能であることが分かった。しかしながら、実際の判定でも基準が難しいとされる部分の判定については、CNN による画像判定でも課題が残った。今後は更に教師データを増加させた場合や CNN のハイパーパラメータに最適化を行った場合の精度検証、複数の CNN を用いて判定を行うなど、新たなシステムを構築し精度の検証をしていくと共に判定精度の向上を図る必要がある。

【参考文献】

- [1] 地震調査研究推進本部 HP：南海トラフで発生する地震 (URL:https://www.jihin.go.jp/regional_seismicity/rs_kaiko/k_nankai/, 2019.7.7 閲覧)
- [2] 内閣府 HP：防災情報のページ (URL:<http://www.bousai.go.jp/taisaku/pdf/shishin011.pdf>, 2019.7.7 閲覧)
- [3] 日本建築学会：東日本大震災合同調査報告、第 6 章：阪神・淡路大震災と東日本大震災の ICT 関連被害比較と震災のための情報システムのあり方、日本建築学会、pp.247-256, 2016.9
- [4] 肥田剛典, 八百山太郎, 高田毅士：ディーブラーニングによる地震時建物被害判定と被害の地理的分布推定手法, 日本建築学会大会学術講演梗概集 (中国), 2017.8
- [5] 内閣府防災情報のページ, 災害に係る住家の被害認定基準運用指針 参考資料 (損傷程度の例示), 地震による被害「木造・プレハブ」(http://www.bousai.go.jp/taisaku/pdf/h3003sonsho_1.pdf)
- [6] 上岡洋平, 田中聡, : 深層学習による建物被災写真を用いた建物被害判別モデルの検討, 日本建築学会大会学術講演梗概集 (東北), 2018.9
- [7] 被災建築物応急危険度判定研究会：被災建築物応急危険度判定マニュアル, 日本建築防災協会, 1998.1
- [8] 国土技術政策総合研究所：災害調査報告書 (国内), (<http://www.nilim.go.jp/lab/bbg/saigai/index.html>, 2020.3.30 閲覧)
- [9] 今駒清則, 建築ドキュメンタリーシリーズ第 1, 2 集 阪神大震災記録写真集 Vol.1, 2, 3, 4, 5, 1996.1
- [10] Keras : Python の深層学習ライブラリ (<https://keras.io/ja/>, 2020.7.6 閲覧)
- [11] Keras : オプティマイザ (最適化アルゴリズム) の利用方法 (<https://keras.io/ja/optimizers/>, 2020.7.6 閲覧)