

# 車載カメラ画像による道路の地震被害の抽出に向けた検討

## Extraction of Road Damage after an Earthquake using Images Captured by a Car-Mounted Camera

○瀬崎 陸<sup>1</sup>, 丸山 喜久<sup>2</sup>, 永田 茂<sup>3</sup>

Riku SEZAKI, Yoshihisa MARUYAMA and Shigeru NAGATA

<sup>1</sup>千葉大学 大学院融合理工学府

Graduate School of Science and Engineering, Chiba University

<sup>2</sup>千葉大学 大学院工学研究院

Graduate School of Engineering, Chiba University

<sup>3</sup>鹿島建設株式会社 技術研究所

Kajima Technical Research Institute

The 2016 Kumamoto earthquake caused serious damage to road network, and it prevented from emergency responses. Since the road network is an important infrastructure for the quick restoration activity after disasters, it is necessary to grasp the damage situation of the road network at an early stage. This study tries to detect the earthquake-induced damage to the roadways using the images captured by a vehicle-mounted camera. A deep learning module CAFFE (convolutional architecture for fast feature embedding) was employed to develop an image classifier

**Keywords** : deep learning, Convolutional Neural Net (CNN), CAFFE, 2016 Kumamoto earthquake

### 1. はじめに

2016年4月に発生した熊本地震は、震源に近い熊本市、益城町、西原村などに甚大な被害をもたらした。とくに4月16日に発生したマグニチュード7.3の地震の影響で被害はより深刻化し、多数の人的被害に加えて道路被害が多発する<sup>1)</sup>など、復旧、復興への妨げとなった。道路ネットワークは被災後の迅速な復旧活動に重要なライフラインであるため、地震後早期に道路状況を把握する必要がある。現在、道路被害を確認する方法は、自治体職員による目視が中心<sup>2)</sup>であり、時間がかかり、職員への負担も大きい。

近年、画像認識の手段として、ディープラーニングが注目されている<sup>3)</sup>。従来のピクセルベースの手法は、道路被害の判別に特徴量や閾値を設定する必要があり、さらにこの設定値が検出精度に大きな影響を与える。一方、ディープラーニングは特徴量まで学習するため、人の手で特徴量を抽出する必要がなく、特に画像認識の分野で従来よりも高い精度を示すようになってきている。

そこで本研究では、2016年4月17日、20日に熊本地震の被災地である益城町などで撮影した車載カメラ画像を用いて、ディープラーニングに基づく判別モデルを作成した。その判別モデルを用いて、判別モデル作成時には使用しなかった画像を判別させ精度の検証を行った。

### 2. ディープラーニングに基づく学習の流れ

本研究では車の正面から撮影した連続写真に被害の有無をラベル付けしたデータセットを作成し、深層学習による学習を行った。

#### (1) 対象地域と使用データ

本研究の対象地域は熊本県益城町周辺であり、熊本地震発生後の4月17日と4月20日にグローバル・サーベ

イ株式会社によって撮影されたものを使用した。車載カメラの画像を、道路閉塞（建物の倒壊などにより、車の通行が困難な場合）、道路被害（ひび割れ、段差など）、無被害の3種類に目視によって分類し、それらを正規化した。自然画像を入力とする本研究では、入力画像の全体的な明るさやコントラストの違いをうまく吸収する必要があり、画像が撮影されるときカメラの露出（シャッタースピードなど）や太陽光の当たり方で、画像全体の明るさやコントラストは大きく変化する。この問題を解決するため、予め全画像を正規化した。正規化の後、256×256ピクセルに解像度を変更し、道路路面のみを判別させるために下四分の一に切り取り使用した(図1)。また、教師用画像に関しては学習用画像を増やすため、左右反転させ元の画像の2倍の枚数にした<sup>4)</sup>。



図1 解析に使用した画像の例

画像を3種類に分類する際には、まず6人が31,951枚の画像を目視で道路閉塞、道路被害、無被害に分類した。6人中5人以上が道路閉塞、道路被害と分類した画像を道路閉塞、道路被害としてディープラーニングに使用することとした。また、無被害に関しては6人全員がそのように目視で分類した画像を用いた。

このようにした結果、道路閉塞と分類された画像は512枚、道路被害は673枚、無被害は14,346枚となった。この中から、各分類350枚ずつをランダムに教師用画像として抽出した。これらを左右反転させ、道路閉塞、道路被害、無被害それぞれ700枚ずつ、計2100枚を教師

用画像とした。残りの道路閉塞 162 枚，道路被害 323 枚，無被害 13996 枚を判別精度の評価画像として使用した。

## (2) 深層学習のフレームワーク

教師用画像を深層学習させることにより判別モデルの作成を行う。本研究では，ディープラーニングのフレームワークとして CAFFE (convolutional architecture for fast feature embedding) を用いた<sup>5)</sup>。ネットワークの構成はリファレンスモデルの cifar10 をベースとして一部改変したものを使用した(図 2)。ここでは，入力層が 1 層，畳み込み層が 5 層，プーリング層が 5 層，全結合層が 2 層，出力層が 1 層とした。



図 2 本研究で使したネットワーク

畳み込み層<sup>6)</sup>(図 3)では，画像の局所領域において式(1)に示す畳み込み処理を行い，入力データに対しフィルタをかけることで情報の圧縮を行う。ここで $\phi$ は活性化関数， $w$ はフィルタ， $K$ は  $N-1$  層目の特徴マップ， $b$  はバイアスである。

$$h_j^n = \phi\left(\sum_{k=1}^K h_k * w_k^n\right) + b_k \quad \text{式(1)}$$

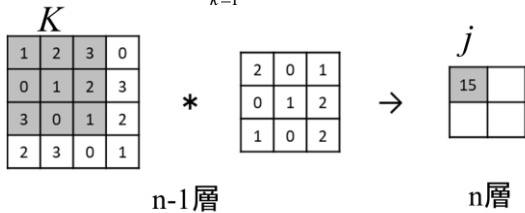


図 3 フィルタによる画像の畳み込み

プーリング層(図 4)は式(2)で表され，通常畳み込み層の直後に設置される。局所領域の情報を一部捨て，情報を限定することで画像中の物体の位置の違いを吸収する。これらのネットワークを多段に重ねることにより，中間ノードに画像を構成する基底のイメージが抽出される。

全結合層<sup>7)</sup>とは，畳み込み層とプーリング層の繰り返しの後に隣接層のユニットが全結合した層である。一般に複数連続して配置される。最後の出力層は，順伝搬型ネットワークと同様で，目的がクラス分類の本研究においては，最後の出力層は式(3)のソフトマックス関数となる。ここで， $y_k \equiv z_k^{(L)}$  は出力， $k$  が判別したいユニットである。出力層  $l=L$  の各ユニット  $k(=1, \dots, k)$  の総

入力，一つ下の層  $l=L-1$  の出力をもとに， $u_k^{(L)} = W^{(L)} z^{(L-1)} + b^{(L)}$  と与えられる。

$$h_j^n(x, y) = \max h_j^{n-1}(\bar{x}, \bar{y}) \quad \text{式(2)}$$

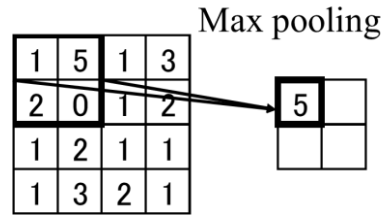


図 4 プーリング層

$$y_k \equiv z_k^{(L)} = \frac{\exp(u_k^{(L)})}{\sum_{j=1}^K \exp(u_j^{(L)})} \quad \text{式(3)}$$

## (3) 計算条件

一般に繰り返し計算回数を増やせば増やすほど精度は向上するが，本研究では 10000 回を超えても精度の向上が見られなかったため 10000 回とした。

学習率<sup>8)</sup>は，式(4)で求められる重みの最適化であり，以下の勾配法を利用することによって求められる。

$$x_{i+1} = x_i - lr \frac{\partial f}{\partial y} \quad \text{式(4)}$$

$x_i$  はある時点での変数の値であり， $\frac{\partial f}{\partial y}$  は勾配， $x_{i+1}$  は次の変数の値， $f$  は損失関数， $lr$  が学習率である。学習率は，一回の学習でどれだけパラメータを更新するかを決めるものである。式(4)は一回の更新式であり，これを繰り返し行う。つまり変数の値の更新を何度か繰り返すことによって，徐々に損失関数の値を減らしていく。本研究では学習率を 0.0001 と仮定した。

活性化関数<sup>9)</sup>とは閾値を境に出力が切り替わる関数であり，本研究では収束の速さや安定性の観点から ReLU 関数を使用した。ReLU 関数は入力が 0 を超えていれば，その入力をそのまま出力し，0 以下ならば 0 を出力する関数であり，以下の式(5)で表される。

$$h(x) = x(x > 0), 0(x \leq 0) \quad \text{式(5)}$$

学習を行う際には，教師用画像を学習用とテスト用に分けて用いる。学習用画像で判別モデルを作成し，その精度をテスト用画像で実験することを繰り返すことによって，判別モデルの精度を向上させる。本研究では学習用画像とテスト用画像の枚数比を 6:1 とした。学習 100 回ごとに損失関数を出力する。また，学習 500 回ごとにテストを行い，損失関数と正確性を表示させ，グラフ化した。損失関数はニューラルネットワークの性能の悪さを表す指標であり，本研究ではクラス分類用の損失関数として用いられる交差エントロピー誤差<sup>10)</sup>を使用した。エントロピー誤差は式(6)で表される。ここで， $n$  はサンプル番号， $K$  はクラス数， $y_{nk}$  は  $n$  サンプル目のクラス  $k$  の出力データ(確率)， $t_{nk}$  は  $n$  サンプル目のクラス  $k$  の教師データ(0 または 1)である。式(6)によって各データの損失関数の値が算出され，それらを合計し全体の損

失関数の値として使用した。

$$E_n = -\sum_{k=1}^K t_{nk} \log y_{nk} \quad \text{式 (6)}$$

バッチサイズ<sup>11)</sup>とは画像をひとまとめにしてコンピューターで計算するうえで処理を楽にするものである。バッチ処理することによって、画像一枚あたりの処理時間を大幅に短縮することが可能である。また、ニューラルネットワークの計算において、データ転送がボトルネックになる場合は、バッチ処理を行うことで、バス帯域の負荷を軽減することが可能になる。本研究ではバッチサイズを 32 として判別モデルの作成を行った。

#### (4) 判別モデルの検証

判別モデルの構築の後、そのモデルの判別能力を検証する。(1)の道路閉塞 162 枚、道路被害 323 枚、無被害 13996 枚の精度評価用画像をすべて判別させ、その判別精度を検証する。1枚の画像につき道路閉塞、道路被害、無被害の確率が算出されるので、最も大きい確率が付与された分類をその画像の判別結果とした。さらに、本研究に使用した画像すべてに緯度経度が付加されているので、これを基に地理情報システム (GIS) 上に表示させ精度を可視化した。

### 3. 学習過程と判別結果

図5に学習500回ごとのテストの正確性、図6に損失関数、図7に学習100回ごとの損失関数を示す。一般に損失関数は学習時、テスト時共に0付近に漸近することが理想である。本研究の結果は、テスト時の損失関数が増大する傾向にあり、過学習の傾向がある<sup>12)</sup>。しかし、テスト時の正確性に関しては7割付近まで上昇している。

構築した判別モデルに精度評価用画像を判別させた結果を表1にまとめる。全体の正答率は53.9%となった。道路閉塞、道路被害は約70%の確率で判別できたものの、無被害を無被害と正しく判別する精度が低かった。100%の確率で正しく無被害とした画像の例を図8に、正しく道路被害とした画像の例を図9に、正しく道路閉塞とした画像の例を図10に示す。

以上の結果をGIS上に表示した。図11に道路の被害状況 (目視による判別)、図12に判別モデルによる結果を示す。無被害が道路閉塞、道路被害と誤判別されていることが分かる。

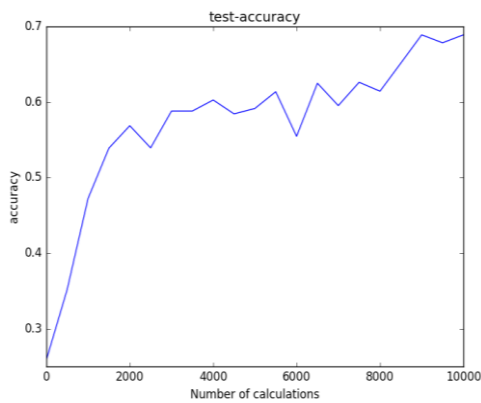


図5 テスト時の正確性

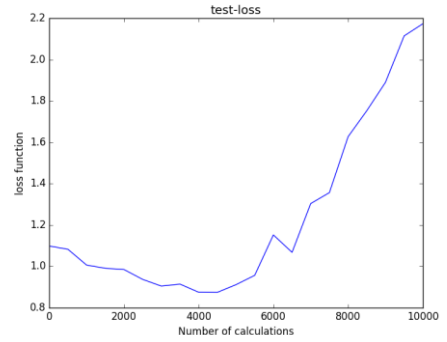


図6 テスト時の損失関数

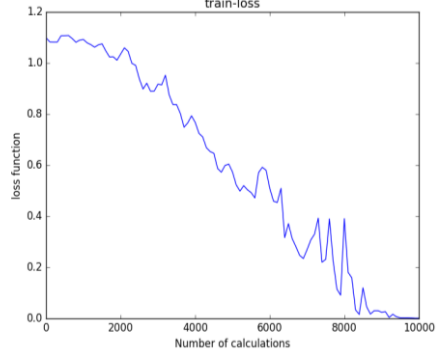


図7 学習時の損失関数

表1 画像の判別精度

	総数	正答枚数	正答率
道路閉塞	162	120	74.1%
道路被害	323	225	69.7%
無被害	14014	7476	53.3%



図8 無被害と正しく判別した画像例



図9 道路被害と正しく判別した画像例



図10 道路閉塞と正しく判別した画像例



図11 道路の被害状況 (目視判別の結果)



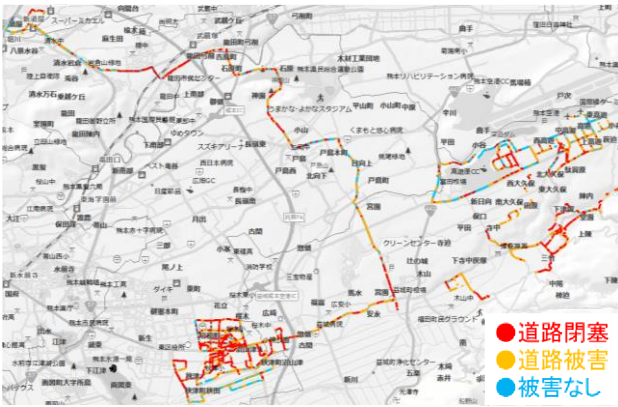


図 12 判別モデルが推定した被害状況

#### 4. 判別モデルの考察

本研究の結果では、無被害である画像を無被害と判別する能力が低かった。しかしながら、道路閉塞、道路被害、無被害すべてにおいて、判別精度が50%を超えていることから、本モデルを基礎として教師用画像を増やすなど改善を行えば精度向上が見込まれる。

無被害の画像を詳しく分析すると、無被害にもかかわらず100%の確率で道路閉塞と誤判別された画像が38枚、道路被害と誤判別された画像が84枚、無被害と正しく判別された画像が604枚存在した。これらの画像の中でセンターラインがない画像は、道路閉塞と誤判別されたものが28枚(73.7%)、道路被害と誤判別されたものが59枚(70.2%)、無被害で44枚(7.3%)であった。このことから、センターラインがないと道路閉塞または道路被害と誤判別される傾向があるといえる。

図13は無被害にも関わらず道路閉塞と判別された画像の例、図14は無被害にも関わらず道路被害と判別された画像の例である。センターラインがない道路は道幅が非常に狭く、交通量も少ない道路と想像される。そのため、整備が充分でなく、センターラインのある道路とは平常時の状況が異なるものと考えられる。センターラインがある画像のみを学習させるなどの基準をつくることで判別器の精度向上が見込まれる。



図 13 道路閉塞と誤判別した例



図 14 道路被害と誤判別した例

#### 5. まとめ

本研究は熊本地震後に撮影された車載カメラ画像を用いて、道路閉塞、道路被害、無被害を判別するモデルを深層学習で構築し、その精度を検証した。本研究では約53.9%の判別精度となったが、道路閉塞と道路被害に限

ればある程度の精度で判別できた。教師用画像データを見直すことで精度向上が期待できる。具体的には教師用画像を増やしたり、センターラインの有無を考慮するなど、教師用画像の選定をさらに細分化することである。また、ネットワークの構成や学習率などを最適化することでも精度の向上は計れると考えている。

#### 参考文献

- 1) 国土交通省 熊本地震による被災及び復旧状況 : <http://www.mlit.go.jp/common/001135910.pdf>
- 2) 的確な震後対応のための被災イメージ活用と被災状況の迅速な把握技術の開発 : <http://www.mlit.go.jp/chosahokoku/h17giken/program/kadai/pdf/shitei/shi2-01.pdf>
- 3) 浅川伸一 : Pythonで体験する深層学習, コロナ社, 2016.
- 4) 石井友, 松岡昌志, 牧紀男, 堀江啓, 田中聡, 中村良介 : 現地写真の深層学習による建物被災度判定の試み, 日本リモートセンシング学会第60回学術講演会, Paper No. U06, 2016.
- 5) Deep Learningハンズオン勉強会 : <http://www.slideshare.net/yasuyukisugai/deep-learningcaffe>
- 6) 田村誠信, 濱上知樹 : 深層学習を用いたモチーフ分析にもとづく小屏風画像の特徴分析, 人文科学とコンピューターシンポジウム, pp. 247-252, 2015
- 7) 岡谷貴之 : 深層学習, 講談社, 2015.
- 8) 武井宏将 : 初めてのディープラーニング, リックテレコム, 2016.
- 9) Slide Share Deep Learningと画像認識~歴史・理論・実践~ : [https://www.slideshare.net/nlab\\_utokyo/deep-learning-40959442](https://www.slideshare.net/nlab_utokyo/deep-learning-40959442)
- 10) 藤田一弥, 高原歩 : 実装ディープラーニング, オーム社, 2016
- 11) 斎藤康毅 : ゼロから作るDeep Learning, オライリージャパン, 2016
- 12) 機械学習モデルの汎化を妨げる過学習とは? 旗揚げ画像の二値分類を例に : <https://www.altus5.co.jp/blog/machine-learning/2016/11/28/over-learning/>