(43) 畳み込みニューラルネットワークを用いた橋梁部位の領域抽出の研究

岩下　将也[1]・山口　治[2]・大塚　義一[3]・梅田　貴大[4]・角田　賢明[4]

[1]正会員　株式会社奥村組　技術研究所（〒300-2612 茨城県つくば市大砂 387）
[2]正会員　株式会社奥村組　東日本支社リニューアル技術部（〒108-8381 東京都港区芝 5-6-1）
E-mail:osamu.yamaguchi@okumuragumi.jp
[3]正会員　株式会社奥村組　土木本部（〒545-8555 大阪市阿倍野区松崎町 2-2-2）
[4]非会員　株式会社ジャストイノベーション企画部（〒225-0012 神奈川県横浜市青葉区あざみ野南 2-4-1）
E-mail:justinnovation@just-Ltd.co.jp

インフラ構造物の維持管理・更新にかかる労務の効率化が喫緊の課題となっている．AI を用いてインフラ構造物の損傷を検出する研究事例が多数出ているが，構造物の損傷の原因や対策を考える上では，損傷の部位が重要な情報になる．本研究では橋梁構造物を対象に AI を用いて構造物の 10 種の部位を検出する実験を行った．10 種いずれかの橋梁部位を含む画像を畳み込みニューラルネットワークによって学習し，画像から部位の検出を試みたところ，転移学習を用いたモデルで$mIoU$47%の性能を得た．部位によって検出性能は異なり，特に画像の中で占める割合が小さい部位は，$mIoU$が低下する結果となった．本稿では，これらの実験方法，実験結果，考察について報告する．

Key Words:*bridge, structure parts, image segmentation, convolutional neural network, transfer learning*

1.　はじめに

インフラ構造物の老朽化は加速度的に進んでおり，維持管理・更新の必要性は，今後さらに増大していくと言われている．労働人口が減り続ける中，維持管理・更新にかかる労務の効率化は喫緊の課題となっている．

維持管理・更新にかかる労務の効率化の実現に向け，近年著しく性能が向上している AI による画像認識技術を適用する研究事例が増えている．深層学習と呼ばれる高度なニューラルネットワークを用いれば，従来では困難であった人間の直感に近い画像認識が実現でき損傷部位の目視点検の補助や代替が期待できる．既に構造物の損傷を矩形領域（バウンディングボックス）で検出する手法 [1)]，損傷の程度やその形状を領域（セグメンテーション）で検出する手法 [2)] などで，多くの研究が報告されており，これらの技術を利用した目視点検の代替，補助が期待される．

点検・診断業務においては，目視等で確認された損傷をもとに，損傷の危険性の評価，発生原因の推定，対策の検討など，維持・更新に向けた対応を講じる必要がある．このような二次的な判断をするためには，損傷の種類，位置，形状だけでなく，その損傷がどの構造部位に発生しているかという背景情報が重要である．**図-1** に示すように構造物の部位を AI によって検出することができれば，損傷の検出結果と合わせることで，損傷の原因，危険性や対策などの二次的な判断の推定に AI が活用できる可能性がある．本研究では，診断・点検業務の効率化を目的とした基礎的研究として，橋梁構造物を対象に，畳み込みニューラルネットワークを用いて画像から部位領域を検出する実験を行った．本稿では，実験方法と推定性能について報告をする．

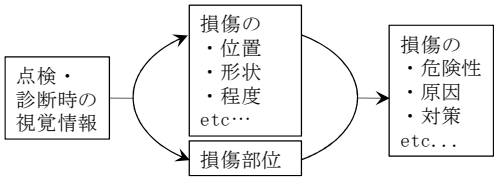


図-1　損傷と損傷部位からの損傷評価

2.　橋梁の部位と検出対象

インフラ構造物の中でも，橋梁はトンネルや管きょと異なり，屋外構造物であること，撮影地点の制約の都合

などから，構造物を比較的全景で撮影する状況が多く，画像内に複数種の部位が含まれることが多い．

国土交通省の定める「橋梁の定期点検要領」[3] では，橋梁の部位・部材として，主桁などを含む 29 種の上部構造，梁部などを含む 8 種の下部構造，支承本体を含む 6 種の支承部，高欄などを含む 8 種の路上部位・部材，その他，排水設備等の 6 種，合計 57 種の部位・部材が規定されており，多数の部位・部材が構造物全体に配置されている．

同要領では，損傷とその程度の判定，損傷の対策区分の判定，橋梁の健全性評価の各方法についても定められている．損傷の対策区分は損傷の発生部位によっても変わり，例えば床版に遊離石灰や鉄筋露出がある場合，「緊急の対策が必要」と判断される．部位は損傷の危険性や対策を考える上で，重要な参照情報となっている．

本実験では，同要領で規定される 57 種の橋梁部位・部材の中から，構造上の重要性，画像上の視認性の観点から，**表-1** に定める 10 種の部位を検出対象とした．

3. 実験方法

(1) 教師あり学習

畳み込みニューラルネットワークを用いた物体検出は，広義には教師あり学習の 1 つであり，データセットとアルゴリズムを基に，推論モデルを自動的に生成する．

本実験における教師あり学習の進め方を**図-2** に示す．**表-1** の橋梁部位を最低 1 つ含む橋梁部位の撮影画像（以下，橋梁部位画像）550 枚を教師データとし，440 枚の訓練・検証データと，110 枚の評価データに分割する．これらの画像から，10 種の部位，もしくはそれ以外の計 11 種の領域に分けた正解画像を作成する（アノテーション）．訓練・検証データの橋梁部位画像と正解画像を基に，橋梁部位画像から正解画像を出力する計算モデルを生成する．（この過程を学習と呼ぶ）．生成された計算モデルに評価データの橋梁部位画像を入れ，部位画像を出力する．出力された画像と評価データの正解画像を比較し，計算モデルの性能を評価する．

なお形状の領域検出の考え方には，物体ごとに領域を分けるインスタンスセグメンテーションと，種別ごとに領域を分けるセマンティックセグメンテーションの 2 種があるが，構造物の部位は固有の物体として個別に判断できないものが多いことから，本実験ではセマンティックセグメンテーションを用いている．

表-2 に収集した橋梁部位画像の諸元を，**図-3** に作成した正解画像のサンプルを合わせて示す．画像の中に複数の部位があることから，部位数は合計 2,682 となった．

表-1 検出対象の橋梁 10 部位

上部構造		1. 主桁	2. 縦桁
		3. 横桁	4床版
下部構造	橋脚	5. 柱部･壁部	6. 梁部
	橋台	7. 胸壁	8. 竪壁
支承部		9. 支承本体	
路上		10. 高欄	

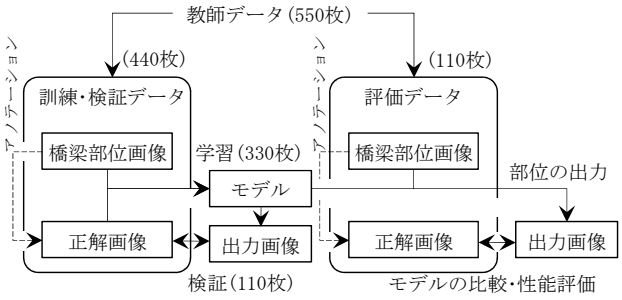
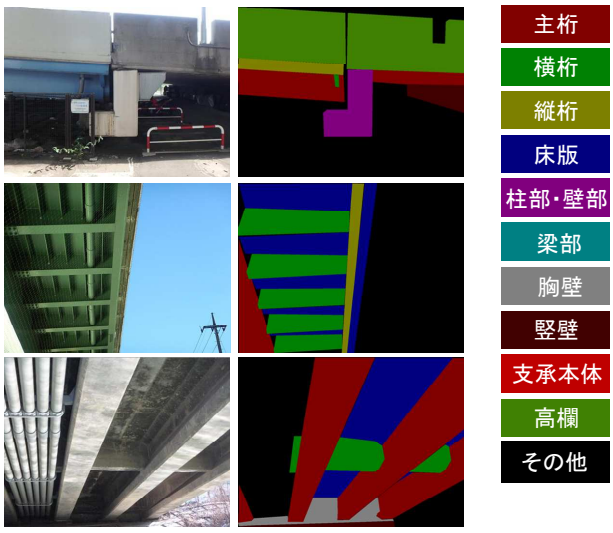


図-2 本実験での教師あり学習の進め方

表-2 全撮影画像 550 枚の諸元

	枚数	構成率※1 [%]	出現率※2 [%]	占有率※3 [%]
1. 主桁	481	17.9	87.5	22.7
2. 横桁	209	7.8	38.0	2.5
3. 縦桁	93	3.5	16.9	1.5
4. 床版	321	12.0	58.4	10.5
5. 柱部･壁部	159	5.9	28.9	3.2
6. 梁部	221	8.2	40.2	8.4
7. 胸壁	126	4.7	22.9	2.0
8, 竪壁	199	7.4	36.2	6.8
9. 支承本体	124	4.6	22.5	2.7
10. 高欄	199	7.4	36.2	3.4
11. その他	550	20.5	100.0	36.0
データ総数	2,682	100		100

※1 構成率=部位の画像枚数／データ総数（2,682）
※2 出現率=部位の画像枚数／画像総数（550）
※3 占有率=部位面積／部位出現画像総面積



橋梁部位画像　　正解画像

図-3 アノテーション済み画像サンプル

図-4　適用した水増しパターン

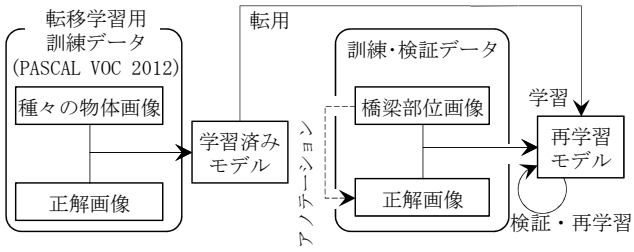


図-5　転移学習のイメージ

(2)　モデルの決定と学習手法

セマンティックセグメンテーションのための畳み込みニューラルネットワークのモデルは，複数報告されているが，本研究では 2016 年に発表された Pyramid Scene Parsing Model（以下，PSPNet[4]）を用いた．従来のセマンティックセグメンテーションモデルと比較し，広域に分布する特徴をより高性能に検出できる．点検診断時の画像撮影では，構造物の局所的な画像の中に部位が広範に分布している状況が多くあり，広域に分布する部位の形状を検出するモデルとして適していると考えた．

(3) 学習手法

これまで説明した考え方で生成したモデルを「モデル 1」とし，これに対し以下の 2 つの学習手法を適用して性能向上を試みた．

a)　水増しの利用

訓練データの画像に対し種々の変換を行い，モデルの汎化性能を高める手法である．本実験では，訓練データの学習時に図-4 に示す 3 パターンの画像変換をランダムに行った．水増しデータを合わせ学習したモデルを，「モデル 2」とする．

b)転移学習の利用

異なるデータセットを用いた学習済みモデルを，別の問題に転用し，部分的な学習をすることで，モデルの性能向上を狙う手法である．図-5 に転移学習の考え方を示す．特に教師データが十分でない場合に，推論性能の向上と学習時間の低減が期待できる手法である．本実験では，22 種の物体をセマンティックセグメンテーションにてラベル付けしている画像データセット Pascal VOC 2012Semantic Segmentation Dataset[5]（図-6）を用い，転移学習をした．水増しおよび転移学習をしたモデルを「モデル 3」とする．

図-6　PascalVOC 2012 のデータセットサンプル[5]

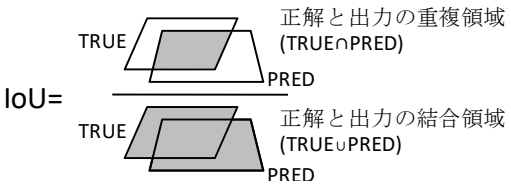


図-7 IoU の考え方

表-3　各モデルの学習時間と mIoU

	水増し	転移学習	mIoU[%]
モデル1	なし	なし	25.1
モデル2	あり	なし	30.2
モデル3	あり	あり	47.0

(4)　モデルの性能評価方法

生成された 3 つのモデルに対し，評価用の橋梁部位画像を入力し，各部位の画像を出力する．出力された部位と，正解の部位を比較し，生成されたモデルの性能を評価する．モデルの性能評価にあたっては，*mIoU*（Mean Intersection Over Union）を用いた．IoU とは図-7 に示す通り，正解領域と出力領域の一致率を数値化したもので，本実験では画像番号i部位番号jをパラメータとした式(1)によって求める．これを全評価画像 110 枚で平均した部位ごとの$mIoU_j$を式(2)で算出，この結果をさらに全部位で平均した$mIoU$で式(3)を算出し，モデルの性能を評価した．

$$IoU_{i,j} = \frac{PRED_{i,j} \cap TRUE_{i,j}}{PRED_{i,j} \cup TRUE_{i,j}} \tag{1}$$

$$mIoU_j = \frac{1}{I}\sum_{i=1}^{I} IoU_{i,j} \tag{2}$$

$$mIoU = \frac{1}{J}\sum_{j=1}^{J} mIoU_j \tag{3}$$

ここに，

i, j：画像番号，部位番号

I, J: 評価画像総数(= 110)，部位総数(= 11)

$PRED_{i,j}$：画像iにおける部位jと出力した画素

$TRUE_{i,j}$：画像iにおける部位jの正解画素

4.　実験結果

(1)　各モデルの性能

表-3 に各モデルの $mIoU$ を示す．水増しと転移学習

を適用したモデル 3 で，$mIoU$が最大となった．参考として，本実験で収集したの教師データの数（550 枚）は一般的な深層学習で扱うデータ数（数千～数万枚以上）と比較して，少量であり特に転移学習が有効に機能したものと考えられる．

(2) モデル 3 の考察

a) 部位ごとの$mIoU_j$

図-8 にモデル 3 の式(2)で算出される部位ごとの$mIoU_j$と部位ごとの占有率平均値の対応を示す．図中に$mIoU$を合わせて示す。横桁，縦桁，高欄，胸壁で$mIoU_j$が平均を下回っているが、これらの部位は占有率も低いことがわかる．占有率は画像内の部位の占める割合であり、これが低いことが検出に影響を与えていると考えられる．桁は橋梁の下面にあり、必然的に遠方からの撮影となること，高欄や胸壁のように部位自体の面積が全体的に小さいことなどが影響していると推察される．

b) 出力結果

出力結果のうち正解画像と良好な対応が取れている画像サンプルを**図-9** に，対応が不良な画像サンプルを**図-10** に示す．支承本体のような特徴的な部位が画像内にわかりやすく収まっている画像，橋梁の全景がわかる画像では，正解とよく対応している．しかし，局所的な部位のみが大きく映っている画像，検出対象以外の対象物（配管）が大きく映っている画像，日射によるコントラストがある画像などで対応が悪くなっている．診断，点検業務では晴天・曇天に限らず，局所的な部位の撮影をすることも多いため，このような画像に対する性能を向上させていく必要があり，今後の課題である．

5. まとめ

橋梁の損傷の原因，対策などを推定する参照情報として，画像から部位を検出することを目的とし，AI を用いて 10 種の橋梁部位の領域を出力する実験を行った．水増しと転移学習を用いた場合に$mIoU$は 47.0%と最大となり，特に転移学習の有効性が確認された. ただし，損傷の背景データとして用いるには実用的な性能とは言えず，継続した研究が必要である．特に，占有率の小さい部位に着目した再学習や，損傷画像の撮影する点検・診断員の見知から見た対策などを模索していきたい．

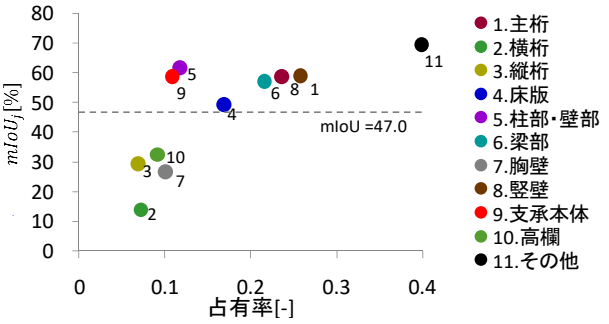


図-8 モデル 3 の各部位ごとの占有率と$mIoU_j$

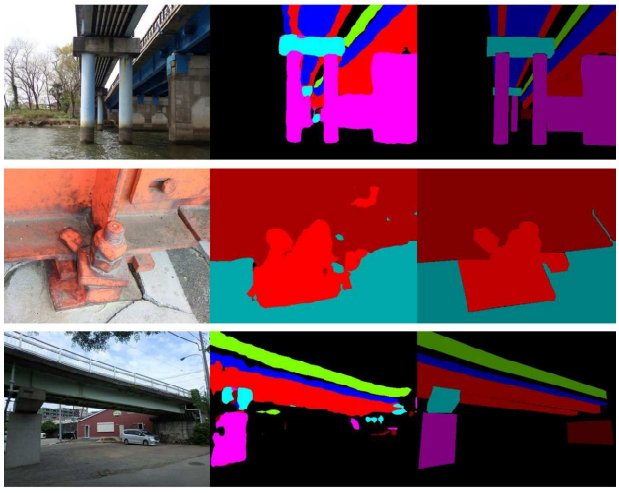


図-9 正解画像と良好な対応の出力画像サンプル

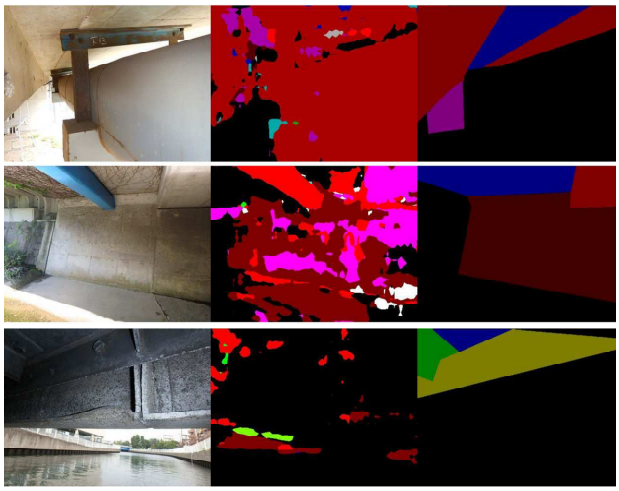


図-10 正解画像と対応が不良な出力画像サンプル

参考文献

1) 例えば，畳み込みニューラルネットワークに基づくコンクリート表面のひび割れ検出システム:野村泰稔，土木学会，2017，（入手 2018.5.4）
2) 例えば，目視点検の損傷画像による鉄筋露出セグメンテーションの転移学習：中島道浩，他:人工知能学会大会，2019，（入手 2019.6.10）
3) 国土交通省：橋梁の定期点検要領，pp. 10-13，2019，（入手 2019.5.20）
4) Hengshuang, Z., Jianping, S., Xiaojuan, Q., Xiaogang, W., Jiaya, J.: Pyramid scene parsing network, *IEEE Conference on computer vision and pattern recognition*, pp. 6230–6239, 2017,（入手 2018.6.1）
5) Everingham, M., Van Gool, L. K. I., Williams, C., Winn, J., Zisserman, A.: The PASCAL Visual Object Classes （VOC） Challenge, *International Journal of Computer Vision vol. 88*, no. 2, pp. 303-338, 2010,（入手 2018.6.1）