

博士論文

深層学習によるコンクリート道路橋の  
橋梁定期点検要領に準拠した変状検出に関する研究

A Study on Concrete Deterioration Detection Based on Periodic Bridge  
Inspection Guideline for Concrete Road Bridges by Deep Learning

2020年11月

青島 亘佐

山口大学大学院創成科学研究科



# 深層学習によるコンクリート道路橋の 橋梁定期点検要領に準拠した変状検出に関する研究

## 【論文の要旨】

青島 亘佐

我が国の膨大な社会インフラは、高度経済成長期に建設された施設を中心に、近い将来一斉に老朽化の目安とされる建設後 50 年に到達し、そのような施設の割合が今後加速度的に高まっていくことが予見されている。さらに、熟練技術者の退職や生産年齢人口の減少により、今後の専門技術者の不足も懸念されている。そのような社会的な潮流の中、平成 26 年に全国の道路構造物のうち、橋梁やトンネル等については、近接目視による 5 年に 1 回の定期点検が法律で定められた。近年のデフレが進行する社会情勢や財政事情において、この法律改正による市町村への負担の増大の影響は大きく、主要な道路構造物に対する点検は、これまで以上に膨大な予算およびマンパワーが必要となった。維持管理コストの縮減は喫緊の課題であり、特に膨大な施設数である道路橋は、社会インフラ全体の維持管理コストに占める比率が高く、縮減の必要性が高い。そこで、点検→診断→措置→記録の工程で構成されるメンテナンスサイクルの起点となる「点検」に焦点を当てて、コストの縮減に寄与する技術の研究を行った。

本研究は、橋梁点検の省力化のために、コンクリート変状の画像からの検出に取り組んだ研究成果を取りまとめたものである。本論文は全 6 章から成り、各章の主な内容は以下のとおりである。

「第 1 章 序論」では、我が国における社会インフラの維持管理の現状と課題をまとめるとともに、コンクリートの画像からの変状検出に関する既往の研究動向の整理を行った。その上で本研究の目的と着眼点を整理し、本論文の構成と概要について述べた。

「第 2 章 深層学習による変状画像のクラス分類」では、深層学習によるコンクリートの変状検出の基礎検討として、変状画像のクラス分類について検討を行った。また、深層学習の課題の一つに挙げられるブラックボックス問題への懸念から、テストデータに対するクラス分類実施時の判断根拠の可視化による確認にも取り組んだ。検討の結果、深層学習による物体認識技術に転移学習等の精度向上の方策を用いることで、コンクリ

ートの変状の画像について、画像内の適切な部分に着目した上で、変状の種類のみならず変状のランクの区分まで含めてクラス分類を精度よく行えることが検証できた。

「第3章 深層学習による物体検出およびセグメンテーション」では、難易度が低いタスクであるクラス分類が適切に行われたことを受けて、クラスの分類に加えて対象とする物体の領域を特定することから難易度が数段高いタスクとなる、物体検出手法による検出とセグメンテーション手法による検出について検討を行った。本研究の主目的である、セグメンテーションによるピクセル単位での変状の検出に向けて、複数の手法についての比較評価を行い、コンクリートの変状の検出に対して適性の高い手法を選定するとともに、精度よく領域を検出する方法を検討した。具体的には、矩形領域で変状を検出する物体検出を前段で行って、その矩形領域に対して物体検出時に特定された種類の変状と背景との2クラス分類のセマンティックセグメンテーションを実行する。この手法により、多クラス分類のセマンティックセグメンテーションを一括して行った場合よりも、実際の点検における評価に沿った適切な検出結果を得ることが可能となった。

社会インフラの点検という目的の性質上、検出漏れ (False Negative) の発生を最小限に抑制する必要があるが、あらゆる多様性を有するコンクリートの変状画像への物体検出およびセグメンテーションの適用は、まだまだ難易度が高い。そのため、続く第4章、第5章では、第3章で検討を行った本研究の変状検出の軸であるセグメンテーションを補完する技術について検討を行った。

「第4章 変状検出における深度画像活用の検討」では、本研究のセグメンテーション手法の前段部分となる、物体検出による変状領域の検出における不確かさを排除する目的で、ピクセル単位で深度の情報を有する深度画像を併用した検出手法について検討を行った。コンクリート面に一定量以上の凹凸が併発する変状に適用が限られるが、深度という物理量を用いることで、大量の教師データから変状の特徴量と位置を統計的に学習するよりも容易に領域の検出を行えることを検証した。

また、深度画像は、得られる深度の情報と撮影機器の画角を基に1ピクセルあたりの尺度を把握できることから、変状の定量値を把握する処理方法の検討も併せて実施した。深度画像についての一連の検討により、従来のRGB画像に加えて深度画像も取得できるRGB-Dカメラが、目視点検の省力化および効率化に有効なデバイスであることを立証した。

「第5章 深層学習による異常検知手法を用いた変状検出の検討」では、もう一つの補完技術として、収集が容易なコンクリートの正常画像を用いた教師なし学習による異常検知手法について検討した。検討の結果、深層学習の生成モデルを応用した AnoGAN と深層学習の識別モデルを応用した DOC を併用した異常検知により、複数の変状が存在

する広域的な画像に対しても何らかの異常が疑われる領域をスクリーニング的に検出でき、検出漏れを補完可能であることが検証できた。

「第6章 結論」では、第2章から第5章までで得られた検討結果の総括と今後の課題について述べた。

本研究の成果は、深層学習による画像認識技術を用いて、多くの研究が行われてきたひび割れ以外の、剥離・鉄筋露出、漏水・遊離石灰という橋梁定期点検要領に規定された変状も対象として、変状のランクの区分も踏まえた検出を可能とする手法、実務的な運用に向けてエラーを最小化するために併用する補完技術を開発したことである。



# A Study on Concrete Deterioration Detection Based on Periodic Bridge Inspection Guideline for Concrete Road Bridges by Deep Learning

[Abstract]

**Kosuke AOSHIMA**

The vast amounts of social infrastructures in Japan, centering on facilities constructed during the period of high economic growth, are reaching their designed average lifespan of 50 years. It is predicted that the proportion of such facilities will increase at an accelerated rate in the future. Furthermore, due to the retirement of skilled technicians and the decline of the working-age population, a shortage of professional civil engineers is becoming more concerned. Whereas, under such social trends, it was set under the law in 2014 that among road structures nationwide, bridges and tunnels shall be periodic inspected by close visual inspection for every 5 years. The revision of this regulation largely increased the burden on the government of cities, towns and villages, as it requires a larger amount of budget and manpower than ever, especially under the deflation progressing financial situation in recent years. Reducing maintenance costs becomes an urgent issue, especially for road bridges which own an enormous number of facilities that making up a large proportion of the maintenance cost of infrastructures. Therefore, I decided to study the technology that contributes to cost reduction, focusing on the "inspection" that is the starting point of the maintenance cycle consisting of the processes of inspection → diagnosis → countermeasures → recording.

In this study, in favor of labor saving in bridge inspection, the research results of concrete deterioration detection according to images recognition were summarized. This thesis consists of 6 chapters, and the main content of each chapter is as follows.

## **[Chapter 1: Introduction]**

The current status and issues of infrastructure maintenance in Japan were summarized, as well as the previous research results regarding concrete deterioration detection by images recognition. The purpose and focus point of this research and composition of this paper were narrated.

## **[Chapter 2: Classification of deterioration images by deep learning]**

As a basic examination of concrete deterioration detection by deep learning, I examined the classification of deterioration image classes. In addition, because of concern about the black box problem, which is one of the problems of deep learning, I also worked on the confirmation by visualizing the judgment basis when classifying test data. As a result, by using deep learning such as transfer learning for object recognition technology to improve accuracy, the image of the

concrete deterioration is focused on an appropriate part in the image, and it was verified that not only the types but also the rank level of deterioration can be classified.

### **[Chapter3: Object detection and segmentation by deep learning]**

The classification, as an easy task for image recognition, was appropriately performed. Therefore, in Chapter 3, I discussed the object detection and segmentation detection, which are tasks with increased several levels of difficulty due to the identification of the target object's region in addition to the classification. The main purpose of this study is to detect pixel-by-pixel segmentation, a method with high suitability for concrete deterioration detection is selected, and the improvement method for higher detection accuracy was discussed. Specifically, this method conducts object detection in advance that detects the deterioration in bounding box, and then performs the semantic segmentation detection under 2-class classification (deterioration of the type detected by object detection and background which means normal concrete surface). With this method, it is possible to obtain more appropriate detection results according to the evaluation in the actual inspection than when the semantic segmentation of multi-class classification is performed collectively.

On the purpose of application for actual infrastructure inspection, it is necessary to minimize the occurrence of false negatives detection. However, the application of object detection and segmentation detection to concrete with a variety of deterioration by image recognition is still in high difficulty. Therefore, in the following chapters 4 and 5, the complementary technologies for segmentation detection as is the main axis in this study that discussed in Chapter 3 was examined.

### **[Chapter 4: Utilization of depth images for deterioration detection]**

Detection method that uses depth images that have depth information in pixel units in order to eliminate the uncertainty in the deterioration region detection due to object detection was investigated, which is the front stage of the segmentation method. Though the application of this method is limited to the deterioration in which at least a certain amount of unevenness occurs on the concrete surface, it was verified that adopting depth image can detect the bounding box more easily than statistical learning of deterioration features and locations from a large amount of human-labeled training data.

Additionally, since the depth image can get a scale per pixel based on the depth information and the view angle of the camera device, the processing method to evaluate the deterioration quantitatively was also discussed. Through a series of studies on depth images, it was proved that the RGB-D camera, which can acquire the depth image in addition to the conventional RGB image, is an effective device for labor saving and improvement of visual inspection.



### **[Chapter 5: Deterioration detection using anomaly detection method by deep learning]**

As another complementary technology, anomaly detection method by unsupervised learning using images of easy-collecting regular concrete was examined. The anomaly detection method adopted both AnoGAN applying deep learning generative model and DOC applying deep learning discriminant model. It was validated that with this method, regions suspected of abnormalities can be detected by screening in a wide area image with multiple deterioration, and false negative can be complemented.

### **[Chapter 6: Summary]**

This chapter summarizes the study results obtained from Chapter 2 to 5, and describes the needs for further research on the labor-saving technology for bridge inspection.

As the results of this study, by using image recognition technology based on deep learning, it is applicable to detect concrete deterioration defined in the periodic bridge inspection guideline, such as peeling, reinforcement exposure, water leakage, calcium leaching and cracking, as well as the classification and levels of deterioration. Furthermore, the complementary technologies that can minimize errors in practical application were developed.



## 目次

---



# 目次

## 第 1 章 序論

1-1 本研究の背景.....	1
1-1-1 我が国における社会インフラの維持管理の現状と課題.....	1
1-1-2 情報通信技術の活用の要請.....	5
1-2 本研究の目的.....	6
1-3 既往研究の整理.....	8
1-3-1 画像処理技術を用いたコンクリートの変状検出に関する研究.....	8
1-3-2 深層学習技術を用いたコンクリートの変状検出に関する研究.....	9
1-3-3 国外におけるコンクリートの変状検出に関する研究.....	10
1-3-4 研究の着眼点.....	12
1-4 本論文の構成と概要.....	14
1-4-1 本研究の対象.....	14
1-4-2 本論文の構成と概要.....	18
第 1 章 参考文献.....	24

## 第 2 章 深層学習による変状画像のクラス分類

2-1 深層学習によるクラス分類.....	29
2-1-1 クラス分類の概要.....	29
2-1-2 深層学習によるクラス分類.....	31
2-2 変状画像のクラス分類.....	33
2-2-1 検討条件.....	33
2-2-2 汎化性能向上の方策.....	34
2-2-3 分類結果.....	37
2-3 判断根拠の可視化による確認.....	44
2-3-1 Grad-CAM.....	44
2-3-2 可視化結果.....	44
2-4 まとめ.....	51
第 2 章 参考文献.....	52

第 3 章 深層学習による物体検出およびセグメンテーション	
3-1 教師データの構成	53
3-2 深層学習による物体検出	55
3-2-1 物体検出の概要	55
3-2-2 深層学習による物体検出	56
3-3 変状検出への物体検出手法の適用	58
3-3-1 検討条件	58
3-3-2 検出結果	60
3-4 深層学習によるセグメンテーション	64
3-4-1 セグメンテーションの概要	64
3-4-2 セグメンテーションの分類	64
3-5 変状検出へのセグメンテーション手法の適用	66
3-5-1 適用手法の選定	66
3-5-2 セグメンテーションの精度向上のための検討	81
3-6 まとめ	88
第 3 章 参考文献	89
第 4 章 変状検出における深度画像活用の検討	
4-1 深度画像の概要	92
4-1-1 RGB-D カメラの概要	92
4-1-2 深度画像活用の可能性	93
4-2 深度画像を活用した変状検出	95
4-2-1 変状の検出方針	95
4-2-2 変状領域の検出	96
4-3 深度情報による変状の定量値の把握	101
4-3-1 変状の定量値の把握方針	101
4-3-2 定量値把握結果	102
4-4 まとめ	105
第 4 章 参考文献	106
第 5 章 深層学習による異常検知手法を用いた変状検出の検討	
5-1 深層学習による異常検知	107

5-1-1 異常検知の概要 .....	107
5-1-2 検討に用いる異常検知手法 .....	109
5-2 変状検出への異常検知手法の適用 .....	113
5-2-1 検討内容 .....	113
5-2-2 適用結果 .....	115
5-2-3 適用手法によるモデルの性能評価 .....	118
5-3 実用性の検証 .....	122
5-3-1 検証内容 .....	122
5-3-2 検証結果 .....	124
5-4 まとめ .....	127
第 5 章 参考文献 .....	128
第 6 章 総括	
6-1 本研究のまとめ .....	130
6-2 今後の課題 .....	135
第 6 章 参考文献 .....	137

謝辞





## 第 1 章 序論

---

### 第 1 章

- 1-1 本研究の背景
- 1-2 本研究の目的
- 1-3 既往研究の整理
- 1-4 本論文の構成と概要



## 第1章 序論

### 1-1 本研究の背景

#### 1-1-1 我が国における社会インフラの維持管理の現状と課題

##### (1) 社会インフラの維持管理の現状

我が国の膨大な社会インフラは、新設のピークであった高度経済成長期に建設された施設を中心に、近い将来一斉に老朽化の目安とされる建設後50年に到達し、そのような施設の割合が今後加速度的に高まっていくことが予見されている(表1-1参照)。そこで、主要な構造物に対して、定期的な巡視および点検により施設の健全度を把握した上で、対策の内容や時期等を長寿命化計画として策定することが推進されてきた。この長寿命化計画では、計画に基づき優先順位を定めて対策を実施することで、施設の長寿命化とともにトータルコストの縮減を図ること等を維持管理の基本方針としている。しかしながら、近年のデフレが進行する社会情勢や財政事情を反映して、社会インフラのメンテナンスに対する投資や施策等が十分とはいえない状況が長く続いていた。このような状況下で、2012年12月に笹子トンネルにおいて老朽化した天井板が崩落し、9名の方が亡く

表 1-1 建設後 50 年以上経過する社会インフラの割合<sup>1)</sup>

	2018年3月	2023年3月	2033年3月
道路橋 [約73万橋 <sup>注1)</sup> (橋長2m以上の橋)]	約25%	約39%	約63%
トンネル [約1万1千本 <sup>注2)</sup>	約20%	約27%	約42%
河川管理施設(水門等) [約1万施設 <sup>注3)</sup>	約32%	約42%	約62%
下水道管きよ [総延長:約47万km <sup>注4)</sup>	約4%	約8%	約21%
港湾岸壁 [約5千施設 <sup>注5)</sup> (水深-4.5m以深)]	約17%	約32%	約58%

注1) 道路橋約73万橋のうち、建設年度不明橋梁の約23万橋については、割合の算出にあたり除いている。(2017年度集計)

注2) トンネル約1万1千本のうち、建設年度不明トンネルの約400本については、割合の算出にあたり除いている。(2017年度集計)

注3) 国管理の施設のみ。建設年度が不明な約1,000施設を含む。(50年以内に整備された施設については概ね記録が存在していることから、建設年度が不明な施設は約50年以上経過した施設として整理している。)(2017年度集計)

注4) 建設年度が不明な約2万kmを含む。(30年以内に布設された管きよについては概ね記録が存在していることから、建設年度が不明な施設は約30年以上経過した施設として整理し、記録が確認できる経過年数毎の整備延長割合により不明な施設の整備延長を按分し、計上している。)(2017年度集計)

注5) 建設年度不明岸壁の約100施設については、割合の算出にあたり除いている。(2017年度集計)

なるという痛ましい事故が発生し、社会インフラのメンテナンス体制の本格的な見直しを行う契機となった。これを受けて、全国の道路構造物のうち、約70万の橋梁や約1万のトンネル等については、近接目視による5年に1回の定期点検が法律で定められ、平成26年7月1日より施行された<sup>2)</sup>。

## (2) 社会インフラの維持管理の課題

国土交通省管理の橋梁およびトンネルについては、平成26年以前から近接目視による5年に1回の定期点検が行われてきたが、地方公共団体、その中でも特に市町村では、すべての施設に対して近接目視点検を行っている管理者は多くはなかった。また、**図1-1**に示すように、社会インフラ全体において、市町村管理の施設数が占める比率は非常に高い。そのため、上記の法律改正による市町村への負担の増大の影響は大きく、主要な道路構造物に対する点検は、これまで以上に膨大な予算およびマンパワーが必要となった。

施設数が非常に多い道路橋では、国土交通省の橋梁定期点検要領<sup>3)</sup>において規定された、**表1-2**に示す判定区分に基づき、対策の優先順位が定められる。しかし、地方公共団体では、早期あるいは緊急に措置が必要となる判定区分Ⅲ、Ⅳの橋梁でさえも、修繕の着手率および完了率が非常に低い（**図1-2**参照）。地方公共団体においては、点検・診断・修繕等の措置や長寿命化計画等の充実を含む維持管理の業務サイクル（通称；メンテナンスサイクル<sup>4)</sup>）が、良好に機能しているとは言い難いというのが実状である。

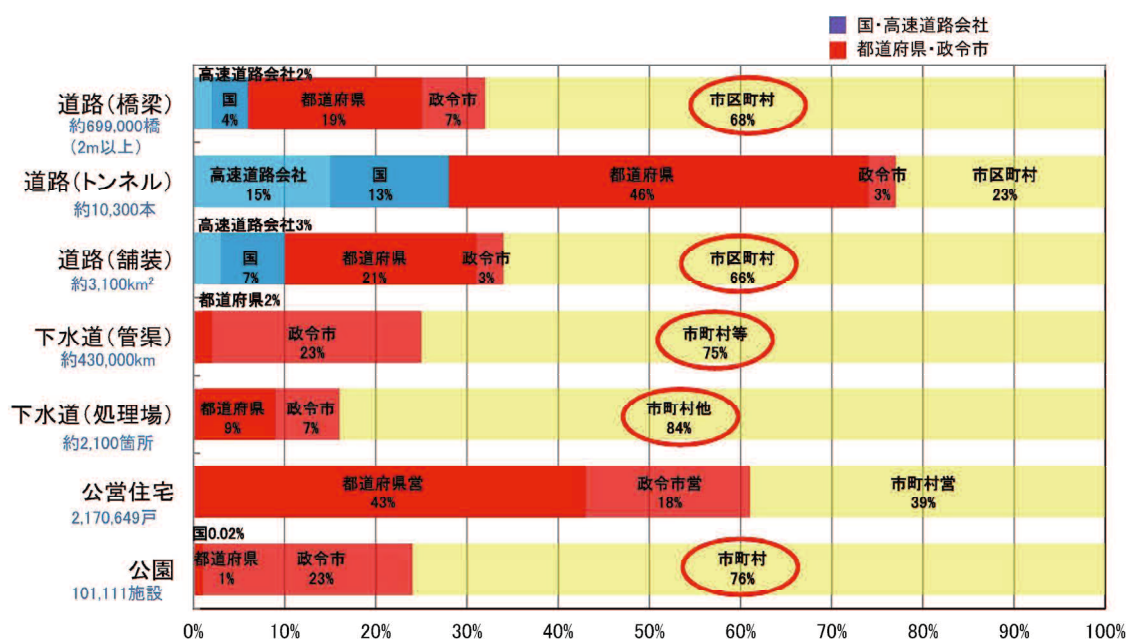


図1-1 各分野の管理者別の施設数の割合<sup>1)</sup>

表 1-2 道路橋の健全性の診断における判定区分<sup>3)</sup>

区分		定義
I	健全	道路橋の機能に支障が生じていない状態。
II	予防保全段階	道路橋の機能に支障が生じていないが、予防保全の観点から措置を講ずることが望ましい状態。
III	早期措置段階	道路橋の機能に支障が生じる可能性があり、早期に措置を講ずべき状態。
IV	緊急措置段階	道路橋の機能に支障が生じている、又は生じる可能性が著しく高く、緊急に措置を講ずべき状態。

判定区分Ⅲ・Ⅳの橋梁における修繕着手・完了率						
管理者	修繕が必要な施設数(A)	修繕着手済み施設数(B)	うち完了(C)	修繕未着手施設数	点検年度	修繕着手率(B/A)、完了率(C/A)
						0% 20% 40% 60% 80% 100%
国土交通省	3,427	1,811 (53%)	617 (18%)	1,616 (47%)	H26	39% 92%
					H27	27% 77%
					H28	12% 62%
					H29	7% 28%
					H30	8% 13%
高速道路会社	2,647	846 (32%)	457 (17%)	1,801 (68%)	H26	55% 78%
					H27	31% 49%
					H28	15% 41%
					H29	10% 25%
					H30	4% 7%
地方公共団体	62,977	12,700 (20%)	7,430 (12%)	50,277 (80%)	H26	26% 35%
					H27	18% 29%
					H28	12% 22%
					H29	4% 11%
					H30	2% 6%
都道府県政令市等	20,586	4,889 (24%)	2,684 (13%)	15,697 (76%)	H26	25% 34%
					H27	21% 35%
					H28	13% 26%
					H29	5% 16%
					H30	3% 9%
市区町村	42,391	7,811 (18%)	4,746 (11%)	34,580 (82%)	H26	27% 35%
					H27	16% 26%
					H28	11% 20%
					H29	4% 9%
					H30	2% 5%
合計	69,051	15,357(22%)	8,504(12%)	53,694(78%)		修繕完了済 修繕着手済

図 1-2 判定区分Ⅲ・Ⅳの橋梁における修繕着手・完了率<sup>5)</sup>

さらに、社会インフラの維持管理を適切に実行していくためには、それらを担う人材確保・育成や、ノウハウの蓄積、技術力向上が不可欠であるが、熟練技術者の退職や生

産年齢人口の減少により、今後の専門技術者の不足が懸念されている。建設業における技能労働者等の推移を図1-3に示し、建設業就業者の高齢化の進行を図1-4に示す。また、現在政府は「働き方改革」で労働時間の短縮を目指しており、社会インフラのメンテナンス分野においても、労働生産性の向上と労働条件の改善等の建設業の魅力向上による人材の確保が喫緊の課題となっている。

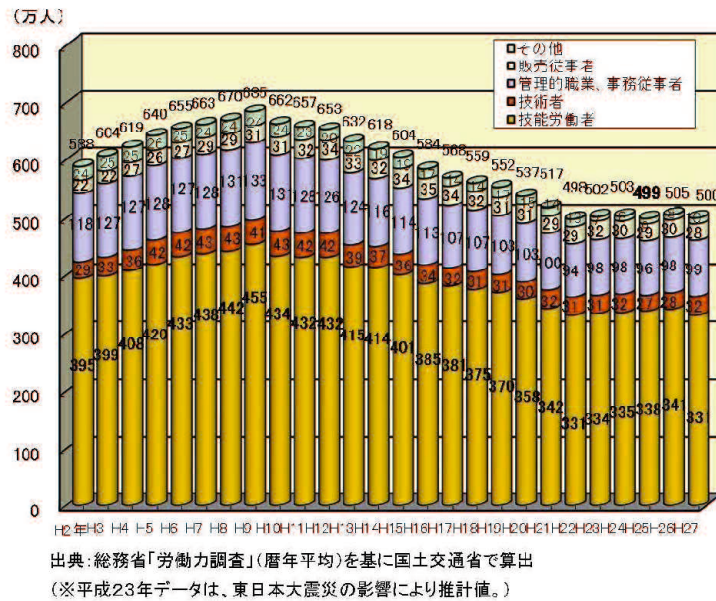


図 1-3 建設業における技能労働者等の推移<sup>6)</sup>

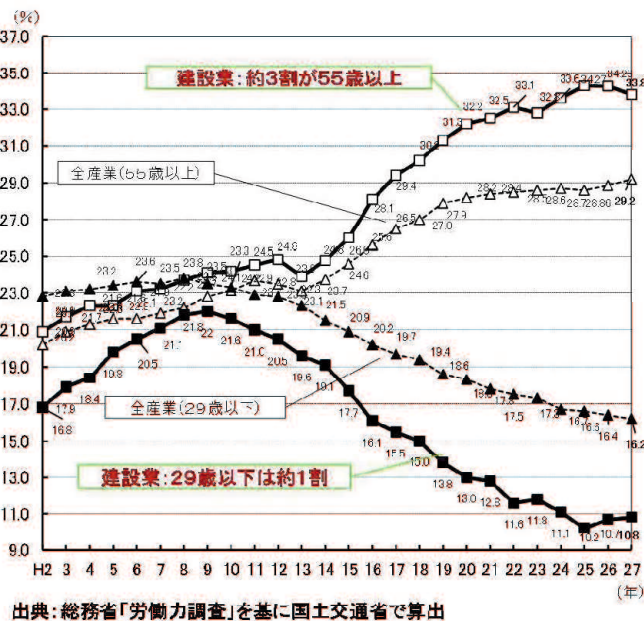


図 1-4 建設業就業者の高齢化の進行<sup>6)</sup>

### 1-1-2 情報通信技術の活用の要請

このような状況のもと、社会インフラの維持管理の課題に対して、効果的な対応を行っていく必要があるが、その対応策の1つとして、ICT（Information and Communication Technology；情報通信技術）に対する期待が近年高まっている。具体的には、維持管理技術の効率化や高度化に資する社会インフラのモニタリング技術、点検支援ロボット等の開発・促進が推進されており、維持管理コストの低減が期待されている。国土交通省では、建設業界の生産性と魅力の向上を目的として、調査・測量から維持管理・更新までの全てのプロセスでICT等を活用する「i-Construction」に取り組んでいるが、点検支援ロボット等の技術の試行も、その一環として進められている。

## 1-2 本研究の目的

1-1節で述べた背景の通り，維持管理コストの縮減は喫緊の課題である．特に，膨大な施設数である道路橋は，社会インフラ全体の維持管理コストに占める比率が非常に高いため，コストの縮減の必要性が高い．また，地方公共団体ではメンテナンスサイクルの実効性の課題がある旨を前述したが，このメンテナンスサイクルの実効性を向上させるためには，点検の信頼性の確保とともにコスト負担の軽減が不可欠である．そこで，点検→診断→措置→記録（→次の点検）の工程で構成されるメンテナンスサイクル（図1-5参照）の起点となる「点検」に焦点を当てて，コストの縮減に寄与する技術の研究を行うこととした．

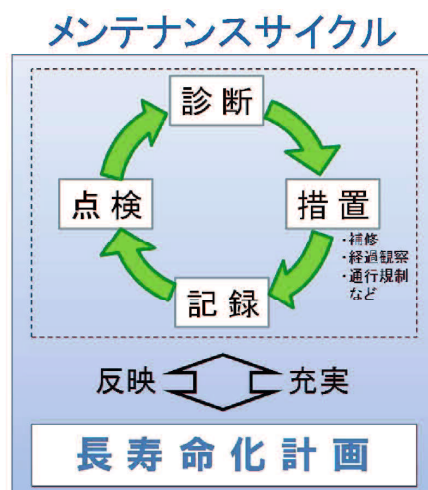


図1-5 メンテナンスサイクル<sup>4)</sup>

構造物の点検を合理的かつ経済的に行っていくための方策として，ICT の活用が推進されており，特に近年著しく進展している AI の一手法である深層学習が注目されている．深層学習に関する技術は，近年著しい進展を遂げており，現在では特定のタスクにおける識別能力は人間と比肩するレベルに達し，適用されるタスク領域が広がり，飛躍的な精度の向上がみられる．特に画像データへの適用は，深層学習の分野の中で最も研究が盛んな分野の一つで，幅広い研究が行われている．

道路橋の点検では，現地での外業では，発見した損傷に対して，部位や部材の最小評価単位ごと，変状の種類ごとに主要な寸法の計測や写真撮影，状況のスケッチ等を行って概況を把握し，変状に関する客観的な事実により設定された区分に基づき程度の分類を行う．また，内業では，スケッチを基にCAD等により変状図を作成し，変状写真や所



見等とともに橋梁単位で点検調書の取りまとめを行い、健全性の診断や、維持・補修等の計画の検討に利用する基礎資料として蓄積される。本研究では、深層学習の技術を用いて、これらの工程のうち、スケッチによる記録や程度の評価、変状図作成の工程の効率化に寄与する技術について検討を行う。道路橋点検の省力化イメージを図1-6に示す。

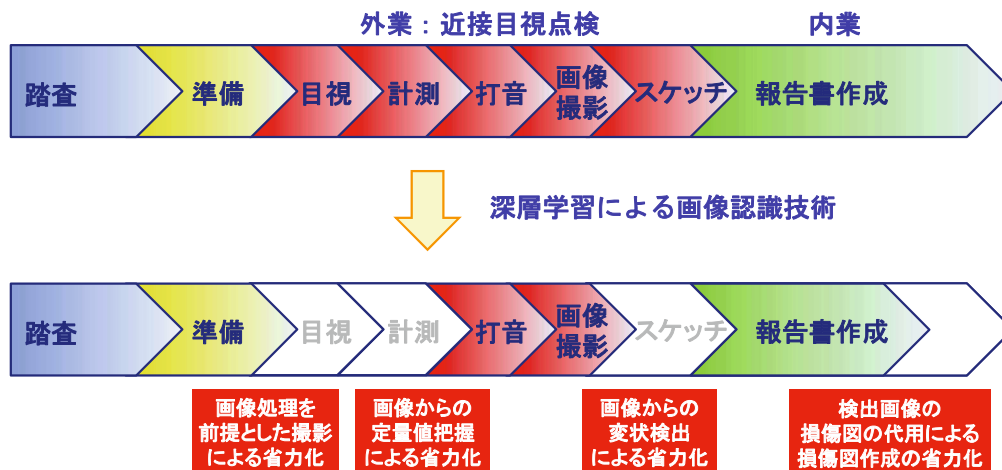


図 1-6 道路橋点検の省力化イメージ

### 1-3 既往研究の整理

構造物の点検を合理的かつ経済的に推進していくという社会的な要請に対して、これまでもコンクリート構造物のひび割れの抽出等において、画像処理技術や最適化手法、機械学習等を応用する研究が盛んに行われてきた。本研究を実施するにあたり、関連する分野の既往研究動向を以下に把握・整理した。

#### 1-3-1 画像処理技術を用いたコンクリートの変状検出に関する研究

コンクリートの画像からの変状検出の先駆けとなったものは、1998年に販売されたコンクリート構造物診断支援ソフトで、画像処理技術を活用し、オペレーターが判断しながらコンクリート表面のひび割れや剥落を半自動で抽出するソフトであった<sup>7)</sup>。その後、外業における変状状況の記録および内業におけるそれらの整理の省力化を目的として多くの技術開発が進められてきた。特に、コンクリートの代表的な変状であるひび割れは、点検時の記録に労力を要することから自動検出に関する研究が数多く行われている。

山口・橋本<sup>8)・9)</sup>は、ひび割れの局所的な直線性と方向性に着目した Gray-scale Hough 変換や形状情報と輝度情報を付加したパーコレーション処理による高精度化検出手法を提案している。藤田ら<sup>10)・12)</sup>は、多重スケール線強調処理および確率的弛緩法を用いて、幅別の分類が可能で、かつパラメータの設定を不要としたひび割れの検出手法を提案している。全ら<sup>13)・14)</sup>は、統計的な閾値の導出による二値化処理、連結要素における幾何学的特徴に着目した候補の絞り込みによるひび割れの検出手法を提案している。そして、ハフ変換により型枠跡およびPコン跡を同定して誤検出結果から除去することで自動検出の精度向上も図っている。画像の輝度以外の要素に着目した手法では、たとえば小山・丸屋ら<sup>15)・17)</sup>が提案した、デジタル画像にガボールフィルタに基づいたウェーブレット変換を適用してウェーブレット係数によりひび割れ判定を行う研究があり、実構造物での検証も行っている。遺伝的プログラミングあるいは遺伝的アルゴリズムを利用した手法では、たとえば西川ら<sup>18)</sup>は、遺伝的プログラミングにより木構造状フィルタ生成システムを構築し、これにより生成されたロバストな複合画像フィルタによるひび割れ検出手法を提案している。また、河村ら<sup>19)・21)</sup>は、対話型遺伝的アルゴリズムを用いた視覚的な画像処理パラメータ調整手法を提案している。

ひび割れ以外の変状を扱った研究として、トンネルの事例では、たとえば石森・塩崎・佐田ら<sup>22)・25)</sup>は、照明による陰影を利用して覆工コンクリート表面のひび割れとともに段

差を検出する手法を提案している。また、鶴飼ら<sup>26)-28)</sup>は、変状に応じた画像処理アルゴリズムを用いることにより、ひび割れのみならず閉合ひび割れや漏水、エフロレッセンスを検出する手法を提案している。この手法は、ラインセンサカメラを用いたトンネル変状検査システムとして、多くの鉄道路線で適用されている。トンネル以外の事例としては、鶴飼・長峯<sup>29)</sup>は、橋梁を対象として、閾値処理を行って抽出した領域と元画像の論理積を用いて、対象領域の絞り込みおよびエッジ抽出等を行い、はつり跡や修復跡を検出する手法を提案している。

### 1-3-2 深層学習技術を用いたコンクリートの変状検出に関する研究

画像からの変状の自動検出において、画像処理技術を主体とした手法では閾値の設定、遺伝的プログラミングあるいは遺伝的アルゴリズムを利用した手法ではフィルタの設定等が必要となる。しかし、これらのパラメータを調整して、多様な画像に対して高い頑健性を有するような設定を行うことは、膨大な労力を要する。

対して、大量なデータからの学習により特徴量を自動獲得する深層学習を用いた画像認識では、頑健性のある検出のためのパラメータの最適解を効率的に得られる。近年では、深層学習を利用した研究も多く見られ、ひび割れを対象とした研究では、たとえば野村ら<sup>30),31)</sup>は、深層学習に基づく物体検出手法によるひび割れ検出を行うとともに、深層生成モデルを用いた教師データの増強により、精度の向上を図っている。山根・全<sup>32)</sup>は、セグメンテーション手法の段階的な適用により、P コン跡および型枠跡の誤検出を防止したひび割れ検出を行っている。榎本ら<sup>33)</sup>は、トンネルを対象として、幅の分類も含めて学習を行った CNN によるひび割れの検出、敵対的生成ネットワークの GAN の 1 種である pix2pix を用いた画像生成によるひび割れの領域検出を行っている。機械学習あるいは深層学習と他の手法を組み合わせたひび割れ検出の研究では、たとえば全・井後による画像解析と Random Forest による機械学習を基にした大局的形状判定手法を組み合わせたひび割れ検出手法<sup>34)</sup>、それを発展させた深層学習と Random Forest によるひび割れ検出手法<sup>35)</sup>がある。さらに、中村・河村ら<sup>36)</sup>による深層学習と対話型遺伝的アルゴリズムによる半自動的抽出手法を組み合わせたひび割れ検出手法、鶴飼<sup>37)</sup>による深層学習を用いた識別モデルから得られたひび割れの存在領域を基に位置や方向性に着目した画像解析を行う検出手法等がある。

ひび割れ以外の変状を扱った研究では、たとえば横山・松本<sup>38)-41)</sup>は、深層学習を用いた画像認識によるひび割れ、エフロレッセンス、チョーク文字の自動検出器を開発し、Web システムとしての実装を行っている。矢吹<sup>42)</sup>は、深層学習に基づく物体検出手法に

より、ひび割れに加えて、鉄筋露出、漏水に該当する遊離石灰 c、遊離石灰の析出に該当する遊離石灰 d、著しい遊離石灰の析出等に該当する遊離石灰 e の 5 つのクラスに分類した上で、矩形領域での検出を行っている。ピクセル単位で領域を推定するセグメンテーション手法による検出では、柏・長井ら<sup>43)</sup>は、コンクリート床版におけるコンクリートのはく離および鉄筋の露出について検出を行っている。中島ら<sup>44)</sup>は、画像内に占める検出対象領域が疎である鉄筋露出の画像を教師画像とした条件において、転移学習を行ったセマンティックセグメンテーションのモデルが有効であることを検証している。

検出手法の主流はセグメンテーションに移行しつつあり、対象とする変状はひび割れ以外の変状にも拡大し始めている、という状況が国内の概ねの潮流である。

### 1-3-3 国外におけるコンクリートの変状検出に関する研究

国内と同様に、国外においてもコンクリートからの変状の検出はひび割れを中心に進められてきた。画像処理によるひび割れ検出の本格的な研究としては、Adbel-Qader ら<sup>45)</sup>による橋梁のひび割れ検出における高速ハールウェーブレット変換、高速フーリエ変換、Sobel フィルタ、Canny フィルタの有効性の比較が挙げられる。他の画像処理技術を用いた研究では、Hutchinson・Chen は、従来の画像処理によるひび割れ検出の課題であったエッジ分析の閾値に関する曖昧さに対して、人間の介入を最小限に抑えて ROC（受信者操作特性）分析を使用した最適な閾値設定の手順を提案し、高速フーリエ変換、Canny フィルタによる検出<sup>46)</sup>を行っている。また、制御された環境での撮影ではない画像のノイズを除去した上でひび割れを検出し、さらにレベルセット表現を用いてひび割れに関する面積と周長および幅の定量値を取得させる研究<sup>47)</sup>も行っている。Pereira ら<sup>48)</sup>は、UAV（無人航空機）で撮影した画像に対して、エッジデバイス上で粒子フィルタを用いてスクリーニング的にひび割れ検出を行っている。

機械学習の技術を用いた検出では、Adbel-Qader ら<sup>49)</sup>は、線形モデリングと PCA（主成分分析）アルゴリズムを使用したひび割れ検出を行っている。Prasanna ら<sup>50)</sup>は、画像内の各ブロックにおいて検出した線分に対して空間的に調整された複数の特徴量を計算し、機械学習により「ひび割れ」と「非ひび割れ」に分類させて検出を行っている。Li ら<sup>51)</sup>は、セグメンテーションのための領域ベースのモデルとノイズ除去のための線形 SVM（サポートベクターマシン）を組み合わせた、コンクリート橋の画像から全自動でひび割れを検出する機械学習ベースのアルゴリズムを開発している。

近年は、深層学習の技術を用いた検出の研究が多数見られ、たとえば Cha ら<sup>52)</sup>は、CNN（畳み込みニューラルネットワーク）を用いて、スライディングウィンドウで画像

の小片単位毎に「ひび割れあり」と「ひび割れなし」の分類を行い、ひび割れが生じている領域の検出を行っている。Kim・Cho<sup>53)</sup>は、Chaらの手法と同様にスライディングウィンドウでCNNを適用させた分類により検出を行っているが、分類のクラス数を5つ（①ひび割れのない健全なコンクリート面、②ひび割れ、③複数の打ち継ぎあるいは縁、④単一の打ち継ぎあるいは縁、⑤植物）に細分化している。また、小片内の4隅におけるひび割れの存在の過小評価の回避のために、4つの小片の1/4ずつを新たな小片として再度スライディングウィンドウで分類を行うオーバーラップウィンドウによる再検証を実施し、精度の向上を図っている。Alipourら<sup>54)</sup>は、FCN(完全畳み込みネットワーク)を用いて、セマンティックセグメンテーションによりピクセルレベルでのひび割れ領域の検出を行っている。Zhangら<sup>55)</sup>は、教師データを大幅に削減して自ら学習を行う概念の自己教師あり学習によるひび割れの検出を行って、教師あり学習によるひび割れ検出と同等の精度が得られることを確認している。

さらに、深層学習の特徴抽出の能力を活用した、ひび割れ以外の変状の検出を行う研究も見られる。Makantasisら<sup>56)</sup>は、CNNを用いて、トンネル表面の画像をピクセル単位で「変状」と「非変状」に分類し、ひび割れに限定しない変状全般の検出を行っている。Nasrollahiら<sup>57)</sup>は、点群データと深層ニューラルネットワークを用いて、コンクリート橋表面の画像をピクセル単位で「変状」と「非変状」に分類し、Makantasisらと同様にひび割れに限定しない変状全般の検出を行っている。Makantasisらの研究とNasrollahiらの研究は、ひび割れに限定しない変状全般を対象としているが、健全なコンクリート面とは異なる領域の検出に留まっており、変状の内容についての情報を得ることはできない。対して、Hoskereら<sup>58)</sup>は、コンクリートのひび割れ、はく離、鉄筋露出に加えて鋼部材の腐食および疲労亀裂、舗装のひび割れの6種類の変状に分類した上でピクセルレベルでの領域を検出するセグメンテーションのネットワークを提案している。Liら<sup>59)</sup>は、深層学習のセマンティックセグメンテーションにより、ひび割れ、はく離、エフロレッセンス、穴の4種類の変状に分類した上でピクセルレベルでの領域の検出を行っている。また、Zhangら<sup>60)</sup>は、深層学習の物体検出により、ひび割れ、ポップアウト、はく離、鉄筋露出の4種類の変状に分類した上で矩形領域での検出を行っている。

検出手法については、深層学習技術が普及した当初は分類のみを行うCNNを用いた研究が多く見られたが、技術の向上により物体検出およびセグメンテーションの精度が向上したことを受けて、画像内の位置の情報も捉える研究が増加している。検出手法の主流は、国内と同様にセグメンテーションに移行しつつあると考えられる。セグメンテーションを適用する研究の増加の背景には、国内外を問わず、任意の形状で領域の検出

が可能で変状の定量値の把握に活用しやすい点にあると考えられる。対象とする変状は、ひび割れが主体であるが、テーマは精度の向上の探求、エッジデバイデスでの適用、3次元情報との組み合わせ等、非常に多岐にわたる。また、ひび割れ以外の変状を対象とした研究も、国内と同様に増え始めている。

### 1-3-4 研究の着眼点

コンクリート表面の変状の検出は、これまでも多くの研究が行われてきたが、その大半がひび割れを対象としたものであり、コンクリートのはく離や鉄筋の露出、遊離石灰の析出等の他の変状までを対象とした研究は少ない。ひび割れは、コンクリートの変状の中でも記録に労力を要する性質の変状であるため、それらの研究により有益な知見が得られ、技術的に多大な貢献がなされてきた。しかし、橋梁の状態を満遍なく把握する橋梁定期点検という観点においては、ひび割れのみの記録では近接目視点検に対する省力化としては十分ではない。構造物の点検要領は、インフラの管理者や施設により対象とする変状およびその状況の評価内容は異なる。変状検出に関して多様な研究が行われているものの、特定の点検要領に準拠した変状の検出までには至っていないため、これらの研究が構造物の点検に広く用いられていないものと考えられる。そこで、本研究では、ひび割れ以外のコンクリートの主要な変状も含めた検出に取り組むこととした。さらに、メンテナンスサイクルにおいて「点検」の次のステップとなる「診断」に供する情報としては、変状の種類のみでは評価に寄与しないため、橋梁定期点検要領における変状の程度の区分も含めて識別させることとした。国内および国外の研究では複数の変状を同時に検出する研究も僅かに見られるが、鉄筋の露出および遊離石灰の析出等を、変状の進展具合を考慮した評価区分によってさらに細分化した条件で、セグメンテーションにより検出する研究は行われていない。本研究の意義は、5年に1回の頻度で行われる橋梁定期点検の省力化のために、コンクリートの変状を橋梁定期点検要領における変状の種類と程度の区分も含めて識別させた上で、画像からピクセル単位の領域で検出するという点にある。

変状の検出に際しては、深層学習による技術を中心に研究を進めた。画像からのひび割れの検出の黎明期に多く用いられた画像処理技術は、単独の技術では実例に基づく学習を行えないという点と、認識のための閾値等を人力により定義しなければならないという課題があった。そこからのパラダイムの変化により、実例に基づく学習を行う機械学習では、学習により特定のタスクを実行できる能力はあるが、抽出する特徴を人力により定義しなければならないという課題が残った。「特徴の設計」から「特徴の学習」へ

とさらなるパラダイムの変化を遂げた深層学習では、データからの学習により、抽出する特徴そのものをコンピューターが自動定義することで、機械学習の課題の克服を可能とした。しかし、その反面、深層学習では、機械学習で必要としたデータ量以上の教師データが必要となり、教師データの収集に多大な労力を要するコンクリートの変状の検出においては、この点が大きな課題となる。また、深層学習は、一定のコントロール化に置かれた条件での画像に対する適用においては、非常に高精度な結果が得られているが、あらゆる多様性を有する条件下の画像に対する適用においては、十分な精度が得られているとは言い難い。そのため、本研究では、限られたデータ量で精度を向上させる手法の検討を行うとともに、実際の運用を見据えて精度を補完する技術についての検討も行った。

1-4 本論文の構成と概要

1-4-1 本研究の対象

(1) 本研究で対象とする変状

一般道路の道路橋に対する定期点検では、国土交通省の橋梁定期点検要領で規定された26損傷（表1-3参照）が、多くの道路管理者により採用されている。地方公共団体の橋梁点検のコスト負担の軽減が本研究の主目的であるが、多くの地方公共団体では、コンクリート橋が管理橋梁に高い比率を締める傾向にある。そのため、本研究では、この26損傷のうち、コンクリートの変状に着目して、画像認識の検討を行うこととした。

なお、甚大な損傷であることから学習に用いる画像が少ない「抜け落ち」や、損傷の有無について画像からの判断が困難な「うき」、床版特有の条件を加味した撮影画像が少なく、「ひびわれ」と外観が類似する「床版ひびわれ」は対象から除外した。また、対象損傷のうち、「ひびわれ」については、ひび割れの幅や間隔の定量的情報がないと画像のみでは人間の目でも損傷程度の分類は困難なため、ひび割れの有無のみの識別を行うこととした。以上の損傷を対象として、本研究では、変状写真に対して、深層学習による画像認識を用いて、変状位置を特定するとともに、変状の種類とその程度の区分の識別を行う。具体的な識別内容は、表1-4に示す「ひびわれ」、「剥離・鉄筋露出」、「漏水・遊離石灰」の程度の区分（ランク）内容に応じた8つのクラスである。各クラスの変状写真の事例を図1-7に示す。

表 1-3 橋梁定期点検要領における着目損傷<sup>3)</sup>

材料	番号	損傷の種類	材料	番号	損傷の種類
鋼	[1]	腐食	その他	[13]	遊間の異常
	[2]	亀裂		[14]	路面の凹凸
	[3]	ゆるみ・脱落		[15]	舗装の異常
	[4]	破断		[16]	支承部の機能障害
	[5]	防食機能の劣化		[17]	その他
コンクリート	[6]	ひびわれ	共通	[10]	補修・補強材の損傷
	[7]	剥離・鉄筋露出		[18]	定着部の異常
	[8]	漏水・遊離石灰		[19]	変色・劣化
	[9]	抜け落ち		[20]	漏水・滞水
	[11]	床版ひびわれ		[21]	異常な音・振動
	[12]	うき		[22]	異常なたわみ
				[23]	変形・欠損
		[24]	土砂詰まり		
		[25]	沈下・移動・傾斜		
		[26]	洗掘		



表 1-4 対象損傷の損傷程度の区分と分類クラスの関係

損傷種類	区分	一般的状況 <sup>3)</sup>	クラス No.
ひびわれ	a	損傷なし	—
	b	最大ひびわれ幅： 小      最小ひびわれ間隔： 小	1
	c	最大ひびわれ幅： 小 最大ひびわれ幅： 中      最小ひびわれ間隔： 大 最小ひびわれ間隔： 小	
	d	最大ひびわれ幅： 中 最大ひびわれ幅： 大      最小ひびわれ間隔： 大 最小ひびわれ間隔： 小	
	e	最大ひびわれ幅： 大      最小ひびわれ間隔： 大	
剥離・鉄筋露出	a	損傷なし	—
	b	—	—
	c	剥離のみが生じている。	2
	d	鉄筋が露出しており、鉄筋の腐食は軽微である。	3
	e	鉄筋が露出しており、鉄筋が著しく腐食又は破断している。	4
※ 漏水・遊離石灰	a	損傷なし	—
	b	—	—
	c	ひびわれから漏水が生じている。 錆汁や遊離石灰はほとんど見られない。	5
	d	ひびわれから遊離石灰が生じている。 錆汁はほとんど見られない。	6
	e	ひびわれから著しい漏水や遊離石灰（例えば、つらら状）が生じている。 漏水に著しい泥や錆汁の混入が認められる。	7 8

※) 打継目や目地部から生じる漏水・遊離石灰についても、ひびわれと同様の扱いとする。

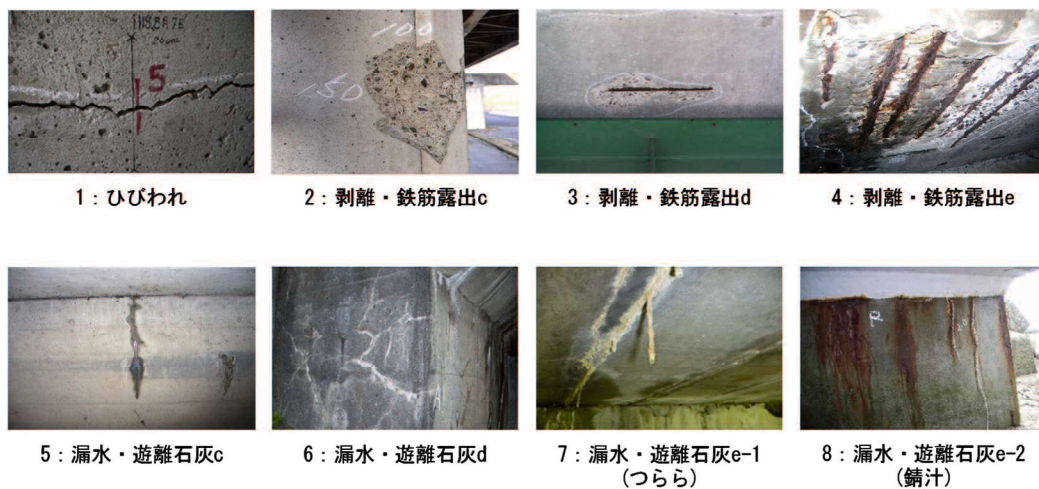


図 1-7 各クラスの変状写真の事例

「橋梁定期点検要領に準拠した変状検出」の既往研究との関係を図 1-8 に示す。

本研究では、国土交通省の橋梁点検要領の変状の分類に沿って教師データの構築を行うが、他の点検要領に沿った橋梁であっても、変状内容の定義が同じ部分については、適用可能である。なお、橋梁以外のコンクリート構造物であっても、変状内容の定義が同様の条件の画像であれば、基本的に適用可能と考えられるが、たとえば護岸等の無筋コンクリートの構造物のひび割れについては、幅が大きく外観的な特徴の相違が生じている可能性もある。そのため、橋梁への適用を基本とする。橋梁の部材としては、RC 橋および PC 橋の上部工と下部工を教師データに含めており、いずれの部材へも適用が可能である。ただし、「剥離・鉄筋露出」の扱いとなる PC 鋼材の露出については、対象としていない。

また、変状の用語は、一般的な学術用語では「ひび割れ」、「はく離」、「浮き」が用いられるが、橋梁定期点検要領を用いたコンクリート橋の点検への適用を基本としているため、混乱を避けて表 1-3 に示す名称に従って、以後「ひびわれ」、「剥離」、「うき」に本論文中の表記を統一する。また、橋梁定期点検要領では着目対象となっている 26 種類の変状を「損傷」と称しているが、初期欠陥・損傷・劣化等の総称を「変状」と称する一般的な定義に従って、「変状」という用語を用いることとする。

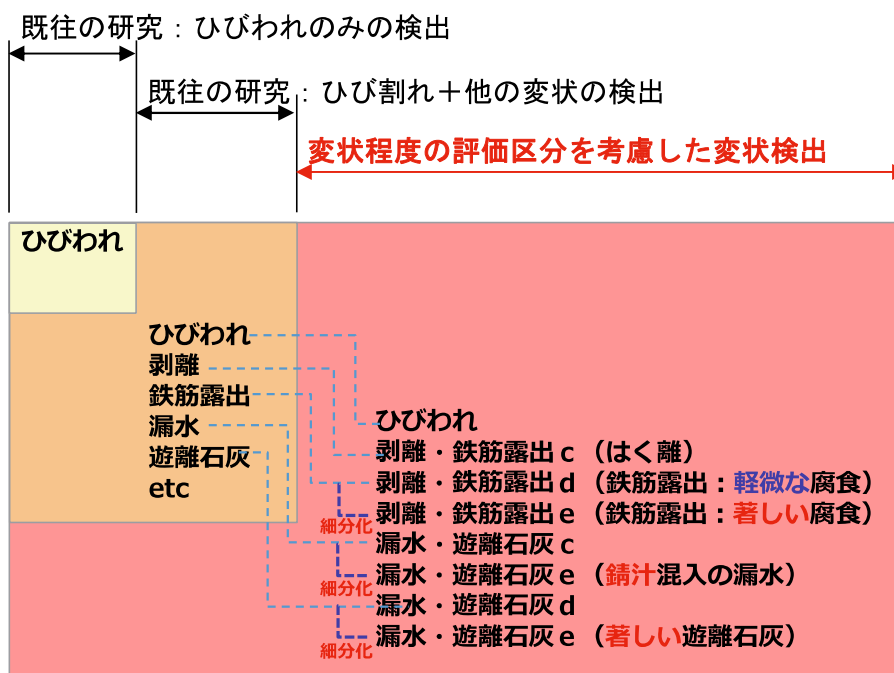


図 1-8 「橋梁定期点検要領に準拠した変状検出」の既往研究との関係

## (2) 教師データに用いる画像

一般的に、深層学習の学習においては、ネットワークの種類や構成等以上に、学習サンプルが重要だと言われている<sup>6)</sup>。本研究では、橋梁の点検や踏査等の際に撮影した画像を用いて検討を行う。これらの画像は、最小限の量は確保できたものの、深層学習における教師データとして使用する目的で撮影した画像ではない。また、日射や影の影響を受けにくく撮影距離が概ね一定であるトンネルのような撮影条件に対して、橋梁は多様な撮影条件が前提となる。これらのことを鑑みると、本研究で用いる撮影画像は、深層学習に用いる教師データとして、以下の点から統一性が欠如しており、質や特徴量の面で適しているとは言い難い。

- ・データ容量を縮減するために、画像については解像度を低くしている
- ・撮影距離が多様（近接して撮影した画像，変状箇所周辺まで含めて撮影した画像等）
- ・撮影時の角度が多様（変状に正対して撮影した画像，点検条件より斜方向から撮影した画像等）
- ・撮影時の照度が多様（日向，日陰，夜間等）

しかし、橋梁の点検においては、画一的なコントロールされた条件で撮影を行うことは困難で、多様な撮影条件とならざるを得ないため、最適な教師データとなり得る撮影条件を規定することは非常に難しい。また、橋梁の点検では多様な撮影条件となることを踏まえると、学習に用いる教師画像は、統一的な撮影条件での画像のみでなく、多様な画像を含んでいる方が、多様なテスト画像に対する汎化性能を発揮できる可能性がある。さらに、学習のための大量のデータが必要とされる深層学習では、データの収集に非常に労力を要することから、何らかの既往のデータを使用することは珍しいことではない。以上のことより、橋梁の点検や踏査等の際に撮影した画像データを用いて、深層学習への適性に関する検証を行うこととした。一般的な点検や踏査等の際の撮影画像を用いて深層学習の検討を行うことは、このような画像の教師データとしての適用性の検証にもつながり、今後のAIの利活用促進を見据えると、その観点からも意義があるものと考えられる。

教師データとした画像は、上下部工関係なく「ひびわれ」、「剥離・鉄筋露出」、「漏水・遊離石灰」に該当する画像を抽出し、「剥離・鉄筋露出」および「漏水・遊離石灰」については、道路橋点検士有資格者により損傷程度の区分の評価を行った。上部工は、RC桁とPC桁の桁種に関係なく教師データに含めたが、PC桁の鉄筋露出の事例が

少なかったため、剥離・鉄筋露出 d および剥離・鉄筋露出 e の画像は、RC 桁と下部工の画像が主体となっている。

### 1-4-2 本論文の構成と概要

#### (1) 論文の構成

本研究は、深層学習による画像認識技術を用いて、構造物の点検の省力化に資する処理方法の構築に向けた検討を行ったものである。画像認識 (Image Recognition) とは、画像に写っている内容を理解することで、画像中の物体を理解して適切なラベルを付与する物体認識と、画像が表現する包括的な状態を理解するシーン認識とに分類される<sup>62)</sup>。物体認識は、さらに固有名詞等で表現されるような特定の物体を認識する特定物体認識と、対象とする物体を一般的な名称で認識する一般物体認識に分類される。本研究では、画像認識技術の中でも、一般物体認識の技術を中心に検討を行う。一般物体認識の主要なタスクを図 1-9 に示す。まず、一般物体認識の中では難易度の低いタスクである「クラス分類」について検証を行い、コンクリートの変状に関する画像のクラスを適切に分類できるか確認した上で、難易度の高いタスクについての検証へ移行する。第 2 段階として外接矩形で画像内の物体位置を特定した上でその領域のクラス分類を行う「物体検出」、第 3 段階としてピクセル単位で画像内の物体の領域を特定した上でその領域のクラス分類を行う「物体領域抽出 (セグメンテーション)」についての検討を行う。なお、「物体検出」および「物体領域抽出」は、画像内の対象とする物体を検出することから、そのように呼称されるが、本研究では画像内のコンクリートの物体における変状部分の領域の検出を行う。本論文の構成を図 1-10 に示す。



図 1-9 一般物体認識の主要なタスク

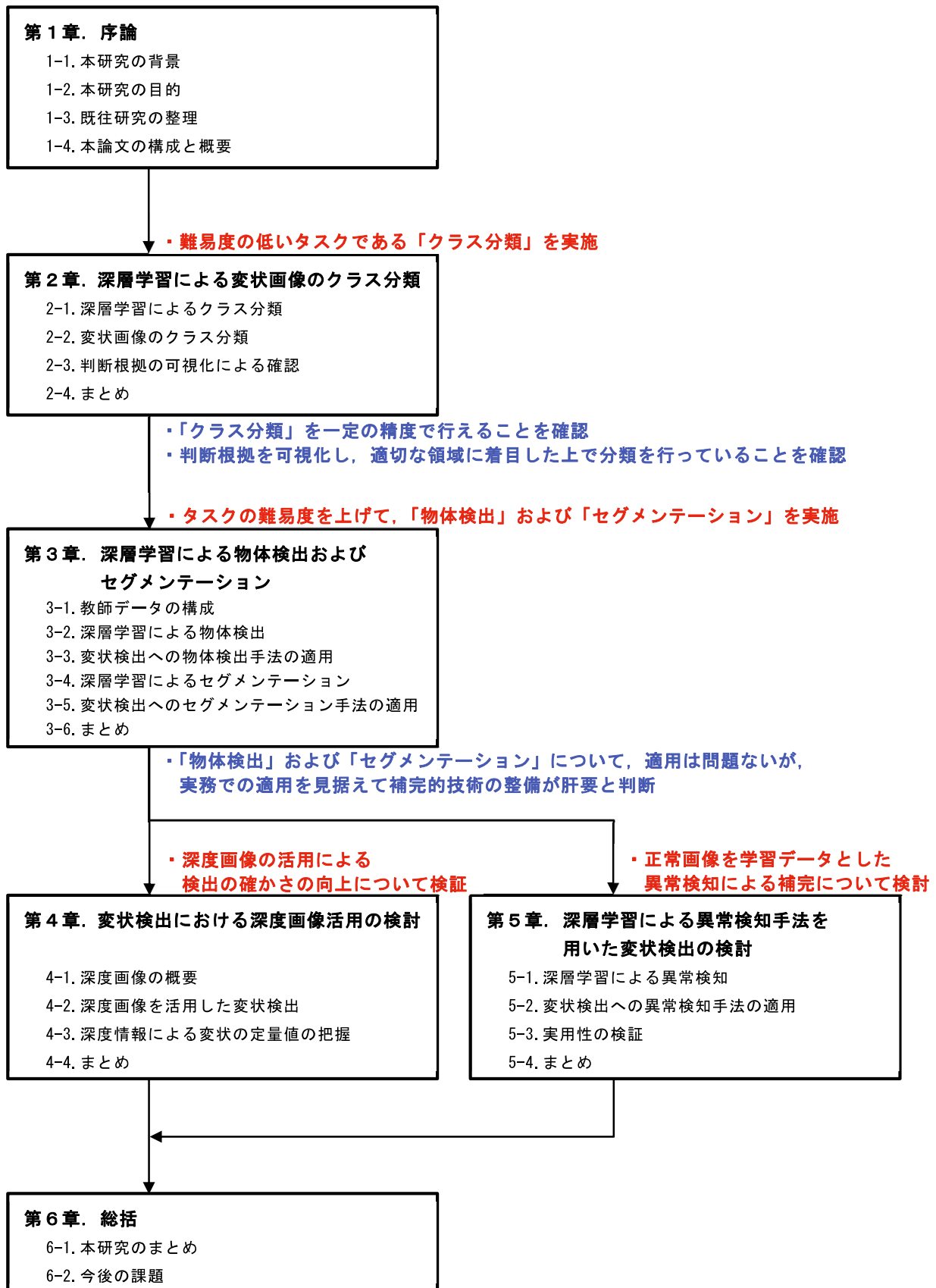


図 1-10 本論文の構成

本研究では、橋梁定期点検要領における変状の種類と程度の区分も含めて識別させた上で、セグメンテーション手法によりピクセル単位の任意の領域形状で、画像からの変状検出を行う。近年では、点検の効率化に向けて自動化技術導入の要請が本格的に高まってきているが、社会インフラの点検という高い精度が要求される性質の分野であるため、実際の導入に向けてのハードルは非常に高い。また、橋梁の変状の検出は、深層学習による画像認識技術の導入が進んでいる製造業や医用画像の分野に比べると、画像の撮影条件や検出対象の変状の形状に統一性が乏しく、教師データの収集の労力も大きいため、難易度が非常に高い。そのため、本論文は、検出のエラーの最小化を目指して、限られたデータ量での精度の向上と、実際の運用を見据えて併用する補完的な技術について検討を行う章構成となっている。

各章の研究内容の関係性を示した、本研究における変状検出の運用方針を図 1-11 に示す。

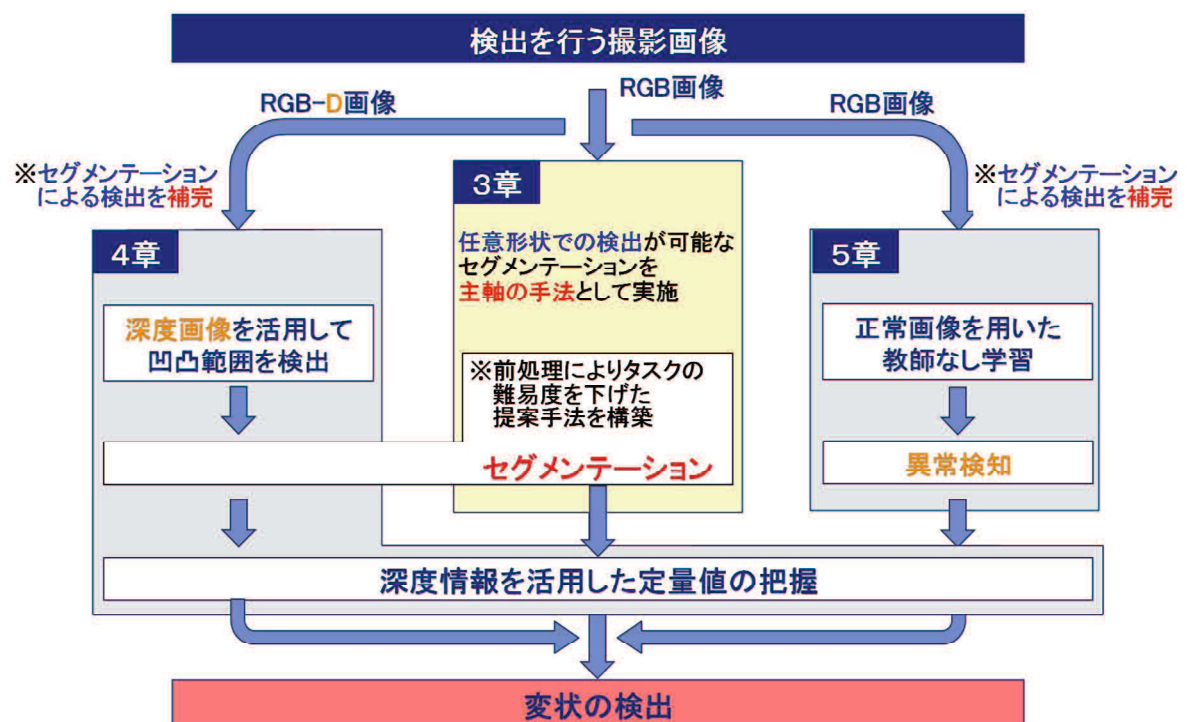


図 1-11 本研究における変状検出の運用方針

### (2) 論文の概要

本論文は全6章で構成しており、各章の概要は以下のとおりである。

#### [第1章 序論]

本章では、本研究を行う背景となった我が国における社会インフラの維持管理の現状と課題を把握し、研究の目的を明確にするとともに、論文の構成と概要を説明する。

#### [第2章 深層学習による変状画像のクラス分類]

本章では、コンクリート面の画像から変状部分を検出するという最終目的に向けて、その実現性の確認のために、一般物体認識の中でも難易度の低いタスクである「クラス分類」について検証を行う。具体的には、コンクリートの変状に関する画像について、変状の種類に加えて変状の状態の区分までを考慮したクラス分類を適切に実行できるか確認を行う。また、クラスの分類に際して画像内で着目した領域を可視化し、判断根拠の確認も併せて行う。

#### [第3章 深層学習による物体検出およびセグメンテーション]

本章では、第2章における検討結果を踏まえ、変状について画像内の位置的な情報も含めた認識を行う、という難易度の高いタスクについての検証へ移行する。前半部分では外接矩形で画像内の物体位置を検出した上でその領域のクラス分類を行う「物体検出」、後半部分ではピクセル単位で画像内の物体の領域を検出した上でその領域のクラス分類を行う「セグメンテーション」についての検討を行う。

橋梁の変状の検出は、画像の撮影条件や検出対象の変状の形状に統一性が乏しく、教師データの収集の労力も大きいため、難易度が非常に高いという観点に立って、続く第4章、第5章では、本研究の変状検出の主軸であるセグメンテーションを補完する技術について検討を行う。

#### [第4章 変状検出における深度画像活用の検討]

本章では、第3章で検討を行った独自のセグメンテーション手法について、前段部分の物体検出の処理における統計的な不確かさを排除する目的で、ピクセル単位で深度(撮影対象までの距離情報)の情報を有している深度画像を活用した検出手法について検討



を行う。また、深度画像から得られる深度の情報を基に、変状の定量値を把握する一連の処理の構築を行う。

**[第5章 深層学習による異常検知手法を用いた変状検出の検討]**

本章では、セグメンテーション手法による変状検出を補完する技術として、健全なコンクリート面である正常画像を用いた、教師なし学習による異常検知手法について検討を行う。

**[第6章 序論]**

本章では、本研究で得られた成果を総括する。

### 第1章 参考文献

- 1) 国土交通省 総合政策局 社会資本整備政策課, 国土交通省 総合政策局 公共事業企画調整課: 社会資本の維持管理に関する取組, 2018.
- 2) 道路法施行規則 省令 第四条の五の五(道路の維持又は修繕に関する技術的基準等), 平成 26 年 3 月 31 日公布
- 3) 国土交通省 道路局 国道・技術課: 橋梁定期点検要領, 2019.
- 4) 社会資本整備審議会 道路分科会 道路メンテナンス技術小委員会: 道路のメンテナンスサイクルの構築に向けて, 2013.
- 5) 国土交通省 道路局 国道・技術課: 橋梁等の平成 30 年度点検結果をとりまとめ～道路メンテナンス年報(一巡目)の公表～, 2019.
- 6) 国土交通省: 社会資本ストックの戦略的維持管理, 平成 28 年度 政策レビュー結果(評価書), 2017.
- 7) 小出博: 特殊用途カメラの研究「画像診断支援システム」開発とその活用, 映像情報, pp. 37-40, 1990.
- 8) 山口友之, 橋本周司: Gray-scale Hough 変換を用いたコンクリート表面のひび割れ検出, 情報処理学会研究報告, pp. 57-64, 2004.
- 9) 山口友之, 橋本周司: 既設コンクリート構造物におけるひび割れ幅の実画像計測, 電気学会論文誌 C, Vol. 127, No. 4, pp. 605-614, 2007.
- 10) 藤田悠介, 中村秀明, 浜本義彦: 画像処理によるコンクリート構造物の高精度なひび割れ自動抽出, 土木学会論文集 F, Vol. 66, No. 3, pp. 459-470, 2010.
- 11) 木元亮輔, 藤田悠介, 河村圭, 浜本義彦: 画像処理によるコンクリート構造物のひび割れ計測の実用化に関する基礎研究, ファジィシステムシンポジウム講演論文集, Vol. 26, pp. 874-879, 2010.
- 12) 藤田悠介, 中村秀明, 浜本義彦: 画像処理によるコンクリート構造物のひび割れ幅の分類, コンクリート工学年次論文集, Vol. 34, No. 1, pp. 1792-1797, 2012.
- 13) 全邦釘, 片岡望, 三輪知寛: 統計的特徴および幾何学的特徴に着目したコンクリート表面ひび割れの画像解析による検出, 土木学会論文集 F3, Vol. 70, No. 2, pp. I\_1-I\_8, 2014.
- 14) 全邦釘, 井後敦史, 橋本和明: 画像解析によるひび割れ検出とハフ変換による型枠跡および P コン跡の同定, コンクリート工学年次論文集, Vol. 37, No. 1, pp. 1777-1782, 2015.

- 
- 15) 武田均, 堀口賢一, 小山哲, 丸屋剛: ウェーブレット変換を用いたコンクリートのひび割れ画像解析手法の開発, コンクリート工学年次論文集, Vol. 28, No. 1, pp. 1895-1900, 2006.
  - 16) 丸屋剛, 堀口賢一, 小山哲, 澤健男: ウェーブレット変換を用いた床版コンクリートのひび割れ調査の実用化, コンクリート工学年次論文集, Vol. 29, No. 2, pp. 691-696, 2007.
  - 17) 小山哲, 丸屋剛, 堀口賢一, 澤健男: ガボールウェーブレット変換を用いたコンクリートのひび割れ画像解析技術の開発, 土木学会論文集 E2, Vol. 68, No. 3, pp. 178-194, 2012.
  - 18) 西川貴文, 吉田純司, 杉山俊幸, 齊藤成彦, 藤野陽三: 木構造状フィルタを用いたコンクリートのクラック抽出のためのロバストな画像処理システム, 土木学会論文集 A, Vol. 63, No. 4, pp. 599-616, 2007.
  - 19) 河村圭, 宮本文穂, 中村秀明, 佐藤亮: 対話型遺伝的アルゴリズムによるデジタル画像からのひび割れ抽出, 土木学会論文集, No. 742/VI-60, pp. 115-131, 2003.
  - 20) 河村圭, 吉野孝亮, Amir Tarighat, 中村秀明: 遺伝的アルゴリズムおよび決定木を用いたひび割れ抽出のための画像処理パラメータ有効範囲の特定, 土木学会論文集 F3, Vol. 69, No. 2, pp. I\_13-I\_23, 2013.
  - 21) 河村圭, 村上慧季, 塩崎正人: 対話型遺伝的アルゴリズムを用いたひび割れ半自動抽出手法の研究, 土木学会論文集 F3, Vol. 71, No. 2, pp. I\_114-I\_122, 2015.
  - 22) 塩崎正人, 菊地典明, 前川了重, 益山茂樹, 石坂哲宏, 佐田達典: 画像計測による照明の陰影を利用した変状抽出手法の研究, 土木学会論文集 F3, Vol. 67, No. 2, pp. I\_119-I\_124, 2011.
  - 23) 石森章之, 佐田達典, 石坂哲宏, 塩崎正人: RGB 単色光と陰影を利用したトンネルコンクリート壁面のひび割れ検知に関する研究, 土木学会論文集 F3, Vol. 69, No. 2, pp. I\_82-I\_89, 2013.
  - 24) 塩崎正人, 石森章之, 菊地典明, 佐田達典: 画像計測による照明の陰影を利用した変状抽出手法の研究, コンクリート工学年次論文集, Vol. 36, No. 1, pp. 2152-2157, 2014.
  - 25) 石森章之, 佐田達典, 塩崎正人: 投光器の照射角度と陰影の幅を利用したコンクリート壁面の段差高計測に関する研究, 土木学会論文集 F3, Vol. 70, No. 2, pp. I\_249-I\_256, 2014.
  - 26) 鶴飼正人: 画像処理によるトンネル構造物ひび割れ検査精度の向上, 鉄道総研報告,
-

- Vol. 24, No. 3, pp. 53-56, 2010.
- 27) 鶴飼正人, 長峯望, 御崎哲一, 坂本保彦: トンネル変状抽出のための高精度画像処理手法の開発, サイバネティクス, Vol. 17, No. 2, pp. 58-62, 2012.
  - 28) 鶴飼正人: 画像処理技術を用いてトンネルを検査する, RRR, Vol. 72, No. 9, pp. 20-23, 2015.
  - 29) 鶴飼正人, 長峯望: 画像処理技術による構造物検出手法の高精度化, 鉄道総研報告, Vol. 28, No. 11, pp. 11-16, 2014.
  - 30) 野村泰稔, 村尾彩希, 阪口幸広, 古田均: 深層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, 土木学会論文集 F6, Vol. 73, No. 2, pp. I\_189-I\_198, 2017.
  - 31) 野村泰稔, 重村知輝: 深層学習に基づく物体検出と生成モデルを用いた構造表面損傷の実時間検出技術の開発, 材料, Vol. 68, No. 3, pp. 250-257, 2019.
  - 32) 山根達郎, 全邦釘: Deep learning による Semantic Segmentation を用いたコンクリート表面ひび割れの検出, 構造工学論文集, Vol. 65A, pp. 130-138, 2019.
  - 33) 榎本真美, 藤原広志, 松原健二, 重田佳幸, 安田亨: AI を用いた道路施設の変状箇所抽出モデルの開発, 建設コンサルタント協会業務研究発表会論文集, Vol. 18, pp. 107-110, 2018.
  - 34) 全邦釘, 井後敦史: Random Forest によるコンクリート表面ひび割れの検出, 土木学会論文集 F3, Vol. 71, No. 2, pp. I\_1-I\_8, 2015.
  - 35) 全邦釘, 嶋本ゆり, 大窪和明, 三輪知寛, 大賀水田生: ディープラーニングおよび Random Forest によるコンクリートのひび割れ自動検出手法, 土木学会論文集 F3, Vol. 73, No. 2, pp. I\_297-I\_307, 2017.
  - 36) 中村優志, 児玉聖治, 河村圭, 村上慧季, 塩崎正人: 深層学習を用いたひび割れ領域の自動識別法に関する研究, 土木学会第 72 回年次学術講演会, pp. 45-46, 2017.
  - 37) 鶴飼正人: 深層学習を用いたトンネル覆工面のひび割れ検出手法の開発, 鉄道総研報告, Vol. 32, No. 5, pp. 5-10, 2018.
  - 38) 横山傑, 松本高志: ディープラーニングを用いたコンクリートのひび割れ自動検出器の開発, 応用力学シンポジウム講演論文集, Vol. 19, pp. 125-126, 2016.
  - 39) 横山傑, 松本高志: Deep Learning によるコンクリートの変状自動検出器の開発と Web システムの実装, 応用力学シンポジウム講演論文集, Vol. 20, pp. 31-32, 2017.
  - 40) 横山傑, 松本高志: SNS を用いた Deep Learning によるコンクリートの変状自動検出器の開発, 土木情報学シンポジウム講演集, Vol. 42, pp. 235-236, 2017.
-

- 
- 41) 横山傑, 松本高志: Deep Learning によるコンクリートの変状自動検出器の開発と Web システムの実装, 土木学会論文集 A2, Vol. 73, No. 2, pp. I\_781-I\_789, 2018.
  - 42) 矢吹信喜: 深層学習を用いたデジタル画像からのコンクリート表面変状の検出, SIP インフラ連携委員会報告 インフラ維持管理への AI 技術適用のための調査研究報告書, pp. 79-85, 2019.
  - 43) 柏貴裕, 龍田斉, 井林康, 長井宏平, PRENDINGER Helmut, RUBIO Juanjo: 畳み込みニューラルネットワークを用いたコンクリート床版の損傷検出, 土木学会第 73 回 年次学術講演会, pp. 31-32, 2018.
  - 44) 中島道浩, 安野貴人, 永富大亮, 野田一弘, 青柳聖, 関口斉治: 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, 人工知能学会全国大会論文集, Vol. 33, 1H3-J-13-03, 2019.
  - 45) Adbel-Qader, I., Abudayyeh, O. and Kelly, M.: Analysis of edge detection techniques for crack identification in bridges, *Journal of Computing in Civil Engineering*, Vol. 17, No. 4, pp. 255-263, 2003.
  - 46) Hutchinson, T. C., Chen, Z.: Improved image analysis for evaluating concrete damage, *Journal of Computing in Civil Engineering*, Vol. 20, No. 3, pp. 210-216, 2006.
  - 47) Chen, Z., Hutchinson, T. C.: Image-based framework for concrete surface crack monitoring and quantification, *Advances in Civil Engineering*, 2010.
  - 48) Pereira, F. C., Pereira, C. E.: Embedded Image Processing Systems for Automatic Recognition of Cracks using UAVs, *IFAC-PapersOnLine*, Vol. 48, No. 10, pp. 16-21, 2015.
  - 49) Adbel-Qader, I., Pashaie-Rad, S., Abudayyeh, O. and Yehia, S.: PCA-based algorithm for unsupervised bridge crack detection, *Advances in Engineering Software*, Vol. 37, No. 12, pp. 771-778, 2006.
  - 50) Prasanna, P., Dana, K. j., Gucunski, N., Basily, B. B., La, H. M., Lim, R. S. and Parvardeh, H.: Automated crack detection on concrete bridges, *IEEE Transactions on Automation Science and Engineering*, Vol. 13, No. 2, pp. 591-599, 2016.
  - 51) Li, G., Zhao, X., Du, K., Ru, F. and Zhang, Y.: Recognition and evaluation of bridge cracks with modified active contour model and greedy search-based support vector machine, *Automation in Construction*, Vol. 78, pp. 51-61, 2017.
  - 52) Cha, Y.-J., Choi, W. and Büyüköztürk, O.: Deep Learning-based Crack Damage Detection Using Convolutional Neural Networks, *Computer-Aided Civil and Infrastructure Engineering*, Vol. 32, No. 5, pp. 361-378, 2017.
-

- 53) Kim, B., Cho, S. : Automated Vision-Based Detection of Cracks on Concrete Surfaces Using a Deep Learning Technique, *Sensors*, Vol. 18, No. 10, 2018.
- 54) Alipour, M., Harris, D. K. and Miller, G. R. : Robust Pixel-Level Crack Detection Using Deep Fully Convolutional Neural Networks, *Journal of Computing in Civil Engineering*, Vol. 33, No. 6, 2019.
- 55) Zhang, K., Zhang, Y. and Cheng, H. D. : Self-Supervised Structure Learning for Crack Detection Based on Cycle-Consistent Generative Adversarial Networks, *Journal of Computing in Civil Engineering*, Vol. 34, No. 3, 2020.
- 56) Makantasis, K., Protopapadakis, E., Doulamis, A., Doulamis, N. and Loupos, C. : Deep convolutional neural networks for efficient vision based tunnel inspection, *IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, pp. 335-342, 2015.
- 57) Nasrollahi, M., Bolourian, N. and Hammad, A. : Concrete bridge surface damage detection using a single - stage detector, *In Proceedings of the CSCE Annual Conference*, pp. CON032-1- CON032-10, 2019.
- 58) Hoskere, V., Narazaki, Y., Hoang, T. and Spencer Jr, B. F. : Vision-based Structural Inspection using Multiscale Deep Convolutional Neural Networks, *arXiv: 1805.01055v1*, 2018.
- 59) Li, S., Zhao, X. and Zhou, G. : Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network, *Computer-Aided Civil and Infrastructure Engineering*, Vol. 34, No.7, pp. 616-634, 2019.
- 60) Zhang, C., Chang, C.-c. and Jamshidi, M. : Concrete bridge surface damage detection using a single - stage detector, *Computer-Aided Civil and Infrastructure Engineering*, Vol. 35, No.4, pp. 389-409, 2020.
- 61) 山下隆義 : イラストで学ぶディープラーニング, 講談社, 2016.
- 62) 原田達也 : 画像認識, 講談社, 2017.

## 第 2 章 深層学習による変状画像のクラス分類

---

### 第 2 章

2-1 深層学習によるクラス分類

2-2 変状画像のクラス分類

2-3 判断根拠の可視化による確認

2-4 まとめ





## 第2章 深層学習による変状画像のクラス分類

---

### 2-1 深層学習によるクラス分類

#### 2-1-1 クラス分類の概要

1-4 節で前述したように、画像認識は物体認識とシーン認識とに分類され、さらに物体認識は特定物体認識と一般物体認識に分類される。クラス分類は、この分類中の一般物体認識のタスクの1種である。対象とする物体に対して、特定物体認識では固有名詞等で表現されるような特定の物体として認識し、一般物体認識では一般的な名称のカテゴリレベルでの認識が行われる。一般物体認識では、入力画像のカテゴリを判定するが、一般的な名称で表されるカテゴリの範囲は極めて広いため、入力画像と全く同じ物体の写真が教師画像になくても、学習による汎化によって、そのカテゴリを認識する必要がある。そのため、特定物体認識よりも一般物体認識の方が、難易度の高いタスクである。一般物体認識の研究は、古くは1960年代より行われており、1990年代後半以降は画像から局所特徴量と呼ばれる特徴量ベクトルを抽出して機械学習を用いる手法が主流となり、2010年代以降は大量のデータからの学習により特徴量を自動獲得する深層学習が主流となっている。

画像のクラス分類は、画像中の物体があらかじめ定義されたカテゴリの中で、どのカテゴリに属するかを求める問題である。画像のクラス分類を含む近年の一般物体認識の研究では、大量のラベル付きのデータ（入力データ  $x$  と正解となる出力データ  $y$  のペア）で構成されたデータセットを用いて、入力  $x$  を出力  $y$  に変換する関数  $f : x \rightarrow y$  を統計的学習により帰納的に推定することが不可欠となっている。画像のクラス分類の場合には、このラベル付きデータは、入力画像  $x$  と正解となるカテゴリのクラス情報  $y$  とのペアで構成されるデータセットとなる。クラス分類では、「学習」の段階で、ラベル付きデータを教師データとして、入力と出力の関係を学習する教師あり学習を行い、入力画像  $x$  から正解クラス  $y$  を出力する関数を帰納的に推定する。そして、「推論」の段階で、「学習」により得られた関数に、学習に用いていない未知の画像をテストデータとして入力し、推論結果として分類したクラスの出力を得る。教師あり学習の概念を図 2-1 に示す。

---

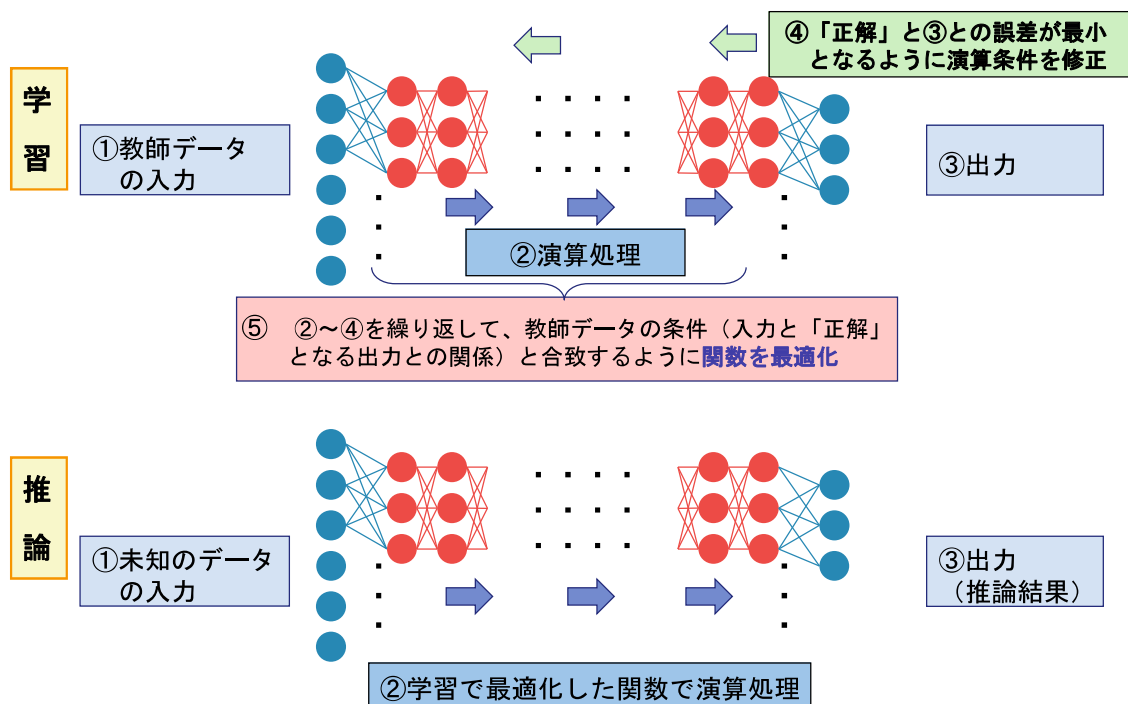


図 2-1 教師あり学習の概念

近年の画像認識においてデータセットが果たす役割は大きく、画像認識のモデルは、データセットと一体となって発展しているともいわれている。また、データセットは、統計的学習を支えるとともに、モデルを評価するベンチマークとしての一面も有しており、ILSVRC<sup>1)</sup>を代表とする多くの画像認識のコンペティションが開催されてきた。ILSVRCは、2010年から2017年の期間に行われた画像認識のコンペティションで、120万枚・1,000クラスから構成される大規模画像データセット ImageNet<sup>2)</sup>のデータを用いて、画像のクラス分類や物体検出等についてスコアを競い合うものである。画像のクラス分類では、与えられた画像に対して1,000のカテゴリから識別を行い（識別の候補上位5クラスに含まれれば正答とみなされる）、その識別エラー率により評価がなされる。ILSVRCにおける画像クラス分類の歴代優勝者の成績を図 2-2 に示す。このコンペティションで2012年に深層学習を用いた手法が圧勝したことを契機として、画像のクラス分類の研究は、深層学習をベースとした手法が主流となっていった。また、2015年には、同様のタスクを人間が行った場合の識別エラー率5.1%を下回るスコアが出たことから、AIが人間を超えたという報道もなされた。

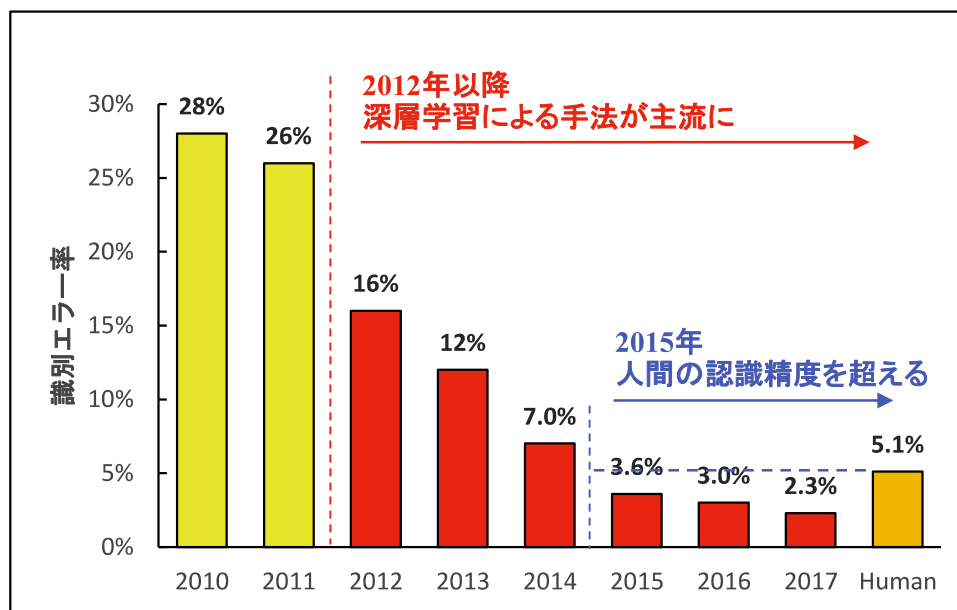


図 2-2 ILSVRC における画像クラス分類の歴代優勝者の成績

### 2-1-2 深層学習によるクラス分類

深層学習とは、コンピュータに人間のような学習能力を獲得させる技術である機械学習の1種で、脳の神経回路を数理的にモデル化したニューラルネットワーク（図 2-3 参照）の中間層を多層化したアルゴリズムである。画像のクラス分類では、2012 年の ILSVRC で、トロント大学の Hinton 教授らのグループが、深層学習の手法の1つである CNN (convolutional neural network)<sup>3)</sup>を用いて、2 位以下のチームに大差を付けて圧勝したことで注目を集めた<sup>4)</sup>。2010 年前後は、前述した ImageNet 等の大規模画像データセットの普及に加えて、CPU (Central Processing Unit) の高速化と並列化および GPU (Graphics Processing Unit) による計算機能力の著しい向上等により、深層学習の技術が適用可能となる環境が整った時期である。その後も、CNN は、画像認識の分野を中心に高い性能を発揮している。CNN の一般的なネットワーク構成を図 2-4 に示す。CNN は、順伝播型ニューラルネットワークの一種で、中間層の畳み込み層 (convolutional layer) とプーリング層 (pooling layer) を複数つなぎ合わせた深い階層を持ち、表現学習を行って最適となる特徴を抽出するネットワークである。畳み込み層では画像の局所的な特徴抽出、プーリング層では画像の局所ごとの特徴をまとめた上でのダウンサンプリングが行われる。従来の機械学習では人間の手により抽出する特徴を定義する必要があったの

## 第2章 深層学習による変状画像のクラス分類

に対して、深層学習では人間の手を介さずに、ネットワークの学習を通して、特定概念を特徴づける特徴量を自動抽出できる。この点が、深層学習が画像認識の分野で既存手法を著しく上回る精度を実現した要因である。

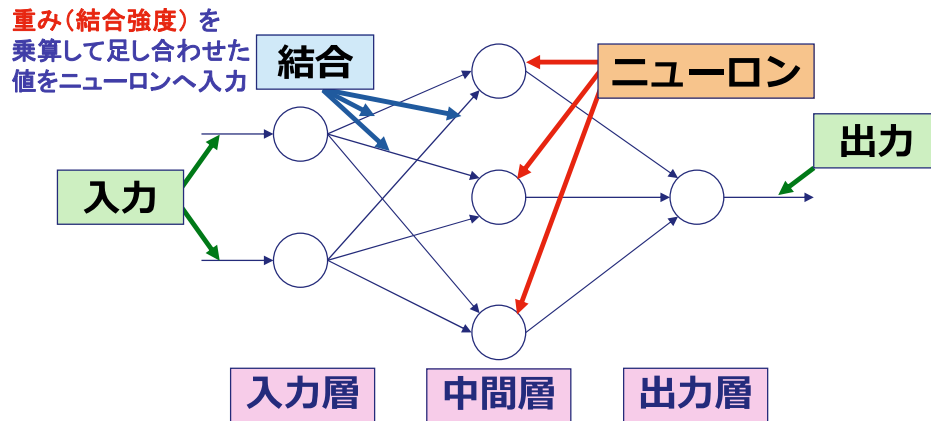


図 2-3 ニューラルネットワーク

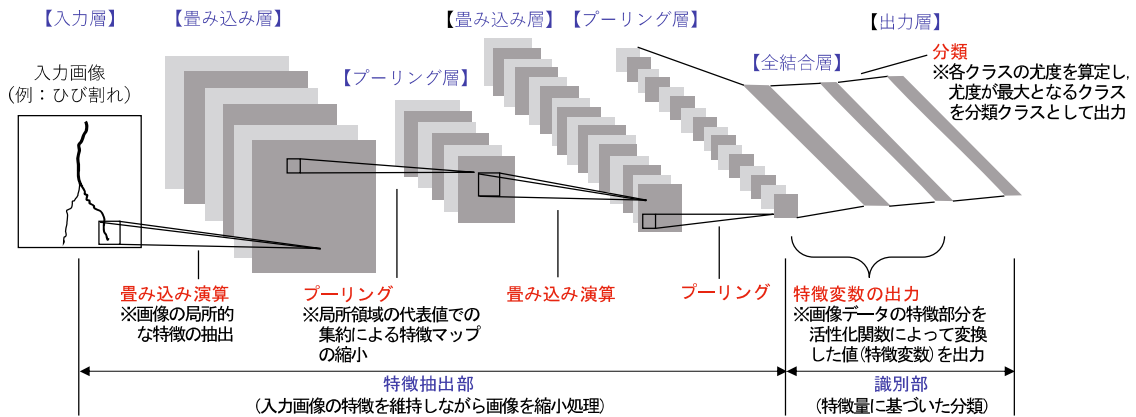


図 2-4 CNN の構成イメージ例 (文献<sup>5)</sup>図 2 を改変)

## 2-2 変状画像のクラス分類

### 2-2-1 検討条件

本研究における画像のクラス分類は、CNN のモデルの中から、VGGNet<sup>6)</sup>を用いた。VGGNet は、2014 年の ILSVRC において、クラス分類の部門の識別エラー率 2 位となったモデルである。畳み込み層とプーリング層から成るシンプルな構成の CNN で、本研究では畳み込み層 16 層+全結合層 3 層=19 層の構成となる VGG19 を使用した。

クラス分類で用いるデータセットは、コンクリート部材における変状を撮影した画像に対して、道路橋点検士有資格者により変状の正解クラスのラベル付を行なって作成した、画像とラベルが対になったデータである。なお、クラス分類では、画像がどのクラスに該当するかの one hot で表現される分類を行うため、学習に用いた画像は、単一のクラスの変状のみが存在する画像を選定している。

検証に際しては、限られたデータの中で検証の信頼性を向上させるため、図 2-5 に示す K-分割交差検証により評価を行った。K-分割交差検証では、データセットを K 個に分割し、そのうちの  $(K - 1)/K$  をモデル自体の調整（学習）に用いる教師（training）データとし、残りの  $1/K$  を学習の結果の評価に用いる検証（validation）データとする。本研究では、 $K=5$  としてデータセットを 5 分割し、全データの 80% を教師データ、20% を検証データとし、5 回の検証結果を平均して分類結果とする。また、テストに用いるデータとして、データセットとは別に、学習にも検証にも使用していないテスト（test）データも用意した。

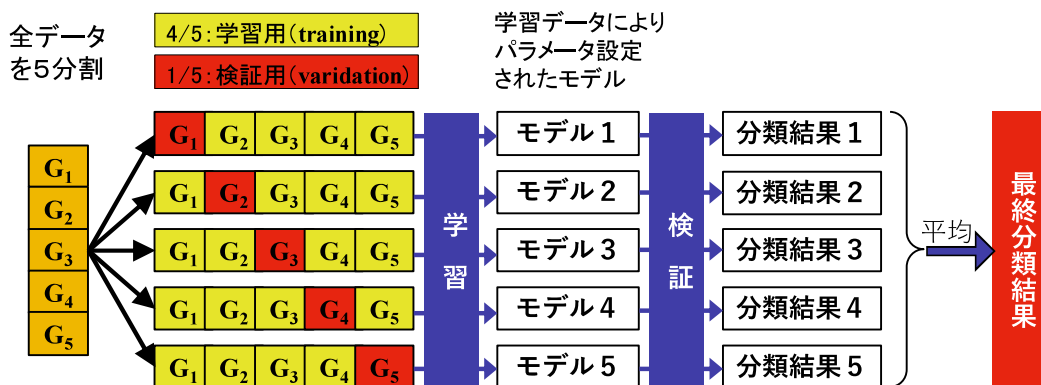


図 2-5 クラス分類における交差検証の概要

### 2-2-2 汎化性能向上の方策

#### (1) 過学習

CNN に代表される深いニューラルネットワークを用いたモデルは、一般に多数のパラメータを有する複雑なモデルほど表現能力は高いが、より多数の教師データがなければ過学習に陥りやすいとされている。過学習とは、教師データに過剰に適合したパラメータが出来上がってしまい、教師データに対しては非常に高い分類精度となるが、教師データと異なる未知のデータに対しては精度が低く適合できていない(汎化できていない)状態に陥る現象である。本タスクのデータセットは、各クラスにつき約 150 画像でデータ量としては十分ではないため、データの偏り等による過学習が懸念される。変状画像のクラス分類に際して、このような過学習の抑制および分類精度の向上を図る必要がある。

本研究では、データ量の不足を補い、過学習の抑制および分類精度向上を図る方策として、擬似的なデータを作成してデータ量を増加させる「データ拡張(Data Augmentation)」, 他の関連した問題のデータや学習結果を再利用する「転移学習(Transfer Learning)」を行い、効果を検証する。

#### (2) データ拡張 (Data Augmentation)

画像認識において、学習させた認識モデルの性能を向上させるためには、認識モデルの設計や学習方法も重要であるが、分類性能に一番影響を及ぼすのは教師データである。高度なモデルを用いても、教師データセットが想定していないデータの認識は不可能である。そのため、十分な汎化性能を得るために、目的のタスクに対して十分に網羅的で、かつ一定量以上の教師データを揃える必要がある。

データ拡張は、今回のような限られたサンプル数で深層学習を行う場合に、保有するデータを基に擬似的なデータを作成してデータ量の拡張を図り、データ不足を補って認識の頑健性を高める手法である。今回のクラス分類では、回転・平行移動・傾斜・拡大・縮小・水平反転を組み合わせ、保有データの18倍の拡張を行った。データ拡張の実施例を図2-6に示す。

また、新たなデータ拡張手法であるmixup<sup>7)</sup>についても検討を行なった。mixupとは、異なるクラス間の組み合わせも含むランダムに選んだ2つの教師データをマージし、新たな教師データを生成する手法である。具体的には、画像データ  $x$  とラベル  $y$  のペア  $(x_1, y_1)$ ,  $(x_2, y_2)$  から、式(1),(2)により新たな画像  $x$  とラベル  $y$  を作成する。なお、画像に対する正解ラベルは、one-hot表現 (正解クラスが1でその他のクラスは0となるビット

列) になっているものとする. mixupの実施例を図2-7に示す. mixupでは, それぞれのクラスの補間画像を生成することができることから, 劣化損傷のように時系列的に徐々に変化して, クラスの境界の設定が難しい事象の画像に対しては有効となる可能性があると考えた.

$$x = \lambda \cdot x_1 + (1 - \lambda) \cdot x_2 \quad (1)$$

$$y = \lambda \cdot y_1 + (1 - \lambda) \cdot y_2 \quad (2)$$

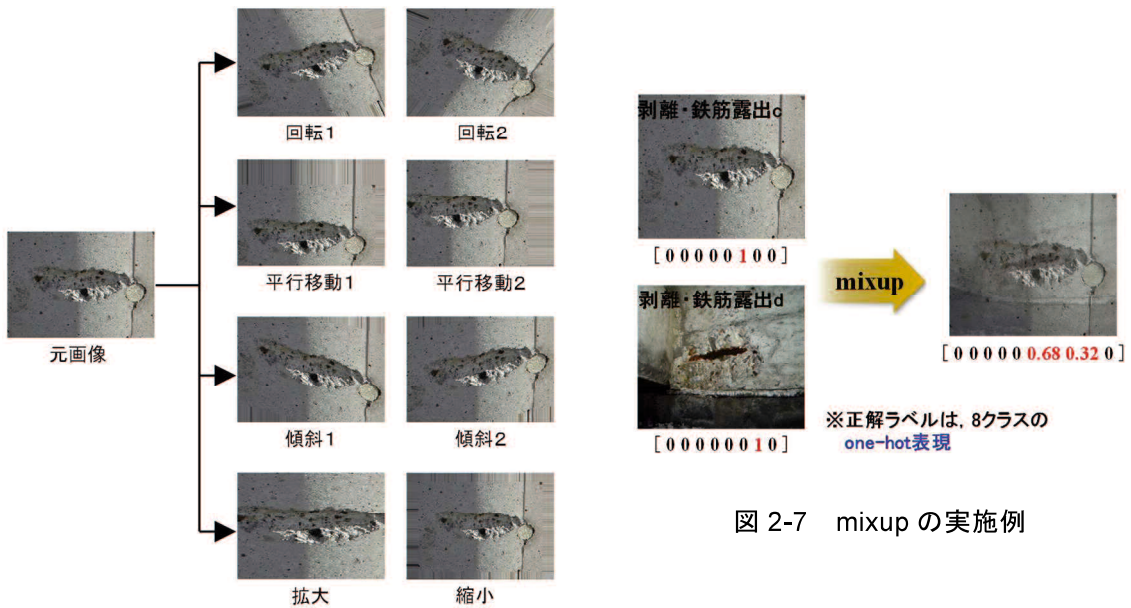


図 2-7 mixup の実施例

(元画像+8パターン) × 水平反転 = 18倍

図 2-6 データ拡張の実施例

### (3) 転移学習 (Transfer Learning)

転移学習とは, 既に学習済みのモデルの知識や特徴を再利用し, 学習時間の削減や教師データ不足等を補う手法である. CNNにおける転移学習の手法は, 主に図2-8に示す2つがある. 一つは, 学習済みモデルの出力層以外の部分を, 層間の結合重みは変更せずに特徴抽出器として使用し, 分類等のためにSVM (サポートベクターマシン) 等の機械学習モデルへの入力を行う手法である. もう一つは, Fine-tuningと呼ばれる手法で, 初期値は学習済みモデルの層間の結合重みをそのまま使い, 途中の層から学習し直して当該タスクに適した重みに再調整する手法である.

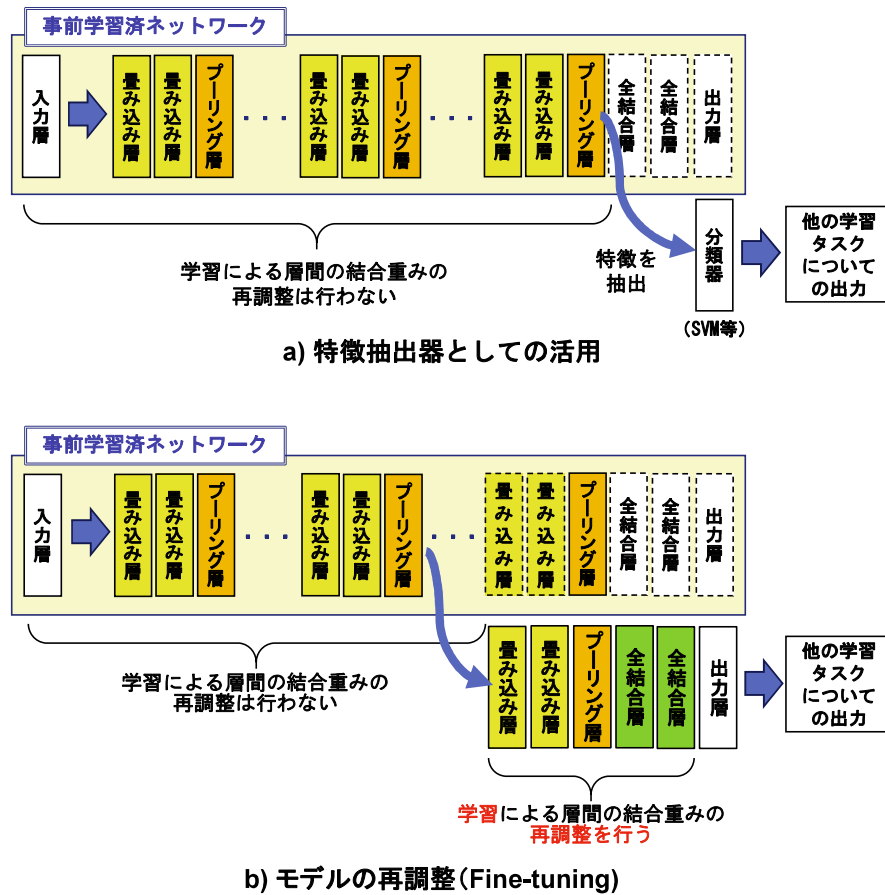


図 2-8 転移学習の概要

本研究では、大規模画像データセットImageNetを学習したモデルをもとにFine-tuningを行った。CNNでは、入力層に近い層ではエッジ等の低次の汎用的な特徴が抽出されるのに対し、出力層に近い層では当該タスクの教師データに特化した特徴が抽出される傾向にあるとされている<sup>8)</sup>。そのため、出力層に近い層において学習を行い、モデルの再調整を行なった。本研究におけるFine-tuningの適用を図2-9に示す



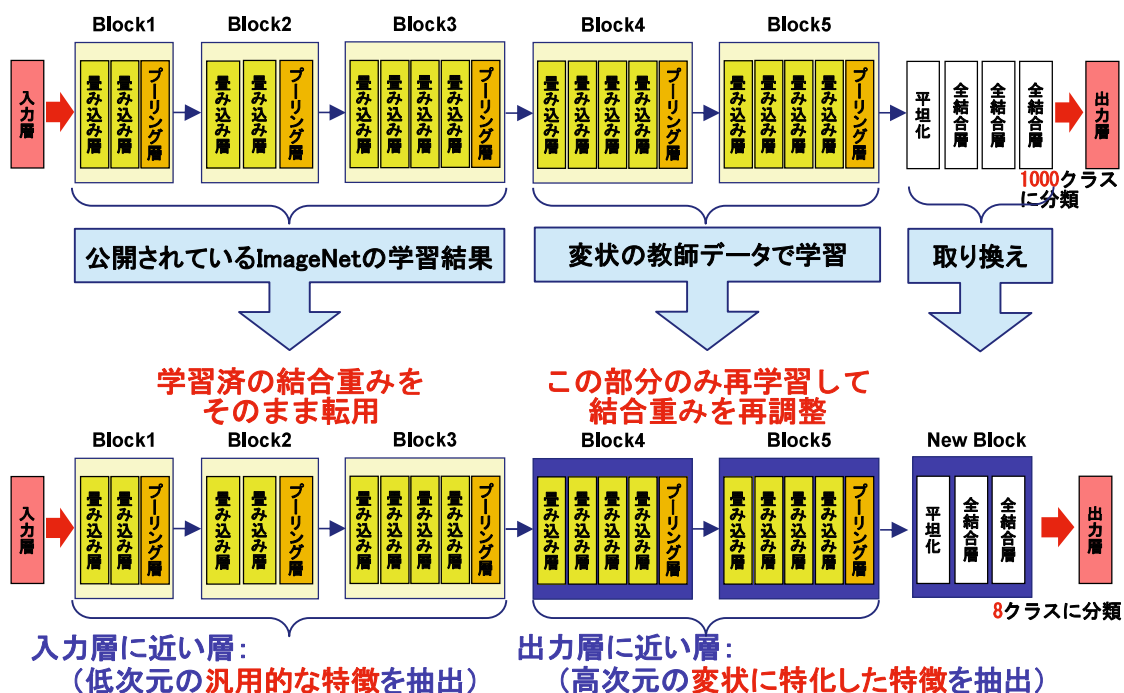


図 2-9 本研究における Fine-tuning の適用

### 2-2-3 分類結果

クラス分類は、入力画像について未処理の状態、従来のデータ拡張あるいはmixupを行なった状態と、Fine-tuningの実施を組み合わせる表2-1に示す計5ケースの学習を行い、汎化性能の向上について検証を行った。CNNによる分類では、図2-4中の出力層において、8クラスそれぞれの判定確率が出力として得られ、その確率が最大となるクラスを予測結果のクラスとしている。各ケースの分類結果は、図2-10に示す混同行列の形式で整理

表 2-1 クラス分類の検討ケース

No.	処理条件	検討条件		
		Fine-tuning	データ拡張 (従来)	データ拡張 (mixup)
ケース 1	未処理	—	—	—
ケース 2	データ拡張 (従来) のみ	—	●	—
ケース 3	Fine-tuning のみ	●	—	—
ケース 4	Fine-tuning + データ拡張 (従来)	●	●	—
ケース 5	Fine-tuning + データ拡張 (mixup)	●	—	●

		予測クラス	
		Positive (予測結果が正)	Negative (予測結果が負)
正解 クラス	Positive (正解クラスが正)	真陽性 TP (True Positive)	偽陰性 FN (False Negative)
	Negative (正解クラスが負)	偽陽性 FP (False Positive)	真陰性 TN (True Negative)

図 2-10 混同行列 (Confusion Matrix)

を行った。混同行列とはクラスの分類結果をまとめたもので、正しく分類できた割合（再現率 [真陽性率] =  $TP / (TP + FN)$ ）、誤って分類した割合（偽陰性率 =  $FN / (TP + FN)$ ）を知ることができる。各ケースの分類結果を混同行列の形式で整理したものを表2-2～2-6に、各ケースのクラス別再現率を表2-7に示す。また、それぞれの変状のケース別再現率を図2-11～2-18に示す。図2-11～2-18によると、Fine-tuningの実施により、再現率の大幅な向上が見られた。事前学習に用いているImageNetのデータセットは、一般的な物体のデータセットでコンクリートの変状等の特殊なクラスは含まれていない。しかし、入力に近い層で行われる汎用的な特徴の抽出は、全く異なるクラスであるコンクリートの変状に対する学習においても十分な効果が得られている。また、一般に、CNNの学習は初期値依存性が強く、特に教師データが少ない場合は、過学習の抑制の観点からも適切な初期値を与えることが重要とされている<sup>8)</sup>。この点からも、再現率の大幅な向上に寄与したものと考えられる。

データ拡張については、Fine-tuningとの組み合わせでは表2-7および図2-16に見られるように漏水・遊離石灰dで再現率が若干低下したが、概ね性能の向上が見られ、有効性が確認できた。

また、mixupについては、判断に曖昧性を含むように学習するという特性がある。この特性から、漏水・遊離石灰および剥離・鉄筋露出のように同種の変状で程度の区分によりクラスがさらに細分化される条件に対しては、その有効性が期待されたが、従来のデータ拡張ほどの効果は得られなかった。今後、同種の変状のクラス内のみの組み合わせに限定して使用する等、活用方法を工夫した条件での検証を重ねていきたい。

変状の種類によって若干のばらつきはあるが、全5ケースでは、ケース4のFine-tuning + データ拡張の条件が、大半の変状および全変状の平均で再現率が最高値となった。漏

水・遊離石灰cについては、8クラスの中で最も再現率が低く、ケース4でも78%であった。ひびわれからの漏水は、「漏水・遊離石灰c」という評価とともに「ひびわれ」としても評価を行う変状で、画像としても双方の特徴を有するため、どちらか一方のクラスとしての分類は難しい画像があったためと考えられる。また、漏水・遊離石灰e-2錆汁との誤識別も多く、壁面等を液体が上から下に流下するという共通点に起因したものと考えられる。なお、漏水・遊離石灰cは、ケース3とケース4との間での再現率の向上が最も大きかったクラスで、Fine-tuning後のデータ拡張の効果が大きいと考えられることから、教師データの量を増やすことにより、精度の向上が図れると考えられる。

表 2-2 ケース 1 (未処理) の分類結果

		分類結果 (%)								
		ひびわれ	剥離・鉄筋露出			漏水・遊離石灰				
			c	d	e	c	d	e-1 つらら	e-2 錆汁	
正解クラス	ひびわれ	61.1	6.7	8.9	0.0	5.5	14.4	0.0	3.3	
	剥離・鉄筋露出	c	31.1	21.1	20.0	4.4	5.6	13.3	1.1	3.3
		d	11.1	10.0	54.5	6.7	2.2	8.9	1.1	5.5
		e	0.0	0.0	7.8	81.1	5.6	0.0	0.0	5.5
	漏水・遊離石灰	c	14.5	15.5	22.2	4.4	14.4	11.1	3.3	14.5
		d	42.2	12.2	16.7	0.0	4.5	22.2	0.0	2.2
		e-1	0.0	25.9	16.7	24.1	5.6	9.3	9.3	9.3
		e-2	8.9	5.6	21.1	16.7	16.7	3.3	0.0	27.8

■ : True Positiveに該当するゾーン (数値は再現率を意味する)  
 赤字 : 各クラスの分類結果の比率の最高値

表 2-3 ケース 2 (データ拡張 (従来)) の分類結果

		分類結果 (%)								
		ひびわれ	剥離・鉄筋露出			漏水・遊離石灰				
			c	d	e	c	d	e-1 つらら	e-2 錆汁	
正解クラス	ひびわれ	66.7	3.3	5.5	0.0	4.4	16.7	1.1	2.2	
	剥離・鉄筋露出	c	22.2	32.2	15.5	1.1	0.0	11.1	12.2	5.5
		d	1.1	8.9	50.0	7.8	4.4	7.8	5.6	14.4
		e	0.0	0.0	4.4	63.4	1.1	0.0	3.3	27.8
	漏水・遊離石灰	c	7.8	7.8	21.1	2.2	10.0	11.1	7.8	32.2
		d	20.0	18.9	6.6	3.3	3.3	35.6	10.0	2.2
		e-1	3.7	16.7	11.1	13.0	9.3	14.8	16.7	14.8
		e-2	0.0	3.3	11.1	12.2	15.6	0.0	1.1	56.7

■ : True Positiveに該当するゾーン (数値は再現率を意味する)  
 赤字 : 各クラスの分類結果の比率の最高値

表 2-4 ケース 3 (Fine-tuning) の分類結果

		分類結果 (%)								
		ひびわれ	剥離・鉄筋露出			漏水・遊離石灰				
			c	d	e	c	d	e-1 つらら	e-2 錆汁	
正解クラス	ひびわれ	91.3	0.0	5.3	0.0	2.7	0.0	0.0	0.7	
	剥離・鉄筋露出	c	0.0	89.3	0.7	1.3	3.3	2.0	0.0	3.3
		d	1.3	3.3	88.7	2.0	0.7	1.3	2.0	0.7
		e	0.0	0.0	4.0	94.0	1.3	0.0	0.7	0.0
	漏水・遊離石灰	c	8.0	0.0	8.7	1.3	58.0	13.3	4.7	6.0
		d	0.0	5.3	2.7	0.7	2.6	87.3	1.3	0.0
		e-1	0.0	1.1	2.2	3.4	3.4	11.1	75.6	3.3
		e-2	0.0	0.0	0.0	4.0	2.7	0.7	3.3	89.3

■ : True Positiveに該当するゾーン (数値は再現率を意味する)  
 赤字 : 各クラスの分類結果の比率の最高値

表 2-5 ケース 4 (Fine-tuning+データ拡張 (従来)) の分類結果

		分類結果 (%)								
		ひびわれ	剥離・鉄筋露出			漏水・遊離石灰				
			c	d	e	c	d	e-1 つらら	e-2 錆汁	
正解クラス	ひびわれ	96.7	0.7	0.7	0.0	0.0	0.0	2.0	0.0	
	剥離 ・ 鉄筋露出	c	0.0	97.4	0.0	0.0	1.3	1.3	0.0	0.0
		d	0.0	2.6	88.7	5.3	3.3	0.0	0.0	0.0
		e	0.0	0.0	2.0	96.7	0.0	0.0	0.0	1.3
	漏水 ・ 遊離石灰	c	5.3	0.0	2.7	1.3	78.0	4.7	0.0	8.0
		d	2.0	3.3	0.0	0.7	2.6	86.0	5.3	0.0
		e-1	0.0	1.1	1.1	0.0	3.4	11.1	83.3	0.0
		e-2	0.0	0.0	0.0	4.0	0.0	0.7	4.0	91.3

■ : True Positiveに該当するゾーン (数値は再現率を意味する)  
 赤字 : 各クラスの分類結果の比率の最高値

表 2-6 ケース 5 (Fine-tuning+mixup) の分類結果

		分類結果 (%)								
		ひびわれ	剥離・鉄筋露出			漏水・遊離石灰				
			c	d	e	c	d	e-1 つらら	e-2 錆汁	
正解クラス	ひびわれ	92.7	0.7	2.6	0.7	2.0	1.3	0.0	0.0	
	剥離 ・ 鉄筋露出	c	0.0	93.4	0.0	0.7	2.6	2.7	0.0	0.7
		d	1.3	3.3	85.3	3.3	2.7	2.0	0.0	2.0
		e	0.0	0.7	4.0	94.7	0.0	0.0	0.0	0.7
	漏水 ・ 遊離石灰	c	8.0	1.3	9.3	2.0	64.7	7.3	0.7	6.7
		d	0.7	9.3	2.0	1.3	4.7	75.3	6.0	0.7
		e-1	2.2	1.1	7.8	5.6	2.2	11.1	66.7	3.4
		e-2	0.0	0.0	0.0	6.0	2.6	2.6	2.0	86.7

■ : True Positiveに該当するゾーン (数値は再現率を意味する)  
 赤字 : 各クラスの分類結果の比率の最高値

表 2-7 各ケースのクラス別再現率

ケース	再現率 (%)								
	ひびわれ	剥離・鉄筋露出			漏水・遊離石灰				平均
		c	d	e	c	d	e-1 つらら	e-2 錆汁	
1	61.1	21.1	54.5	81.1	14.4	22.2	9.3	27.8	36.4
2	66.7	32.2	50.0	63.4	10.0	35.6	16.7	56.7	41.4
3	91.3	89.3	88.7	94.0	58.0	87.3	75.6	89.3	84.2
4	96.7	97.4	88.7	96.7	78.0	86.0	83.3	91.3	89.8
5	92.7	93.4	85.3	94.7	64.7	75.3	66.7	86.7	82.4

赤字 : 各クラスの再現率の最高値

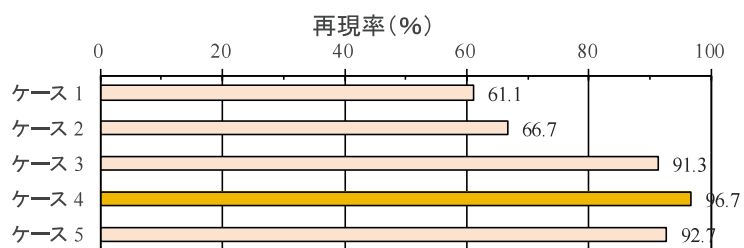


図 2-11 ひびわれのケース別再現率

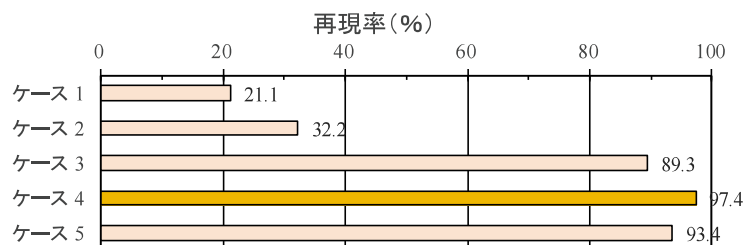


図 2-12 剥離・鉄筋露出 c のケース別再現率

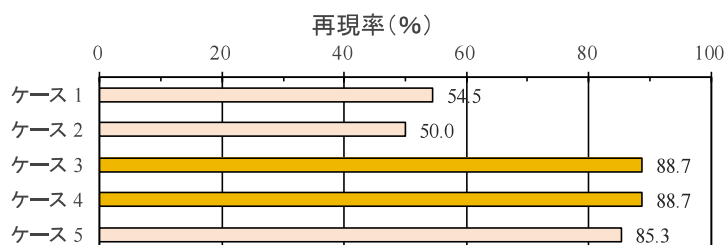


図 2-13 剥離・鉄筋露出 d のケース別再現率

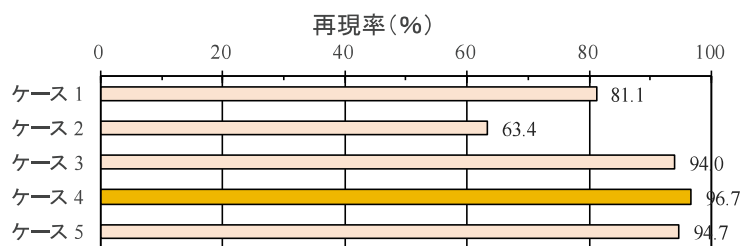


図 2-14 剥離・鉄筋露出 e のケース別再現率

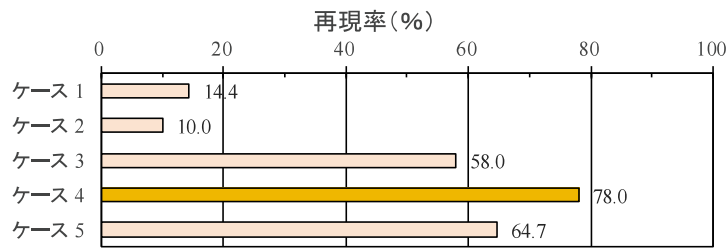


図 2-15 漏水・遊離石灰 c のケース別再現率

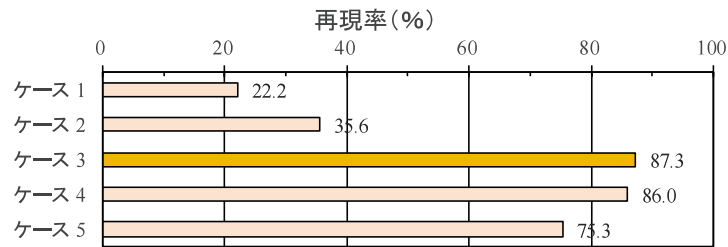


図 2-16 漏水・遊離石灰 d のケース別再現率

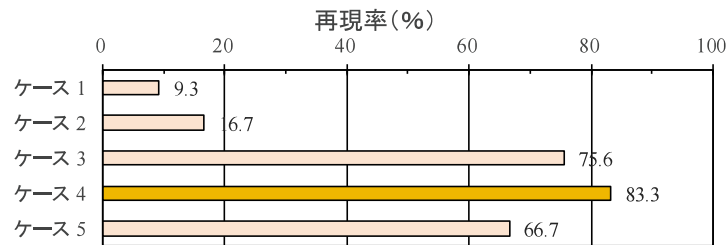


図 2-17 漏水・遊離石灰 e-1 つららのケース別再現率

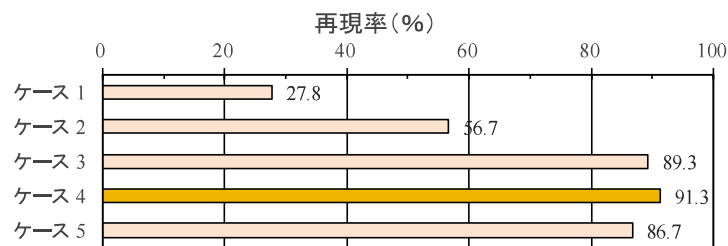


図 2-18 漏水・遊離石灰 e-2 錆汁のケース別再現率

### 2-3 判断根拠の可視化による確認

#### 2-3-1 Grad-CAM

次に、クラス分類を行う際に着目した領域の可視化を行い、分類についての検証を行った。深層学習全般における課題の一つに、学習から得られたモデルにより分類、予測、異常検知等を行なった際に、結果の根拠を説明できないというブラックボックス問題が挙げられる。画像認識の分野においては、このような課題に対して、判断根拠を理解するための手法が多数提案されている。それらの提案手法の中で、本研究では、クラスの分類において、着目している領域を可視化するという特徴を有する Grad-CAM (Gradient-weighted Class Activation Mapping)<sup>9)</sup>を用いて検証を行なった。Grad-CAM は、テストデータを CNN へ入力した際に、クラス分類に至るまでの各特徴マップの寄与の度合いを算出し、その重みを逆伝搬していくという手法により、ヒートマップのような図で、クラス分類に大きく寄与した領域すなわち着目領域を図化する技術である。

#### 2-3-2 可視化結果

各変状のクラスのテストデータに対してクラス分類を行った際の、Grad-CAM による判断根拠の可視化結果を表 2-8 に示す。剥離・鉄筋露出 e 以外は、Fine-tuning を行っていないケース 1, 2 と Fine-tuning を行なったケース 3~5 では、明確に傾向が異なっており、2-2 節における再現率の傾向と整合する。

ケース 3, 4, および剥離・鉄筋露出 d と漏水・遊離石灰 d 以外のケース 5 では、変状部分に強く反応して、クラスの特定に寄与していることが明瞭に表れており、適切に学習が行われた上で分類されていることが裏付けられた。剥離・鉄筋露出 d と漏水・遊離石灰 d は、ケース 3, 4 では正解クラスに対する判定確率がほぼ 100%であるのに対して、mixup でデータを拡張したケース 5 では判定確率の順位では 2 番目で 50%以下となっており、着目している領域も適切でなく、判定が曖昧となったことが明確に見て取れる。剥離・鉄筋露出も漏水・遊離石灰も、c ランクおよび e ランクはケース 5 においてケース 4 からの極端な判定確率の低下は見られないが、d ランクは極端に判定確率が低下している。ケース 5 において剥離・鉄筋露出 d と漏水・遊離石灰 d の判定結果が低調となったのは、中間の区分の d であることから、上位の区分および下位の区分との mix を含むデータ拡張となったことにより、曖昧さを含んだ判定となったものと考えられる。

データ拡張を行う前後で、ケース 1 からケース 2 で 8 クラス中 5 クラス、ケース 3 からケース 4 で 8 クラス中 6 クラスについて、図 2-11~2-18 に示す再現率が向上してい



る。ただし、ケース2についてのテストデータに対する Grad-CAM の適用結果では、剥離・鉄筋露出 e 以外はクラスの識別に関係ない部分に着目しており、判定確率も正解のクラスについて突出した値とまではなっていない。ケース2におけるデータ拡張の効果は、図 2-11～2-18 に示す数値上では出ているが、初期値依存性が強いとされている CNN において、コンクリート変状の画像におけるあらゆる多様性から十分に特徴を抽出できるモデルとなっていないものと考えられる。Fine-tuning を行った段階のケース3で、ようやくコンクリート変状の画像の特徴を適切に抽出できるモデルとなったことが、テストデータに対する Grad-CAM の適用結果、正解のクラスについて判定確率が突出した値となったことから裏付けられている。ケース3とケース4と間の変化は、モデルの質が一定レベル以上になった上での再現率の数値の変化であり、本質的な頑健性の向上の伴った効果だと考えられる。

以下に、8種類のクラスのそれぞれについて詳述する。

### (1) ひびわれ

ケース1, 2では、画像全体に着目しており、ひびわれの幅が太い部分については逆に着目領域から除外しており、正解クラスに対する判定確率はケース1で5番目、ケース2で4番目という低調な結果である。

Fine-tuning を行なったケースでは、ケース3, 4, 5のいずれも変状の領域に的確に着目しており、正解クラスに対する判定確率はケース3で97.23%、ケース4で99.92%、ケース5で94.70%といずれも高い判定確率での1位の順位であり、適切な分類が行われている。ケース3, 4, 5では、ひびわれの部分が最も影響した領域となっており、適切な領域に着目した上で、「ひびわれ」と分類していることを裏付ける可視化結果である。

### (2) 剥離・鉄筋露出 c

ケース1では、画像全体に着目しており、正解クラスに対する判定確率は最高値となったものの28.16%と低く、分類の判定に妥当性はない。ケース2では、変状と関係ない領域に着目しており、正解クラスに対する判定確率は3番目であるが、ケース1と同様に分類の判定に妥当性はない。

Fine-tuning を行なったケースでは、ケース3, 4, 5のいずれも変状の領域を的確に着目しており、正解クラスに対する判定確率はケース3で98.87%、ケース4で100%、ケース5で88.81%といずれも高い判定確率での1位の順位であり、適切な分類が行われて

いる。コンクリートの剥離部分の変状領域のみを着目しており、適切な領域に着目した上で、「剥離・鉄筋露出 c」と分類していることを裏付ける可視化結果である。

### (3) 剥離・鉄筋露出 d

ケース 1 では、全く関係ない領域に着目しており、正解クラスに対する判定確率は 5 番目であった。ケース 2 では、剥離・鉄筋露出の部分も着目しているが、大半は全く関係ない領域に着目しており、正解クラスに対する判定確率は最高値となったものの 27.75%と低く、分類の判定に妥当性はない。

Fine-tuning を行なったケースでは、ケース 3, 4 ともに変状の領域を的確に着目しており、正解クラスに対する判定確率はケース 3 で 99.94%、ケース 4 で 100%といずれも高い判定確率での 1 位の順位であり、適切な分類が行われている。さらに鉄筋の部分が最も影響した領域となっており、鉄筋の露出と鉄筋周辺のコンクリートの剥離部分に着目した上で、「剥離・鉄筋露出 d」と分類していることを裏付ける可視化結果である。ケース 5 では、鉄筋ではなくコンクリートの剥離部分の領域に着目しており、正解クラスに該当する確率は「剥離・鉄筋露出 c」の 52.51%に次ぐ 2 番目で 41.14%であった。「剥離・鉄筋露出 d」は変状の領域に、「剥離・鉄筋露出 c」と同様にコンクリートの剥離部分を含んでいることから、mixup によるデータ拡張の結果、特徴量の捉え方が曖昧になったものと思われ、従来のデータ拡張に対する優位性は見られなかった。

### (4) 剥離・鉄筋露出 e

他の変状と異なり、Fine-tuning を行なっていないケース 1, 2 においても、変状の領域に着目しており、正解クラスに対する判定確率はケース 1 で 61.46%、ケース 2 で 91.11%といずれも高い判定確率での 1 位の順位であり、適切な分類が行われている。

Fine-tuning を行なったケース 3, 4, 5 においても、変状の領域を的確に着目しており、正解クラスに対する判定確率はケース 3 で 100%、ケース 4 で 100%、ケース 5 で 99.94%といずれも高い判定確率での 1 位の順位であり、適切な分類が行われている。さらに鉄筋の腐食が著しい部分が最も影響した領域となっており、鉄筋の著しい腐食と鉄筋周辺のコンクリートの剥離部分に着目した上で、「剥離・鉄筋露出 e」と分類していることを裏付ける可視化結果である。

### (5) 漏水・遊離石灰 c

ケース1では、画像全体に着目しており、正解クラスに対する判定確率は4番目であった。ケース2では、変状と関係ない領域に着目しており、正解クラスに対する判定確率は6番目で、ケース1と同様に分類の判定に妥当性はない。

Fine-tuningを行なったケースでは、ケース3, 4, 5のいずれも変状の領域を的確に着目しており、正解クラスに対する判定確率はケース3で71.56%、ケース4で99.56%、ケース5で68.55%といずれも高い判定確率での1位の順位であり、適切な分類が行われている。コンクリートから水が滲み出た変状領域のみを着目しており、適切な領域に着目した上で、「漏水・遊離石灰 c」と分類していることを裏付ける可視化結果である。

### (6) 漏水・遊離石灰 d

ケース1では、画像全体に着目しており、正解クラスに対する判定確率は2番目となったものの、分類の判定に妥当性はない。ケース2では、遊離石灰の析出部分を一部着目しているが、大半は全く関係ない領域に着目しており、正解クラスに対する判定確率は3番目となったものの、ケース1と同様に分類の判定に妥当性はない。

Fine-tuningを行なったケースでは、ケース3, 4ともに変状の領域を的確に着目しており、正解クラスに対する判定確率はケース3で98.54%、ケース4で98.23%といずれも高い判定確率での1位の順位であり、適切な分類が行われている。遊離石灰の析出部分が最も影響した領域となっており、遊離石灰に的確に着目した上で、「漏水・遊離石灰 d」と分類していることを裏付ける可視化結果である。ケース5では、遊離石灰の析出部分を一部着目しているが、大半は全く関係ない領域に着目しており、正解クラスに該当する確率は「漏水・遊離石灰 c」の45.25%に次ぐ2番目で42.83%であったものの、分類の判定に妥当性はない。mixupによるデータ拡張の結果、特徴量の捉え方が曖昧になったものと思われ、従来のデータ拡張に対する優位性は見られなかった。

### (7) 漏水・遊離石灰 e-1 (つらら)

ケース1では、変状以外の全く関係ない領域を中心に画像全体を着目しており、正解クラスに対する判定確率は7番目であった。ケース2では、遊離石灰の析出部分を一部着目しているが、大半は全く関係ない領域に着目しており、正解クラスに対する判定確率は4番目であった。

Fine-tuning を行なったケースでは、ケース 3, 4, 5 のいずれも変状の領域を的確に着目しており、正解クラスに対する判定確率はケース 3 で 61.37%, ケース 4 で 100%, ケース 5 で 95.73%といずれも高い判定確率での 1 位の順位であり、適切な分類が行われている。遊離石灰の析出部分の中でも、細く尖ったつららの形状の特徴的な部分が最も影響した領域となっており、適切な領域に着目した上で、「漏水・遊離石灰 e-1 (つらら)」と分類していることを裏付ける可視化結果である。

### (8) 漏水・遊離石灰 e-2 (錆汁)

ケース 1 では、変状以外の全く関係ない領域に着目しており、正解クラスに対する判定確率は 2 番目となったものの、分類の判定に妥当性はない。ケース 2 では、漏水の部分を一部着目しているが、大半は全く関係ない領域に着目しており、正解クラスに対する判定確率は 2 番目となったが、ケース 1 と同様に分類の判定に妥当性はない。

Fine-tuning を行なったケースでは、ケース 3, 4, 5 のいずれも変状の領域を的確に着目しており、正解クラスに対する判定確率はケース 3 で 78.31%, ケース 4 で 97.19%, ケース 5 で 87.60%といずれも高い判定確率での 1 位の順位であり、適切な分類が行われている。漏水の部分の中でも、錆汁の茶色い部分が最も影響した領域となっており、適切な領域に着目した上で、「漏水・遊離石灰 e-2 (錆汁)」と分類していることを裏付ける可視化結果である。

表 2-8(1) 各クラスのテストデータに対する Grad-CAM による可視化結果 (その1)

テスト画像	判断根拠の可視化結果				
	ケース 1	ケース 2	ケース 3	ケース 4	ケース 5
ひびわれ	8.15%	12.80%	97.23%	99.92%	94.70%
剥離・鉄筋露出c	25.72%	36.61%	0.01%	0.01%	0.40%
剥離・鉄筋露出d	34.61%	7.02%	2.51%	0.03%	0.11%
剥離・鉄筋露出e	0.52%	0.06%	0.00%	0.00%	0.33%
漏水・遊離石灰c	11.26%	2.28%	0.04%	0.01%	4.18%
漏水・遊離石灰d	5.94%	14.69%	0.14%	0.01%	0.10%
漏水・遊離石灰e-1つらら	4.00%	25.27%	0.04%	0.00%	0.10%
漏水・遊離石灰e-2錆汁	9.81%	1.27%	0.03%	0.01%	0.09%
ひびわれ	9.27%	5.23%	0.01%	0.00%	0.53%
剥離・鉄筋露出c	28.16%	16.07%	98.87%	100.00%	88.81%
剥離・鉄筋露出d	2.73%	4.30%	0.02%	0.00%	4.02%
剥離・鉄筋露出e	0.23%	0.39%	0.00%	0.00%	0.07%
漏水・遊離石灰c	22.05%	8.34%	0.00%	0.00%	2.99%
漏水・遊離石灰d	22.34%	15.69%	0.00%	0.00%	0.12%
漏水・遊離石灰e-1つらら	9.30%	30.16%	0.01%	0.00%	1.37%
漏水・遊離石灰e-2錆汁	5.92%	19.81%	1.09%	0.00%	2.07%
ひびわれ	6.27%	1.42%	0.00%	0.00%	0.82%
剥離・鉄筋露出c	19.51%	17.79%	0.01%	0.00%	52.51%
剥離・鉄筋露出d	10.95%	27.75%	99.94%	100.00%	41.14%
剥離・鉄筋露出e	1.59%	0.69%	0.00%	0.00%	0.30%
漏水・遊離石灰c	17.95%	8.13%	0.00%	0.00%	1.08%
漏水・遊離石灰d	3.52%	22.79%	0.00%	0.00%	0.39%
漏水・遊離石灰e-1つらら	12.11%	17.31%	0.00%	0.00%	2.77%
漏水・遊離石灰e-2錆汁	28.10%	4.12%	0.06%	0.00%	0.99%
ひびわれ	2.44%	0.00%	0.00%	0.00%	0.00%
剥離・鉄筋露出c	0.70%	0.04%	0.00%	0.00%	0.00%
剥離・鉄筋露出d	4.99%	0.57%	0.00%	0.00%	0.04%
剥離・鉄筋露出e	61.46%	91.11%	100.00%	100.00%	99.94%
漏水・遊離石灰c	3.70%	0.28%	0.00%	0.00%	0.00%
漏水・遊離石灰d	0.06%	0.01%	0.00%	0.00%	0.01%
漏水・遊離石灰e-1つらら	6.94%	2.84%	0.00%	0.00%	0.00%
漏水・遊離石灰e-2錆汁	19.70%	5.14%	0.00%	0.00%	0.00%

※1 画像中の赤色の領域がクラス分類に最も影響した領域  
 ※2 ■は判定確率が最も高かったクラス、■は判定確率が2番めに高かったクラス  
 ※3 赤字は、正解に該当するクラスについての判定結果

第2章 深層学習による変状画像のクラス分類

表 2-8(2) 各クラスのテストデータに対する Grad-CAM による可視化結果 (その2)

テスト画像	判断根拠の可視化結果				
	ケース1	ケース2	ケース3	ケース4	ケース5
判定確率					
ひびわれ	26.95%	10.69%	0.00%	0.00%	1.12%
剥離・鉄筋露出c	22.33%	20.71%	0.02%	0.00%	0.55%
剥離・鉄筋露出d	8.81%	3.93%	0.00%	0.00%	0.42%
剥離・鉄筋露出e	0.29%	0.14%	0.02%	0.00%	5.99%
漏水・遊離石灰c	8.87%	3.60%	71.56%	99.56%	68.55%
漏水・遊離石灰d	21.24%	12.11%	28.00%	0.43%	15.15%
漏水・遊離石灰e-1つらら	6.44%	46.61%	0.13%	0.01%	7.74%
漏水・遊離石灰e-2錆汁	5.07%	2.22%	0.25%	0.00%	0.48%
判定確率					
ひびわれ	32.27%	71.23%	0.01%	0.00%	2.67%
剥離・鉄筋露出c	19.50%	13.83%	0.07%	0.02%	0.42%
剥離・鉄筋露出d	9.88%	1.79%	0.01%	0.00%	3.22%
剥離・鉄筋露出e	0.20%	0.02%	0.00%	0.00%	1.03%
漏水・遊離石灰c	6.59%	0.94%	0.56%	0.02%	45.25%
漏水・遊離石灰d	24.42%	10.67%	98.54%	98.23%	42.83%
漏水・遊離石灰e-1つらら	3.55%	1.25%	0.61%	1.70%	3.67%
漏水・遊離石灰e-2錆汁	3.59%	0.27%	0.20%	0.03%	0.93%
判定確率					
ひびわれ	29.83%	6.10%	0.00%	0.00%	1.46%
剥離・鉄筋露出c	18.96%	27.39%	0.01%	0.00%	0.26%
剥離・鉄筋露出d	11.87%	29.62%	0.00%	0.00%	0.19%
剥離・鉄筋露出e	0.32%	0.20%	0.00%	0.00%	0.41%
漏水・遊離石灰c	8.14%	5.33%	0.00%	0.00%	0.69%
漏水・遊離石灰d	21.76%	17.93%	38.53%	0.00%	0.40%
漏水・遊離石灰e-1つらら	4.46%	10.27%	61.37%	100.00%	95.73%
漏水・遊離石灰e-2錆汁	4.66%	3.16%	0.08%	0.00%	0.86%
判定確率					
ひびわれ	6.09%	1.34%	0.18%	0.00%	3.76%
剥離・鉄筋露出c	16.32%	1.92%	0.28%	0.00%	0.36%
剥離・鉄筋露出d	28.11%	17.34%	7.78%	0.02%	1.93%
剥離・鉄筋露出e	2.15%	0.46%	0.09%	0.00%	0.65%
漏水・遊離石灰c	18.97%	44.39%	12.86%	2.79%	3.27%
漏水・遊離石灰d	3.60%	2.32%	0.48%	0.00%	0.44%
漏水・遊離石灰e-1つらら	4.93%	0.83%	0.03%	0.00%	1.99%
漏水・遊離石灰e-2錆汁	19.84%	31.40%	78.31%	97.19%	87.60%

- ※1 画像中の赤色の領域がクラス分類に最も影響した領域
- ※2 ■は判定確率が最も高かったクラス, ■は判定確率が2番めに高かったクラス
- ※3 赤字は, 正解に該当するクラスについての判定結果

## 2-4 まとめ

本章では、深層学習によるコンクリートの変状検出の基礎検討として、変状画像のクラス分類について検討を行った。本章における検討により、以下の点を明らかにした。

- (1) 深層学習を目的とした撮影ではない点検時および踏査時の撮影画像を教師データに用いても、深層学習による画像のクラス分類において、変状の種類と程度の区分も含めて識別させた上で高い精度で分類が行えることを明らかにした。
- (2) コンクリートの変状に関する教師データが十分な量でない場合でも、一般的な画像についての事前学習済みモデルを用いた転移学習（Fine-tuning）により、分類精度の格段の向上が図れることを明らかにした。
- (3) 深層学習の課題の一つにブラックボックス問題が挙げられるが、テストデータに対してGrad-CAMにより分類時の着目領域の可視化を行なった結果、適切な部分に着目した上でクラス分類の判断を行っていることを把握した。

上記の結果より、深層学習による物体認識技術を用いることで、コンクリートの変状の画像について、画像内の適切な部分に着目した上で、変状の種類のみならず程度の区分も含めて精度よく分類を行えることが検証できた。領域の検出を行う物体検出およびセグメンテーションは、クラス分類よりも難易度が高いタスクであるが、領域の検出と並行的あるいは段階的にクラス分類を行う工程を経る。本章の結果を踏まえて、次章では難易度が高くなる物体検出およびセグメンテーションに関して、コンクリートの変状検出の適用性について検討を行う。

また、深層学習によるクラス分類は、経験が浅い技術者が現地での撮影画像を基に損傷程度の判断を行う際の、支援ツールとしても活用できると考えられる。

## 第2章 参考文献

- 1) スタンフォード大学, プリンストン大学 : ImageNet Large Scale Visual Recognition Challenge (ILSVRC), <http://image-net.org/challenges/LSVRC/>, (最終閲覧日 : 2020.03.01)
- 2) Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L. : ImageNet: A large-scale hierarchical image database, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 248-255, 2009.
- 3) LeCun, Y., Boser, B., Denker, J., Henderson, D., How-ard, R., Hubbard, W. and Jackel, L. : Backpropagation applied to handwritten zip code recognition, *Neural Computation*, Vol. 1, No. 4, pp. 541-551, 1989.
- 4) Krizhevsky, A., Sutskever, I. and Hinton, G. E. : ImageNet classification with deep convolutional neural networks, *Neural Information Processing Systems*, pp. 1097-1105, 2012.
- 5) LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. : Gradient-Based Learning Applied to Document Recognition, *in Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278-2324, 1998.
- 6) Simonyan, K., Zisserman, A. : Very deep convolutional networks for large-scale image recognition, *arXiv: 1409.1556v6*, 2015.
- 7) Zhang, H., Cisse, M., Dauphin, Y. N. and Lopez-Paz, D. : mixup: Beyond Empirical Risk Minimization, *arXiv: 1710.09412v2*, 2018.
- 8) 中山英樹 : 深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習, 電子情報通信学会技術研究報告, Vol. 115, No. 146, pp. 55-59, 2015.
- 9) Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D. : Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, *IEEE International Conference on Computer Vision (ICCV)*, pp. 618-626, 2017.



## 第 3 章 深層学習による物体検出およびセグメンテーション

---

### 第 3 章

3-1 教師データの構成

3-2 深層学習による物体検出

3-3 変状検出への物体検出手法の適用

3-4 深層学習によるセグメンテーション

3-5 変状検出へのセグメンテーション手法の適用

3-6 まとめ



---

## 第3章 深層学習による物体検出およびセグメンテーション

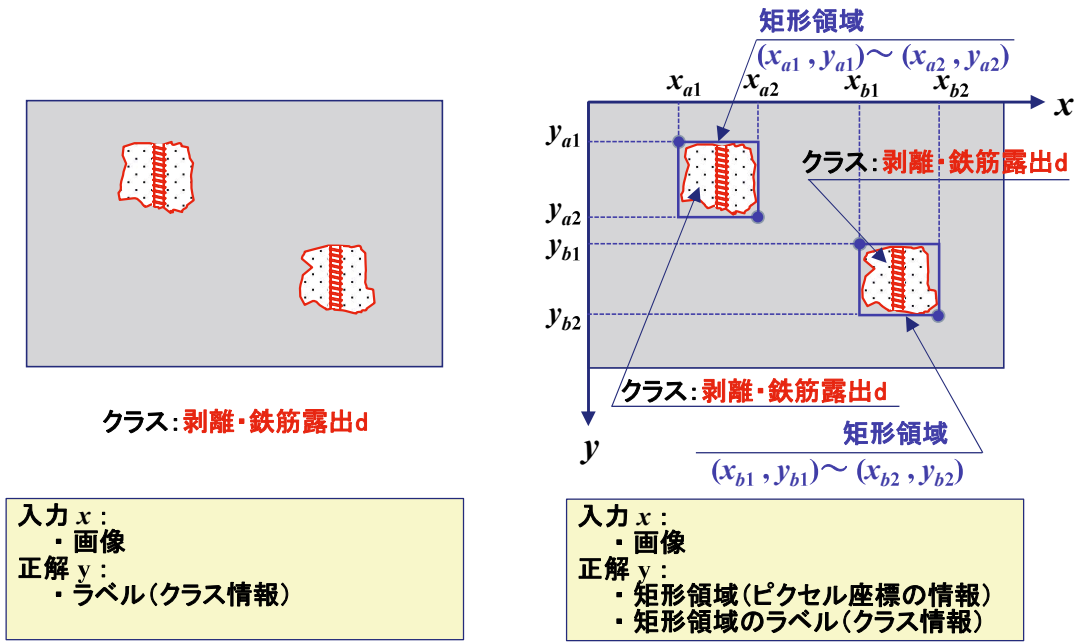
---

### 3-1 教師データの構成

第2章において、変状のクラス分類については、一定の成果が得られた。画像からのコンクリート変状の検出のためには、さらに画像内のどの部分に変状が存在しているかを特定させる必要がある。そこで、本章では、物体検出手法とセグメンテーション手法を用いて、変状の領域までを特定する「変状箇所の検出」のタスクについて検討を行う。

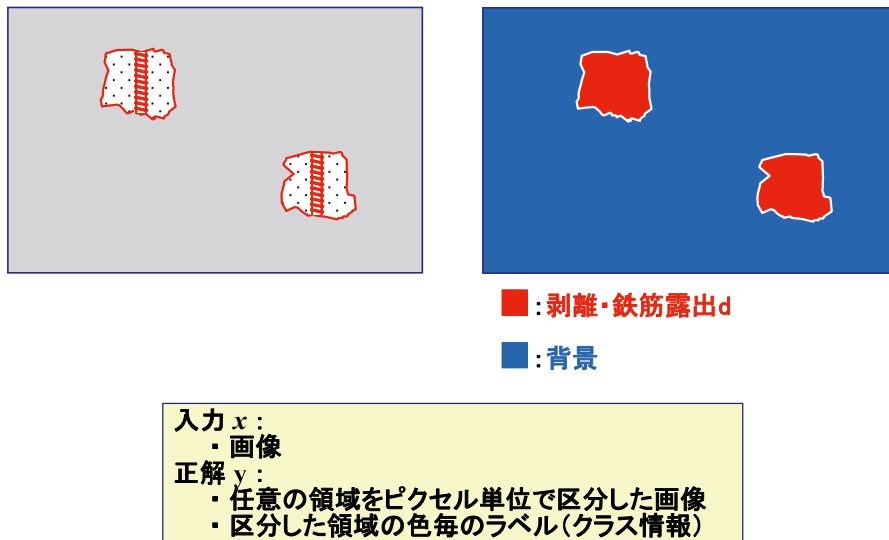
クラス分類の教師データは、入力  $x$  の画像に対して、クラス情報のラベルを正解  $y$  としたペアで構成される。そして、統計的学習により、画像  $x$  の入力に対して、クラス  $y$  を出力する関数を帰納的に推定していた。一方、物体検出の教師データは、入力  $x$  の画像に対して、矩形領域（ピクセル単位の座標情報）とその矩形領域のクラスを正解  $y$  としたペアで構成される。また、セグメンテーションの教師データは、入力  $x$  の画像に対して、任意の領域をピクセル単位で区分した画像とその区分領域の色毎のクラスを正解  $y$  としたペアで構成される。そして、統計的学習により推定する関数は、画像  $x$  の入力に対して、クラス分類と同様のクラスに加えて矩形領域あるいはピクセル単位で区分した領域を  $y$  として出力する関数となり、タスクの難易度が数段上がる。図 3-1 に、それぞれのタスクにおける教師データの構成について示す。

教師データの作成に際しては、データに対して関連する情報を注釈として付与するアノテーションと呼ばれる作業を行う必要がある。入力  $x$  の画像に対して、ペアとなる正解  $y$  の情報を注釈として付与する。深層学習の教師あり学習では、図 2-1 に示した「学習」の段階の画像の入力および関数  $f : x \rightarrow y$  の帰納的推定、「推論」の段階の画像の入力および推論結果の出力は自動的に行われるが、このアノテーションの作業は人力による作業が必要となる。アノテーションの労力は、クラス分類、物体検出、セグメンテーションと、タスクの難易度とともに増大し、特に物体検出の矩形領域の座標情報、セグメンテーションの任意領域の区分画像の作成は膨大な労力を要する。



a) 物体認識 (変状のクラス分類)

b) 物体検出 (変状箇所の検出とそのクラス分類)



c) セグメンテーション (変状の領域抽出とそのクラス分類)

図 3-1 画像認識における教師データの構成

## 3-2 深層学習による物体検出

### 3-2-1 物体検出の概要

物体検出 (object detection) とは、画像内の対象とする物体を、その物体に外接する矩形領域で囲むとともにクラスの分類を行い、画像内のどこにどのような物体が写っているかという、位置とクラスを同時に認識する問題である。この外接する矩形領域は、バウンディングボックス (bounding box) と呼ばれている。物体検出の研究は、問題設定とアプローチの観点から、「前景 (対象物体) と背景を分離する問題」、「三次元形状等の情報が既知である特定物体を検出する問題」、「検出対象のクラスに属する物体を検出する問題」の3つに大別できるとされている<sup>1)</sup>。近年の深層学習が主流となった物体検出の研究では、この3番目の「検出対象のクラスに属する物体を検出する問題」が中心的なテーマとなっており、本研究で対象とするコンクリート構造物の変状検出も、これに該当する。

物体検出の研究は、顔や人を検出対象とした研究を中心に古くから取り組まれてきた。1990年代後半以降は、クラス分類と同様に、画像から高次元の特徴量ベクトルを抽出して機械学習手法による分類を行う手法が主流となった。分類の前段となる物体候補領域の提案は、ある特定のサイズの領域を一定のピクセル単位でスライドさせて領域を提案するスライディングウィンドウ法、あらかじめ物体らしさを評価して領域の候補数を絞り込む選択的検索法等により行われた。その後2010年代以降は、深層学習による手法が主流となっている。

深層学習による物体検出の近年の潮流は、R-CNN (Regions with CNN features)<sup>2)</sup>に依るところが大きい。R-CNNは、2005年から2012年に開催された画像認識のコンペティション PASCAL VOC Challenge (Pattern Analysis, Statistical Modelling and Computational Learning Visual Object Classes Challenge)<sup>3)</sup>の画像データセットに対する2012年時点のmAP (平均適合率 AP の全クラスの平均) の最高値を30%以上向上させたことで脚光を浴びた物体検出手法である。2013年には、より大規模な画像データセットを用いる ILSVRC<sup>4)</sup>においても、物体検出のタスクに関するコンペティションが開始され、次々に新たな手法が開発されるとともに、検出の精度や速度が向上し続けている。図3-2に、ILSVRCにおける物体検出の歴代優勝者の成績について示す。

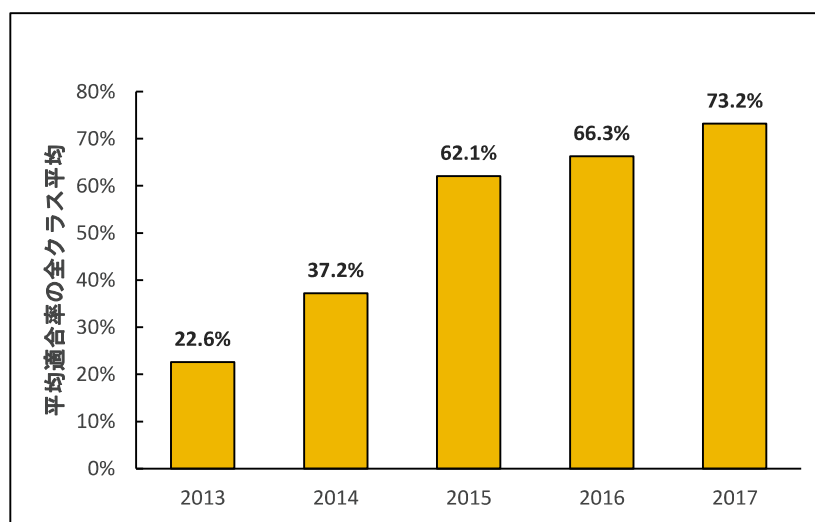


図 3-2 ILSVRC における物体検出の歴代優勝者の成績

### 3-2-2 深層学習による物体検出

深層学習による物体検出は、R-CNN に端を発し、以後多くの手法が提案されている。R-CNN の手順としては、Selective Search<sup>5)</sup>等の物体候補領域の探索手法により得られた領域の画像を、それらを一つずつ CNN に入力し、それぞれの領域について取り出した特徴量を SVM で適切なクラスに分類する工程を経て、検出が行われる。R-CNN 以後も文献<sup>6)-12)</sup>等にあるように、より高速で高精度な物体検出手法が提案されている。深層学習による物体検出の主な手法について表 3-1 に示す。

これらの手法のアプローチとしては、大きくは分類問題と回帰問題に分けられる。R-CNN, SPP net<sup>6)</sup>, Fast R-CNN<sup>7)</sup>, Faster R-CNN<sup>8)</sup>は分類問題として検出を行う手法である。候補領域を提案した後に、その候補領域が分類器を通過してクラス分類され、候補領域の補正を行って矩形領域 (bounding box) が特定される。これに対して、YOLO<sup>9)</sup>, SSD<sup>10)</sup>, YOLO v2<sup>11)</sup>, YOLO v3<sup>12)</sup>は回帰問題として検出を行う手法である。候補領域とクラスを予測した後に、矩形領域を直接生成し、それに対する補正を行って領域の特定を行う。このアプローチの特徴の一つは、候補領域だけでなく、画像全体を利用して周辺の状態も同時に学習するため、背景の誤認識を抑制できる点である。このことにより、変状周辺のコンクリート躯体面の誤認識を最小限とした、精度の高い検出が期待できる。

その後も、今なお新しい手法は次々と開発されており、有名なところでは、M2Det<sup>13)</sup>, EfficientDet<sup>14)</sup>, YOLO v4<sup>15)</sup>等が発表されている。

表 3-1 深層学習による物体検出の主な手法

	深層学習の内容			物体検出の工程			備考
	入力	使用するネットワーク	出力	物体候補領域の抽出	クラスの分類	領域の特定	
[1] R-CNN <sup>2)</sup>	個々の領域画像	CNN	CNN特徴量	Selective Search <sup>5)</sup> 等の領域探索手法により、クラスに依存しない候補領域を提案。	CNNにより個々の候補領域内の画像の特徴を抽出し、それを分類器(線形SVM)に入力してクラスを予測	CNN特徴量から矩形領域のパラメータ(中心位置, 幅, 高さ)へ回帰	各々の候補領域(約2000個)に対して個別にCNNによる特徴抽出を行うため、速度が遅い。
[2] SPP-net <sup>6)</sup>	画像全体	CNN	CNN特徴量	Selective Search <sup>5)</sup> 等の領域探索手法により、クラスに依存しない候補領域を提案。	CNNにより画像全体の複数スケールの特徴量を抽出し、各領域候補の特徴をSPPによってベクトル化し、最適なスケールの特徴量を分類器(線形SVM)に入力してクラスを予測	CNN特徴量から矩形領域のパラメータ(中心位置, 幅, 高さ)へ回帰	画像全体に対してCNNによる特徴抽出を行うことで、[1]における解析時間の課題を解消。
[3] Fast R-CNN <sup>7)</sup>	・画像全体 ・候補領域	CNN	CNN特徴量	Selective Search <sup>5)</sup> 等の領域探索手法により、クラスに依存しない候補領域を提案。	ROI(Region of Interest;関心領域)プーリングによって候補領域毎に異なるCNN特徴量を固定長の特徴量に変換した上で、FCN(Fully Convolutional Network;完全畳み込みネットワーク)に入力して学習によりマルチタスク損失を最小化させてクラス分類と矩形領域への回帰を同時に行う。		画像全体に対してCNNによる特徴抽出を行うことで、[1]における解析時間の課題を解消。
[4] Faster R-CNN <sup>8)</sup>	画像全体 CNN特徴量 CNN特徴量 候補領域 スコア	CNN RPN FCN	CNN特徴量 候補領域 スコア 候補領域 スコア	RPN(Region Proposal Network;領域提案ネットワーク)での学習により、クラスに依存しない候補領域を提案。	ROI(Region of Interest;関心領域)プーリングによって候補領域毎に異なるCNN特徴量を固定長の特徴量に変換した上で、FCN(Fully Convolutional Network;完全畳み込みネットワーク)に入力して学習によりマルチタスク損失を最小化させてクラス分類と矩形領域への回帰を同時に行う。		候補領域の抽出も深層学習により行うことで、[1], [2], [3]の課題である「候補領域の冗長的な抽出」「項目毎の不連続性」を解消し、速度と精度の向上を実現。
[5] YOLO <sup>9)</sup>	画像全体	CNN	クラス 矩形領域	---	画像を分割した各グリッドに対して、学習により領域と信頼度(候補領域の精度と物体の有無を反映)、そしてクラスを予測。領域の予測とクラスの予測を結合し、閾値処理等を行ってクラス分類と矩形領域を確定。		[4]に対して速度が大幅に向上したが、精度が若干低下している。
[6] SSD <sup>10)</sup>	画像全体	FCN	クラス 矩形領域	---	学習によりネットワークの中間の各段階で、各々のスケールのCNN特徴マップに対して領域とクラスの予測を行い、各領域で生成する領域内の物体の存在に関するスコアにより物体の形にマッチするように領域形状の調整を行って領域を確定。		[5]より高速で、[4]より高精度な性能を実現。
[7] YOLO v2 <sup>11)</sup>	画像全体	FCN	クラス 矩形領域	---	基本的に[5]を踏襲。 [5]に対して、データセットの矩形を事前にクラスタリングしての学習、マルチスケールトレーニング等により性能を改善。		性能改善により、[4]の精度および[6]の速度を上回る性能を実現。
[8] YOLO v3 <sup>12)</sup>	画像全体	FCN	クラス 矩形領域	---	基本的に[7]を踏襲。 [7]に対して、Softmax関数からロジスティック回帰に変更しての分類、3つの異なるスケールでの領域の予測等により性能を改善。		性能改善により、[7]よりも速度が低下したが、精度の向上を実現。

: 深層学習を用いている工程

### 3-3 変状検出への物体検出手法の適用

#### 3-3-1 検討条件

##### (1) 適用手法

物体検出の適用手法について、分類と回帰の二つのアプローチについて比較評価を行う。分類問題として検出を行う手法についてはFaster R-CNNを、回帰問題として検出を行う手法についてはYOLO v3をそれぞれ代表とし、これら二つの手法を用いて変状検出の検討を行うこととした。学習に際しては、変状のクラス分類において得られた知見より、双方の手法ともに、データ拡張および学習済みモデルから得られる結合重みの転用を行う。

Faster R-CNN は、R-CNN と Fast R-CNN と同様に、物体の候補領域の探索の後に、候補領域のクラス分類を行う 2 段階の検出手法であるが、図 3-3 に示す RPN (Region Proposal Network ; 領域提案ネットワーク) の導入により、クラス分類だけでなく物体候補領域の探索も深層学習で行う点が、先行の 2 手法とは異なる。入力画像全体を CNN に入力して得た特徴量である特徴マップを RPN に入力した後に、k 個の固定枠 (Anchor) をラスタスキャンさせて物体の候補領域を探索する。画像全体の特徴マップと RPN から得た候補領域の特徴量 (物体らしさのスコアと画像上の検出座標) を FCN (Fully Convolutional Network ; 完全畳み込みネットワーク) に入力し、分類クラスとそのクラスである確率および矩形領域を獲得する。

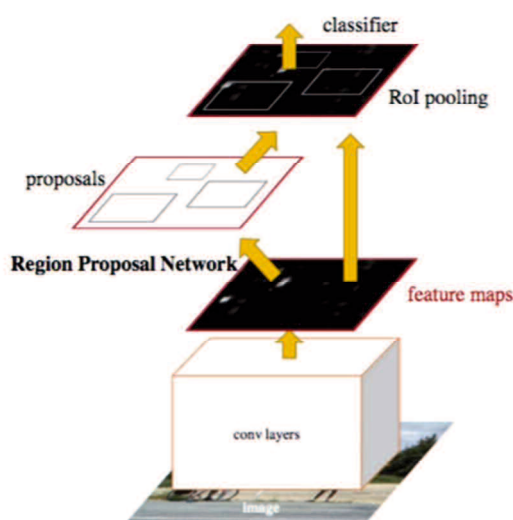
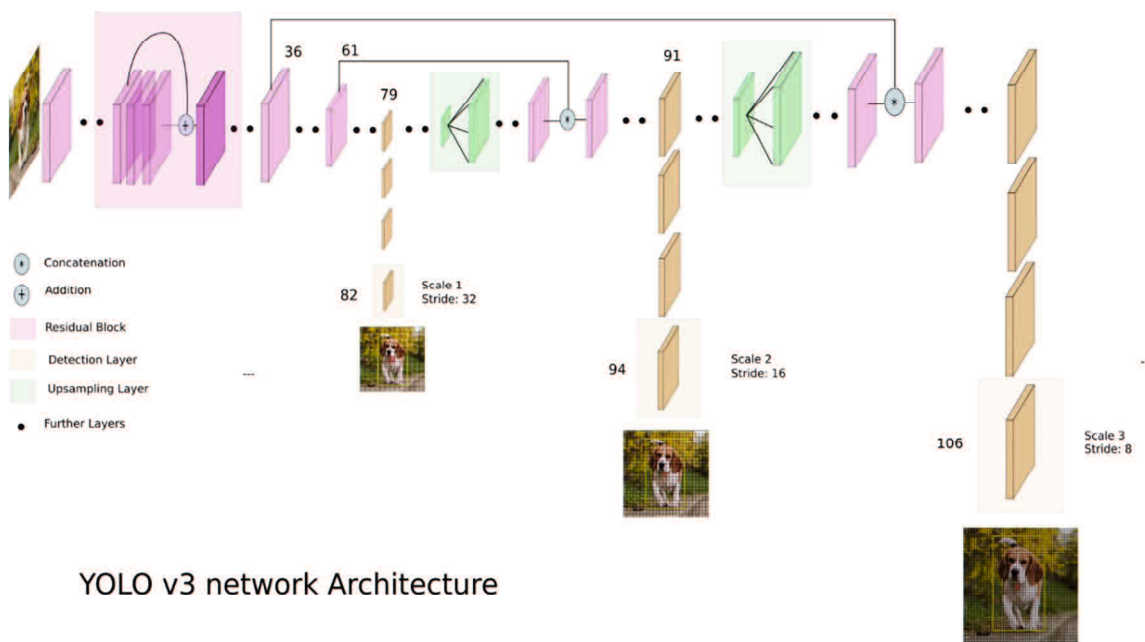


図 3-3 Faster R-CNN の手法概要図 <sup>16)</sup>



YOLO v3 は、物体の候補領域の探索と候補領域のクラス分類を1段階で同時に行う検出手法である。入力した画像をグリッドセルに分割し、各セル中心を中心とする物体候補領域（Anchorbox）を配置し、候補領域の特徴量（物体らしさのスコアと画像上の検出座標）から候補領域の補正を行う。物体を含むそれぞれの Anchorbox において、ロジスティック回帰を使用して各クラススコアを予測して、最終的に、分類クラスとそのクラスである確率および矩形領域を獲得する。YOLO は、大まかに定義したグリッドに沿って検出を行うために物体のスケールの変化に弱いという欠点があったが、YOLO v2 からの主要改善点の一つ、3つの異なるスケールでの検出により、YOLO v3 では小さい物体にも対応が可能となっている。YOLO v3 の手法概要図を図 3-4 に示す。



YOLO v3 network Architecture

図 3-4 YOLO v3 の手法概要図 <sup>17)</sup>

## (2) 教師データ

物体検出における教師データでは、クラス分類時の教師データである画像と正解ラベルに加えて、物体が存在する矩形領域の座標情報が必要となる。変状箇所の検出の検証に際して、変状のクラス分類に用いたデータセットと同じ画像に対して、あらかじめ道路橋点検士有資格者により、正解の矩形領域（Ground-truth bounding box）と個々の領域についての変状の正解クラスの設定を行った。変状が細長い領域で斜め方向に分布する

ような場合には、正解の矩形領域を複数に分割して変状が矩形領域内の大半を占める状況となるように設定を行った。データ拡張は、画像中心の鉛直方向軸に対して左右の反転を行って、2倍に拡張した。

#### (3) テストデータ

比較は、図 3-5 に示すテストデータに対する検出結果により評価を行う。画像は、クラス分類と同様の画像であり、画像中の赤枠が正解の矩形領域である。また、入力に対する出力は、クラス分類時は予測クラスとその判定確率であったが、物体検出では、それらに予測の矩形領域 (Predicted bounding box) が加わる。



図 3-5 テストデータ

#### 3-3-2 検出結果

Faster R-CNNとYOLO v3による変状検出結果は、正解クラスと予測クラスとの正誤、予測クラス判定確率および正解の矩形領域と予測の矩形領域の2つの領域の重なり具合により対比を行った。クラスの判定に際しては、CNNによるクラス分類と同様に、8クラスのそれぞれに対する判定確率が算定されるが、その判定確率が最大となるクラスが予測クラスである。2つの領域の重なり具合は、IoU (Intersection over Union) という図3-6に示す位置特定の評価指標を算定することにより行った。通常、IoUが0.5以上であれば、良好な予測が行われたと判断することが多い。

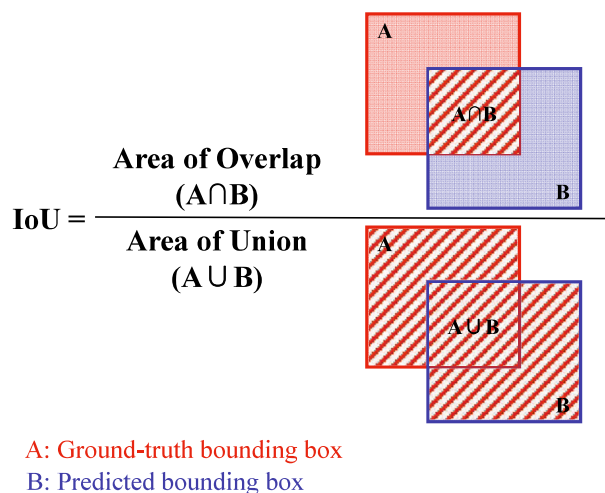


図 3-6 IoU (Intersection over Union) の考え方

Faster R-CNNとYOLO v3によるコンクリート変状検出結果を図3-7に示す。変状の検出を行った結果、いずれのテストデータにおいても、Faster R-CNNとYOLO v3の双方の手法で、予測クラスが正解クラスと一致した。今回用いたテストデータに対しては、検出漏れ (False Negative) は生じなかったが、漏水・遊離石灰cの画像に対するFaster R-CNNによる検出において、主桁側面の汚れと考えられる部分を漏水・遊離石灰dとする誤検出 (False Positive) があつた。正解クラスと予測クラスとの正誤、予測クラス判定確率については、手法間で大きな差は生じなかったが、8種類の変状のうち3種の変状でYOLO v3の予測クラス判定確率の方が5%以上高い数値となり、YOLO v3の方が若干優位な結果であった。

IoUについては、漏水・遊離石灰cのみFaster R-CNNの方が高い値で、漏水・遊離石灰e-1と剥離・鉄筋露出dは同等の値となったが、それら以外の5種の変状については、YOLO v3がFaster R-CNNを大きく上回つた。IoUに差が生じた5種の変状の検出結果のうち剥離・鉄筋露出cを除く4種の変状では、Faster R-CNNによる予測の矩形領域は、正解の矩形領域に対して狭小な範囲となっている。ひびわれについては、幅の細くなる画像の上方部分、漏水・遊離石灰e-2については漏水部分が細くなつていゝ変状の上方部分が検出できていないことにより、IoUの値に差が生じている。漏水・遊離石灰e-1については、細長い領域で斜め方向に分布している変状に対して、変状が矩形領域内の大半を占めるように正解の矩形領域を複数に分割して設定していた。そのため、双方の手法ともに広域的に検出を行つてIoUの評価としては低調であつたが、Faster R-CNNはつらら状の遊離石灰

のサイズが小さくなる画像の右下部分を、同様に検出できていない。YOLO v3の主要な改善点の一つとして、3つの異なるスケールで検出を行う点が挙げられる。このことにより、より小さな物体の検出の精度の向上が図られていることから、YOLO v3の矩形領域の特定の精度が優位になったと考えられる。

また、人や顔および車等の一般的によく行われる物体検出では、背景と検出対象が明らかに異なるが、コンクリート変状の検出を行う場合は、背景であるコンクリート躯体面と検出対象である変状に外観上の極端な違いはない。そのため、分類と回帰とで、背景の誤認識に関して優劣が生じるのではないかと考えていたが、双方の手法ともにコンクリート躯体面に対する誤認識は無かった。

変状のクラス分類では、アップで撮影した単一のクラスの変状のみが存在する画像を対象に、該当するクラスの分類を行った。本研究では、変状のクラス分類と同様の画像に対して変状箇所の検出を行ったが、実際の点検業務への適用に際しては、広域的な画像内に混在する複数の変状について、それぞれの領域と該当するクラスを適切に予測できなければならない。この点からも、多数の物体の検出に対応が可能で、3つの異なるスケールで検出を行うことができるYOLO v3は、変状の検出に対する適性が高いと考えられる。しかしながら、物体検出は難易度が高いタスクであるため、多様なデータを教師データおよびテストデータに追加して、検出をより頑健にするための検討を重ねていくことが肝要である。

		Faster R-CNN		YOLO v3	
ひびわれ	出力画像				
	予測クラス	ひびわれ	正誤	ひびわれ	正誤
	IoU	判定確率 90.5%	○	判定確率 98.5%	○
		0.77		0.87	
剥離・鉄筋露出 c	出力画像				
	予測クラス	剥離鉄露c	正誤	剥離鉄露c	正誤
	IoU	判定確率 99.4%	○	判定確率 99.5%	○
		0.80		0.93	
剥離・鉄筋露出 d	出力画像				
	予測クラス	剥離鉄露d	正誤	剥離鉄露d	正誤
	IoU	判定確率 100.0%	○	判定確率 99.5%	○
		0.89		0.88	
剥離・鉄筋露出 e	出力画像				
	予測クラス	剥離鉄露e	正誤	剥離鉄露e	正誤
	IoU	判定確率 100.0%	○	判定確率 99.8%	○
		0.53		0.67	
漏水・遊離石灰 c	出力画像				
	予測クラス	漏水遊離c	正誤	漏水遊離c	正誤
	IoU	判定確率 96.0%	○	判定確率 98.1%	○
		0.78		0.46	
漏水・遊離石灰 d	出力画像				
	予測クラス	漏水遊離d	正誤	漏水遊離d	正誤
	IoU	判定確率 90.4%	○	判定確率 97.0%	○
		0.57		0.92	
漏水・遊離石灰 e-1 つらら	出力画像				
	予測クラス	漏水遊離e-1	正誤	漏水遊離e-1	正誤
	IoU	判定確率 97.2%	○	判定確率 97.1%	○
		0.43		0.45	
漏水・遊離石灰 e-2 錆汁	出力画像				
	予測クラス	漏水遊離e-2	正誤	漏水遊離e-2	正誤
	IoU	判定確率 94.1%	○	判定確率 99.1%	○
		0.49		0.75	

□ : Ground-truth bounding box (正解の矩形領域), □ : Predict bounding box (予測の矩形領域)

図 3-7 Faster R-CNN と YOLO v3 によるコンクリート変状検出結果

### 3-4 深層学習によるセグメンテーション

#### 3-4-1 セグメンテーションの概要

セグメンテーション (segmentation) とは、特徴や意味が類似した部分領域に画像を分割する処理である。近年の画像処理では、画像内の対象とする物体と背景の領域を分割してピクセル単位で認識するとともにクラスの分類を行い、画像内のどこにどのような物体が写っているかという、領域とクラスを同時に認識するタスクを実質的に意味する。

セグメンテーションの手法については、古くから種々の処理手法が提案されてきた。1990年代頃までは、二値化処理やエッジ抽出処理等の基本的な画像処理技術が用いられてきたが、1990年代後半以降は統計的分類器やグラフカット<sup>18)</sup>等の最適化理論に基づく処理が主流となった。そして、ILSVRCでCNNが注目を集めた2012年、セグメンテーションでもCNNにより得られた特徴マップを用いる手法<sup>19)</sup>が提案され、以降は深層学習による手法が主流となっている。難易度の高いタスクであるセグメンテーションは、高い精度を実現するには時間を要すると考えられていたが、深層学習による手法が提案されて以降、他のタスクと同様に格段の精度の向上を実現している。

#### 3-4-2 セグメンテーションの分類

セグメンテーションは、セマンティックセグメンテーション (semantic segmentation)、インスタンスセグメンテーション (instance segmentation)、パノプティックセグメンテーション (panoptic segmentation) の3つに分類される。深層学習を用いた手法によりセマンティックセグメンテーションを高精度で実現できるようになって以降、より複雑なインスタンスセグメンテーションに関する研究も盛んになった。また、2018年に、セマンティックセグメンテーションとインスタンスセグメンテーションを統合した新たなタスクとして、パノプティックセグメンテーションが導入された。

セマンティックセグメンテーションは、物体のクラスごとに領域分割を行うが、同じクラスの物体が隣接して存在する場合、それらを個別には認識できない(図 3-8(b)参照)。それに対して、インスタンスセグメンテーションは、各物体インスタンスを検出してセグメンテーションするタスクで、同じクラスの物体が隣接して存在する場合でも、物体の1つずつを境界まで切り分けて、それらを個体ごとに領域分割する(図 3-8(c)参照)。また、インスタンスセグメンテーションは、対象とする物体に対してのみ領域の検出を行うが、セマンティックセグメンテーションは、対象の物体以外の背景に対しても不定形の領域の検出を行う。そして、パノプティックセグメンテーションは、インスタンス

セグメンテーションと同様に個体ごとの領域分割を行った上で、セマンティックセグメンテーションと同様に空や道路等の物体以外の背景に対しても不定形の領域の検出を行う(図3-8(d)参照)。セマンティックセグメンテーションは、画像の各ピクセルへの意味的なクラスの割り当てを目的としているのに対して、インスタンスセグメンテーションは、同一のクラスにおける異なるインスタンスの区別を目的としており、同じセグメンテーションであってもアプローチの異なるタスクである。

コンクリートの変状の検出では、異なるクラスの変状が隣接して存在する場合は、それらを個別に領域分割して認識する必要があるが、同じクラスの変状が隣接するピクセルに存在していても、個体ごとの境界は不明瞭であり、個別に認識することの必要性は低い。そのため、人認識や、車の自動運転への適用をイメージした交通シーンのような、インスタンスセグメンテーションにより個体ごとの領域分割を行う必要性が高いタスクとは動機が異なる。以上のことより、コンクリートの変状の検出は、セマンティックセグメンテーションによる手法を基本として検討を行うこととした。

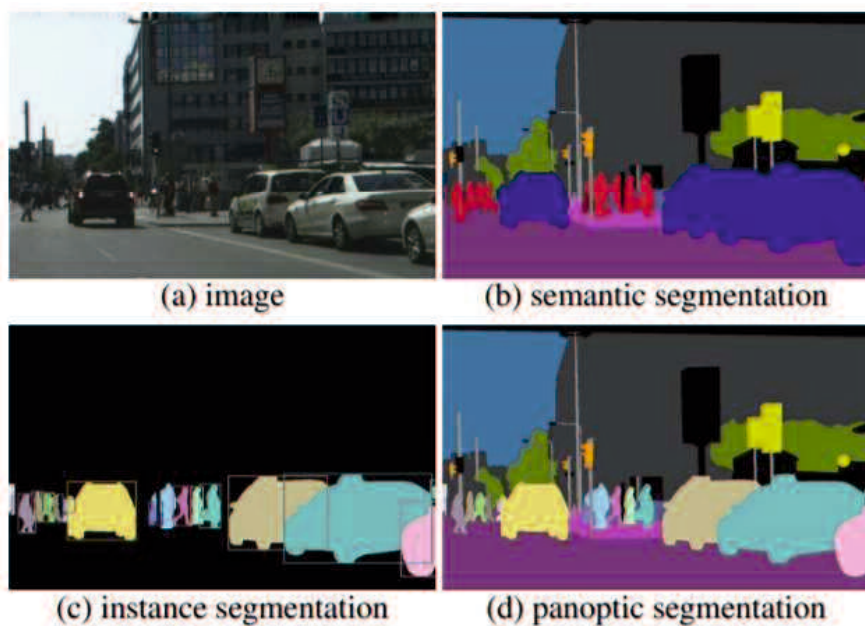


図 3-8 セグメンテーションの分類<sup>20)</sup>

### 3-5 変状検出へのセグメンテーション手法の適用

#### 3-5-1 適用手法の選定

##### (1) 比較評価を行う手法

文献<sup>21)</sup>では、セマンティックセグメンテーション手法を、アプローチの方針等により年代順に「深層学習以前」、「深層学習初期」、「FCN 以後」、の3つに大きく分類している。

FCN (Fully Convolutional Network)<sup>22)</sup>は、セマンティックセグメンテーションの技術的な分岐点となった手法で、革新的な点が3つある。1点目は全結合層が存在しないネットワーク構成とした点で、入力画像のサイズを固定する制約がなくなったため任意の画像サイズへの対応が可能となった。

2点目は、畳み込み層の出力を任意の解像度の高密度ピクセルにアップサンプリングして畳み込み処理を行う逆畳み込み層 (deconvolutional layer) の採用である。アップサンプリングとは、元画像のサイズでセグメンテーション結果を出力するために、畳み込み層とプーリング層で縮小した特徴マップの拡大を行う処理である。通常の畳み込みは、プーリングによるダウンサンプリングの前に行われるが、アップサンプリングの後に畳み込みを行うことから、逆畳み込みと呼ばれている。

3点目は、畳み込みの繰り返しにより局所的な特徴を抽出できる反面、プーリングの繰り返しにより全体的な位置情報が失われていくという CNN の性質に対して、スキップ接続を採用した点である。プーリングを行うことによって、位置に関する変動を吸収して小さなフィルターで大域的な特徴を取れるようになり受容野が広がるが、位置に関

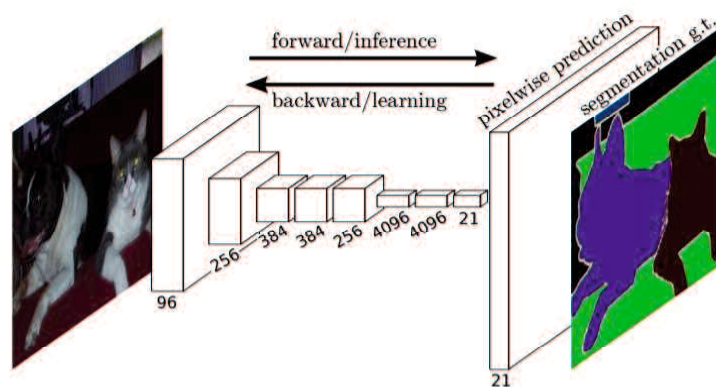


図 3-9 FCN のネットワーク構造<sup>22)</sup>



する詳しい情報を失ってしまう。そのため、入力層に近い層では鮮明であった位置情報が、出力層に近い層では曖昧になっていく。プーリング層による位置情報の不明瞭化は、クラス分類では物体の位置やサイズの影響をあまり受けない頑健なパターン認識に寄与するが、セグメンテーションではピクセル単位での領域抽出の能力の低下につながる。そこで、接続をスキップしてプーリング層をバイパスさせることで、深い層へ全体的位置情報を伝搬させ、最終層で統合する処理を行っている。

これらの3つの点は、FCN 以後においても、多くの手法で継承された。本研究で行う適用手法の比較評価は、FCN 以後の手法の中から、特徴的な手法である SegNet<sup>23)</sup>、UNet<sup>24)</sup>、PSPNet<sup>25)</sup>の3手法を対象として行う。以下に各手法の概要について記す。

#### a) SegNet

SegNet の特徴的な点は、Encoder-Decoder 構造の採用である。Encoder-Decoder 構造は、入力した画像を Encoder（符号化器）により低次元の中間表現に変換した後に、Decoder（復号化器）により入力画像に何らかの処理を施した画像を生成するネットワークである。SegNet は、Encoder ネットワークとそれに対応する Decoder ネットワーク、最後に各ピクセルのクラスを予測する分類層を経て最終的なセグメンテーションが生成される構成になっている（図 3-10 参照）。

Encoder では、畳み込み層による画像の局所特徴の抽出と、プーリング層による最大プーリングでのダウンサンプリングが実行され、空間的な次元を徐々に減少させて情報を圧縮したスケール不変な特徴を後続の層に伝播していく。

Decoder では、Encoder によって得られた物体の種類と大まかな位置情報を持つ低解像度の特徴マップに対して、物体の詳細と空間的な次元を徐々に復元し、元の解像度の特徴マップにおける個々の物体が対応するピクセル値への対応付けを行う。具体的には、低解像度の特徴マップをアップサンプリングし、その後に畳み込み処理を行うことによって、物体の形をより明瞭に補完してセグメンテーション結果の出力を行う。

なお、SegNet のスキップ接続では、Encoder での最大プーリング中に保存された最大プーリングインデックス（どのピクセルが最大値であったかの位置情報）が、Decoder に転送され、アップサンプリングに反映される。FCN では、中間層の特徴マップ全体を保存して転送していたため、メモリの使用量が大きかったが、SegNet では最大プーリングインデックスのみの転送とすることにより、高速化および省メモリ化を図っている。

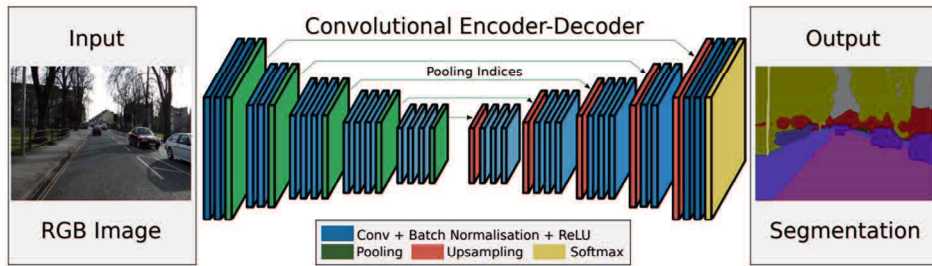


図 3-10 SegNet のネットワーク構造 <sup>23)</sup>

b) U-Net

U-Net も SegNet と同様の Encoder-Decoder 構造である。SegNet のスキップ接続では、最大プーリングインデックスを転送して位置に関する情報を補充しているのに対して、U-Net では特徴マップ全体を転送して位置に関する情報を補充している。特徴マップ全体を転送するため、U-Net は SegNet に比べてモデルが大きくなり、より多くのメモリを必要とする。また、FCN のスキップ接続では、違うサイズの層の位置情報を最終的な特徴マップに統合していたのに対して、U-Net は、Encoder の各中間層で出力される特徴マップを、対応するサイズの Decoder の層の特徴マップに直接連結している。(図 3-11 参照)

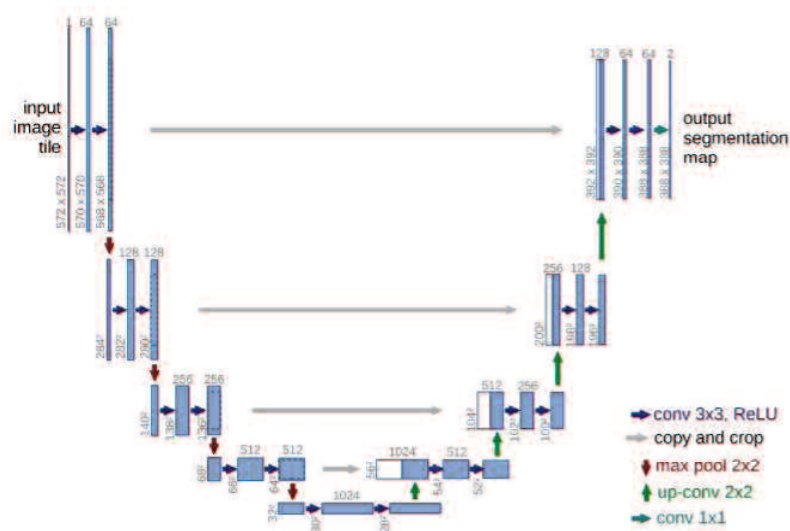


図 3-11 U-Net のネットワーク構造 <sup>24)</sup>

c) PSPNet

PSPNet (Pyramid Scene Parsing Network) も、前述の2手法と同じ Encoder-Decoder 構造である。特徴は、解像度を落とさずに受容野を広げるための工夫として、スキップ接続ではなく Dilated 畳み込み<sup>26)</sup>を用いている点と、Encoder と Decoder の間に PPM (Pyramid Pooling Module) を追加している点である。

セグメンテーションでは、脈絡的な情報が得られる大域的特徴も重要な要素となるため、プーリングによるダウンサンプリングを行って CNN の受容野を広げるとともに、特徴マップの解像度を低下させない処理が必要となる。プーリングでは、空間的に情報を要約していくことで受容野を広げて広範囲の情報を取り込むことができるようになるが、解像度は低下し全体的位置情報も曖昧となる。そのため、前述の2手法では、スキップ接続を用いて入力層に近い層の特徴を再利用できるようにすることで、失われた解像度と全体的位置情報を回復させている。一方、PSPNet では、Dilated 畳み込みを用いている。Dilated 畳み込みは、フィルタの間隔が広いことから受容野を広げることが可能で、プーリングを必要としないことから解像度は下がる。この Dilated 畳み込みで抽出した特徴マップが、PPM に入力される。

PPM では、Encoder で抽出した特徴マップに対して、複数のスケールで最大プーリングを行ってそれぞれのスケールで捉えた特徴マップを得る。PPM で得た複数の特徴マップは、畳み込み処理を行った後に Encoder で得た元の特徴マップと同一のサイズにアップサンプリングされ、連結層で元の特徴マップと連結し、大域的な情報と局所的な情報の両方を持った特徴マップとなる。最終的に、この連結した特徴マップに対して畳み込みを行ってセマンティックセグメンテーションの結果を得る。

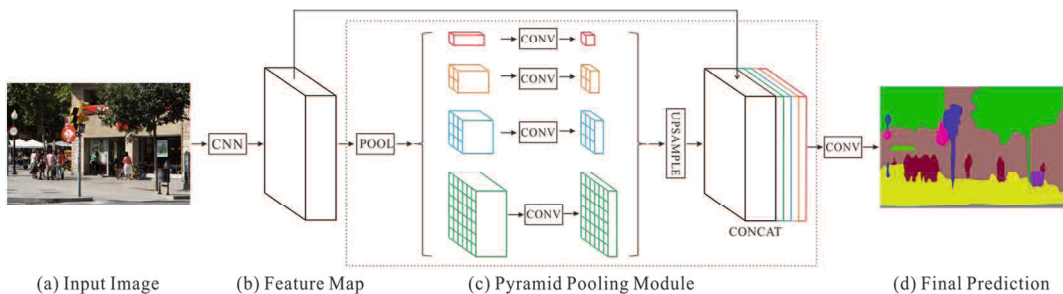


図 3-12 PSPNet のネットワーク構造<sup>25)</sup>

(2) 検討条件

a) 教師データ



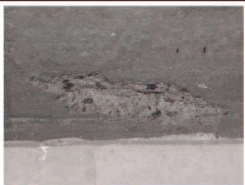


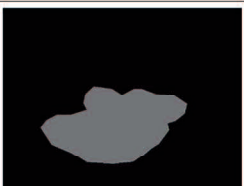



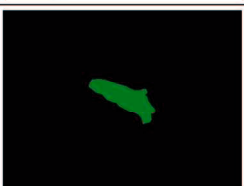





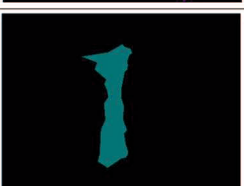
セグメンテーションの教師データは，入力  $x$  の画像に対して，任意の領域をピクセル単位で区分した画像とその区分領域の色毎のクラスを正解  $y$  としたペアで構成される．データセットのうち，約 20%を比較評価のためのテストデータとした．各クラスのデータ数を表 3-2 に示す．

区分の領域を分割した正解画像は，labelme<sup>27)</sup>というインターネット上で公開されているツールを用いて作成を行った．各クラスの変状部分以外は，背景 1 クラスのみとして，背景の内容に応じたクラスの細分化は行っていない．教師データの例を表 3-3 に示す．なお，画像の前処理として，入力画像に対する正規化処理，入力画像および正解画像ともに 240×240 ピクセルへの画像サイズの統一を行っている．

表 3-2 各クラスのデータ数

クラス	データ数		
	教師データ	テストデータ	計
ひびわれ	120	30	150
剥離・鉄筋露出c	120	30	150
剥離・鉄筋露出d	120	30	150
剥離・鉄筋露出e	120	30	150
漏水・遊離石灰c	120	30	150
漏水・遊離石灰d	120	30	150
漏水・遊離石灰e-1 つらら	120	30	150
漏水・遊離石灰e-2 錆汁	120	30	150

表 3-3 教師データの例

クラス	入力画像	正解画像	正解画像のRGB値[R,G,B]	
			変状	背景
ひびわれ			[128, 0, 0]	[ 0, 0, 0]
剥離・鉄筋露出c			[ 0, 0,128]	[ 0, 0, 0]
剥離・鉄筋露出d			[128,128,128]	[ 0, 0, 0]
剥離・鉄筋露出e			[128, 64, 0]	[ 0, 0, 0]
漏水・遊離石灰c			[ 0,128, 0]	[ 0, 0, 0]
漏水・遊離石灰d			[128,128, 0]	[ 0, 0, 0]
漏水・遊離石灰e-1 つらら			[128, 0,128]	[ 0, 0, 0]
漏水・遊離石灰e-2 錆汁			[ 0,128,128]	[ 0, 0, 0]

b) データ拡張

教師データについては、十分なデータ量ではないため、学習に際してデータ拡張を行った。データ拡張の有効性を検証するために、一切の処理を行わず元々の画像データ数で行った「未処理」、各種画像処理によるデータ拡張を行った「データ拡張(1)」、さらに画像枚数はデータ拡張(1)と同様で、ランダムにガンマ変換で画像の明るさを変化させた「データ拡張(2)」の3ケースで比較を行った。データ拡張の処理内容を表 3-4 に示す。

画像処理の内容は、回転については-30度～+30度の範囲、平行移動については画像の縦横のサイズの割合でそれぞれ0.0～0.3の範囲、拡大・縮小については倍率0.7～1.3の範囲で、それぞれランダムに2パターンずつ変化させた画像を作成した。データ拡張(2)のガンマ変換は、0.5～1.5の範囲でランダムに画像の明るさを変える処理を行った。以上の処理で元画像+6パターンとしたものを、左右反転させて14倍に拡張を行った。なお、これらの画像処理において、画像内に空白部分が発生した場合は、補間を行わずにそのままの状態とした。データ拡張の実施例を図 3-13 に示す。

表 3-4 データ拡張の処理内容

	処理内容
未処理	処理なし (既存のデータ数)
データ拡張(1)	(元画像+回転2パターン+平行移動2パターン+拡大・縮小2パターン) ×左右反転 → 14倍
データ拡張(2)	(元画像+回転2パターン+平行移動2パターン+拡大・縮小2パターン) ×左右反転 → ガンマ変換 → 14倍

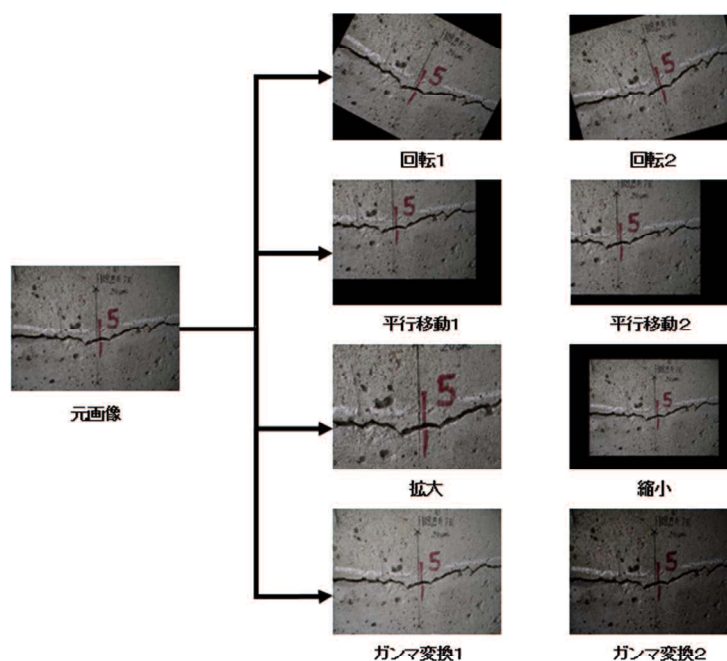


図 3-13 データ拡張の実施例

c) その他

学習方法は、第2章で述べた K 分割交差検証を用いることとし、 $k=3$  として教師データの  $2/3$  を学習に、 $1/3$  を検証データとして、3回の交差検証を行って平均によりパラメータを決定する。最終的に決定したパラメータを用いたモデルにより、テストデータに対する推論を行う。

検討ケースは、剥離・鉄筋露出および漏水・遊離石灰について、変状の程度のランクも含めて適切な分類が可能かという点に着目するために、表 3-5 に示すように変状の種類別に行った。CASE1 は「ひびわれ」を、CASE2 は「剥離・鉄筋露出」を、CASE3 は「漏水・遊離石灰」を対象としている。

表 3-5 検討ケース

クラス	CASE1 (2クラス分類)	CASE2 (4クラス分類)	CASE3 (5クラス分類)
ひびわれ	●		
剥離・鉄筋露出c		●	
剥離・鉄筋露出d		●	
剥離・鉄筋露出e		●	
漏水・遊離石灰c			●
漏水・遊離石灰d			●
漏水・遊離石灰e-1 つらら			●
漏水・遊離石灰e-2 錆汁			●
背景	●	●	●

比較評価を行う各手法の主要なパラメータは、表 3-6 に示す内容で設定を行った。平等な条件の設定が困難であったため、Fine-tuning は行っていない。

表 3-6 各手法の主要なパラメータ

項目	SegNet	U-Net	PSPNet
Epoch数	30	30	30
最適化アルゴリズム	AdaDelta	Adam	MomentumSGD
損失関数	交差エントロピー誤差	交差エントロピー誤差	交差エントロピー誤差
活性化関数(最終層)	Softmax関数	Softmax関数	Softmax関数

### (3) 適用手法の比較

#### a) 評価指標

セグメンテーション結果についての評価は、Mean IoU (Mean Intersection over Union) を用いて行われるのが通例である。IoU は、式(1)で表されるオーバーラップ率で、ピクセル単位の分類予測の正解率(Accuracy)を意味する。個々の画像における IoU をクラス毎に平均した Average IoU を、さらに全クラスで平均した値が Mean IoU となる。ただし、本研究においては、分類するクラス数が少なく、高い値となる背景についての IoU が平均値を算定する際に支配的となるため、Mean IoU で評価を行うと変状領域の検出結果についての評価が不明瞭となる。そのため、各変状の IoU に着目して評価を行うこととした。

$$IoU = \frac{TP}{(TP + FP + FN)} \quad (1)$$

ここで、 $TP$ : 「対象とする変状」のクラスのピクセルを、正しく「対象とする変状」のクラスに分類したピクセル数

$FP$ : 「対象とする変状以外」のクラスのピクセルを、誤って「対象とする変状」のクラスに分類したピクセル数

$FN$ : 「対象とする変状」のクラスのピクセルを、誤って「対象とする変状以外」のクラスに分類したピクセル数



## b) 評価結果

各手法の各変状における比較結果を表 3-7～3-9 に、各手法の 2 ケースのデータ拡張の最大値の比較結果を表 3-10 に示す。Mean IoU は、いずれの変状においても、U-Net が最も高い結果となった。個別の変状に着目すると、漏水・遊離石灰の c および d のみ PSPNet が最も IoU が高い結果となったが、それ以外の変状は U-Net が最も高い結果となった。各手法のデータ拡張(2)の条件でのセグメンテーション結果の例を図 3-14 に示す。IoU の数値だけでなく、コンクリートという一様な領域に対して境界が不明瞭な変状領域を捉えるという本タスクにおいて、U-Net の領域の表現力の優位性が見て取れる。

データ拡張は、表 3-7～3-9 に示すように、PSPNet のひびわれ、SegNet の剥離・鉄筋露出 e、SegNet および U-Net の漏水・遊離石灰 e-2 以外については、IoU の飛躍的な向上が見られた。U-Net の教師データ条件別のセグメンテーション結果の例を示した図 3-15 においても、その効果が明確に現れている。平行移動、回転、拡大・縮小および左右反転により位置情報を変化させたことが、位置情報が重要な要素となるセマンティックセグメンテーションにおいて、汎化性能の向上につながったものと考えられる。また、テストデータに用いた画像は、撮影条件が多様で、画質や明度等がさまざまであることから、ガンマ変換で教師画像の明るさを変化させることが有効となると考えたが、データ拡張(1)とデータ拡張(2)とでは、手法あるいは対象の変状により優劣が異なり、明確な傾向は見られなかった。全ケースで Mean IoU が最高となった U-Net においては、ひびわれではデータ拡張(1)、剥離・鉄筋露出ではデータ拡張(2)、e-2 を除く漏水・遊離石灰ではデータ拡張(2)が、それぞれ IoU が高い結果となった。教師データに対して、入力の前処理として正規化処理を行い、学習中はバッチ正規化を行っているため、ガンマ変換の有無が結果にあまり影響しなかったものと思われる。

表 3-7 CASE1（ひびわれ）についての比較評価結果

手法	項目	IoU		
		未処理	データ拡張 (1)	データ拡張 (2)
SegNet	背景	0.995	0.995	0.995
	ひびわれ	0.000	0.035	0.032
	Mean IoU	0.498	0.515	0.514
U-Net	背景	0.995	<b>0.996</b>	<b>0.996</b>
	ひびわれ	0.019	<b>0.298</b>	0.163
	Mean IoU	0.507	<b>0.647</b>	0.580
PSPNet	背景	0.995	0.995	0.995
	ひびわれ	0.000	0.000	0.000
	Mean IoU	0.498	0.498	0.498

※赤字：3手法中の最大値

表 3-8 CASE2（剥離・鉄筋露出）についての比較評価結果

手法	項目	IoU		
		未処理	データ拡張 (1)	データ拡張 (2)
SegNet	背景	0.852	0.880	0.890
	c	0.049	0.405	0.415
	d	0.050	0.386	0.397
	e	0.564	0.629	0.664
	Mean IoU	0.379	0.575	0.592
U-Net	背景	0.435	0.888	<b>0.902</b>
	c	0.000	0.501	<b>0.560</b>
	d	0.000	0.506	<b>0.557</b>
	e	0.253	0.649	<b>0.704</b>
	Mean IoU	0.172	0.636	<b>0.681</b>
PSPNet	背景	0.834	0.868	0.880
	c	0.130	0.383	0.496
	d	0.185	0.474	0.421
	e	0.481	0.634	0.675
	Mean IoU	0.408	0.590	0.618

※赤字：3手法中の最大値

表 3-9 CASE3（漏水・遊離石灰）についての比較評価結果

手法	項目	IoU		
		未処理	データ拡張 (1)	データ拡張 (2)
SegNet	背景	0.922	0.938	0.939
	c	0.073	0.280	0.325
	d	0.079	0.200	0.150
	e-1	0.001	0.310	0.323
	e-2	0.411	0.488	0.448
	Mean IoU	0.297	0.443	0.437
	U-Net	背景	0.922	0.944
c		0.081	0.335	0.344
d		0.001	0.282	0.353
e-1		0.000	0.257	<b>0.361</b>
e-2		0.439	<b>0.532</b>	0.502
Mean IoU		0.289	0.470	<b>0.501</b>
PSPNet		背景	0.916	0.933
	c	0.025	0.347	<b>0.349</b>
	d	0.010	0.313	<b>0.363</b>
	e-1	0.000	0.158	0.203
	e-2	0.269	0.458	0.396
	Mean IoU	0.244	0.442	0.449

※赤字：3手法中の最大値

表 3-10 各手法の IoU の最大値の比較結果

変状	項目	IoU		
		SegNet	U-Net	PSPNet
CASE1 ひびわれ	背景	0.995	<b>0.996</b>	0.995
	ひびわれ	0.035	<b>0.298</b>	0.000
	Mean IoU	0.515	<b>0.647</b>	0.498
CASE2 剥離・ 鉄筋露出	背景	0.890	<b>0.902</b>	0.880
	c	0.415	<b>0.560</b>	0.496
	d	0.397	<b>0.557</b>	0.474
	e	0.664	<b>0.704</b>	0.675
	Mean IoU	0.592	<b>0.681</b>	0.618
CASE3 漏水・ 遊離石灰	背景	0.939	<b>0.945</b>	0.935
	c	0.325	0.344	<b>0.349</b>
	d	0.200	0.353	<b>0.363</b>
	e-1	0.323	<b>0.361</b>	0.203
	e-2	0.488	<b>0.532</b>	0.458
	Mean IoU	0.443	<b>0.501</b>	0.449

※■：データ拡張(1)，■：データ拡張(2)，赤字：3手法中の最大値

	入力画像	正解画像	セグメンテーション結果		
			SegNet	U-Net	PSPNet
ひびわれ					
	—	—	IoU : 0.000	<b>IoU : 0.121</b>	IoU : 0.000
剥離・鉄筋露出 c					
	—	—	IoU : 0.796	<b>IoU : 0.992</b>	IoU : 0.000
剥離・鉄筋露出 d					
	—	—	IoU : 0.800	<b>IoU : 0.821</b>	IoU : 0.623
剥離・鉄筋露出 e					
	—	—	IoU : 0.537	<b>IoU : 0.804</b>	IoU : 0.562
漏水・遊離石灰 c					
	—	—	IoU : 0.184	IoU : 0.377	<b>IoU : 0.378</b>
漏水・遊離石灰 d					
	—	—	IoU : 0.377	<b>IoU : 0.717</b>	IoU : 0.478
漏水・遊離石灰 e-1 つらら					
	—	—	IoU : 0.610	IoU : 0.628	<b>IoU : 0.660</b>
漏水・遊離石灰 e-2 錆汁					
	—	—	IoU : 0.359	<b>IoU : 0.765</b>	IoU : 0.414

■ : ひびわれ , ■ : 剥離・鉄筋露出c , ■ : 剥離・鉄筋露出d  
■ : 剥離・鉄筋露出e , ■ : 漏水・遊離石灰c , ■ : 漏水・遊離石灰d  
■ : 漏水・遊離石灰e-1, ■ : 漏水・遊離石灰e-2

図 3-14 テストデータに対するセグメンテーション結果の例

	入力画像	正解画像	U-Netのセグメンテーション結果		
			データ拡張なし	データ拡張1	データ拡張2
ひびわれ					
	—	—	IoU : 0.022	<b>IoU : 0.225</b>	IoU : 0.121
剥離・鉄筋露出c					
	—	—	IoU : 0.000	IoU : 0.900	<b>IoU : 0.992</b>
剥離・鉄筋露出d					
	—	—	IoU : 0.000	<b>IoU : 0.822</b>	IoU : 0.821
剥離・鉄筋露出e					
	—	—	IoU : 0.597	IoU : 0.579	<b>IoU : 0.804</b>
漏水・遊離石灰c					
	—	—	IoU : 0.000	IoU : 0.019	<b>IoU : 0.377</b>
漏水・遊離石灰d					
	—	—	IoU : 0.000	IoU : 0.691	<b>IoU : 0.717</b>
漏水・遊離石灰e-1					
	—	—	IoU : 0.000	<b>IoU : 0.637</b>	IoU : 0.628
漏水・遊離石灰e-2					
	—	—	IoU : 0.255	IoU : 0.718	<b>IoU : 0.765</b>

■ : ひびわれ , ■ : 剥離・鉄筋露出c , ■ : 剥離・鉄筋露出d  
■ : 剥離・鉄筋露出e , ■ : 漏水・遊離石灰c , ■ : 漏水・遊離石灰d  
■ : 漏水・遊離石灰e-1, ■ : 漏水・遊離石灰e-2

図 3-15 U-Net によるセグメンテーション結果の例

SegNet と U-Net は、ともに Skip 接続を採用した Encoder-Decoder 構造であるが、スキップさせて Decoder に渡す情報が、SegNet は最大プーリングインデックスのみであるのに対して、U-Net は対応するサイズの特徴マップ全体である。そのため、U-Net の方が、メモリの使用量は大きくなるが、Decoder において物体の詳細と空間的な次元の復元に使用する情報量は多い。SegNet は主に風景画像、U-Net は主に医用画像をそれぞれ対象として研究されたネットワークである。風景のようにそれぞれ独立した物体が混在する画像に比べて、医用画像はセグメンテーションする領域の境界が不明瞭である。そのため、U-Net では、メモリの使用量等よりもピクセルのディティールを補った特徴の抽出を行うことを重視したモデル設計の対応が取られたものと考えられる。このことは、コンクリートの変状検出が、コンクリートという一様な領域に対して境界が不明瞭な変状領域を捉えるという面で、医用画像におけるセグメンテーションと共通することから、SegNet よりも U-Net の方が適性の高い結果となったものと考えられる。また、図 3-14 の出力例の剥離・鉄筋露出 c、漏水・遊離石灰 d および漏水・遊離石灰 e-2 で見られるように、SegNet は U-Net よりもクラスの誤認識が多く見られる傾向があった。

PSPNet は、大域的な情報を重視して、クラス間の関係性や目立たないクラスの物体を含む小領域にも注意を払わせる等の性質がある。そのため、一様なコンクリート面における領域の検出よりも、より複雑な画像の方が、その特性を発揮するものと考えられる。PSPNet は ResNet<sup>28)</sup>と呼ばれる非常に深い層のネットワークで構成されていることから高い精度の検出を期待したが、結果は U-Net の方が優位となる結果となった。事前に、U-Net の改良版で U-Net をベースに ResNet を組み合わせた FusionNet<sup>29)</sup>での検出を試行したが、層の浅い旧来の U-Net の方が優位な結果であった。データが少量で、シンプルな画像が対象の場合には、ResNet の有効性が十分に発揮されていないと推察される。このことは、層の深さが必ずしも精度に直結するわけではなく、教師データの量やセグメンテーションを行う画像の性質等のタスクの条件に応じて適切な手法の選定が必要ということを示唆していると考えられる。また、図 3-14 の出力例で見られるように、IoU で U-Net を上回っているクラスも含めて、漏水・遊離石灰 e-2 のつらら状の部分等のような複雑な輪郭に対しては、ディティールの詳細な表現が、U-Net よりも劣る傾向があった。

以上より、3 手法の中では、U-Net が本教師データによるコンクリートの変状領域の検出に対する適性が最も高いと考えられることから、以降の検討は U-Net を用いて行うこととした。

変状別の IoU では、剥離・鉄筋露出が 0.6 前後、漏水・遊離石灰が 0.4 前後であったの

---

に対して、ひびわれについては非常に低調な結果であった。これは、ひびわれが、他の面的な変状とは異なり、領域面積が微小で極端にピクセル数が少ないことから、背景との間にデータ量の大きな偏りが生じていることが影響していると考えられる。ひびわれについては、学習時に重み付けを行うなど、データ量の偏りを解消する必要がある。

### 3-5-2 セグメンテーションの精度向上のための検討

#### (1) 検討条件

本項では、3手法で比較評価を行った結果、最も精度の良い結果となった U-Net を用いて、精度向上のための各種検討を行う。教師データは 3-5-1 項と同様とし、データ拡張の方針は 3-5-1 項におけるデータ拡張(1)により行う。

#### (2) 分類クラス数についての検証

3-5-1 項では、変状の種類別にセグメンテーションを実行して比較評価を行ったが、実務的には一括して 8 クラスの変状と背景の 9 クラス分類を行うこととなる。しかし、3-5-1 項における結果では、4 クラス分類であった剥離・鉄筋露出よりも 5 クラス分類であった漏水・遊離石灰の方が IoU は低い値となっており、分類問題としては、クラス数が多い方が精度は落ちると考えられる。そこで、8 クラスの変状と背景の 9 クラス分類としてセグメンテーションを行う場合と、個々の変状と背景の 2 クラス分類としてセグメンテーションを行う場合とで比較を行い、分類クラス数の影響について検証を行った。検討ケースを表 3-11 に示す。

表 3-11 分類クラス数についての検証における検討ケース

クラス	CASE								
	1	2	3	4	5	6	7	8	9
ひびわれ	●	●							
剥離・鉄筋露出c	●		●						
剥離・鉄筋露出d	●			●					
剥離・鉄筋露出e	●				●				
漏水・遊離石灰c	●					●			
漏水・遊離石灰d	●						●		
漏水・遊離石灰e-1 つらら	●							●	
漏水・遊離石灰e-2 錆汁	●								●
背景	●	●	●	●	●	●	●	●	●

クラス数についての比較評価結果を表 3-12 に示す。漏水・遊離石灰 d 以外は、2 クラス分類の方が高い結果となっている。ただし、背景に限った場合は、9 クラス分類でも IoU は 0.925 という結果であり、変状と背景の分離はほぼ適切に行えている。変状の領域内において、該当する変状以外の他のクラスの変状との誤認識の比率が高い傾向である。クラスの誤認識の例を図 3-16 に示す。

表 3-12 分類クラス数についての比較評価結果

クラス	Average IoU	
	9クラス分類	2クラス分類
ひびわれ	0.024	<b>0.632</b>
剥離・鉄筋露出c	0.522	<b>0.773</b>
剥離・鉄筋露出d	0.477	<b>0.639</b>
剥離・鉄筋露出e	0.694	<b>0.704</b>
漏水・遊離石灰c	0.301	<b>0.512</b>
漏水・遊離石灰d	<b>0.458</b>	0.454
漏水・遊離石灰e-1 つらら	0.322	<b>0.447</b>
漏水・遊離石灰e-2 錆汁	0.501	<b>0.521</b>






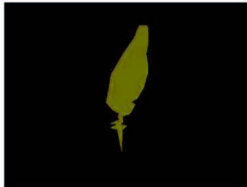


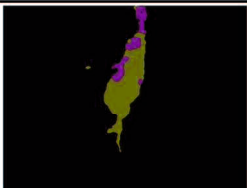





		漏水・遊離石灰 d	漏水・遊離石灰 e-1	剥離・鉄筋露出 d
入力画像				
正解画像				
セグメンテーション結果	9クラス分類			
	2クラス分類			
<p> <span style="color: red;">■</span> : ひびわれ , <span style="color: blue;">■</span> : 剥離・鉄筋露出c , <span style="color: gray;">■</span> : 剥離・鉄筋露出d  <span style="color: brown;">■</span> : 剥離・鉄筋露出e , <span style="color: green;">■</span> : 漏水・遊離石灰c , <span style="color: olive;">■</span> : 漏水・遊離石灰d  <span style="color: purple;">■</span> : 漏水・遊離石灰e-1, <span style="color: teal;">■</span> : 漏水・遊離石灰e-2                 </p>				

図 3-16 9 クラス分類によるクラスの誤認識の例

たとえば剥離・鉄筋露出であれば、d ランクであっても e ランクであっても、鉄筋が露出した箇所の周囲は c ランクと同様のコンクリートが剥落した領域がある。また漏水・遊離石灰であれば、錆汁による e ランクであっても、錆汁の周囲に c ランクと同様の漏水や、d ランクと同様の遊離石灰の析出が見られる領域がある。このような経年的な劣化によりクラスが変化する事象では、異なるクラス間で外観上共通する領域が存在することが多いため、専門技術者でもクラスの境界の判断に悩むことも少なくない。しかし、通常、点検業務においては、連続する領域内で細かな分類を行うことはなく、一番進行しているランクをその領域のランクとして評価を行う。そのため、連続する領域内に複数のクラスが混在する 9 クラス分類の出力は、専門技術者が行う分類とは異なることになる。さらに、多クラス分類の場合には、クラス間のデータ量の偏りにも留意が必要となる。

以上より、IoU についての評価、異なるクラス間で外観上共通する領域が存在するという課題への対処の面から、変状と背景の2クラス分類でセグメンテーションを行うことが望ましいと判断した。なお、ひびわれについては、背景との間のデータ量の偏りを解消するために、クラスの重み付けを行った結果、表 3-7 において 0.298 であった IoU が 0.632 となり、大幅な向上が見られた。

#### (3) 実用性の検証

実際の点検画像では、画像内に複数種類の変状が存在する場合がある。変状と背景の2クラス分類でセグメンテーションを行うためには、セグメンテーション適用の前段階で、変状のクラスが絞り込まれている必要がある。そこで、まずはじめに、物体検出で変状箇所の矩形領域での検出とクラス分類を行い、矩形領域に対して2クラス分類のセグメンテーションを行う、2段階での検出を行うこととした。これにより、工程が増えて、手間や時間は掛かるが、精度の向上が期待できる。深層学習の分野は、技術革新のスピードが速く、物体検出およびセグメンテーションの新たな手法が次々に提案されている。2段階の検出は、それぞれの段階で、その時点での最適な手法の選択を行いやすいというメリットもある。そこで、本研究では、物体検出に YOLO v3、2クラス分類のセグメンテーションに U-Net を採用する手法を提案した。提案手法によるセグメンテーションの流れを図 3-17 に示す。提案手法におけるセグメンテーションは、大半の領域を変状が占める画像への適用となるため、変状部分を切り出した画像を教師データとして学習を行った。なお、物体検出の段階で変状の領域全体を網羅できない場合も考えられるため、検出した矩形領域よりも拡大した範囲に対してセグメンテーションを行うこととした。



図 3-17 提案手法によるセグメンテーションの流れ

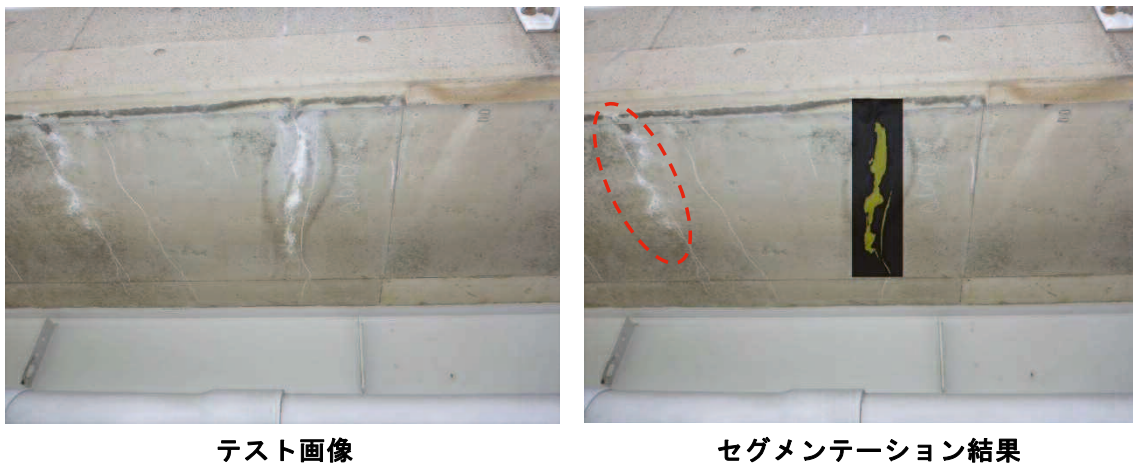
3-5-1 項で用いた変状種類別の U-Net, 9 クラス分類の U-Net, 提案手法の検出結果の比較例を図 3-18 に示す. 前出の 2 手法では, 複数のクラスが混在する結果であるのに対して, 提案手法では, 単一のクラスの変状として検出を行い, IoU も向上させている. 提案手法を適用した結果, 良好な領域検出結果が得られ, 有効性が確認できた.

	入力画像	正解画像	セグメンテーション結果		
			U-Net データ拡張1 (変状種類別)	U-Net (9クラス分類)	YOLO v3 + U-Net
ひびわれ					
	—	—	IoU : 0.225	IoU : 0.163	<b>IoU : 0.729</b>
剥離・鉄筋露出 c					
	—	—	IoU : 0.918	IoU : 0.916	<b>IoU : 0.934</b>
剥離・鉄筋露出 d					
	—	—	IoU : 0.632	IoU : 0.451	<b>IoU : 0.841</b>
剥離・鉄筋露出 e					
	—	—	IoU : 0.686	IoU : 0.574	<b>IoU : 0.757</b>
漏水・遊離石灰 c					
	—	—	<b>IoU : 0.936</b>	IoU : 0.792	IoU : 0.921
漏水・遊離石灰 d					
	—	—	IoU : 0.541	IoU : 0.698	<b>IoU : 0.837</b>
漏水・遊離石灰 e-1 つらら					
	—	—	IoU : 0.459	IoU : 0.237	<b>IoU : 0.721</b>
漏水・遊離石灰 e-2 錆汁					
	—	—	IoU : 0.747	IoU : 0.716	<b>IoU : 0.786</b>

■ : ひびわれ , ■ : 剥離・鉄筋露出c , ■ : 剥離・鉄筋露出d , ■ : 剥離・鉄筋露出e  
■ : 漏水・遊離石灰c , ■ : 漏水・遊離石灰d , ■ : 漏水・遊離石灰e-1, ■ : 漏水・遊離石灰e-2

図 3-18 セグメンテーション結果の比較例

後段の処理のセグメンテーションについては、2クラス分類とすることで難易度を大幅に緩和させて精度の向上を実現した。そのため、前段の物体検出部分の精度により検出が左右される。たとえば、物体検出の過程において検出漏れとなった領域に対しては、セグメンテーションが実行されず、誤検出となった領域に対しては、適切でないセグメンテーションが実行されてしまう可能性がある。特に、大半の教師データが画像中心付近に対象が位置する画像であるため、画像縁端部における検出、複数の変状が混在する画像に対する検出への留意が必要と考えられる。図3-19では、画像中心付近から外れた部分の変状に対して、前段の物体検出で検出できていないため、セグメンテーションがなされていない。複数の変状が混在するような広域的な画像への適用を見据えて、スライディングウィンドウをオーバーラップさせる形での検出の処理を行うことを考えている。




 : 物体検出の段階で検出されず、セグメンテーションが実行されなかった領域

図3-19 提案手法による検出漏れの例

### 3-6 まとめ

本章では、深層学習によるコンクリートの変状検出の検討として、物体検出手法による検出とセグメンテーション手法による検出について検討を行った。本章における検討により、以下の点を明らかにした。

- (1) 物体検出では、点検時および踏査時の撮影画像を教師データにした条件で、クラスと矩形領域の双方について妥当な予測結果が得られた。Faster R-CNNとYOLO v3の2手法の比較では、YOLO v3の方が領域の特定に関する精度において優位な結果となり、YOLO v3のコンクリート変状検出についての適性の高さを確認した。
- (2) セマンティックセグメンテーションでは、SegNet, U-NetおよびPSPNetの3手法の比較において、U-Netが領域の特定に関する精度において最も優位な結果となり、限られたデータ量でのコンクリート変状のセグメンテーションについて、U-Netの適性の高さを確認した。また、2章のクラス分類では、データ拡張のみでは十分な精度の向上にはつながらなかったが、セマンティックセグメンテーションではデータ拡張が精度の向上に大きく寄与することを明らかにした。
- (3) 対象とする変状と背景の2クラス分類で実行するという独自の方法でコンクリート変状のセグメンテーションを行うことで、精度の向上が図れ、実際の専門技術者が行う点検に即した変状領域の検出を実現した。具体的には、物体検出ベースのセグメンテーションとして、YOLO v3により検出した矩形領域に対して、U-Netにより該当するクラスの変状と背景の2クラス分類のセグメンテーションを行った。

深層学習による物体検出技術およびセグメンテーション技術を併用することで、コンクリートの変状の画像について、変状の種類および程度の分類まで含めて領域の検出を行えることが検証できた。ただし、クラス分類においてはAIが人間を超えたとまでいわれる評価を得たが、領域の特定も併せて行う物体検出あるいはセグメンテーションにおいては、まだまだ人間と同等のレベルには至っていない。論文等において、大規模画像データセットを用いて整えられた実験環境下では一定の精度で検出を行えているが、データ量が十分でない独自の画像データセットで高精度に行うには難易度が高いため、さらなる工夫を行って精度の向上を図る必要がある。

### 第3章 参考文献

- 1) 山内悠嗣, 山下隆義, 藤吉弘亘 : 画像からの統計的学習手法に基づく人検出, 電子情報通信学会論文誌 D, Vol. J96-D, No. 9, pp. 2017-2040, 2013.
- 2) Girshick, R., Donahue, J., Darrell, T. and Malik, J. : Rich feature hierarchies for accurate object detection and se-mantic segmentation, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 580-587, 2014.
- 3) Everingham, M., Gool, L. V., Williams, C. K. I. and Winn, J. : The PASCAL visual object classes (VOC) challenge, *Int. J. Comput. Vis.*, Vol. 88, No. 2, pp. 303-338, 2010.
- 4) スタンフォード大学, プリンストン大学 : ImageNet Large Scale Visual Recognition Challenge (ILSVRC), <http://image-net.org/challenges/LSVRC/>, (最終閲覧日 : 2020.03.01)
- 5) Uijlings, J. R., van de Sande, K. E., Gevers, T. and Smeulders, A. W. : Selective search for object recognition, *International Journal of Computer Vision (IJCV)*, Vol. 104, No. 2, pp. 154-171, 2013.
- 6) He, K., Zhang, X., Ren, S. and Sun, J. : Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition, *European Conference on Computer Vision (ECCV)*, pp. 346-361, 2014.
- 7) Girshick, R. : Fast R-CNN, *IEEE International Conference on Computer Vision (ICCV)*, pp. 1440-1448, 2015.
- 8) Ren, S., He, K., Girshick, R. and Sun, J. : Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *Neural Information Processing Systems*, Vol. 28, pp. 91-99, 2015.
- 9) Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. : You only look once: Unified, real-time object detection, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788, 2016.
- 10) Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. and Berg, A. : SSD: Single Shot MultiBox Detector, *European Conference on Computer Vision (ECCV)*, pp. 1-15, 2016.
- 11) Redmon, J. and Farhadi, A. : YOLO9000: Better, Faster, Stronger, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6517-6525, 2017.
- 12) Redmon, J. and Farhadi, A. : YOLOv3: An Incremental Improvement, *arXiv: 1804.02767v1*, 2018.

- 13) Zhao, Q., Sheng, T., Wang, Y., Tang, Z., Chen, Y. and Cai, L. : M2Det: A Single-Shot Object Detector Based on Multi-Level Feature Pyramid Network, *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33, No. 1, pp. 9259-9266, 2019.
- 14) Tan, M., Pang, R. and Le, Q. V. : EfficientDet: Scalable and Efficient Object Detection, *arXiv: 1911.09070v4*, 2020.
- 15) Bochkovskiy, A., Wang, C. Y. and Liao, H.-Y. M. : YOLOv4: Optimal Speed and Accuracy of Object Detection, *arXiv: 2004.10934v1*, 2020.
- 16) Ren, S., He, K., Girshick, R. and Sun, J. : Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *arXiv: 1506.01497v3*, 2016.
- 17) Kathuria, A : What's new in YOLO v3?, <https://towardsdatascience.com/yolo-v3-object-detection-53fb7d3bfe6b>, (最終閲覧日 : 2020.03.01)
- 18) Boykov, Y., Jolly, M.-P. : Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images, *IEEE International Conference on Computer Vision (ICCV)*, pp. 7-14, 2001.
- 19) Farabet, C., Couprie, C., Najman, L. and LeCun, Y. : Learning hierarchical features for scene labeling, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 35, No. 8, pp. 1915-1929, 2013.
- 20) Kirillov, A., He, K., Girshick, R. , Rother, C. and Dollar, P. : Panoptic Segmentation, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9404-9413, 2019.
- 21) Ulku, I., Akagunduz, E. : A Survey on Deep Learning-based Architectures for Semantic Segmentation on 2D images, *arXiv: 1912.10230v1*, 2019.
- 22) Long, J., Shelhamer, E. and Darrell, T. : Fully convolutional networks for semantic segmentation, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431-3440, 2015.
- 23) Badrinarayanan, V., Kendall, A. and Cipolla, R. : SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 12, pp. 2481-2495, 2017.
- 24) Ronneberger, O., Fischer, P. and Brox, T. : U-Net: Convolutional Networks for Biomedical Image Segmentation, *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 234-241, 2015.
- 25) Zhao, H., Shi, J., Qi, X., Wang, X. and Jia, J. : Pyramid Scene Parsing Network, *IEEE*



- Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6230-6239, 2017.
- 26) Yu, F., Koltun, V. : Multi-Scale Context Aggregation by Dilated Convolutions, *arXiv: 1912.10230v3*, 2016.
- 27) Kentaro, W : labelme: Image Polygonal Annotation with Python, <https://github.com/wkentaro/labelme>, (最終閲覧日 : 2020.05.01)
- 28) He, K., Zhang, X., Ren, S. and Sun, J. : Deep Residual Learning for Image Recognition, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016.
- 29) Quan, T. M., Hildebrand, D. G. C. and Jeong, W.-K. : FusionNet: A deep fully residual convolutional neural network for image segmentation in connectomics, *arXiv: 1612.05360v2*, 2016.



## 第 4 章 変状検出における深度画像活用の検討

---

### 第 4 章

4-1 深度画像の概要

4-2 深度画像を活用した変状検出

4-3 深度情報による変状の定量値の把握

4-4 まとめ



## 第4章 変状検出における深度画像活用の検討

---

### 4-1 深度画像の概要

3-5節において、物体検出とセマンティックセグメンテーションを組み合わせた手法について、コンクリートの変状領域の検出に対する有効性を確認できた。この組み合わせを行った処理の後段部分のセグメンテーションについては、2クラス分類として難易度を緩和させることで精度の向上を実現した。そのため、前段の物体検出部分の精度により検出が左右されることとなる。そこで、本章では、教師データに基づく統計的学習以外の、物理量を用いた他のアプローチにより、物体検出の代替手段の検討を行うこととした。

#### 4-1-1 RGB-D カメラの概要

3次元情報を取得する代表的な手法としてレーザースキャナが挙げられるが、高価な専用機材を必要とし、設置位置や角度をGPSによって厳密に指定する必要がある。その他の手法としては、レーザースキャナと比べて精度は落ちるが、操作が容易である深度センサが挙げられる。この深度センサを搭載したカメラは、RGB-Dカメラと呼ばれており、従来のRGB画像に加えて、画像内の各ピクセルが深度（距離）に関する測定値の情報をもつ深度画像を取得できる。また、RGB画像と深度（Depth）画像を対にした画像を、RGB-D画像と呼んでいる。

RGB-Dカメラの低価格化と普及が急速に進むきっかけとなったのは、2010年にMicrosoft社の家庭用ゲーム機Xbox360用の周辺機器として登場したKinectで、2012年にはPCに接続して使用できるタイプの機種も発売された。近年では、一層の低価格化や小型化が進んでおり、スマートフォンにも搭載される等、急速に普及が進んでいる。RGB-Dカメラの種類を表4-1に示す。

表 4-1 RGB-D カメラの種類

手法	分類	測量方式	備考
Passive Stereo	受動的 手法	三角測量	<ul style="list-style-type: none"> <li>・受動的な測定器であるカメラだけを用いて深度を計測。</li> <li>・2つのRGBカメラ等により、三角測量の原理で計測を行う。</li> <li>・2つの画像での対応づけを行う特徴点等が少ない場合は適用が困難（対応点問題）。</li> </ul>
Active Stereo	能動的 手法	三角測量	<ul style="list-style-type: none"> <li>・カメラに加えて、赤外光等を投影する装置を利用してパターンを投影して深度を計測。</li> <li>・対象に投影された光を画像中から検出するので、Passive Stereo法で生じる対象点問題は解消される。</li> <li>・遠距離では光路のずれが小さくなるため、精度が低下する。</li> </ul>
Time of Flight (ToF)		同軸測量	<ul style="list-style-type: none"> <li>・物体に光等を当てて、戻ってくるまでの飛行時間により深度を計測。</li> <li>・深度と可視画像が同軸。</li> <li>・時間の誤差は、距離の影響をさほど受けないため、遠距離の場合の精度の低下は軽微。</li> </ul>

本検討では、Active Stereo 方式の RGB-D カメラの中から、屋外での使用も可能な Intel RealSense を用いた。

#### 4-1-2 深度画像活用の可能性

深度センサはさまざまな局面で活用されており、近年ではロボット等において接触の回避のための目の役割を担うことも多い。土木分野の研究事例では、たとえば北川ら<sup>1)</sup>は、RGB-Dカメラによりパノラマ画像とパノラマ深度画像を生成し、構造物の色情報と3次元情報の経年変化を検出するシステムを開発している。蒲地ら<sup>2)</sup>は、TOF方式の深度センサを用いて非接触式のセンサを試作し、温度と変位の関係を計測して鉄道トンネルにおける計測の適用性を検証している。滝澤ら<sup>3)</sup>は、深度画像により付帯構造物を固定する取付ボルトの脱落箇所を抽出するシステムを開発している。菊田ら<sup>4)</sup>は、TOF方式の深度センサを用いて、トンネル切羽変状を定量的に監視するシステムの開発を行っている。中村ら<sup>5)</sup>は、TOF方式の深度センサを用いて、切羽から30m離れた位置で計測可能な切羽監視システム開発のため基礎研究を行っている。

本検討では、深度センサを用いて得られるピクセル毎の深度の情報を、コンクリート面の平面の連続性が損なわれた凹凸のある領域の把握に用いる。これは、大量の教師データから変状の特徴量と位置を統計的に学習するよりも容易で、再現性も高く普遍的な領域の検出が行えると考えられる。また、深度の情報から逆算して1ピクセルあたりの尺度の把握を行えることから、画像内の変状規模の把握のための定量値の取得にも供することが可能である。

凹凸範囲の検出は、一定量以上の平面的な変化が生じていないと検出できないため、微細な変状であるひびわれや、平面的な変化が乏しい漏水・遊離石灰cおよび漏水・遊離石灰e（錆汁）については、適用は困難である。また、漏水・遊離石灰dについては、遊離石灰の析出箇所の平面的な隆起が生じている場合に限って適用可能である。凹凸範囲の検出は適用可能な変状に限られるのに対して、定量値の把握は、1ピクセルあたりの尺度を基に画像内の対象物の換算を行うシンプルな処理であるため、セグメンテーションにより領域を特定できれば、平面的な起伏に関係なく適用が可能である。変状内容に応じた深度画像適用の分類を表4-2に示す。

表 4-2 変状内容に応じた深度画像適用の分類

変状内容	深度画像の活用		備考
	深度に基づく検出	定量値の把握	
ひびわれ	×	○	
剥離・鉄筋露出c	○	○	
剥離・鉄筋露出d	○	○	
剥離・鉄筋露出e	○	○	
漏水・遊離石灰c	×	○	遊離石灰の析出箇所の隆起の状況によっては、深度に基づく検出も適用可能
漏水・遊離石灰d	△	○	
漏水・遊離石灰e-1 つらら	○	○	
漏水・遊離石灰e-2 錆汁	×	○	

## 4-2 深度画像を活用した変状検出

### 4-2-1 変状の検出方針

コンクリートの剥離や鉄筋露出，遊離石灰の著しい析出が生じたコンクリート面は，健全なコンクリート面では保たれている平面の連続性が損なわれることとなる．鉄筋露出が生じたコンクリート面の3次元イメージを図4-1に示す．RGB画像では，断面 a-a における画素値の情報が得られるのみであるが，RGB-D カメラにより深度画像も併せて取得する場合，断面 b-b あるいは c-c における変状の発生によるコンクリート面の凹凸を3次元的に捉えることができる．深度画像の情報を用いて，物理的にコンクリート面に対する凹凸範囲を把握した後に，その範囲に対してクラス分類を行うという流れで，タスクの難易度を下げた上で物体検出の代替とすることが可能となる．深度画像を活用した変状領域の検出の流れを図4-2に示す．

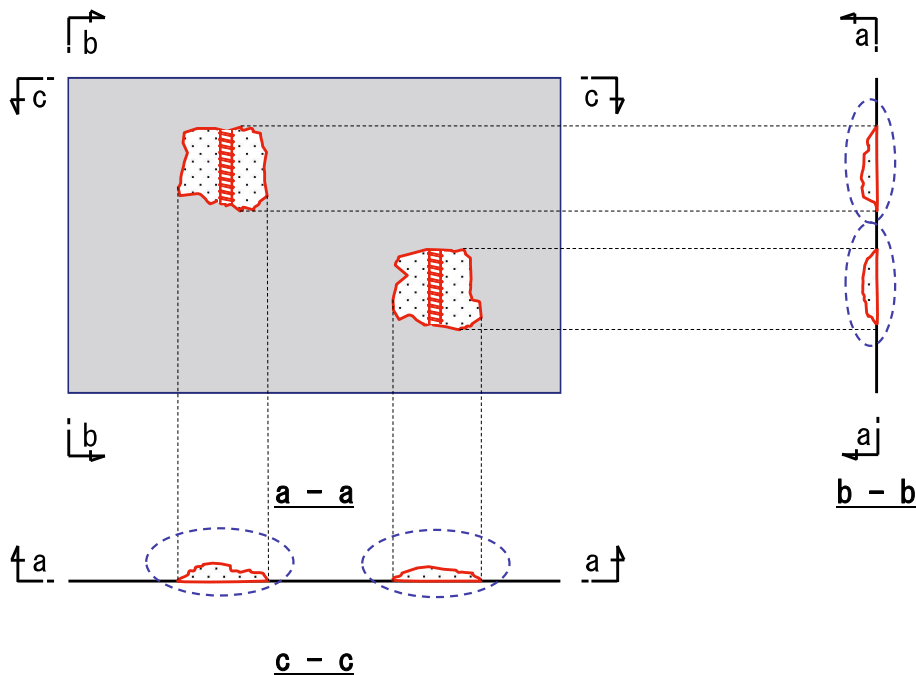


図4-1 鉄筋露出が生じたコンクリート面の3次元イメージ



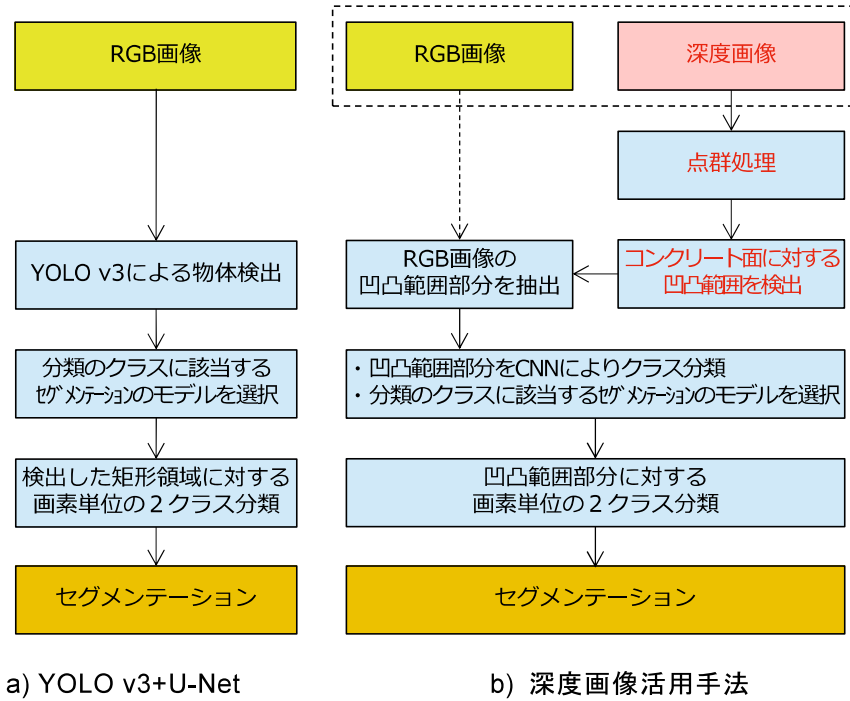


図 4-2 深度画像を活用した変状領域の検出の流れ

#### 4-2-2 変状領域の検出

##### (1) テストデータ

テストデータに用いた剥離・鉄筋露出 d の RGB-D 画像を図 4-3 に示す。RGB-D カメラでは、RGB カメラの位置と深度センサの位置が空間的にずれているため、深度画像の位置を RGB 画像の位置に合わせる調整を行っており、そのために深度画像左端部に空白部分が生じている。

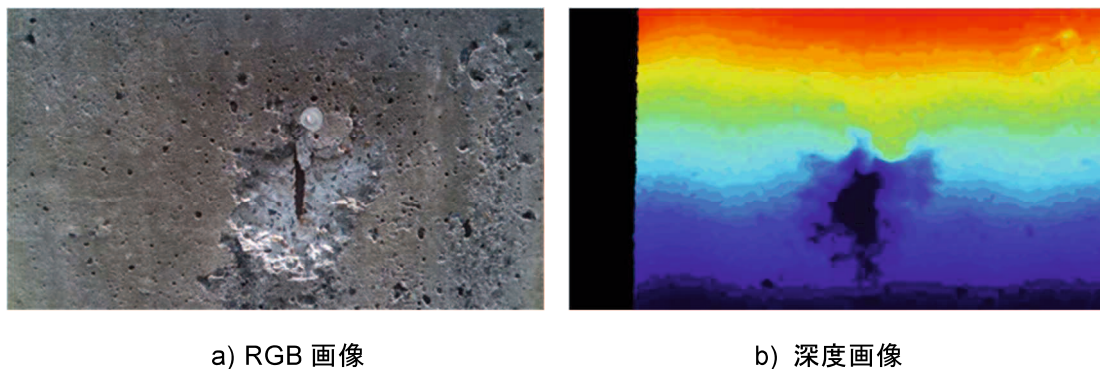


図 4-3 テストデータ

### (2) 点群処理および平面検出

凹凸範囲を検出するために、最初に RGB-D 画像を点群データに変換する。変換した点群データを図 4-4 に示す。

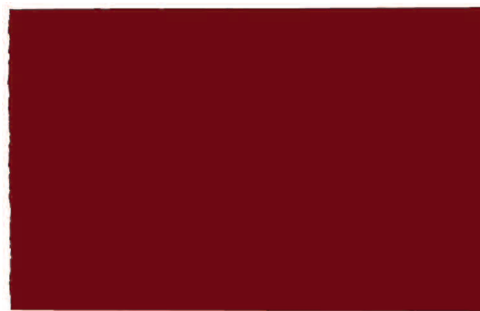
次に、Open3D<sup>6)</sup>を用いて平面の検出を行う。Open3D は、C++と Python の 2 つの言語に対応した、3D データを扱うオープンソースのライブラリである。平面検出のアルゴリズムは、外れ値が多い条件でも頑健なパラメータ推定を行う RANSAC (Random Sample Consensus)<sup>7)</sup>を用いた。RANSAC は、与えられた観測値に外れ値が含まれている可能性を踏まえ、外れ値の影響を抑制してパラメータ推定を行うロバスト推定の一様である。平面検出における処理の流れを以下に示す。

- ① 点群からランダムに  $n$  点のサンプルを抽出
- ② サンプルした点を使って平面方程式  $ax+by+cz+d=0$  のパラメータ  $a, b, c, d$  を推定
- ③ 推定した平面モデルと他の点群との差分が閾値以下となる数を評価
- ④ ①～③を繰り返し、評価値が最も高くなるパラメータにより平面を選定
- ⑤ 平面に合致しない点を削除して平面を検出

閾値を変化させて平面検出を行った結果を図 4-5 に示す。閾値を、1mm, 5mm, 10mm, 50mm の 4 段階で変化させた結果、平面検出結果から 5mm での設定が妥当と判断した。



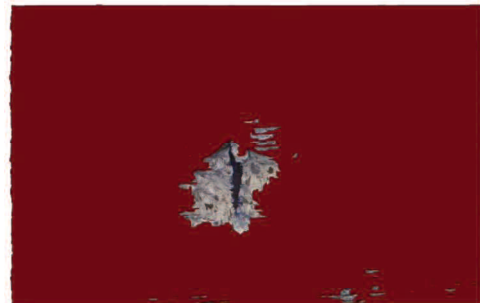
図 4-4 点群データ変換結果



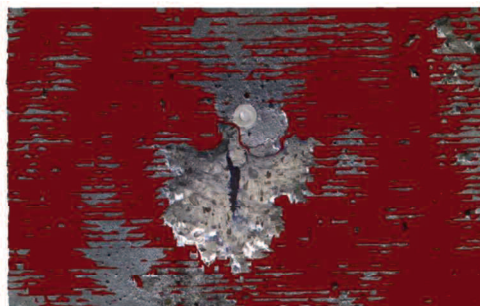
a) 閾値 50mm



b) 閾値 10mm



c) 閾値 5mm



d) 閾値 1mm

図 4-5 平面検出結果

### (3) 変状領域の検出

次に、検出した平面を除去して得られる凹凸範囲の切出画像に対して、2章で構築したCNNによるクラス分類を行った。RANSACにおける閾値を5mmと設定したことにより、5mm未満の凹凸範囲は検出されないため、画像の切り出し範囲は、得られた範囲を基に縦方向に2倍、横方向に2倍に拡大することとした。切出画像およびGrad CAM<sup>8)</sup>による着目領域の可視化画像を図4-6に示す。8クラスの変状の中から判定確率94.1%で剥離・鉄筋露出dと判定された。健全なコンクリート面に対する凹凸範囲を把握することで、物体検出よりもタスクの難易度が1ランク低いクラス分類の適用が可能となり、タスクの難易度を緩和させた上で物体検出と同様のフェーズを完了することができた。

次のフェーズとして、分類結果に該当するクラスのセグメンテーションのモデルを用いて、該当クラスの変状と背景との2クラス分類によるセマンティックセグメンテーションを行う。セグメンテーションには、3章と同様にU-Netを用いた。同じ切出画像に対して、分類結果である剥離・鉄筋露出dに該当する変状と背景との2クラス分類によるセグメンテーションを行った。3章のYOLO v3+U-Netと深度画像活用手法のそれぞれのセグメンテーション結果を図4-7に示す。深度画像活用手法のセグメンテーション結果については、平面除去により得られた範囲の4倍の面積となる範囲に対してセグメンテーションを実施している。その結果、変状領域周辺の表面気泡等の部分も含めて剥離・鉄筋露出dと判定したために、YOLO v3+U-Netよりも若干広い範囲を変状領域として捉えている。ただし、仮にYOLO v3+U-Netの物体検出の段階で変状の矩形領域を捉えられない場合であっても、この深度画像活用手法は凹凸を捉えることで次の段階のセグメンテーションに移行することができるため、多段階の処理となるが、検出の確かさに寄与することができる。

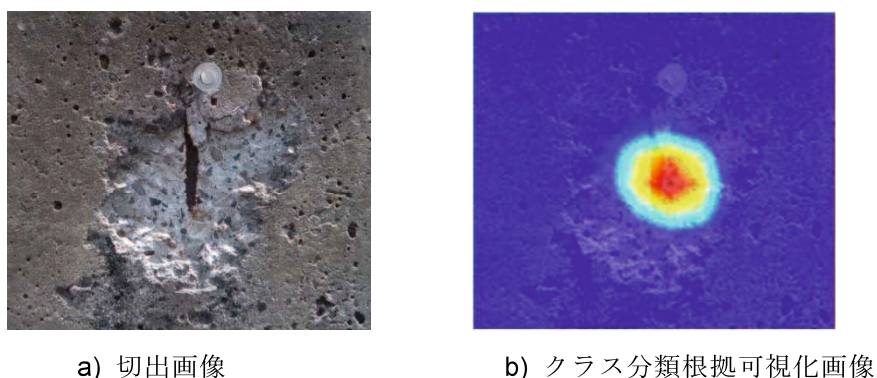


図4-6 切出画像およびクラス分類結果


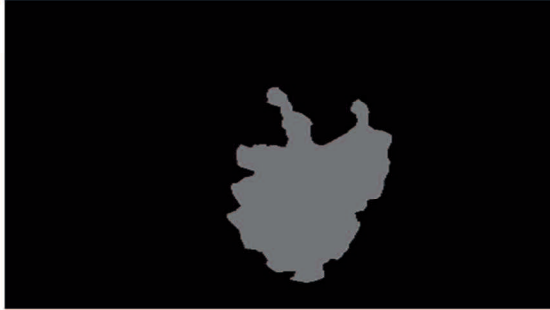


		剥離・鉄筋露出d
入力画像		
正解画像		
セグメンテーション結果	YOLO v3 + U-Net	
		IoU : 0.557
	深度画像併用手法	
		IoU : 0.393

図 4-7 セグメンテーション結果

### 4-3 深度情報による変状の定量値の把握

本節では、深度画像から得られる深度情報を活用して、画像内の変状規模の把握のための定量値の取得についての検討を行う。

#### 4-3-1 変状の定量値の把握方針

RGB-D カメラでは、深度センサで得られる深度の情報とカメラの画角から、画像範囲の寸法を把握することが可能である。深度の観測値による画像寸法の換算イメージを図 4-8 に示す。画像中心の深度と水平画角および垂直画角から、画像の全幅と全高の実際の寸法を算定する。この算定した画像の実寸を基に、変状の領域についての定量値の把握を行う。変状の定量値の把握の流れを図 4-9 に示す。

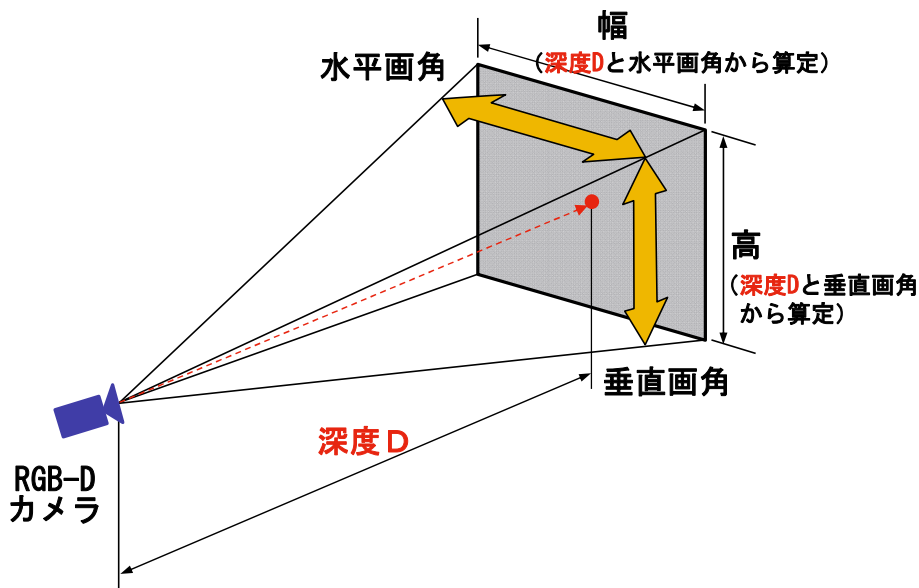


図 4-8 深度の観測値による画像寸法の換算イメージ

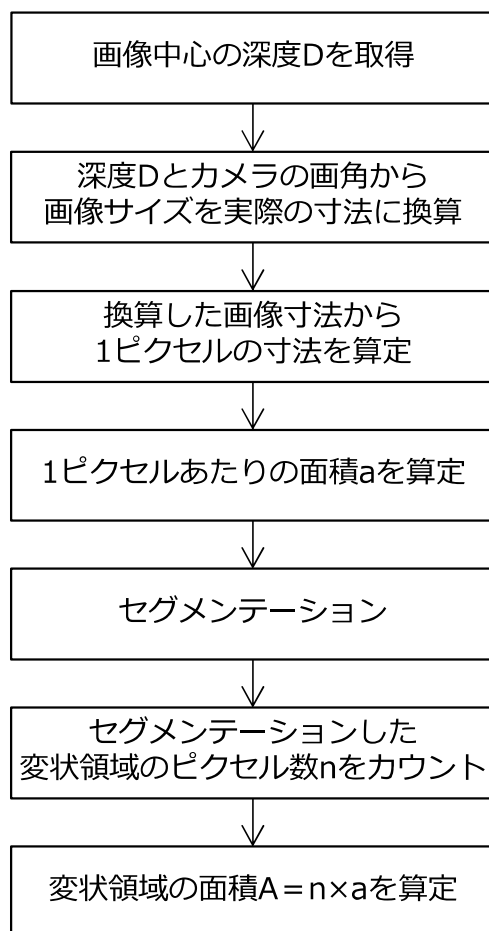


図 4-9 変状の定量値の把握の流れ

#### 4-3-2 定量値把握結果

画像内の深度は、画像中心とその周辺4点の計5点に対して値の取得を行った。周辺4点は、コンクリート面に対して正対した面における寸法値を得るための補正に活用する測点である。画像内における深度の観測位置と深度値を図4-10に示す。画像については、水平方向の2点間については1mm、鉛直方向の2点間については20mmの正対面に対する傾斜があったが、補正值の算定結果では表4-3に示すように垂直方向も水平方向も1.000であった。なお、鉛直方向の若干の傾斜は、図4-3の深度画像においても色相の変化で現れている。

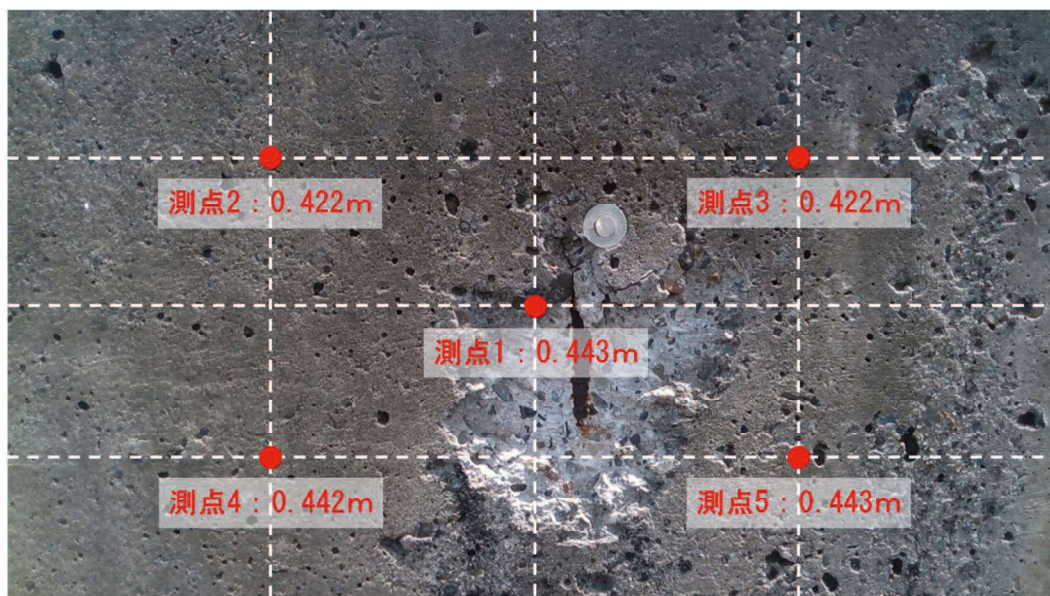


図 4-10 画像内における深度の観測位置と深度値

表 4-3 1ピクセルあたりの面積の換算結果

画像中心 までの深度	画角		画像の 実測サイズ		1pixあたり の寸法		正対面に 対する補正	1pixあたり の面積
0.443 m	水平	69.4°	幅	0.613 m	幅	0.48 mm	1.000	0.23 mm <sup>2</sup>
	垂直	42.5°	高	0.345 m	高	0.48 mm	1.000	

剥離・鉄筋露出 d の領域は，図 4-11 a) に示すように水平方向に 20cm で，概ね鉛直方向が 25cm，面積で 50,000mm<sup>2</sup> 弱程度である．それに対して，セグメンテーションした領域はピクセル数 233,424 で，表 4-3 に示す 1ピクセルあたりの面積により換算すると 53,700mm<sup>2</sup> であり，変状周辺の気泡等の部分も変状領域と判定したため，数値はやや大きいですが，変状の概ねの規模が算出できる精度で面積を把握できている．凹凸範囲の検出は，平面的な起伏の変化が大きい変状に適用が限られるが，定量値の把握は起伏の変化に関係なく適用することが可能である．教師データの充実等によりセグメンテーションの精度の一層の向上は必要であるが，変状の定量値の把握は RGB-D カメラを使うことにより可能である．





a) 寸法入り画像



変状領域のピクセル数 : 233,424  
換算面積  $233424 \times 0.23 \doteq 53,700\text{mm}^2$

b) 面積算定結果

図 4-11 変状の定量値把握結果

### 4-4 まとめ

本章では、第3章で検討を行った独自のセグメンテーション手法について、前段部分の物体検出の処理における統計的な不確かさを排除する目的で、ピクセル単位で深度の情報を有する深度画像を活用した物体検出の代替手段について検討を行った。また、深度画像から得られる深度の情報を基に、変状の定量的把握を行う処理手法を構築した。本章における検討により、以下の点を明らかにした。

- (1) 凹凸を有する変状では、深度画像を活用することによって変状領域を特定することが可能であることを明らかにした。この結果により、構造物点検の効率化への深度画像活用の道筋が描けた。
- (2) 深度画像の活用により、深度と画角を基に1ピクセルあたりの尺度を換算し、変状の定量的把握が可能であることを明らかにした。

上記の結果より、RGB-Dカメラが、目視点検の省力化および効率化に有効なデバイスであることが確認できた。市町村管理の橋梁に多数見られる小規模なコンクリート橋梁等に対しては、レーサースキャナ等の高価な専用機材の適用ではコストの縮減に寄与しない場合も考えられるため、RGB-Dカメラの適用が有効と考えられる。

コンクリート面の平面の連続性が損なわれた凹凸のある領域の把握は、微細な変状であるひびわれや、平面的な変化が乏しい漏水・遊離石灰<sup>c</sup>および漏水・遊離石灰<sup>e</sup>（錆汁）については適用が困難であるが、平面的な変化が伴う変状に対しては有効で、教師データ等に左右されない普遍的な領域検出が可能である。また、デバイス技術の今後の発展や撮影方法等によっては、ひびわれの検出が実現できる可能性もある。

変状の定量値の把握は、セグメンテーションにより領域を特定できれば、平面的な起伏に関係なく適用が可能である。点検結果を基にして行う診断においては、変状の規模の情報も重要であり、定量値の把握は不可欠である。深度情報を活用した変状の定量的把握は、シンプルな処理により行う手法であり、実務的な作業においても十分に活用できると考えられる。

また、本検討では、深度画像を活用した手法としてテストデータにRGB-D画像を用いたが、将来的には教師データにもRGB-D画像を用いて、学習により画素値と深度の双方の特徴量を抽出することも有効と考えられる。

## 第4章 参考文献

- 1) 北川悦司, 田中成典, 塚田義典, 安彦智史, 若林克磨: 距離画像センサを用いた構造物の経年変化の検出と計測に関する研究, 土木学会論文集 F3, Vol. 68, No. 2, pp. I\_85-I\_100, 2012.
- 2) 蒲地秀矢, 大木裕久, 津野究, 藤田輝一: レーザー距離計と無線センサを用いたトンネルひび割れ計測の一検討, 土木学会第 73 回年次学術講演会, pp. 895-896, 2018.
- 3) 滝澤彰宏, 齊藤岳季, 神谷弘志: 深度画像を用いたトンネル付帯構造物の劣化箇所抽出, 土木学会第 74 回年次学術講演会, VI-769, 2019.
- 4) 菊田遥子, 京免継彦, 瀬谷正巳, 新川隆朗, 前田新一: TOF カメラによるトンネル切羽変状監視システム開発の試み, 土木学会第 73 回年次学術講演会, pp. 717-718, 2018.
- 5) 中村隆史, 請関大海, 塩崎正人, 河村圭: デプスカメラを用いたトンネル切羽監視システムの開発における基礎研究, 土木学会第 74 回年次学術講演会, VI-735, 2019.
- 6) Zhou, Q.-Y., Park, J. and Koltun, V. : Open3D: A Modern Library for 3D Data Processing, *arXiv: 1801.09847v1*, 2018.
- 7) Fischler, M. A., Bolles, R. C. : Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM*, Vol. 24, No. 6, pp. 381-395, 1981.
- 8) Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D. : Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, *IEEE International Conference on Computer Vision (ICCV)*, pp. 618-626, 2017.



## 第 5 章 深層学習による異常検知手法を用いた変状検出の検討

---

### 第 5 章

5-1 深層学習による異常検知

5-2 変状検出への異常検知手法の適用

5-3 実用性の検証

5-4 まとめ



## 第5章 深層学習による異常検知手法を用いた変状検出の検討

---

### 5-1 深層学習による異常検知

3-6節で前述したように、コンクリート変状に対する物体検出手法あるいはセグメンテーション手法による検出は、難易度が高い。そのため、検出の精度の向上を図るために、種々の工夫を行う必要がある。教師あり学習である物体検出およびセグメンテーションでは、精度の向上の方策として教師データの増強が考えられるが、教師データの収集については以下の点が課題となる。

- a) 教師データとなる適切な変状画像の不足
- b) 教師データの画像内の変状領域のアノテーションを行う、データセット作成の労力
- c) 教師データから得られる特徴量と類似しない未知の変状との遭遇の可能性

これらの課題に左右されることなく、第3章で検討を行った独自のセグメンテーション手法を補完する手法として、4章では深度画像を活用した手法について検討を行った。そして、本章では正常画像を用いた教師なし学習による異常検知手法について検討を行う。正常画像を用いた教師なし学習による異常検知では、上記の課題のうち、a) については、変状画像ではなく健全なコンクリート面の正常画像を学習に用いることにより解消できる。また、c) についても教師あり学習で変状の特徴量を学習するわけではないので、変状以外の正常画像との何らかの相違点を検知する可能性はある反面、未知の変状を検知できる可能性もあるため、概ね解消できると考えられる。ただし、変状の種類分類までは行えないため、物体検出手法あるいはセグメンテーション手法の補完を目的とした検討と位置付けられる。また、本章における検討では、b) の課題の解消も可能となるように、教師なし学習に用いる正常画像について、着目領域等のアノテーションを伴わない手法を用いることとした。

#### 5-1-1 異常検知の概要

##### (1) 異常検知の概要

情報工学分野における「異常」とは、大多数の正常なデータとは振る舞いが異なるデ

---

一タだと捉えられる。異常検知の基本は、「正常」となるモデルをデータからつくり、そのモデルから外れるものを「異常」として識別することである<sup>1)</sup>。正常と異常を識別するためには何らかの判定基準が必要となる。深層学習による異常検知は、データパターンの自動学習により、この判定基準を見出すことで実現している。近年の異常検知では、深層学習を用いることで、データ内の複雑な構造を捉えることが可能となり、画像等の高次元で複雑なデータへも適用範囲が広がっている。

## (2) 異常検知への深層学習の適用

深層学習による異常検知については、多種多様な手法が提案されている。文献<sup>2)</sup>ではモデルのタイプの観点において、「教師あり学習」、「半教師あり学習」、「教師なし学習」、「ハイブリッド」、「One-Classニューラルネットワーク」の5つに分類を行っている。各タイプの概要を表5-1に示す。

表 5-1 深層学習に基づく異常検知手法の各モデルタイプの概要  
(文献<sup>2)</sup>の記載内容を整理)

モデルタイプ	深層学習における学習内容	特徴	主な長所	主な短所	備考
[1] 教師あり学習モデル	ラベル付きデータセットから分離境界を学習	データのクラス分類に依存している	<ul style="list-style-type: none"> <li>・[2],[3]よりも正確</li> <li>・分類ベースの手法の場合、推論フェーズにおける速度が速い</li> </ul>	<ul style="list-style-type: none"> <li>・「異常」クラスについての正確なラベルが必要</li> <li>・特徴空間が非常に複雑であるため、非線形の場合、異常データの分離が困難</li> </ul>	多クラス分類も含む
[2] 半教師あり学習モデル	「正常」データ周囲の識別境界を学習	「異常」を区別するための識別属性を保持する	<ul style="list-style-type: none"> <li>・ラベル付きデータ（通常1クラス）を使用することにより、[3]よりも性能の大幅な向上が期待できる</li> </ul>	<ul style="list-style-type: none"> <li>・抽出した階層的特徴が、希少な「異常」事例を表すものでない場合、過剰適合を起こしやすい</li> </ul>	1クラス分類も含む
[3] 教師なし学習モデル	「正常」データの固有の特性を学習	「正常」データの特性の説明と理解に重点を置いている	<ul style="list-style-type: none"> <li>・「正常」データの共通点の識別により、異常検知が容易</li> <li>・ラベル付きデータが不要のため、異常検知のための費用対効果が高い</li> </ul>	<ul style="list-style-type: none"> <li>・複雑で高次の空間では、データ内共通性の学習が困難</li> <li>・データのノイズ等に対する感度が敏感で、[1],[2]よりも正確性が低くなる場合がある</li> </ul>	
[4] 深層ハイブリッド（深層学習+従来手法）モデル	特徴抽出器におけるロバストな特徴量抽出を学習	深層学習モデルにより抽出した特徴量を、SVM分類器等の従来型アルゴリズムに入力する2段階学習を行う	<ul style="list-style-type: none"> <li>・特徴抽出器は、特に高次元領域において「次元の呪い」を大幅に軽減できる</li> <li>・縮小入力次元で動作するモデルの場合、計算効率が高い</li> </ul>	<ul style="list-style-type: none"> <li>・一般的な損失関数を使用するため、特徴抽出器の表現学習に寄与せず、最適な異常検知とならない</li> </ul>	
[5] 1クラスニューラルネットワークモデル	「異常」からすべての「正常」を分離するための識別境界を学習	1クラス分類の目的に沿った、超平面や超球面等の豊富なデータ表現を抽出する	<ul style="list-style-type: none"> <li>・推論のためにデータを記憶する必要がないので、メモリの複雑性が非常に低い</li> </ul>	<ul style="list-style-type: none"> <li>・高次元の入力データの場合、学習時間とモデル更新時間が長くなる</li> </ul>	

画像に対する異常検知では、正常画像であるデータの分布に適合する観測値と、適合しない観測値との区別を行う。コンクリートの変状画像をラベル付き教師データとする物体検出あるいはセグメンテーションも、広義においては深層学習による教師あり異常



検知といえる。

社会インフラのコンクリート構造物は、形状が多様で変状画像の撮影条件等の統一が困難なため、画像データから何らかの法則や傾向といったパターンを高精度に見出すのは、難易度が非常に高い。さらに、構造物の変状は、規模や形状および発生部位等の多様性を有していることから、教師データから得られる特徴量に類似しない未知の異常に遭遇する可能性もあり、その場合は検出が困難になる。そのため、「教師あり学習モデル」以外の、データの収集およびデータセットの構築が容易な正常データを用いた学習の有効性に関する検証を行うことは、今後の深層学習等の活用による効率化の促進を見据えると意義があると考えられる。

### 5-1-2 検討に用いる異常検知手法

正常データを用いた学習による異常検知手法のうち、画像を対象とする異常検知の代表的な手法として、教師なし学習の Anomaly Detection with Generative Adversarial Networks<sup>3)</sup> (以下, AnoGAN と称す.) や 1 クラス分類の Deep One-Class Classification<sup>4)</sup> (以下, DOC と称す.) が挙げられる。本章では、この 2 手法を用いて、物体検出あるいはセグメンテーション適用時の検出漏れ (False Negative) に対する補完の実現性について検証を行った。

深層学習による異常検知は、**図 2-1** に示した教師あり学習と同様に、データからの学習により最適なモデルを構築する「学習フェーズ」と、そのモデルに学習データ以外の未知のテストデータを入力して異常検知を行う「推論フェーズ」の 2 つの工程に分けられる。2 手法の異常検知の流れを、上段側に学習フェーズ、下段側に推論フェーズに分割した形で**図 5-1** に示す。

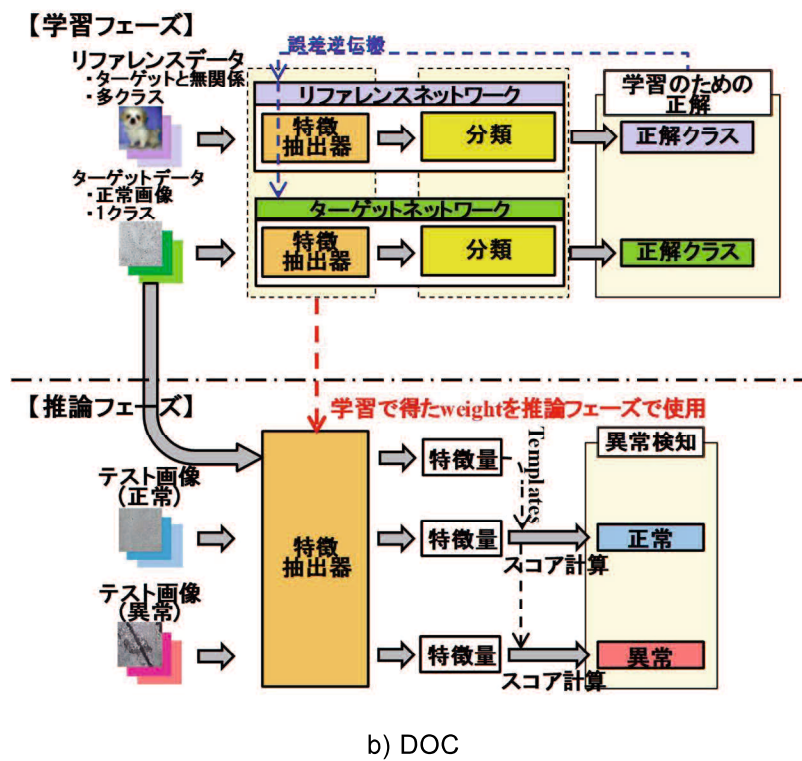
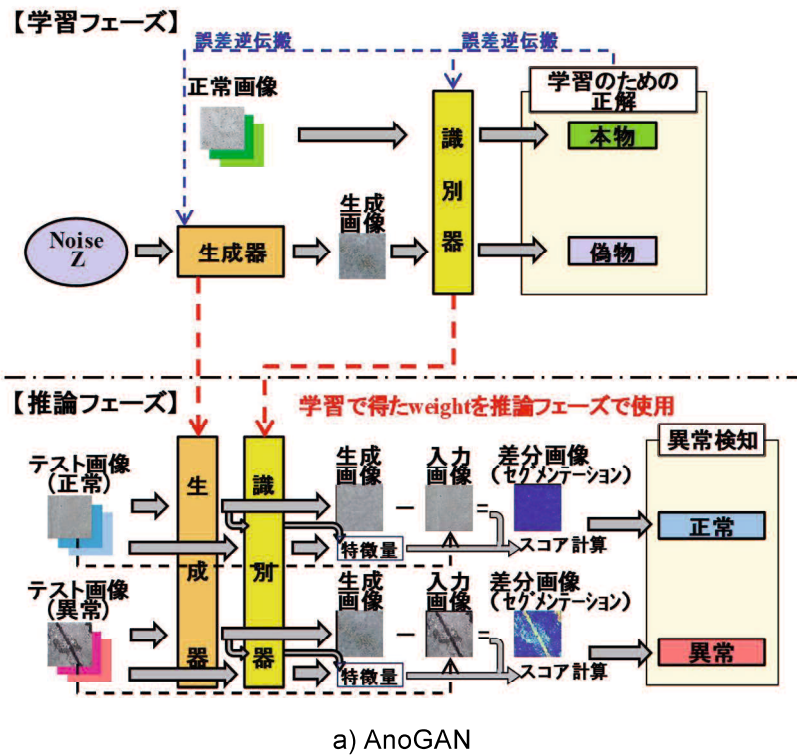


図 5-1 適用手法の異常検知の流れ

### (1) AnoGAN

AnoGANは、敵対的生成ネットワークと呼ばれるGAN (Generative Adversarial networks)<sup>5)</sup>を、一番最初に異常検知に取り入れた手法である。GANでは、学習に用いる本物の画像に類似した偽物の画像を生成する生成器のモデルと、本物と偽物を識別する識別器のモデルを同時に学習させる。従来のGANでは、生成器は識別器が偽物を本物と間違える確率を、識別器は本物と偽物を正しく識別する確率を、それぞれ最大化させるような損失関数により反復学習を行って、リアリティのある偽物画像を生成している。一方、AnoGANは、**図5-1 a)**の学習フェーズに示すようにモデルの構造は従来のGANと同様であるが、損失関数がGANとは異なる。AnoGANでは、正常画像である入力画像と生成器が生成した画像との視覚的非類似性、入力画像と生成画像を識別器に入力したときの中間層における特徴量の非類似性の2項の加重和を損失関数として学習を行う。そして、生成器は正常画像の特徴空間の分布を適切に捉えた画像生成の能力を、識別器は正常画像の特徴表現を適切に踏まえた特徴抽出器としての能力を、それぞれ習得する。推論フェーズでは、未知のサンプルの入力に対して、対応する正常画像を生成器により生成させて、入力画像と生成画像の視覚的非類似性と、識別器により抽出された特徴量の非類似性によりスコアを算定し、異常を識別する。

### (2) DOC

DOCは、1クラス分類を応用した異常検知手法である。主に異常画像検出や特異点検出での使用に特化しており、AnoGANのように画像内の異常領域をセグメンテーションする機能は有していない。教師あり学習における多クラス分類のタスクでは、クラス間距離を最大に、クラス内分散を最小にすることを目的として特徴が学習されるが<sup>6)</sup>、正常画像1クラスのみでは、そのような学習は困難である。そのため、DOCでは、**図5-1 b)**の学習フェーズに示すように、ターゲットとする正常画像とは無関係の、多クラスのデータセットであるリファレンスデータを用いる。そして、ターゲットネットワークとリファレンスネットワークの2つのネットワークで並列的に学習を行う。学習は、ターゲットネットワークにおけるコンパクト性損失とリファレンスネットワークにおける説明性損失の2項の加重和を損失関数として行う。ターゲットネットワークでは、ターゲットとする正常クラス内の画像が類似の特徴表現を持つように、クラス内分散を最小に抑えた特徴抽出能力を学習する。リファレンスネットワークでは、ターゲット以外のクラスの画像が異なる特徴表現を持つように、クラス間距離を最大化する分類能力を学習

する。これらにより、「正常」の識別に最適な特徴抽出能力を習得した特徴抽出器が構築される。なお、リファレンスネットワークは、より高い分類能力を獲得するために、大規模な画像データセットを学習したモデルをもとに Fine-tuning を行う。推論フェーズでは、学習により得られた特徴抽出器を用いて、ターゲットデータである正常画像と未知のサンプルから特徴を抽出させて、双方の特徴のマッチングによりスコアを算定し、異常を識別する。

## 5-2 変状検出への異常検知手法の適用

### 5-2-1 検討内容

#### (1) 検討内容

検討は、健全なコンクリート面の正常画像による学習を行った後に、学習に使用しないテストデータに対して異常検知を実施し、その結果を評価することにより行った。学習フェーズでは、図 5-2 に例を示した 1,700 枚の正常画像を学習データとして用いた。以降、本章では、学習データのうち教師あり学習に用いるラベル付けされたデータを「教師データ」、それ以外のラベル付けされていない学習データを「学習データ」と称する。

また、推論フェーズでは、表 5-2 に示す 8 クラスに分類したコンクリート変状の「ひ

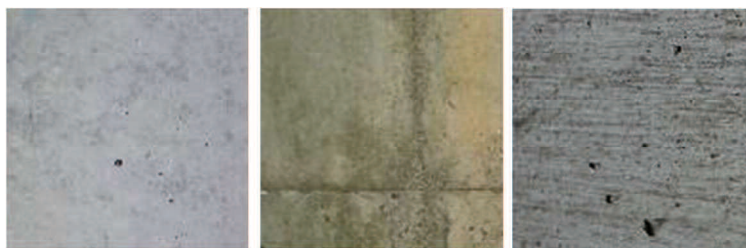


図 5-2 学習データの例

表 5-2 対象変状の程度の区分と分類クラスの関係

損傷種類	区分	一般的状況 <sup>7)</sup>	クラス No.	分類
ひびわれ	a	損傷なし	—	異常
	b	最大ひびわれ幅：小 最小ひびわれ間隔：小	1	
	c	最大ひびわれ幅：小 最大ひびわれ幅：中 最小ひびわれ間隔：大 最小ひびわれ間隔：小		
	d	最大ひびわれ幅：中 最大ひびわれ幅：大 最小ひびわれ間隔：大 最小ひびわれ間隔：小		
	e	最大ひびわれ幅：大 最小ひびわれ間隔：大		
剥離・鉄筋露出	a	損傷なし	—	異常
	b	—	—	
	c	剥離のみが生じている。	2	
	d	鉄筋が露出しており、鉄筋の腐食は軽微である。	3	
	e	鉄筋が露出しており、鉄筋が著しく腐食又は破断している。	4	
※ 漏水・遊離石灰	a	損傷なし	—	異常
	b	—	—	
	c	ひびわれから漏水が生じている。 錆汁や遊離石灰はほとんど見られない。	5	
	d	ひびわれから遊離石灰が生じている。 錆汁はほとんど見られない。	6	
	e	ひびわれから著しい漏水や遊離石灰（例えば、つらら状）が生じている。 漏水に著しい泥や錆汁の混入が認められる。	7 8	
損傷なし	健全なコンクリート面	9	正常	

※) 打継目や目地部から生じる漏水・遊離石灰についても、ひびわれと同様の扱いとする。

びわれ」, 「剥離・鉄筋露出」, 「漏水・遊離石灰」と, 学習データと同様の健全なコンクリート面を含めた全9クラスの画像をテストデータとして用いた. なお, テストデータは, 撮影範囲がコンクリート面のみで, 正対に近い角度で撮影した画像のみとし, 各クラス80枚前後の画像を使用した. このテストデータを入力として, 異常スコアの算定および異常領域のセグメンテーションを行って, 「正常」か「異常」かの識別のテストを行う.

## (2) 各手法の実行条件

AnoGANのモデル構造は, 文献<sup>3)</sup>ではDCGAN<sup>8)</sup>の構造を踏襲しており, 本章における検討でも同様に文献<sup>8)</sup>の構造を踏襲した. また, 最適化アルゴリズムは生成器, 識別器ともにAdam<sup>9)</sup>を用いた.

DOCの識別モデルは, ImageNet<sup>10)</sup>で学習した重みによるFine-tuningを行ったVGG16<sup>11)</sup>を用い, 最適化アルゴリズムは確率的勾配降下法(SGD)<sup>9)</sup>を用いた. リファレンスデータは, 10クラスの60,000のカラー画像で構成されたCIFAR-10<sup>12)</sup>からランダムに抽出した10,000枚を用いた. また, DOCはAnoGANのような画像内の異常領域をセグメンテーションする機能は有していない. そのため, 文献<sup>13)</sup>を参考にして, 現行の特徴抽出器をEncoderに見立てて, Decoderを追加して情報の圧縮と圧縮情報の復元を行うAutoEncoderと同様の機構により, 正常画像を模倣した画像を生成させることとした. そして, AnoGANと同様に差分画像による異常領域のセグメンテーションを実施した. DOCへの異常領域のセグメンテーション機能の付加方法を図5-3に示す.

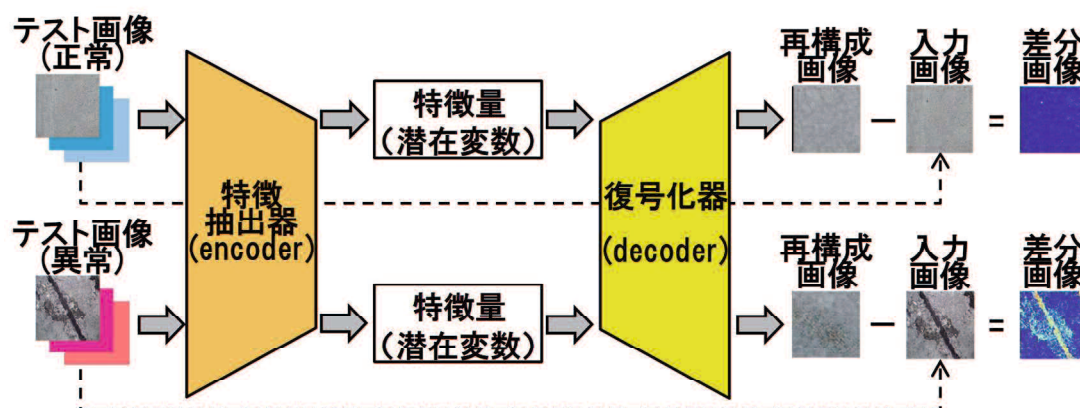


図 5-3 DOC への異常領域のセグメンテーション機能の付加方法

## 5-2-2 適用結果

## (1) 異常スコアの算定結果

異常スコアは、各画像データの異常値のレベルを表す無次元の数値で、AnoGAN と DOC とでは算出方法も値も全く異なる。

AnoGAN における異常スコア  $A(x)$  は、式(1)により算定する。

$$A(x) = (1 - \lambda) \cdot R(x) + \lambda \cdot D(x) \quad (1)$$

$$R(x) = \sum |x - G(z_\gamma)| \quad (2)$$

$$D(x) = \sum |f(x) - f(G(z_\gamma))| \quad (3)$$

ここで、 $R(x)$  は式(2)で計算される入力画像（テスト画像） $x$  と生成画像  $G(z_\gamma)$  との間  
の視覚的非類似性を測定した残差スコアである。 $G(z)$  は、潜在変数  $z$  を生成器  $G$  に入力  
して得られる生成画像を意味する。 $z_\gamma$  は潜在空間の分布からランダムにサンプリングさ  
れた  $z$  について、反復プロセスを経て正常画像に最も類似した生成画像  $G(z_\gamma)$  が得られ  
るように最適化された潜在変数である。 $D(x)$  は式(3)で計算される識別器  $D$  の中間層か  
ら抽出した特徴量に関する入力画像  $x$  と生成画像  $G(z_\gamma)$  との間の非類似性を測定した識  
別スコアで、 $f(\cdot)$  は入力に対応する識別器  $D$  の中間層から特徴量を抽出する関数である。  
なお、式(1)中の  $\lambda$  は文献<sup>3)</sup>を踏襲して 0.1 とした。

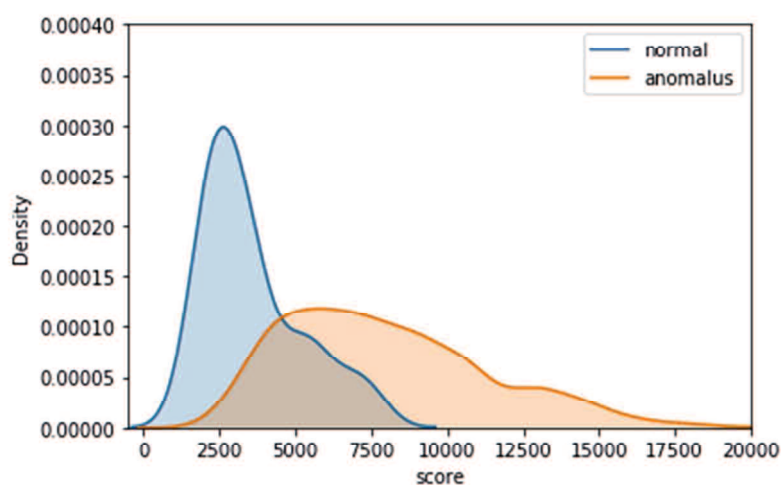
DOC における異常スコア  $S_x$  は、式(4)により算定される。

$$S_x = f(g(x) | g(t_1), g(t_2), \dots, g(t_n)) \quad (4)$$

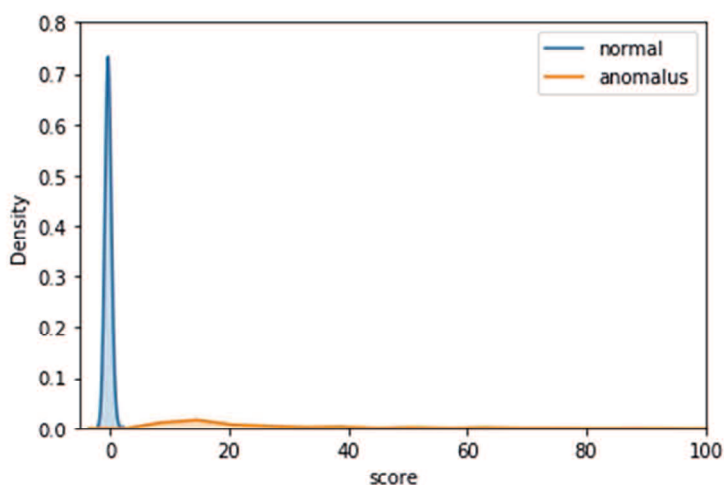
ここで、 $x$  は入力画像（テスト画像）、 $t$  は学習に用いたターゲットデータの正常画像、  
 $f(\cdot)$  はマッチング関数、 $g(\cdot)$  は入力に対応するそれぞれの深層特徴である。本章におけ  
る検討では、外れ値検知技術のひとつである LOF（Local Outlier Factor）<sup>14)</sup> をマッチング  
関数に用いた。LOF は、データ間の距離から空間における局所的なデータの密度を推定  
し、その局所密度に着目して外れ値検知を行う手法である。領域によってデータの分布  
が異なり、特定の閾値を設定することが困難であるような複雑な条件に対しても対応で  
きる点が、LOF の特徴である。

推論フェーズにおける、正常画像 88 枚と異常画像 780 枚のそれぞれの手法の異常ス  
コアの算定結果を図 5-4 に示す。なお、正常画像と異常画像のサンプル数に大きな隔た

りがあり、ヒストグラムではスコアの分布の対比が困難であったため、縦軸を確率密度、横軸をスコアとした確率分布により算定結果を描画した。確率密度関数の推定は、モデルを使用せずに直接的に確率密度関数を推定するノンパラメトリック法の一つである、カーネル密度推定を用いた。DOC では正常画像のスコアが低分散で分布して正常と異常が明瞭に区分できているのに対して、AnoGAN では正常と異常の区分が不明瞭な帯域が広く分布しており、正常と異常の区分の明瞭さにおいて、対照的な結果となった。



a) AnoGAN



b) DOC

図 5-4 異常スコアのカーネル密度推定による分布



(2) 異常領域のセグメンテーション結果

各クラスの異常領域のセグメンテーション結果の一例を表 5-3 に示す. AnoGAN では, 異常領域のセグメンテーションを明瞭に行えており, 正常画像に対しては誤検出を生じさせていない. ただし, 剥離・鉄筋露出 d の画像においては打継目を, 剥離・鉄筋露出 e の画像においてはチョーク跡を, 差分画像においてそれぞれ検出している. 学習画像には打継目やチョーク跡が写っている画像も含めたが, 生成画像ではそのような部分までは表現できていないので, 学習データを充実させて生成画像の熟度を上げていく必要がある.

これに対して, DOC では健全なコンクリート面の画像をリアルに生成できていないことから, 差分画像におけるセグメンテーションは良好に行えていない. DOC は「正常」とそれ以外との距離に着目した特徴量の抽出を行っているため, 異常スコアについては「正常」と「異常」とを概ね分離した形で算定を行えた. しかし, 画像の生成を目的とした特徴量ではないため, 正常画像の特徴空間を十分に捉えることはできていないものと考えられる.

表 5-3 各クラスの異常領域のセグメンテーション結果の例

No.	1	2	3	4	5	6	7	8	9
	異常								正常
クラス	ひびわれ	剥離・鉄筋露出			漏水・遊離石灰				損傷なし
		c	d	e	c	d	e-1(つらら)	e-2(錆汁)	
入力画像									
AnoGAN	生成画像								
	差分画像のヒートマップ								
	異常スコア	2014	2687	6183	5901	6065	10187	8371	5394
DOC	生成画像								
	差分画像のヒートマップ								
	異常スコア	0.25	0.30	2.39	373.71	395.64	24.89	12.29	1.75

5-2-3 適用手法によるモデルの性能評価

本項では、5-2-2 項において異常スコアの算定および異常領域のセグメンテーションを行った AnoGAN と DOC のそれぞれについて性能評価を行い、両手法の特性について整理を行う。

(1) ROC 曲線による評価

最初に、ROC 曲線 (receiver operating characteristic curve) により定まる指標について評価を行った。ROC 曲線は、閾値を変化させたときの偽陽性率 (False Positive Rate : X 軸) と真陽性率 (True Positive Rate : Y 軸) の変化をプロットしたものである。偽陽性と真陽性は、表 5-4 に示す混同行列より設定される指標で、偽陽性率は  $FP / (FP + TN)$ 、真陽性率は  $TP / (TP + FN)$  で算出される。偽陽性率は「正常」が正解であるものを間違えて「異常」と判定した割合、真陽性率は「異常」が正解であるものを正しく「異常」と判定した割合である。理想的なモデルは、間違えて「異常」と判定する偽陽性率が低く、

表 5-4 混同行列

		予測クラス	
		Positive (予測結果が異常)	Negative (予測結果が正常)
正解クラス	Positive (正解クラスが異常)	真陽性 TP (True Positive) ※正しい異常検知	偽陰性 FN (False Negative) ※検知漏れ
	Negative (正解クラスが正常)	偽陽性 FP (False Positive) ※誤検知	真陰性 TN (True Negative)

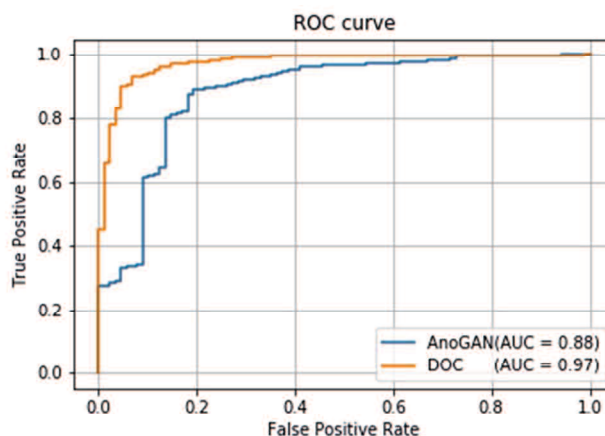


図 5-5 AnoGAN および DOC の ROC 曲線

なおかつ正しく「異常」と判定する真陽性率が高いモデルである。グラフ中の ROC 曲線と X 軸に囲まれた範囲は、AUC (area under curve) と呼ばれる異常検知器の指標となる領域で、0 から 1 の範囲の値となる。この AUC の値が 1 に近いほど、偽陽性率が低くなる閾値であっても真陽性率が高い数値であることを意味し、異常検知の性能が高いモデルとなる。AnoGAN および DOC の ROC 曲線を図 5-5 に示す。AnoGAN が AUC=0.88、DOC が AUC=0.97 でこの検討における異常検知器の性能は、DOC の方が優位であった。

## (2) 抽出特徴量の類似度

次に、t-distributed Stochastic Neighbor Embedding<sup>15)</sup> (以下、t-SNE と称す) により把握した抽出特徴量の類似度について評価を行った。t-SNE は、可視化を目的として次元削減を行う手法で、確率分布に基づくアプローチにより、高次元空間上の類似度を、2 次元または 3 次元の低次元空間上の類似度に埋め込む手法である。この t-SNE を用いて、それぞれの手法の抽出特徴量の類似度を比較した結果を図 5-6 に示す。

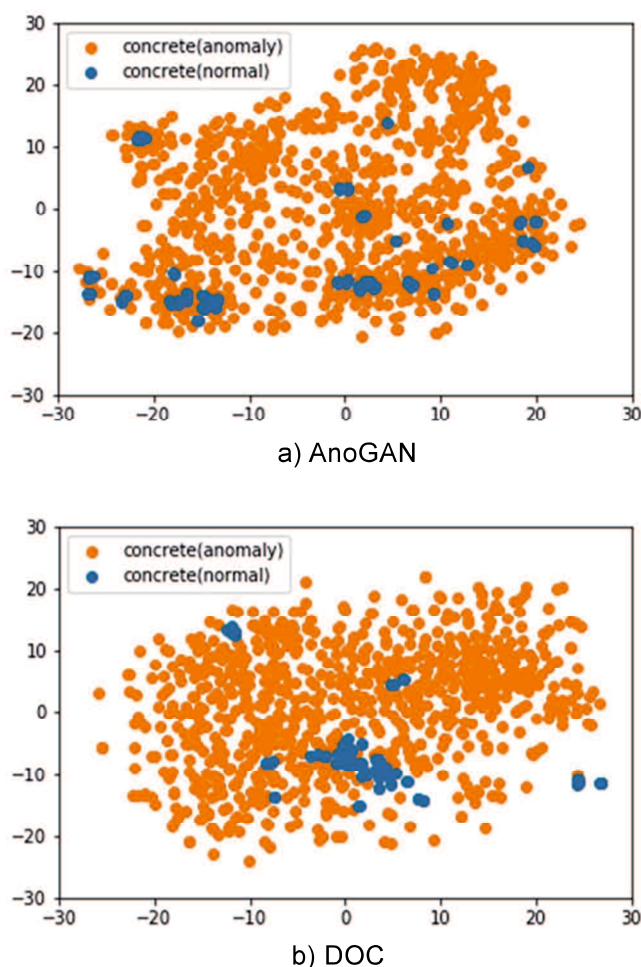


図 5-6 t-SNE による抽出特徴量の可視化結果

DOC は、Fine-tuning を行った上で、リファレンスデータも使って特徴量抽出の学習を行っている。そのため、「正常」と「異常」が比較的分離された状態で特徴量の抽出が行えている。これに対して、AnoGAN は、「正常」のクラス内分散を抑制した状態での特徴量抽出は行えていない。AnoGAN の異常スコアは、視覚的非類似性と特徴的非類似性の2つの項で構成されているが、これにより、特徴的非類似性が適正な値となっておらず、異常スコア上の「正常」と「異常」との区分が不明瞭になったものと考えられる。AnoGAN およびその応用手法<sup>16)-18)</sup>などでは、医用画像や製造業の外観検査への適用事例の報告が多く見られる。これらの事例における画像は、撮影条件や被写体がある程度統一されている。一方、コンクリート構造物の場合には、距離や照度等の撮影条件、コンクリート面の色調や質感等の外観条件等が様々である。そのため、正常画像1クラスによる学習で、複雑で高精度な特徴量の抽出を行うには、より一層のデータの拡充が必要と考えられる。

### (3) 各手法の特性

AnoGAN と DOC を、コンクリート面の変状検出に適用させた結果の評価総括を表 5-5 に示す。画像生成のモデルである GAN をベースとした AnoGAN と距離に着目した特徴量抽出を行う DOC とでは、正常画像の特徴を捉えた画像生成の機能に大きな差がある。そのため、異常領域のセグメンテーションによる変状の検出においては、AnoGAN に優位性がある。しかし、生成画像において、打継目や気泡およびチョーク跡等のような変状以外の相違点については生成しきれておらず、差分画像において異常領域としてセグメンテーションされる結果となっている。そのため、セグメンテーションの前段として、「正常」か「異常」かを本質的に判定する工程が必要である。この点においては、「正常」と「異常」とを明瞭に分離した状態で特徴量の抽出が行えている DOC に優位性がある。

以上より、AnoGAN によるセグメンテーションと DOC の「正常」と「異常」の識別を併用することで、正常画像を用いた異常検知手法による良好なコンクリート変状の検出が実現できると考えた。

表 5-5 AnoGAN と DOC の適用結果の評価総括

		AnoGAN	DOC
異常検知 実行結果	画像における 異常領域の セグメンテーション	正常画像の特徴を捉えたリアルな画像を生成できている。 差分画像において明瞭な異常領域のセグメンテーションを行えている。 (より多様な学習画像を追加することで、生成画像の熟度を向上させることが可能)	正常画像の特徴を捉えたリアルな画像を生成できていない。 差分画像において明瞭な異常領域のセグメンテーションは行えていない。
	異常スコア	「正常」と「異常」のスコアの分布の分離が不明瞭。	「正常」と「異常」のスコアの分布の分離が明瞭。
性能評価 結果	抽出特徴量 の分布	「正常」と「異常」の特徴量の分布の分離が不明瞭。	クラス間距離の最大化、クラス内分散の最小化を実現した形での特徴量抽出が行えている。
	ROC曲線 (AUC)	AUC=0.88で、異常検知器の性能は比較的高い。	AUC=0.97で、異常検知器の性能は非常に高い。
総合評価		異常領域のセグメンテーションは行えているが、スコアおよび特徴量については、「正常」と「異常」を明瞭に分離できていない。	スコアおよび特徴量について、「正常」と「異常」を比較的明瞭に分離できているが、画像生成に適した特徴量の学習は行えていない。

### 5-3 実用性の検証

5-2-2 項において、正常画像を用いた学習に基づく異常検知によるコンクリート変状の検出について、一定の成果が得られた。そして、5-2-3 項では、その結果の評価を行い、有効な検出のための方法論について検討を行った。そこで、5-3 節では、構造物点検における実務的な適用について検証を行う。

#### 5-3-1 検証内容

##### (1) 広域画像への異常検知手法適用の流れ

5-2 節では、比較的近傍で撮影した単一クラスの変状のみが存在する画像を対象に変状の検出を行った。しかし、実際の点検業務への適用に際しては、広域的な画像内に混在する複数の変状について、それぞれを検知する必要がある。広域画像への異常検知手法適用の流れを図 5-7 に示す。まず始めにテスト画像をグリッド状に分割し、それぞれの分割画像（以下、セルと称す）を、AnoGAN と DOC の双方に入力する。AnoGAN では、正常画像のみで学習した生成器によって生成された画像とテスト画像との差分画像によりセグメンテーションを行い、DOC では、学習により得た正常画像の特徴とテスト画像の特徴とのマッチングで異常スコアを算定する。そして、双方の結果を融合させて、コンクリート変状が生じているセルの検出とそのセル内の変状領域のセグメンテーションを行う。

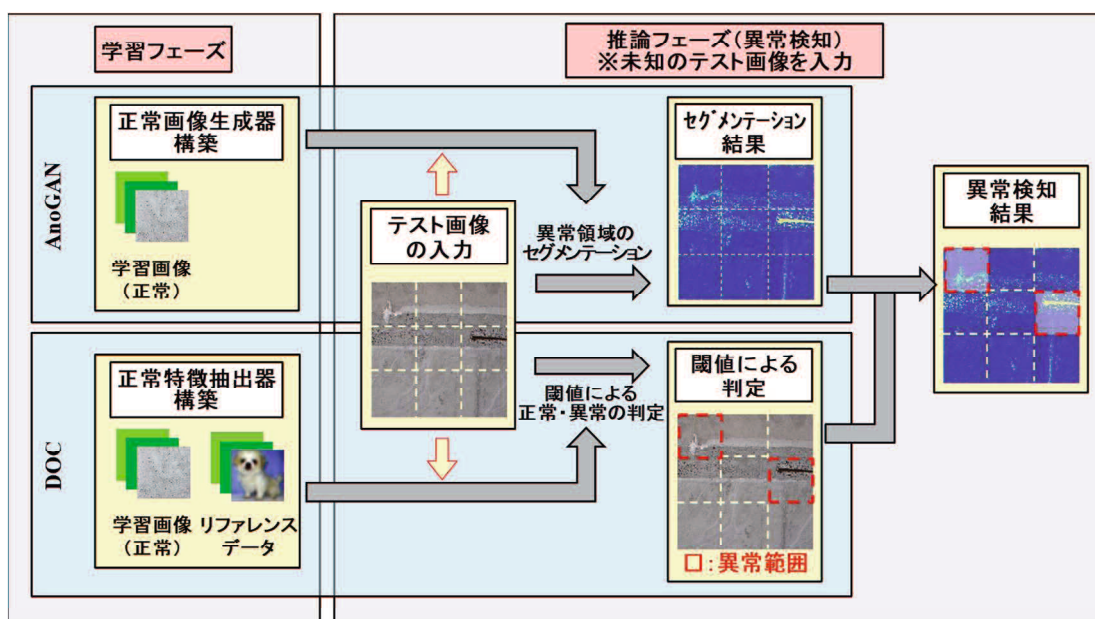


図 5-7 広域画像への異常検知手法適用の流れ

(2) 閾値の算定

DOCによる各セルに対する異常画像検出は、各セルの異常スコアが閾値未満であるか、閾値以上であるかにより行うこととした。閾値は、5-2節のテストデータに対して、閾値を変化させた時の真陽性率と真陰性率の推移の観測結果から設定を行った。閾値を変化させた時の真陽性率と真陰性率の推移を表5-6に示し、カーネル密度推定による正常画像のスコアの分布と閾値の関係を図5-8に示す。閾値を $\mu+1\sigma$ とした条件では、約半分の正常画像において「異常」と誤判定しており、偽陰性率が43.2%であったが、 $\mu+3\sigma$

表 5-6 閾値を変化させた時の真陽性率と真陰性率の変化の推移

閾値	判定	1	2	3	4	5	6	7	8	9	平均 正解率
		異常								正常	
		ひびわれ	剥離・鉄筋露出			漏水・遊離石灰				損傷なし	
	c	d	e	c	d	e(つらら)	e(錆汁)				
$\mu + 1\sigma$	正常	3.3%	0.0%	0.0%	0.0%	2.0%	0.0%	0.0%	0.0%	56.8%	94.6%
	異常	96.7%	100.0%	100.0%	100.0%	98.0%	100.0%	100.0%	100.0%	43.2%	
$\mu + 3\sigma$	正常	30.0%	5.0%	4.0%	0.0%	3.0%	5.0%	1.4%	1.1%	93.2%	93.7%
	異常	70.0%	95.0%	96.0%	100.0%	97.0%	95.0%	98.6%	98.9%	6.8%	
$\mu + 5\sigma$	正常	61.7%	7.0%	12.0%	0.0%	9.0%	21.0%	1.4%	1.1%	97.7%	87.2%
	異常	38.3%	93.0%	88.0%	100.0%	91.0%	79.0%	98.6%	98.9%	2.3%	

■：真陽性率， ■：真陰性率

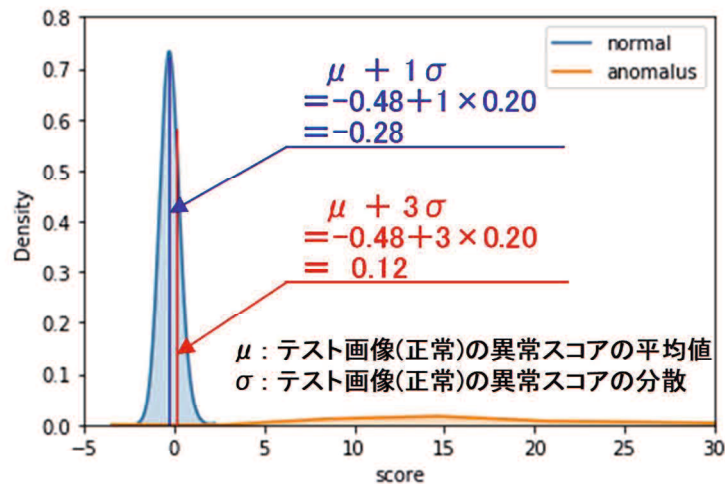


図 5-8 カーネル密度推定による分布と閾値の関係

とした条件では、偽陰性率は6.8%まで低減した。また、コンクリートの変状画像については、閾値を $\mu+1\sigma$ とした条件では、正しく「異常」と判定した真陽性率が全クラスについてほぼ100%で、 $\mu+3\sigma$ まで上げても、ひびわれ以外のクラスは、いずれも95%以上となる。以上の結果より、閾値を $\mu+3\sigma$ とすることが妥当であると判断した。ただし、ひびわれのような微細な変状は、他の面的な広がりを持つ変状と比較して異常スコアが低い値となっている。そのため、閾値 $\mu+1\sigma$ では96.7%であった真陽性率が、 $\mu+3\sigma$ では70.0%まで低下している。微細な変状も検知するためには、構造物の状況等に応じて閾値を下げる、あるいはセルの分割を細かくして感度を上げる等の対応が必要と考えられる。

### 5-3-2 検証結果

PC桁の主桁と横桁で区切られた1区画に、鉄筋露出や遊離石灰の析出が複数発生している画像を対象として、広域画像への適用についての検証を行った。広域画像への適用結果を表5-7に示す。

上段は、左に対象画像、右に道路橋点検士有資格者により変状の領域の正解を付記した画像を、それぞれ表示している。この対象画像の異常検知では、水平方向に9分割、鉛直方向に6分割した全54セルの画像を入力とした。

中段は、2つの異常検知手法のそれぞれの個別の実施結果を表示している。AnoGANの結果では、剥離・鉄筋露出と漏水・遊離石灰の領域とともに、主桁と横桁の境界、横桁と床版の境界もセグメンテーションしている。DOCの結果では、各セルの異常スコアを併記しており、赤字と青字が閾値( $\mu+3\sigma$ )以上のスコアで、赤字が特にスコアが高かった $\mu+5\sigma$ 以上のスコアのセルである。変状が生じている領域は、いずれも $\mu+5\sigma$ 以上のスコアとなっており、異常領域として明瞭に識別を行えている。しかし、画像中の四隅の主桁と横桁の境界部、汚れがある右側の横桁も $\mu+5\sigma$ 以上のスコアと評価している。学習画像には打継目が写っている画像も含めたが、このような斜め方向の継目や汚れが写った画像は含んでいないため、異常スコアを高く評価したものと考えられる。より精度の高い識別を行えるように、学習画像に多様な正常画像を含めて、このような画像に対する感度を下げる必要がある。その他の誤検知(False Positive)したセルは、異常スコアが $\mu+3\sigma$ 以上 $\mu+5\sigma$ 未満であったセルで、閾値は超えたものの極端に高いスコアには評価していない領域である。閾値を上げてこのようなセルを検知させない対応を行うことも可能であるが、誤検知以上に重大な問題となる検知漏れ(False Negative)につながる可能性がある。本章における検討の目的は、物体検出あるいはセグメンテーショ


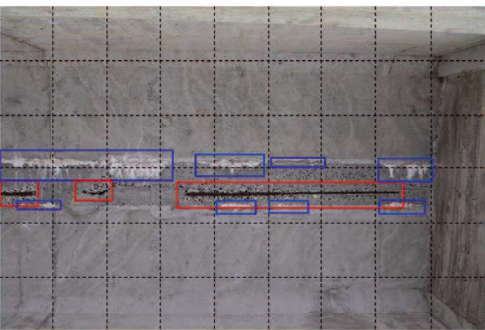
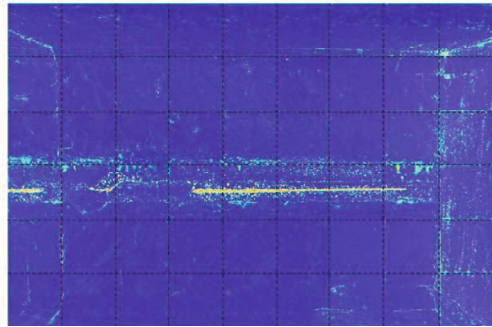
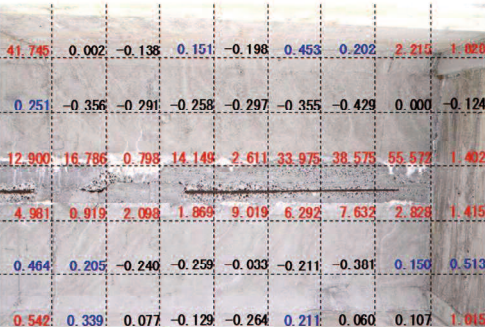
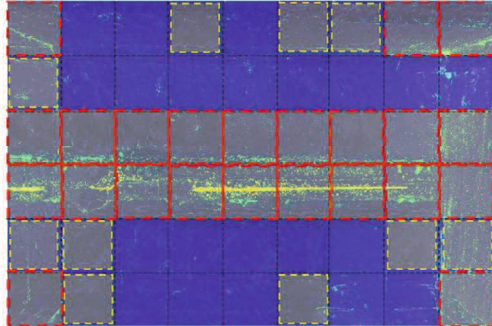
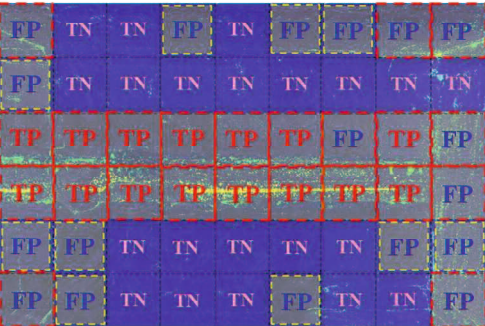


ン適用時の検出漏れに対する補完の実現であるため、閾値に余裕をもたせて誤検知は許容することとした。

下段は、2 手法の評価を融合させた結果である。左側は DOC において閾値を超えたセルの異常領域のセグメンテーション結果を際立たせた画像で、右側はそれに混同行列による分類を併記した画像である。誤検知である「FP」は前述のように数セル見られるが、検知漏れである「FN」は生じさせていない。

以上より、AnoGAN と DOC を併用することで、コンクリート変状の領域を適切に捉えることができた。収集が容易な正常画像を用いた深層学習による異常検知が、物体検出あるいはセグメンテーション適用時の変状の検出漏れを補完可能であることが検証できた。

表 5-7 広域画像に対する異常検知の適用結果

入出力結果	
対象画像	  <p>【正解領域】 □ : 剥離・鉄筋露出d □ : 漏水・遊離石灰d</p>
異常検知個別実施結果	  <p>赤字：異常スコア <math>&gt; \mu + 5\sigma</math> のセル 青字：異常スコア <math>&gt; \mu + 3\sigma</math> のセル</p> <p>AnoGAN DOC</p>
2手法併用による評価	  <p>TP : 真陽性, TN : 真陰性, FP : 偽陽性, FN : 偽陰性</p> <p>検知結果の混同行列による分類</p> <p>異常と判定した領域：スコア <math>&gt; \mu + 5\sigma</math> 異常と判定した領域：スコア <math>&gt; \mu + 3\sigma</math></p>

## 5-4 まとめ

本章では、第3章で検討を行った独自のセグメンテーション手法の補完的技術として、正常画像を用いた学習による異常検知の適用について検討を行った。本章における検討により、以下の点を明らかにした。

- (1) コンクリート面の画像は、異常検知の適用事例が多い医用画像や製造業の外観検査等のような、ある程度統一された画像ではない。しかし、このような画像に対しても、深層学習の生成モデルを応用したAnoGANを用いた異常検知により、コンクリートの変状領域の検出を行えることが確認できた。ただし、特徴量および異常スコアの分布を確認した結果、正常と異常を明瞭に区別する性能は十分な水準ではない。
- (2) 深層学習の識別モデルを応用したDOCは、クラス間距離を最大に、クラス内分散を最小とする特徴の抽出を学習することにより、明瞭に、コンクリートの「正常」と「異常」の識別を行えることを明らかにした。ただし、識別を主目的とした特徴量の抽出であるため、画像に対する変状領域の検出性能は十分な水準ではない。
- (3) AnoGANとDOCそれぞれ単独では、画像からの異常領域の検出性能は不十分であるが、併用することで、コンクリートの変状領域を適切に捉えることが確認できた。複数の変状が混在する広域画像への適用を行った結果、誤検出は発生したが、検出漏れを発生させずに異常領域を検知できた。正常画像を用いた深層学習による異常検知が、物体検出あるいはセグメンテーション適用時の変状検出漏れを補完可能であることを明らかにした。

異常検知の適用について、今後の展望を述べる。打継目やチョーク跡等を含む画像に対して、AnoGANではこれらを表現した画像を生成できなかった。そして、DOCでは斜め方向の継目や汚れを含む画像等に対して、異常スコアを高く評価した。誤検出を抑制するためには、正常画像のデータの量とパターンを増加させて、異常検知の性能を上げることが望ましい。本章における検討の推論フェーズでは、撮影範囲がコンクリート面のみの画像を用いたが、風景等も含んだ画像に対しては、セマンティックセグメンテーションを応用したコンクリート面以外の領域のマスク、あるいは深度画像を用いて一定量以上の深度の領域をマスクさせる等の対応が必要である。また、異常検知では、変状のクラス分類までは行えない。そのため、多段的処理になるが、DOCにより「異常」と判定された領域に対して、難易度の低いタスクであるクラス分類をCNNにより行うことも考えられる。

## 第5章 参考文献

- 1) 井手剛：入門 機械学習による異常検知，コロナ社，2015.
- 2) Chalapathy, R., Chawla, S. : Deep Learning for Anomaly Detection: A Survey, *arXiv: 1901.03407v2*, 2019.
- 3) Schlegl, T., Seebock, P., Waldstein, S. M., Schmidt-Erfurth, U. and Langs, G. : Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, *International Conference on Information Processing in Medical Imaging (IPMI)*, pp. 146-157, 2017.
- 4) Perera, P., Patel, V. M. : Learning Deep Features for One-Class Classification, *arXiv: 1801.05365v2*, 2019.
- 5) Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. : Generative Adversarial Nets., *Advances in Neural Information Processing Systems*, Vol. 27, No. 2, pp. 2672-2680, 2014.
- 6) Belhumeur, P. N., Hespanha, J. P. and Kriegman, D. J. : Eigenfaces vs. fisherfaces: Recognition using class specific linear projection., *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 711-720, 1997.
- 7) 国土交通省 道路局 国道・技術課：橋梁定期点検要領，2019.
- 8) Radford, A., Metz, L. and Chintala, S. : Unsupervised representation learning with deep convolutional generative adversarial networks, *arXiv: 1511.06434v2*, 2016.
- 9) Ruder, S. : An overview of gradient descent optimization algorithms, *arXiv: 1609.04747v2*, 2017.
- 10) Deng, J., Dong, W. and Socher, R., Li, L. J., Li, K. and Fei-Fei, L. : ImageNet: A Large-Scale Hierarchical Image Database, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 248-255, 2009.
- 11) Simonyan, K., Zisserman, A. : Very Deep Convolutional Networks for Large-Scale Image Recognition, *arXiv: 1409.1556v6*, 2015.
- 12) トロント大学コンピュータ科学科：The CIFAR-10 dataset, Alex Krizhevsky's home page, 2018.4.16., <https://www.cs.toronto.edu/~kriz/cifar.html>, (最終閲覧日：2020.03.01)
- 13) @shinmura0：がんばる人のための画像検査機 —可視化の改良—, Qiita, 2019.3.7., <https://qiita.com/shinmura0/items/c2f7a86b156ebc5c5daa>, (最終閲覧日：2020.03.01)
- 14) Breuning, M. M., Kriegel, H. P., Ng, R. T. and Sander, J. : LOF : Identifying Density-Based Local Outliers, *SIGMOD '00 : Proceedings of the 2000 ACM SIGMOD international conference*

*on Management of data*, pp. 93-104, 2000.

- 15) Van der Maaten, L., Hinton, G. : Visualizing Data using t-SNE, *Journal of Machine Learning Research*, Vol. 9, pp. 2579-2605, 2008.
- 16) Zenati, H., Foo, C. S., Lecouat, B., Manek, G. and Chandrasekhar, V. R. : Efficient GAN-Based Anomaly Detection, *arXiv: 1802.06222v2*, 2019.
- 17) Deecke, L., Vandermeulen, R., Ruff, L., Mandt, S. and Kloft, M. : Anomaly Detection with Generative Adversarial Networks, *Open Review.net*, 2018.
- 18) Schlegl, T., Seebock, P., Waldstein, S. M., Langs, G. and Schmidt-Erfurth, U. : f-AnoGAN: Fast unsupervised anomaly detection with generative, *Medical Image Analysis*, Vol. 54, pp. 30-44, 2019.



## 第 6 章 総括

---

### 第 6 章

6-1 本研究のまとめ

6-2 今後の課題





---

## 第6章 総括

### 6-1 本研究のまとめ

これまで、インフラの老朽化の進行への懸念から予防保全型維持管理への移行が推進されてきたが、橋梁の点検作業の効率化については、長い間具体的な対策が取られていなかった。画像からのコンクリートのひびわれ検出は、多くの研究者によって取り組まれ、一部の施設、たとえばトンネルでは覆工画像計測技術<sup>1)</sup>による変状の自動検出技術等が実用化されていたが、橋梁については自動検出等の技術の導入には至っていなかった。

そのような潮流の中、平成25年の国土交通省および経済産業省による「次世代社会インフラ用ロボット開発・導入検討会」の共同設置<sup>2)</sup>、平成30年の土木研究所による「AIを活用した道路橋メンテナンスの効率化に関する共同研究」の開始<sup>3)</sup>等の取り組みがなされ、橋梁点検における自動化技術導入への期待が高まってきている。しかし、変状の自動検出技術の導入に向けては、社会インフラの点検という事業の性質から高い精度が要求されるため、そのハードルをクリアすることが大きな課題となっている。

本研究では、橋梁定期点検要領に準拠した変状検出を行うとともに、上記の課題を包含する点検実務への運用を見据えて、併用する補完的な技術の検討を行った。橋梁定期点検要領に準拠した変状検出は、具体的には、橋梁定期点検要領に規定されたコンクリートの変状の種類と程度の区分も含めて識別させた上で、セグメンテーション手法によりピクセル単位の任意の領域形状で、画像からの検出を行ったものである。既往の研究では、ひびわれに他の変状も加えた複数の変状を同時に検出する研究は少数で、ひびわれのみの検出に関する研究が大多数を占めていた。さらに、鉄筋の露出および遊離石灰の析出を、変状の進展具合を考慮した評価区分によって分類をさらに細分化し、セグメンテーションにより橋梁定期点検要領に準拠した検出を行う報告は、これまでにはなかった。本研究の意義は、この橋梁定期点検要領に準拠した検出を行う点にある。また、深層学習は、一定のコントロール化に置かれた条件での画像へ適用した場合に比べて、あらゆる多様性を有する条件下の画像への適用では精度が得られていない点を鑑み、セグメンテーションを補完する技術として、深度画像の活用と異常検知技術の適用の検討を行った。

---

本研究で得られた結論を各章ごとに要約し、以下に述べる。

**第1章「序論」**では、我が国における社会インフラの維持管理の現状と課題をまとめるとともに、コンクリートの画像からの変状検出に関する既往の研究動向の整理を行った。その上で本研究の目的と着眼点を整理し、本論文の構成と概要について述べた。

**第2章「深層学習による変状画像のクラス分類」**では、深層学習によるコンクリートの変状検出の基礎検討として、点検時および踏査時の撮影画像を教師データに用いて、変状画像のクラス分類について検討を行った。検出を行う際には、領域の特定とその領域内の対象物のクラス分類が行われる。そこで、難易度が高い物体検出およびセグメンテーションについて検討を行う前に、難易度が低いクラス分類を行って、検出の実現性を検証した。また、深層学習の課題の一つに挙げられるブラックボックス問題への懸念から、テストデータに対するクラス分類実施時の判断根拠の可視化による確認にも取り組んだ。分類結果が正しい結果であったとしても、8クラスの変状の分類では1/8の確率で正解となるため、技術者の判断と同様な特徴の捉え方により判断されているかという点を検証する必要があった。これらの検討により以下の点を明らかにした。

- ・深層学習を目的とした撮影ではない点検時および踏査時の撮影画像を教師データに用いても、深層学習による画像のクラス分類において、変状の種類と程度の区分も含めて識別させた上で高い精度で分類が行えることを明らかにした。
- ・コンクリートの変状に関する教師データが十分な量でない場合でも、一般的な画像についての事前学習済みモデルを用いた転移学習 (Fine-tuning) により、分類精度の格段の向上が図れることを明らかにした。
- ・テストデータに対するクラス分類において、Grad-CAMにより分類時の着目領域の可視化を行なった結果、適切な部分に着目した上でクラス分類の判断を行っていることを立証できた。

**第3章「深層学習による物体検出およびセグメンテーション」**では、クラス分類が適切に行われたことを受けて、難易度がより高いタスクとなる、物体検出手法による検出とセグメンテーション手法による検出について検討を行った。そして、本研究の主目的である、セグメンテーションによるピクセル単位での変状の検出に向けて、複数の手法についての比較評価を行い、コンクリートの変状の検出に対して適性の高い手法を選定するとともに、精度を向上させる方法を検討した。これらの検討により以下の点を明らかにした。

- ・物体検出では、点検時および踏査時の撮影画像を教師データにした条件で、クラス

と矩形領域の双方について妥当な推論結果が得られた。Faster R-CNNとYOLO v3の2手法の比較では、YOLO v3の方が領域の特定に関する精度において優位な結果となり、YOLO v3のコンクリート変状検出についての適性の高さを確認した。

- ・セマンティックセグメンテーションでは、SegNet, U-NetおよびPSPNetの3手法の比較において、U-Netが領域の特定に関する精度において最も優位な結果となり、限られたデータ量でのコンクリート変状のセグメンテーションについて、U-Netの適性の高さを確認した。また、2章のクラス分類では、データ拡張のみでは十分な精度の向上にはつながらなかったが、セマンティックセグメンテーションではデータ拡張が精度の向上に大きく寄与することを明らかにした。
- ・対象とする変状と背景の2クラス分類で実行するという独自の方法でコンクリート変状のセグメンテーションを行うことで、精度の向上が図れ、実際の専門技術者が行う点検に即した変状領域の検出を実現した。具体的には、物体検出ベースのセグメンテーションとして、YOLO v3により検出した矩形領域に対して、U-Netにより該当するクラスの変状と背景の2クラス分類のセグメンテーションを行った。

社会インフラの点検という目的の性質上、変状検出における検出漏れ(False Negative)の発生を最小限に抑制する必要があるが、あらゆる多様性を有するコンクリートの変状画像への物体検出およびセグメンテーションの適用は、まだ難易度が高い。しかし、完全なものになるまで導入できないとなると、技術の発展と導入時の課題の探索の機会も失ってしまうこととなる。そのため、当面は、社会インフラの点検への深層学習技術の本格的導入に向けて、運用方法の工夫や補完的技術の併用について検討を行い、段階的な導入を行っていく取り組みが現実的と考える。検討したセグメンテーション手法について、後段の処理のセグメンテーションについては、2クラス分類とすることで難易度を大幅に低下させて精度の向上を実現した。そのため、前段の物体検出部分の精度により検出が左右される。被写体である変状が画像中心にくるように撮影を行った画像を教師データに用いている影響で、画像中心付近から外れた部分の変状に対しては、検出精度が低下している傾向が見られた。そのため、複数の変状が混在するような広域の画像への適用を見据えて、スライディングウィンドウをオーバーラップさせる形での検出処理を行うことを考えている。このようなテクニック上の処理に加えて、4章、5章では本研究の変状検出の主軸であるセグメンテーションを補完する技術について検討を行った。

第4章「変状検出における深度画像活用の検討」では、第3章で検討を行った独自のセグメンテーション手法について、前段部分の物体検出の処理における統計的な不確かさを排除する目的で、ピクセル単位で深度の情報を有する深度画像を活用した物体検出の

代替手段について検討を行った。また、深度画像から得られる深度の情報を基に、変状の定量的把握を行う処理手法を構築した。これらの検討により以下の点を明らかにした。

- ・凹凸を有する変状では、深度画像を活用することによって変状領域を特定することが可能であることを明らかにした。この結果により、構造物点検の効率化への深度画像活用の道筋が描けた。
- ・深度画像の活用により、深度と画角を基に1ピクセルあたりの尺度を換算し、変状の定量的把握が可能であることを明らかにした。
- ・従来のRGB画像に加えて、画像内の各ピクセルの深度に関する測定値の情報をもつ深度画像を取得できるRGB-Dカメラが、目視点検の省力化および効率化に有効なデバイスであることが確認できた。

第5章「深層学習による異常検知手法を用いた変状検出の検討」では、第4章と同様のセグメンテーション手法を補完する技術として、教師なし学習による異常検知手法を用いた変状検出の検討を行った。この検討により以下の点を明らかにした。

- ・AnoGANとDOCそれぞれ単独では、画像からの異常領域の検出性能は不十分であるが、併用することで、コンクリートの変状領域を適切に捉えることが確認できた。
- ・複数の変状が混在する広域画像への適用を行った結果、誤検出は発生したが、検出漏れを発生させずに異常領域を検知できた。正常画像を用いた深層学習による異常検知が、物体検出あるいはセグメンテーション適用時の変状検出漏れを補完可能であることを明らかにした。

次に、教師データの観点から整理を行う。本研究では、深層学習の教師データとしての使用目的で撮影した画像ではなく、橋梁の点検や踏査等の際に撮影した画像を用いて検討を行った。これらの画像は、画像サイズ、撮影時の距離や角度等についての統一性がなく、必ずしも教師データとしての適性が高いデータセットではなかった。しかし、点検の効率化に向けては広域的な画像への適用が想定されることから、極力コンクリート面に正対に近い角度での撮影画像を優先的に選定するとともに、道路橋点検士有資格者がアノテーションを行うことで、可能な限りデータの質を高めた。

また、限られたデータ量で精度を高めるための方策として、クラス分類ではデータ拡張とFine-tuningを実施し、特にFine-tuningが分類の精度の大幅な向上に寄与することを確認した。セグメンテーションによる変状検出では、データ拡張が検出の精度に寄与することを確認したが、加えて前段の処理で物体検出を行って2クラス分類の処理にタスクの難易度を緩和させることで、精度の向上を図った。

---

---

さらに、補完技術として検討した、深度画像を活用した凹凸領域の検出、正常画像による教師なし学習を用いた異常検知には、異常データの教師データ量に左右されないという側面もあり、教師データの量が限られた検討環境下においても有効な技術と考えられる。

点検や踏査等の際に撮影した画像の特徴的な点は、被写体である変状が画像内の中心付近に存在することと、変状周辺がチョーキングされていることである。データ拡張により平行移動を行って被写体の位置をずらす操作、チョーキングがない画像も教師データに若干含めることは行ったが、画像内の変状の位置およびチョーキングの有無により検出精度が影響することも考えられる。そのため、画像内の縁端部に変状が存在する画像、チョーキングが行われていない画像も教師データに含めておくことは頑健性を高めることで有効と考えられる。

教師データについては、国土交通省により、良質で効率的な教師データ整備のあり方や、点検に関するデータの取得・保存・分析・活用を円滑に行うデータ基盤のあり方について検討が行われており<sup>4)</sup>、今後の教師データの充実化の進展が期待される。

### 6-2 今後の課題

本研究において、画像からの変状自動検出について、深層学習等の最新技術の適用を検討することにより、多くの知見を得ることができた。しかし、本研究の目的に対しては、十分な成果を上げることができたものの、実務への適用に向けては検討が必要な課題もあるため、以下に整理して示す。

#### 1) 対象とする変状種類等の拡充

本研究では、橋梁定期点検要領における全26損傷のうちの3種類のコンクリート変状を対象とした検討であったが、コンクリート部材の他の変状、鋼部材における変状についての展開、さらには診断に供するための変状原因の識別への適用等について検討を行う必要がある。

#### 2) 最新技術のキャッチアップ

深層学習の技術は日進月歩であり、新たな手法等が次々に提案されている。物体検出で採用したYOLO v3、セグメンテーションで採用したU-Netに限っても、改良されたネットワークが既に複数提案されている。そのため、最新の技術についても検証を重ねていく必要がある。なお、3章におけるセグメンテーションでは、必ずしも新しい手法が最良ではない結果も得られたため、目的とするタスクへの適性も見極めつつ、手法を選択していかなければならない。

#### 3) 変状自動検出技術の段階的な導入

次世代社会インフラ用ロボット開発・導入検討会では、橋梁に関する重点分野の一つに、「点検要領に基づく近接目視の代替ができる技術・システム」が、当初は設定されていた。しかし、検証結果等に基づいて、後に「点検要領に基づく近接目視の代替・支援ができる技術・システム」に改められた。現状では、変状自動検出技術に限っても、まだ点検を代替できる技術水準にはないと考えられる。そのため、段階的な導入を見据えて、まずは技術者の負担軽減を目的とした「支援」のレベルでの運用を行い、「代替」のレベルでの運用に向けての課題の探索等を行うことが肝要と考えられる。変状自動検出技術の導入方針例を図6-1に示す。

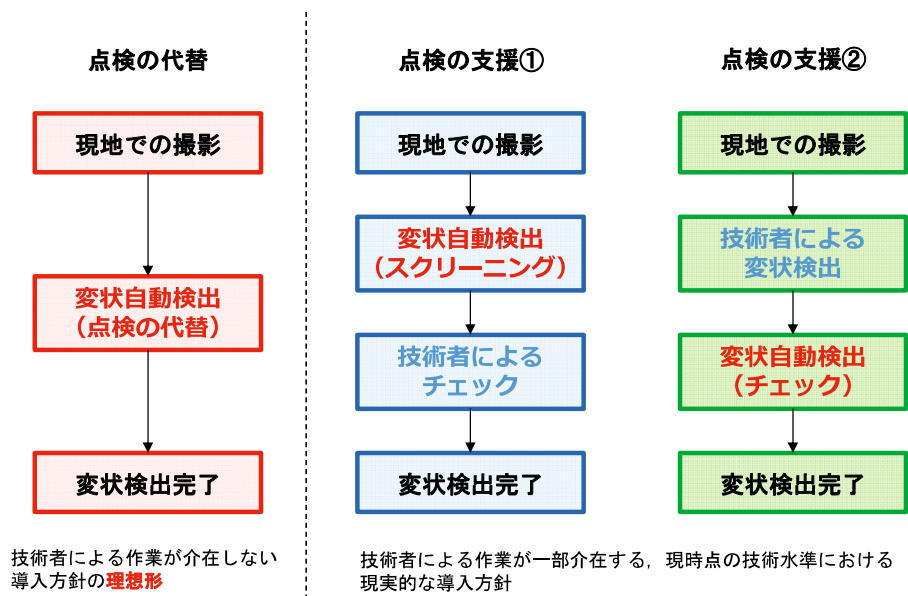


図 6-1 変状自動検出技術の導入方針例

4) 橋梁の点検・診断システムへの拡張

橋梁の点検・診断システムの全体像を図6-2に示す。本研究で行った検討は、システム全体像における「②変状の検出」の部分であり、他の要素技術との入出力等の連携も見据えて、橋梁点検の実務に活用できる技術に昇華していかなければならない。

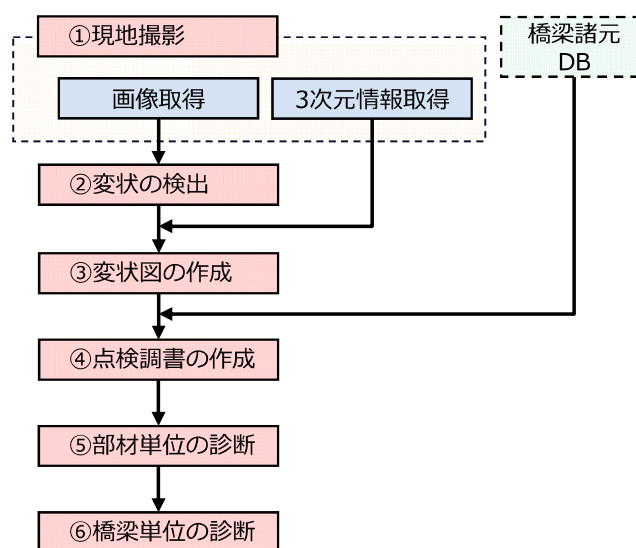


図 6-2 橋梁の点検・診断システムの全体像イメージ

## 第6章 参考文献

- 1) 国土交通省：点検支援技術 性能カタログ（案），pp. 85-104，2019.
- 2) 国土交通省：次世代社会インフラ用ロボット開発・導入検討会，  
[https://www.mlit.go.jp/sogoseisaku/constplan/sosei\\_constplan\\_fr\\_000022.html](https://www.mlit.go.jp/sogoseisaku/constplan/sosei_constplan_fr_000022.html)，（最終閲覧日：2020.05.01）
- 3) 国立研究開発法人 土木研究所：AI を活用した道路橋メンテナンスの効率化に関する共同研究を開始します ～25 者の共同研究者を決定しました～，平成 30 年 11 月 30 日記者発表資料，2018.
- 4) 国土交通省 総合政策局公共事業企画調整課：AI 開発支援プラットフォームの開設準備 WG の設置，平成 30 年 7 月 27 日記者発表資料，2018.



## 謝 辞

本研究の遂行にあたり、本研究に取り組むきっかけを与えてくださり、ご指導をいただきました山口大学大学院創成科学研究科 中村秀明教授に心から感謝の意を表しますとともに、厚く御礼申し上げます。

主査を務めていただいた中村秀明教授には、研究経験の少ない著者に対し、検討方法・学会投稿・口頭発表・本論文の取りまとめに至るまで懇切丁寧なご指導をいただきました。土木構造物の設計や維持管理の業務に携わってきた著者が、情報工学や画像処理に関する知識を全く有していない状態から、画像処理技術に基づく橋梁点検の省力化をテーマに研究したいと申し出た際にも快く応じていただきました。そして、博士の学位取得が目的ではなく取得後も継続して研究活動を行うことの大切さ、未経験分野の研究に対する取組姿勢など、親身にご教授いただきました。ここに記して深甚なる謝意を表します。

本論文を審査していただきました山口大学大学院創成科学研究科 進士正人教授、麻生稔彦教授、河村圭准教授、吉武勇准教授、中村秀明教授には、論文の内容を細部までご精査いただき、適切な助言を賜りました。心より御礼申し上げます。

山口大学工学部 知能情報工学科卒業生(現山口県庁)の堀晃輔氏、山口大学大学院 創成科学研究科修了生(現新光産業株式会社)の河村伸哉氏、山口大学大学院 創成科学研究科博士前期課程の山本拓海氏には、深層学習のプログラミングなどについて、情報工学の知識が乏しい著者をご支援いただきました。株式会社富士ピー・エスの内野英宏氏、徳光卓氏、八木洋介氏、吉田光秀氏には、コンクリート橋梁についての薫陶を賜るとともに、コンクリートの変状画像の撮影の場を提供いただきました。長野県工業技術総合センターの新村諭氏には、異分野からの視点で深層学習についての様々なご助言をいただくとともに、異常検知の基本をご教示いただきました。全邦釘委員長をはじめとする、土木学会構造工学委員会 構造工学での AI 活用に関する研究小委員会の委員の皆様には、委員会活動を通じた様々な形で刺激をいただき、研究意欲を醸成することができました。株式会社福山コンサルタントの常門大祐氏には、データの前処理のツールの作成などに度々時間を割いてご協力いただきました。皆様には心より感謝申し上げます。

そして、著者が籍を置く株式会社福山コンサルタントの技術顧問である東北大学 鈴木基行 名誉教授からは、適切な助言や励ましを頂戴し、また、役員・社員の方々にも絶大なる応援を賜りましたことを感謝申し上げます。特に、本研究の趣旨をご理解いただいて、博士後期課程入学をご推挙いただきました株式会社福山コンサルタント 上席理事 中野聡氏には、心より感謝する次第です。

中村秀明教授から社会人ドクターへのお誘いを受け、株式会社福山コンサルタントのご理解により実現できた本研究ですが、この論文を仕上げるまでには、上述した以外の多くの方々からのご指導・ご支援を賜りました。改めて、皆様に深甚なる謝意を表します。

社会インフラの維持管理に関する研究活動を行っていく中で、改めて我々が抱えている社会的問題の大きさを痛切に感じるようになりました。今後も社会インフラにおける課題に積極的に関わっていきたいと思います。

2020年11月

青島 亘佐