

機械学習による画像分類技術を用いた耐候性鋼材におけるさびの外観評価

九州工業大学大学院 学生会員 ○光野正大
 横浜国立大学/大崎総合研究所 正会員 櫻井彰人
 九州工業大学大学院 正会員 山口栄輝

1. 研究背景および目的

耐候性鋼材は、鋼材表面に緻密なさび層（保護性さび）を形成することで、腐食の進展を抑制する。しかしながら、架設後の周辺環境によっては保護性さびが発生しない場合があり¹⁾、定期的な点検及び健全度の評価が必要となる。簡易的な評価手法として、セロファンテープ試験（表-1）があるが、その外観評価基準は必ずしも定量的ではなく、評価は容易ではない²⁾。そこで、本学 H30 年度の研究では AI の画像認識技術である畳み込みニューラルネットワーク（Convolutional Neural Network, 以下 CNN）を用いた外観評価法の検討が行われた。セロファンテープ試験の 5 段階評価をもとに、評点 1, 2 を安全管理上注意が必要なさび「C」、評点 3, 4, 5 を現状問題のないさび「S」と定義して AI に学習させた。しかしながら、本来 C であるさび画像を S と判定する安全管理上許容されない誤答を含むなど課題が残った³⁾。そこで本研究では、平均外観粒径 5mm という評価基準による判定が難しいさびを新たに N と定義し 3 項分類問題にするとともに、複数 CNN モデルによるアンサンブルを実施することで、簡便性と実用性を有する外観評価法を改めて提案することを目的とする。

表-1 さびの外観評価と状態²⁾

評点	さびの状態
5	さび粒子は細かいが、均一性に欠ける
4	さびの平均外観粒径は1mm程度で細かく均一である。
3	さびの平均外観粒径は1~5mm程度で細かく均一である。
2	さびの平均外観粒径は5~25mm程度で細かく均一である。
1	さびは層状で厚いが、剥離がある

2. データセットの構築

本研究では、耐候性鋼橋 28 橋にて採取されたセロファンテープ試験の資料をスキャンした画像データ（以下さび画像）を用いる。各橋梁にて採取枚数は異なり、さび画像の全データ数は 28 橋の合計で 1998 枚である。

2.1 アノテーション

クラスラベルを C, N, S の 3 項目と定義し、さび画像 1998 枚について本学にて独自にアノテーション作業

を行った。作業時にはさび画像をモニター上に実寸大表示し、さび粒子の外観粒径を実際に測定して判定している。クラスラベルの定義については表-1 を基に次の通りとした。C：平均外観粒径 5mm 以上のさび粒子が多数分布し明らかに評点 1, 2 であるもの。N：平均外観粒径 5mm 程度のさび粒子が分布しているが微細な粒子が多く分布する等、評点 2, 3 の判定が難しいもの。S：平均外観粒径 5mm 以下の微細なさび粒子が分布し明らかに評点 3, 4, 5 であるもの。また、アノテーション作業後のクラスラベル数を表-2 に示す。

表-2 アノテーション結果

クラスラベル	C	N	S	合計
枚数	195	128	1675	1998

2.2 データ編成

画像分類問題では、分類対象の他に背景についても配慮する必要がある。そこで、資料採取場所に起因する背景への影響を排除するべく、画像データを橋梁ごとにまとめて管理しデータ編成を検討した。全データ数が 1998 枚と少なく学習モデルの検証時には 5 分割交差検証（図-1）を行うため、クラス間データ数分布を近づけたグループを作成し、教師データを 5 グループ（各 5 橋）、テストデータを 1 グループ（3 橋）と編成した。

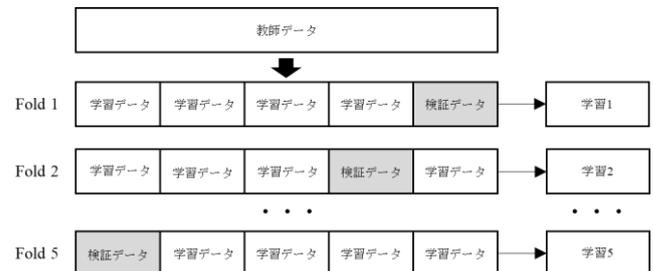


図-1 5 分割交差検証

3. CNN モデルの構築および学習

畳み込み層構造が異なる 3 種の CNN モデル（a~c）を作成し、各モデルに対して 2 種類の Pooling 方法を採用して、6 つの異なる CNN モデルを作成した。（表-3）

表-3 各モデルの層構造と Pooling 方法の組合せ

モデルNo	1	2	3	4	5	6
層構造パターン	a	b	c	a	b	c
Pooling方法	MaxPooling			AveragePooling		

学習に際しては、学習データにおけるクラス分布不均衡を解消するべく、データ数の少ない C, N について反転処理によるデータ拡張を実施しデータ数を 4 倍に増やした。各モデルについて、5 分割交差検証を実施し、テストデータを分類した結果を表-4 に示す。

表-4 テストデータ分類結果 (平均値)

モデルNo	1	2	3	4	5	6
Test_Accuracy	0.960	0.967	0.956	0.965	0.963	0.958
危険側誤判定数	0.8	1.8	3.2	3.6	5.0	6.0
層構造パターン	a	b	c	a	b	c
Pooling方法	MaxPooling			Average Pooling		

5 分割交差検証によって作成された 5 ケースの平均分類精度 (Test_Accuracy) がいずれのモデルについても非常に高く横並びの性能と見えるが、安全管理上許容されない危険側誤判定数に着目した場合、Max Pooling を採用したモデルにおいて誤判定数が少なく、より実用場面に則した性能であると考えられる。しかしながら、本研究では所謂グレーゾーンとして判定の難しいさび画像を N と新たに定義しているため、本来危険側誤判定については可能な限り排除するべきである。

4. アンサンブルの実施

前述したように、各 CNN モデルは単体でも高い分類性能を有する。しかしながら、モデル単体では分類性能向上に限界があった。そこで、分類性能向上のために、複数モデルによるアンサンブル (図-2) を実施した。

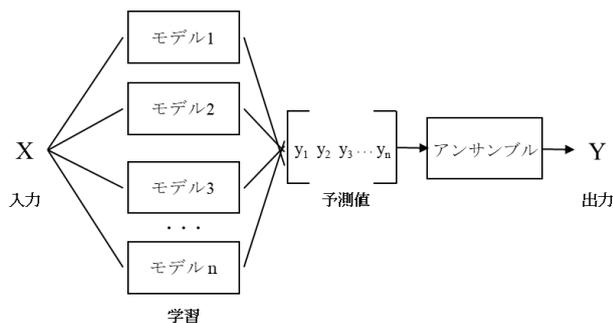


図-2 複数モデルのアンサンブル

アンサンブルに際しては、同一データに対する分類結果の相関が低いモデルを複数選択する必要がある。対象データは学習、検証に用いていないテストデータとし、6 モデルの交差検証 5 回それぞれを 1 ケースと捉え 30 ケースの分類結果の相関を調査したところ、ケース間の相関が非常に高くアンサンブルの効果が期待できないと考えられた。そこで、作成済みの 30 ケースに対して相関が低い分類結果を示す CNN モデルを新たに 3 種作成した。これらのモデルについては、表-4 の結果

より採用する Pooling 方法を Max Pooling とした。これにより、全 9 モデル、45 ケースとなった分類結果から、危険側の誤分類が最も多いものを 1 ケース (No.28) 選択し、それに対して相関が低い 4 ケースを併せることで分類結果の補完を目指した。(表-5) また、それらによる多数決アンサンブルを行った結果を表-6 に示す。

表-5 抽出ケース間における分類結果の相関

ケースNo	28	5	21	35	45
28	1	0.931	0.916	0.927	0.942
5	0.931	1	0.981	0.972	0.972
21	0.916	0.981	1	0.972	0.963
35	0.927	0.972	0.972	1	0.982
45	0.942	0.972	0.963	0.982	1

表-6 アンサンブル後のテストデータ分類結果

ensemble		予測値		
		C	N	S
正解	C	27	0	0
	N	2	18	0
	S	0	2	148

表-6 より、複数モデルによるアンサンブルを実施することで、危険側の誤分類を排除することができたと確認できる。また、分類精度についても 0.980 と、表-4 に示すいずれのモデルよりも優れた結果となった。

まとめ

本研究の結果から、さびの外観評価を実用場面に則した 3 項分類に読み替え、複数 CNN モデルによるアンサンブルを実施することにより、危険側の誤判定を排除するとともに、非常に高い分類性能を実現することができると言える。また、今回のアンサンブルでは、各 CNN モデルにおける層化 5 分割交差検証の結果を抽出して使用している。このことから、抽出して使用した 5 ケースにおいては、モデルおよび教師データの独立性が高く、非常に高い分類性能を実現した場合においても、単体の CNN モデルで分類精度を追求する場合に比べて汎化性能の確保が十分にできたと考える。

謝辞

本研究の一部は、日鉄エンジニアリングの助成を受けて実施したものである。ここに記して感謝致します。

参考文献

- 1) 山口栄輝ら：九州・山口地区における耐候性鋼材橋梁の実態調査，土木学会論文集 A, Vol.62, No.2, pp.243-254, 2006.
- 2) 公益財団法人 日本道路協会：鋼道路橋防食便覧，2014.
- 3) 川崎雄貴：AI による耐候性鋼材のさびの外観評価，九州工業大学卒業研究，2019.