

セロテープ試験試料の機械学習によるさび性状評価に関する検討

山口大学大学院 学生会員○松尾弘樹

山口大学大学院 学生会員 井上龍一

山口大学大学院 正 会 員 麻生稔彦

1. はじめに

耐候性鋼橋梁の長期使用のためには、さび性状の定期的な評価が重要である。その評価手法の一つであるセロテープ試験は、測定機器を必要としない簡便な手法であるが、目視評価のため、評価者の経験による評価の差が懸念されることから、定量的な評価手法として、深層学習によるセロテープ試験試料のさび評価手法が提案されている¹⁾。そこで本研究では、セロテープ試験試料に深層学習を適用し、比較的目的視評価の困難な評点4以上、評点3、評点2以下の3クラスへの分類および、評点1〜5の5クラスへの分類について検討する。

2. 深層学習の設計

2.1 概要

セロテープ試験は、表-1に示す5つの評点で評価されるが、本研究では、深層学習を用い、評点2以下・3・4以上の3クラス分類と、評点1〜5の5クラス分類について、セロテープ試験の評点評価モデルを作成する。学習には、表-2に示す、実橋梁でのセロテープ試験試料412試料を用い、学習に用いる教師データと、学習後のモデルの評価精度の検証に用いるテストデータとしておよそ8:2となるよう無作為に抽出する。

2.2 学習モデルの構築

学習モデルの構築には、画像判読の分野において広く用いられている畳み込みニューラルネットワーク(Convolutional Neural Network: CNN)を用いる。ここでCNNとは、モデルの中間層に、画像の局所的な特徴を抽出する畳み込み層と、画像のダウンサンプリングを行うプーリング層を、多層につなぎ合わせ表現学習をするニューラルネットワークである。学習には、表-2に示す教師データより作成したデータセットを用いる。しかし、試料数が十分ではないため、一般的に少試料数での学習の際に用いられる、FineTuningとデータ拡張を実施した。FineTuningとは、既存の学習済みモデルのパラメータを再利用し、モデルの後半の層のみ再学習させることで、当該タスクに適したパラメータに再調整する手法である。本研究では学習済みモデルとしてVGG16²⁾を用いた。また、データ拡張とは、元画像を加工することで教師データ数を増やす手法である。写真-1にデータの加工前および加工後の画像の例を示す。本研究では、回転・反転・平行移動・色調変化をランダムに組み合わせ、元画像を加工した。表-3にデータ

拡張後の試料の内訳を示す。学習回数は3クラス、5クラスともに50回とした。増加させた試料で学習をし

キーワード 耐候性鋼橋梁、さび評価、深層学習

連絡先 〒755-8611 山口県宇部市常盤台 2-16-1 山口大学工学部社会建設工学科 TEL 0836-85-9323

表-1 さび外観評価

評点5	さびは少なく、比較的明るい色調を呈する。
評点4	さびは1mm程度以下で細かく、均一である。
評点3	さびは1mm〜5mm程度で粗い状態。
評点2	さびは5mm〜25mm程度で、うろこ状の剥離あり。
評点1	さびは層状の剥離あり。

表-2 使用試料の内訳

	5クラス分類		3クラス分類		全試料数
	教師	テスト	教師	テスト	
評点5	44	11	148	38	55
評点4	104	27			131
評点3	91	23	91	23	114
評点2	78	20	88	24	98
評点1	10	4			14

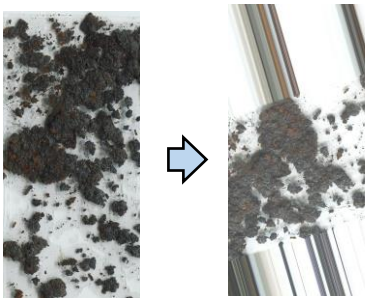


写真-1 加工前(左)と加工後(右)の画像の例

表-3 データ拡張後の試料の内訳

	5クラス分類		3クラス分類	
	教師	テスト	教師	テスト
評点5	1010	11	1042	38
評点4	1039	27		
評点3	1000	23	1547	23
評点2	1014	20	1056	24
評点1	1005	4		

た評点評価モデルを用い、全 85 試料のテストデータを評価することで、技術者目視評点との一致率から評点評価モデルの評価精度の検証をする。

3. 学習結果

3. 1 3クラス分類

表-4 に 3 クラス分類の結果を示す。各評点において技術者目視評点との一致率は 8 割を超えており、全テストデータでは 89. 4%(76/85)であった。このことより、3 クラス分類は、試料のスクリーニングを前提とすれば実用可能な一致率であるとする。写真-2 に技術者目視評点と機械学習評点が不一致となった試料の例を示す。写真は、技術者目視評点 5(4 以上)で機械学習評点 2 以下と評価された。この画像は、細かな塵埃が付着しており、それを 1 つの大きなさび粒子として認識していると考えられるため、深層学習では危険と判断されたと推測する。

3. 2 5クラス分類

表-5 に 5 クラス分類における技術者評価との一致率を示す。評点 1 および評点 5 の一致率が低い。また、全テストデータの一致率は 78. 8%(67/85)であった。評点 1 および評点 5 は、データ拡張前の元画像数が他の評点の試料に比べて少なく、データに偏りがあるものと考えられる。

一致率の低かった 5 クラス分類について、学習回数による一致率を検討する。図-1 は 5 クラス分類の学習精度を示している。学習回数が 30 回を超えると学習精度はほぼ一定となり、精度が向上しないと推測される。学習回数を 10 回から 50 回まで 5 回ずつ変化させ学習し、学習後の評点評価モデルでテストデータを評価することで一致率を算出した。図-2 に学習回数と一致率の関係を示す。学習回数 45 回で最も高い一致率を示したが、他の学習回数でも差はなく、この範囲では学習回数による一致率の向上は見られない。そのためデータ拡張での加工枚数を変え、一致率の向上を検討する必要があるだろう。

4. まとめ

深層学習を用いた、セロテープ試験試料のさび性状評価について、スクリーニングを前提とすれば 3 クラス分類は概ね可能であることを示した。5 クラス分類については、さらなる試料の採取によるデータの偏りの低減、また、データ拡張での加工枚数を変え、評価精度を検討する必要があると考える。

参考文献

- 1) 川崎雄貴, 山口栄輝, ディープラーニングを用いた耐候性鋼材のさびの外観評価, 土木学会全国大会第 75 回年次学術講演会, I-05, 2020.
- 2) Simonyan, K. and Zisserman, A., Very Deep Convolutional Networks for Large-Scale Image Recognition, International Conference on Learning Representations, pp.1-14, 2015

表-4 3 クラス分類の技術者評価との一致率

		機械学習評点		
		4以上	3	2以下
目視 評点	4以上	94.7 (36/38)	2.6 (1/38)	2.6 (1/38)
	3	13.0 (3/23)	82.6 (17/23)	4.3 (1/23)
	2以下	0	12.5 (3/24)	87.5 (21/24)

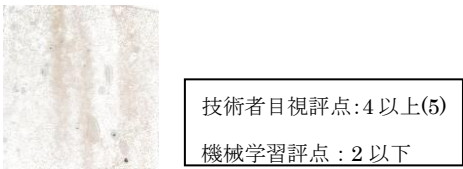


写真-2 技術者の評点と不一致であった試料

表-5 5 クラス分類の技術者評価との一致率

		機械学習評点				
		5	4	3	2	1
技術者 目視 評点	5	63.6 (7/11)	36.4 (4/11)	0	0	0
	4	3.7 (1/27)	92.6 (25/27)	3.7 (1/27)	0	0
	3	8.7 (2/23)	8.7 (2/23)	78.3 (18/23)	4.3 (1/27)	0
	2	0	0	20.0 (4/20)	80.0 (16/20)	0
	1	0	0	0	75.0 (3/4)	25.0 (1/4)

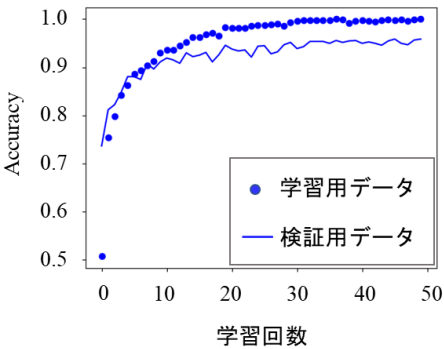


図-1 5 クラス分類の学習精度

