

耐候性鋼橋梁のさび外観評価における深層学習の適用に関する検討

鋼構造研究室 河原達哉

1. はじめに

耐候性鋼橋梁のさび外観評価¹⁾は、訓練を受けた技術者によって、現地におけるさび粒子の大きさやばらつき等を目視で判断し決定されるため、主観を含み、加えて、人材不足の中で、これらの判断を適切に実施可能な熟達した技術者の養成に大きく課題を残す。近年、これらの課題を解決するため、深層学習を用いたさび外観評価支援に関する研究^{2) 3)}が進められている。

本研究では、深層学習を用いたさび外観評価における再現性の担保と識別性能の向上を目的とし、さび厚に基づくさび状態の評価と対応する画像の組を用いた深層学習により、さび外観評価を試みた。また、撮影条件に対する汎化性能評価を行った。

2. 評価因子の整理

さび外観評価は、表-1 に示す外観的な特徴（目視外観）を目安に決定される¹⁾。評点3、評点4並びに評点5の状態である場合には、そのまま引き続き使用できるが、評点2の状態にある場合には、外観の変化を経過観察、評点1の状態にある場合には、補修塗装を行って腐食の進行を止める。このとき、残存板厚が小さい場合には当て板などによる補修も必要になる。

3. 深層学習の概要

(1) 教師データ

本研究では、新潟県内の5橋梁93部位（ウェブ及び下フランジ）において測定したさび厚及び撮影した鋼材表面近接画像を使用する。橋梁内の部位に100 mm×100 mmの正方領域を設け、その内部のさび厚測定と近接画像の撮影を

行った。写真-1に教師データの一部を示す。

さび厚は、電磁誘導式膜厚計により、正方領域内を隈無く12点計測し、最大値と最小値を除いた10点の平均値としている。また、近接画像は、マクロ撮影専用レンズ（MacromaX LM-2）により、鋼材表面から100 mm離れた位置において、正面から撮影した（以下、一眼レフ画像）。

(2) 深層学習の概要

深層学習では、ネットワークアーキテクチャを通じて、画像の特徴抽出を行い、入力データに対する予測値を出力する（順伝播）。このとき、

表-1 さびの外観評点¹⁾

状態	評点	目視外観	さび厚
正常	5	さびの量は少なく、比較的明るい色調を呈する。	200 μ m 未満
	4	さびの大きさは1 mm程度以下で細かく均一である。	400 μ m 未満
	3	さびの大きさは1~5 mm程度で粗い。	400 μ m 未満
要観察	2	さびの大きさは5~25 mm程度のうろこ状である。	800 μ m 未満
異常	1	さびは層状の剥離がある。	800 μ m 以上

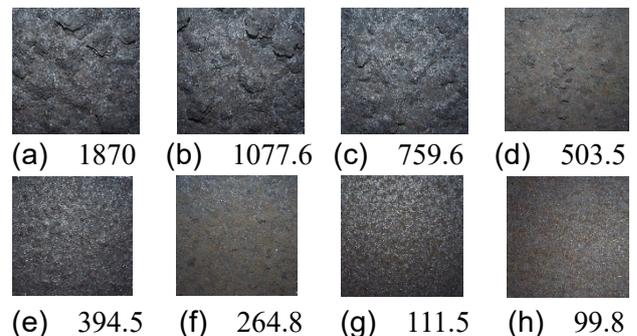


写真-1 教師データ (数値はさび厚, 単位: μ m)

出力された予測値と正解値の差（誤差）から損失関数（以下、loss）を定義し、loss の最小値（最適解）を得るように、勾配を用いて各層の重みパラメータを更新する（逆伝播）。重みパラメータの更新は、最適化アルゴリズムに従って行われ、モデルの構築においては、ネットワークアーキテクチャ、最適化アルゴリズム、並びに最適化アルゴリズムのハイパーパラメータである学習率（以下、lr: learning rate）の調整により、効率的な学習や最適解の精度向上を促す。

また、一般に深層学習では、いくつかのデータをまとめて入力し、それぞれの勾配を計算したあと、その勾配の平均値を用いて重みパラメータの更新を行う。データを用いて重みパラメータを1度更新するまでのことを1 iteration と呼び、データセット全体を使用することを1 epoch と呼ぶ。このとき、1 iteration で扱うデータ数を Batch size と表す。

(3) K-分割交差検証法に基づく学習

データを学習用（Training data）、検証用（Validation data）、テスト用（Testing data）に分割し、モデルを構築する。本研究では、A 橋、B 橋、C 橋の3橋梁からランダムで抜き出した59部位のデータを学習に用い（Training data と Validation data）、残りの17部位のデータ及びD 橋、E 橋の17部位のデータを Testing data とした。また、学習に用いる59部位のデータは、特徴が維持される拡張（回転（×4）及び反転（×2））を試み、Training data を354枚、Validation data を118枚として、K = 4でK-分割交差検証法を行った。また、Testing data のうち、A 橋、B 橋、C 橋のデータを Testing data 1、D 橋、E 橋のデータを Testing data 2 とした。

4. 外観評点の推定

本研究では、ネットワークアーキテクチャとして、文献 2), 3)において使用されている畳み

込みニューラルネットワーク（以下、CNN）に加え、Vision Transformer（以下、ViT）を用いる。本研究では、学習モデルの転移学習により特徴抽出を行った。この学習モデルを用いた特徴抽出は、別タスクで学習された知識を新たな識別タスクに対しても適合できるようにする手法で、文献 2), 3)等でも多用されている。CNN においては、ImageNet で事前学習された VGG16 及び ResNet50、ViT においては ImageNet-21k で事前学習された ViT-B/16 によりモデルを構築した。また、最適化アルゴリズムは、効率的な重みパラメータの設定に長けている RMSprop と Adam を対象とした。

(1) 学習率の調整

本研究では、小さなデータセットに対し、Batch Size = 8 で固定し、最適な lr を検証した。提案論文に記載され、広く一般的と認識される lr を鑑み、RMSprop では、 10^{-2} 、 10^{-3} 、 10^{-4} 、 10^{-5} 、 10^{-6} 、 10^{-7} 、Adam では、 10^{-3} 、 10^{-4} 、 10^{-5} 、 10^{-6} 、 10^{-7} 、 10^{-8} の各6パターンを検証した。本検証では、試験的に学習を 100 epoch で固定し、Validation data に対する loss の変動を確認した。

表-2 に 100 epoch 終了時の Validation data に対する loss を示す。表-2 より、 $lr = 10^{-5}$ 及び $lr = 10^{-6}$ の場合において、100 epoch 終了時の loss が十分に低減されていることが確認される。図-1 に $lr = 10^{-5}$ 及び $lr = 10^{-6}$ の場合における、epoch 毎の Validation data に対する loss の推移を示す。loss の変動は他の Case においても同様な傾向であった。図-1 より、ネットワークアーキテクチャ、最適化アルゴリズム問わず、100 epoch 終了時の loss が十分に低減され、かつ、loss 変動が安定した $lr = 10^{-6}$ が最も優れていることが分かる。

(2) モデルの比較

前節、lr の調整において、本データセットに

表-2 100 epochs 終了時の Validation data に対する loss (4 Case の平均)

モデル		lr							
		10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	
CNN	VGG16	RMSprop	1.240	2.257	2.009	0.203	0.144	0.638	—
		Adam	—	1.240	0.608	0.062	0.062	0.438	1.269
	ResNet50	RMSprop	8.111	3.640	0.525	0.178	0.193	0.904	—
		Adam	—	0.931	1.342	0.107	0.202	0.866	1.551
ViT	ViT-B/16	RMSprop	1.240	1.296	0.839	0.000	0.060	0.466	—
		Adam	—	1.242	0.487	0.005	0.013	0.465	1.200

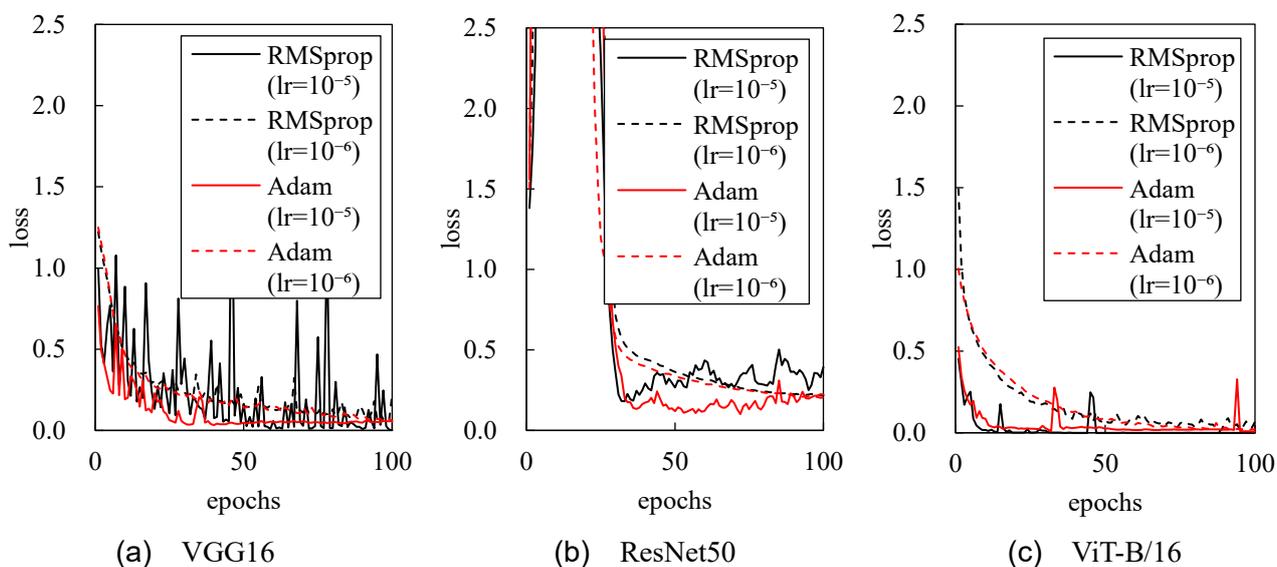


図-1 loss の推移 (Case1)

対して、ネットワークアーキテクチャ、最適化アルゴリズム問わず、 $lr=10^{-6}$ が効率性と安定性に優れることを示した。そこで、以下の検証では、Batch Size = 8, $lr=10^{-6}$ に設定した。また、学習を終了する条件として、上限 epoch を設けず、Validation data に対する loss が 10 epoch 連続で改善されなかった場合に学習を終了することとした。

表-3 に学習終了時の Validation data に対する loss を示す。表-3 より、loss の大きさは、学習モデルでは、ViT-B/16 < VGG16 < ResNet50, 最適化アルゴリズムでは、Adam < RMSprop となり、ViT-B/16 と Adam が優れていることが分かる。

(3) 評価指標及び識別結果

評点 5 及び評点 3 / 評点 4 を陽性 (Positive), 評点 2 及び評点 1 を陰性 (Negative) として、以下の式(1)から(3)を定義し、モデルを評価する。

$$\text{再現率} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{適合率} = \frac{TP}{TP + FP} \quad (2)$$

$$F \text{ 値} = \frac{2 \times \text{再現率} \times \text{適合率}}{\text{再現率} + \text{適合率}} \quad (3)$$

ここに、TP : True Positive, TN : True Negative, FP : False Positive, FN : False Negative である。

表-4 に Testing data 1 及び Testing data 2 に対する識別結果を示す。なお、表-4 に示す数字は、

表-3 学習終了時の Validation data に対する loss

モデル	loss (学習終了時の epoch)					
	Case1	Case2	Case3	Case4	Average	
VGG16	RMSprop	0.034 (137)	0.103 (90)	0.049 (116)	0.049 (126)	0.059 (117)
	Adam	0.019 (159)	0.117 (78)	0.017 (148)	0.044 (145)	0.049 (133)
CNN ResNet50	RMSprop	0.230 (115)	0.158 (121)	0.122 (180)	0.107 (188)	0.154 (151)
	Adam	0.188 (174)	0.065 (237)	0.169 (150)	0.121 (128)	0.136 (172)
ViT ViT-B/16	RMSprop	0.025 (63)	0.089 (44)	0.000 (78)	0.002 (75)	0.029 (65)
	Adam	0.001 (171)	0.012 (116)	0.008 (90)	0.006 (110)	0.007 (122)

表-4 テスト用データの識別結果

正解 予測	評点 5		評点 4 / 評点 3		評点 2		評点 1	
	Testing data 1	Testing data 2						
	評点 5	24	28	4	2	0	0	0
評点 4 / 評点 3	4	0	13	18	2	0	0	0
評点 2	0	0	3	0	9	10	0	8
評点 1	0	0	0	0	1	2	8	0

表-5 識別指標

モデル	指標	再現率	適合率	F 値
	Testing data 1	0.938	0.957	0.947
	Testing data 2	1.000	1.000	1.000

4Case の合計である。また、表-5 に式(1)から(3)で定義した指標でモデルを評価した際の結果を示す。表-4 及び表-5 より学習に用いた橋梁における部位及び学習に用いていない橋梁における部位問わず、十分な識別精度を有していることが確認される。特に、適合率が 95%に達しており、陰性である部位を陽性である部位と誤検出しない識別が行われており、維持管理の際にも有用であることが分かる。

5. 汎化性能に関する検討

(1) スマートフォン画像に対する識別性能

再調査において取得した一眼レフ画像によ

りモデル (学習モデル: ViT-B/16, 最適化アルゴリズム: Adam ($lr=10^{-6}$)) の再構築を行い、テスト用データとして、一眼レフ画像及びスマートフォンカメラにより、ラフに正面から撮影した画像 (以下、スマートフォン画像) を用いた。このうち、スマートフォン画像においては、撮影したままのスマートフォン画像及び RGB 強度が一眼レフ画像と同様になるように調整した画像 (以下、スマートフォン画像 RGB) を用いた。

モデルによるテスト用データの識別指標を、表-6 から表-8 に示す。表-7 より、スマートフォン画像の識別においては、評価指標において、

引き続き高い識別性能を示すものの、表-6 に示す一眼レフ画像の識別に比べ、識別性能が低下することが確認された。一方で、表-8 より、RGB 強度を調整した画像の識別においては、一眼レフ画像の識別と遜色ない結果が確認された。これらより、ラフに撮影した画像においてもさび外観評価が可能である一方で、前述したノイズや撮影機器に起因した色空間の差異が識別性能に影響することを示唆する。

(2) 撮影姿勢に対する評価

本研究では、橋梁内の部位に設けた 100 mm×100 mm の正方領域にワールド座標系 (X_w, Y_w, Z_w) とカメラの焦点から画像面に垂直に伸ばした光軸と画像面に平行な 2 方向によるカメラ座標系 (X_c, Y_c, Z_c) により撮影姿勢 (回転角及び回転軸) を定義する。再々調査の実施により、任意の撮影姿勢から取得したデータから、回転角及び回転軸が大域的に分布するように、各評点 5 部位から 20 枚程度、合計 100 枚のテスト用データを抽出した。

表-9 に回転角の大きさ (回転軸を無視) に対する識別結果を示す。ここに、維持管理上の措置の必要性有無の一致を正解と記述し、高精度範囲は、4 Case すべてで正解、中精度範囲は、3 Case あるいは 2 Case で正解、低精度範囲は 1

Case あるいは 0 Case で正解である。表-9 より、陰性部位 (評点 1 及び評点 2) 及び陽性部位 (評点 3/4 及び評点 5) に共通し、回転角の増大に伴い識別性能が低下する傾向が確認されるが、特に、陰性部位において影響が顕著である。陰性部位においては、概ね 22°前後の低い回転角で最も識別が安定し、回転角が 42°程度に達すると誤認識が増え、50°程度になると、識別精度がさらに低下し、識別が困難であることが分かる。一方で、陽性部位においては、低精度範囲の画像は存在せず、回転角に対する影響はほとんどないと考察できる。

図-2 に陰性部位における回転軸の比率に対する識別結果を示す。表示範囲は、高精度範囲と低精度範囲の平均を反映し、正解と不正解が混同する回転角 22~50°である。図-2 より、Z 軸に対する回転が卓越すると、識別精度が低下する傾向にあることが確認できる。

Z 軸の回転は撮影者によって制御可能であることから、表-10 に Z 軸の回転成分が小さい場合 (回転軸の比率が全体の 1%未満) の回転角の大きさに対する識別結果 (陰性部位) を示す。表-10 より、Z 軸の回転を含むデータに比べ、各範囲の回転角の平均値が 2~4°程度増大し、また、標準偏差が 2~3°程度減少していることが分かる。

表-6 識別指標 (一眼レフ画像)

データ \ 指標	再現率	適合率	F 値
Testing data 1	1.000	0.941	0.970
Testing data 2	1.000	0.970	0.985

表-7 識別指標 (スマートフォン画像)

データ \ 指標	再現率	適合率	F 値
Testing data 1	0.844	0.964	0.900
Testing data 2	0.844	1.000	0.915

表-8 識別指標 (スマートフォン画像 RGB)

データ \ 指標	再現率	適合率	F 値
Testing data 1	1.000	0.941	0.970
Testing data 2	0.906	1.000	0.951

表-9 回転角に対する評価（回転軸を無視）

評点	回転角 [°]	高精度範囲	中精度範囲	低精度範囲
評点 1	平均	22.4	42.8	49.6
	標準偏差	15.2	15.1	14.4
評点 2	平均	22.0	42.7	47.2
	標準偏差	15.2	15.9	14.1
評点 3/4	平均	37.0	38.0	-
	標準偏差	18.7	21.0	-
評点 5	平均	36.9	40.1	-
	標準偏差	19.3	23.5	-

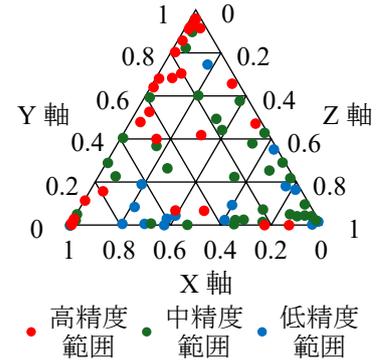


図-2 回転軸に対する評価（陰性部位）

表-10 回転角に対する評価（回転軸を考慮）

評点	回転角 [°]	高精度範囲	中精度範囲	低精度範囲
評点 1 及び評点 2	平均	22.2	42.8	48.0
(Z 軸の回転を含む)	標準偏差	15.2	15.4	14.2
評点 1 及び評点 2	平均	24.6	47.0	51.8
(Z 軸の回転を含まない)	標準偏差	12.1	11.7	11.7

以上より、Z 軸の回転が制御することで、標準偏差を考慮し、回転角が 40°以下の範囲で安定した識別が期待できる。

6. おわりに

本研究では、さび厚に基づくさび状態の評価と対応する画像の組を用いた深層学習により、さび外観評価を試みた。以下、本研究で得られた知見を示す。

- 1) ハイパーパラメータの最適化を通じ、本研究で検討した条件においては、学習モデル、最適化アルゴリズム問わず、lr=10⁻⁶が最も優れることが分かった。学習における loss は、学習モデルでは、ViT-B/16 < VGG16 < ResNet50、最適化アルゴリズムでは、Adam < RMSprop となり、ViT-B/16 と Adam が優れていることが分かった。
- 2) 構築したモデル（学習モデル：ViT-B/16、最適化アルゴリズム：Adam）に基づく特徴抽出により、有用的なさび外観評価が可能で

あることが示唆された。

- 3) 撮影画像の RGB 強度の調整により、学習時とは異なる撮影機器を用い、ラフに撮影した場合においても、さび外観評価が可能であることが示唆された。
- 4) Z 軸の回転を制御することで、回転角が 40° 程度の範囲内で安定した識別が期待できる。

参考文献

- 1) (社)日本鉄鋼連盟, (社)日本橋梁建設協会：耐候性鋼の橋梁への適用, 2023.3.
- 2) 高田 耕平, 北原 武嗣：深層学習を用いた耐候性鋼橋梁のさび外観評点識別システム, AI・データサイエンス論文集, 1 巻 J1 号
- 3) 佐々木 輝音, 原田 隆郎：曇み込みニューラルネットワークによる耐候性鋼材のさび外観評点判定モデルの汎用性向上に関する基礎的研究, 土木学会論文集, 79 巻 22 号, 2023.