

機械学習による金属錆画像の自動判定

日大生産工(院) ○杜 浩平 行木 靖人

日大生産工 矢野 耕也 大前 佑斗 豊谷 純

日大生産工(学部) 谷平 春人

1. はじめに

本研究の目的は、AI 技術を用いて設備の金属部品の腐食状況を判定し、判定精度と効率を向上させることである。従来の方法では、主に目視検査に頼っており、検査者の経験や環境要因に影響されやすいという課題があった。そこで、ディープラーニング技術の一つである畳み込みニューラルネットワーク

(CNN) を採用し、事前に分類された画像を学習させることで腐食判定を行い、その結果を評価する。

特に先行研究¹⁾では、畳み込み層の数を増やしてもモデルの精度向上には限界があることが示されている。この点を踏まえ、本研究では層数は少ないモデル「AlexNet」を実験した。さらに画像データの前処理、特にトリミング処理が判定精度に与える影響についても検証を行う。

また本報告では、元の画像から金属部品に関係の無い部分を手動でトリミングして、金属部品に関係の無い部分を削除した際に、精度がどれだけ向上するかを検証する。

2. 手法

金属部品の腐食・損傷判定に関する先行研究としては、2024 年に行木ら¹⁾によって CNN の 5 層構造と ResNet18 モデルを使用して、その有効性を明らかにした。本研究では、さらに他のモデルを適用してモデルの違いによる腐食・損傷判定の違いを比較検討する。

畳み込みニューラルネットワーク (CNN) は、画像認識や分類に特化したディープラーニングモデルで、画像から自動的に特徴を抽出し分類を行う。CNN は、畳み込み層、プーリング層、全結合層から構成され、ピクセルのエッジや形状、模様といった特徴を階層的に学習する。これにより、従来の手作業による特徴量設計が不要となり、高い精度での画像認識が可能になる。CNN は、物体認識、顔認識、自動運転技術など幅広い分野で利用されている。

ResNet (Residual Network) は、ディープラーニングモデルが抱える「勾配消失問題」に対応するために開発されたネットワークである。通層が深くなるほど学習が困難であるが、ResNet は「残差ブロック (Residual Block)」を用いることで、スキップ接続 (Skip Connection) を行い、前の層の情報を直接次の層に伝達できるようにした。この仕組みにより、非常に深い層でも学習が可能になり、従来のネットワークと比べて精度が大きく向上した。ResNet18 は、その中でも 18 層の比較的浅いモデルを適用しても先行研究において ResNet18 が CNN よりも精度の高い結果が確認されている¹⁾。

本研究では基礎的な知見を得るために CNN や ResNet よりも以前に提案された AlexNet を用いて、得られる結果の精度を検証することにした。AlexNet は、2012 年に ILSVRC ImageNet Large Scale Visual Recognition Challenge (ILSVRC) で優勝したディープラーニングモデルで、CNN を広く普及させた画期的なモデルである。AlexNet は、5 つの畳み込み層と 3 つの全結合層で構成され、大規模な画像データセットに対して高い認識精度を誇る。

AlexNet は、Rectified Linear Unit (ReLU) という非線形活性化関数を使用し、従来のシグモイド関数と比べて学習を高速化した。また、過学習を防ぐためのドロップアウト

(Dropout) 技術を取り入れることで、モデルの汎用性を向上させた。このモデルは、ディープラーニングの発展に大きな影響を与え、その後のさまざまなネットワークに大きな影響を与えている。

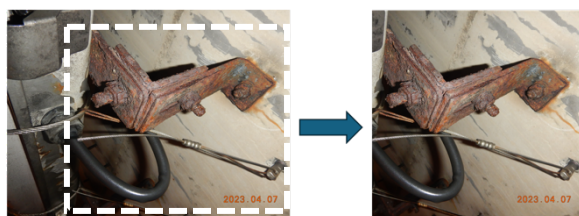
また画像拡張 (Augmentation) という技術である、これはモデルの汎化能力を向上させ、未学習データへの対応力を高めるために重要な手法である。Mikołajczyk et al. らは (2018)²⁾ の研究で、データセットの拡張 (データ増強) を通じて、モデルのパフォーマンスを向上させることができる。

Automatic Classification of Metal Images Using Machine Learning

Haoping TU, Seito NAMEKI, Kouya, YANO, Yuto OMAE,
Jun TOYOTANI and Haruto TANIHIRA

今回は5つの手法を使用した。1つ目は画像サイズを224×224ピクセルに変更する事である。2つ目はPyTorchで必要なテンソル形式への変換である。3つ目は画像をランダムに左右反転させる手法である。4つ目はランダムに回転や拡大・縮小を行い、サイズを80%から120%の範囲で調整する手法である。最後にRGBチャンネルごとに平均と標準偏差で正規化した。

1章で述べたトリミングとは画像の不要部分を取り除く、機械学習の前処理手法である。元の画像の一部を切り取って新しい画像を作成する手法である。トリミングは、画像の中で必要な部分にフォーカスしたり、不要な背景を取り除いたりする目的で使用される。トリミングされた画像は、元のサイズよりも小さくなるため、図1のように特定の領域に注目した画像を作成できる。



(a) トリミング前

(b) トリミング後

図1 トリミング処理した写真のイメージ

3. 検証条件

(1) 使用画像と腐食の判定基準

本研究では、東日本エンジニアリングのメンテナンス基準に基づき、腐食判定を4つのカテゴリに分類した。錆のタイプは4種類があります。0は軽い状態から3は厳しい状態で腐食度に応じて0～3と4つのファイルに分類した。

全体で600枚の画像を用い、そのうち480枚を学習用データ、120枚をテストデータとして分割した。学習用データの画像をトリミング処理した。モデルの汎化性能を評価するためにK分割交差検証を実施し、学習用データ480枚を80%の訓練データ384枚と20%の検証データ96枚に分け、5つのグループで検証を行い、結果を平均化した。

(2) パラメータサーチ

今回の研究では、CNN、ResNet18、AlexNetの3つのモデルを用いて、各モデルの性能を比較検討する。ハイパーパラメータの最適化には

Optuna³⁾を採用し、各モデルに対して少ない計算時間で効果があるパラメータを発見するもので計算資源を有効に活用しながら良好なパラメータセットを探索する。具体的には、150エポックでの訓練を通じて、50回の探索を行い、効率的なハイパーパラメータを発見することを目指す。

このアプローチにより、モデルの性能向上が期待する。探索結果に基づいてハイパーパラメータを調整し、モデルの精度と効率を向上させるように設計している。

(3) モデルの構築

本研究では、先行研究¹⁾に基づいて、5層のCNNと事前学習済みのResNet18を参照してモデルを構築した。

CNNの5層の中はどちらも畳み込み層、バッチ正規化層、プーリング層から形成されており、畳み込み層の出力チャンネル数は次のように設定されている。層1は32、層2は64、層3は256、層4は256、層5は256である。表1は、ハイパーパラメータの最適化を実行した結果であり、その時の設定が示されている。

表1 CNN5層の採択パラメータ

パラメータ	設定範囲	CNN5層
全結合層1のユニット数	256～1024	384
全結合層2のユニット数	128～512	231
ドロップアウト率1	0.1～0.5	0.151026
ドロップアウト率2	0.1～0.5	0.156857
学習率	0.00001～0.01	0.000086
オプティマイザ	Adam RMSprop SGD	RMSprop
バッチサイズ	16, 32, 64	32
モーメンタム	0.5～0.9	0.79802

ResNet18は、He, Kaiming, et al.らは2016年の研究⁴⁾に基づいた基本構造を採用し、今回の研究では事前学習済みのモデルのみを使用している。CNNと同様に、ハイパーパラメータの最適化を行っており、その設定を表2に示す。

AlexNetは、Alex Krizhevskyによる2012年の研究⁵⁾に基づいた基本構造を構築し、畳み込み層も同じ5層の設定とした。第4層の最大プーリング層を未設定、モデルにある活性

化関数 ReLU やドロップアウトを保留するそれに前と同じにハイパーパラメータの最適化を実行畳み込み層は設定されている。層 1 は 108, 層 2 は 142, 層 3 は 462, 層 4 は 442, 層 5 は 396 である。表 3 はパラメータ最適化を行った時の結果が示されている。

表 2 ResNet 18 の採択パラメータ

パラメータ	設定範囲	ResNet 18
学習率	0.00001~0.01	0.0000853
オプティマイザ	Adam RMSprop SGD	SGD
バッチサイズ	16, 32, 64	32
weight decay	0.000001~ 0.01	0.0000531
モーメントム	0.5~0.9	0.82299

表 3 AlexNet の採択パラメータ

パラメータ	設定範囲	AlexNet
全結合層 1 のユニット数	1024~4096	2048
全結合層 2 のユニット数	512~2048	844
ドロップアウト率 1	0.2~0.7	0.35
ドロップアウト率 2	0.2~0.7	0.5
学習率	0.00001~0.01	0.000085
オプティマイザ	Adam RMSprop SGD	Adam
バッチサイズ	16, 32, 64	32

(4) モデル学習の実施と検証

図 2 及び図 3 は, CNN, ResNet18, AlexNet 三つのモデルにおける, 学習曲線に検証データの損失と検証データ精度を二つに分け, 結果を示している。図 2 は検証損失の変化において, ResNet18 ではトリミングある(実線)が最も低い損失を示している。エポックが進むにつれて損失が減少と共に安定している。それに CNN と AlexNet 両方が事前学習なしのモデルが 0.8 前後で停滞する傾向があり, CNN の損失は AlexNet より低く, 安定性も CNN の方が良い。

さらに, ResNet18 にトリミングなしのデータを検証した, 損失も低く安定しているが, トリミング後のモデルよりわずかに損失が高めとなっている。

図 3 は検証データ精度の推移が示されている, ResNet18 でトリミングある(実線)は他のモデルよりも一貫して高い精度を達成しており, 約 0.9 の高い精度に達している。一方, CNN5 と AlexNet は同様の精度で推移しており, 0.7 前後の精度に留まっている。

ResNet18 にトリミングなしの精度も高い 0.85 に近いだが, 振動幅が大きくなっている。

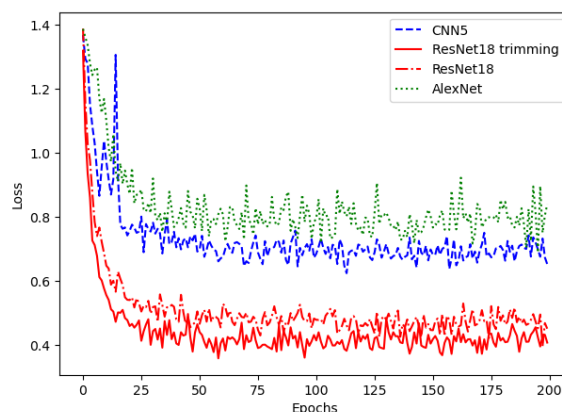


図 2 CNN, ResNet18, AlexNet の Validation loss

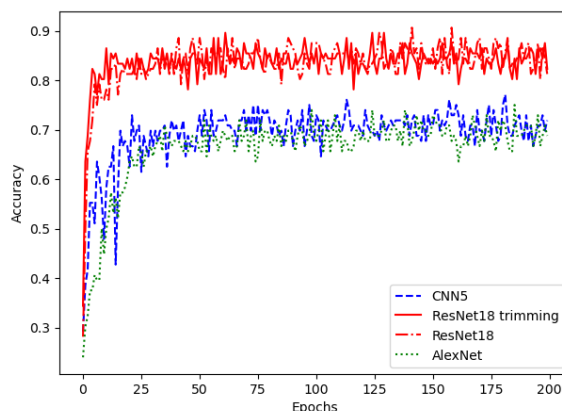


図 3 CNN, ResNet18, AlexNet の Validation Accuracy

これらの結果から, ResNet18 が他のモデルよりも優れた性能を発揮しており, 高い精度と低い損失を維持している事が示されている。そして画像のトリミング処理が効果的であり, 損失をさらに減らせることが確認できる。

4. 結果

図 4 は CNN, ResNet18, AlexNet 三つのモデルの混同行列である。まず CNN の結果はクラス 0, 1, 3 の分類性能は 0.73%程度があり, 逆にクラス 2 の方が精度は低くなって

いる．これを見るとクラス 0 に誤分類される可能性が 13%あった．

ResNet18 はクラス 1 と 3 の分類性能は非常に優れており再現率は 0.93 がある．クラス 2 は精度が多少低い事が示されている．その再現率は 0.60 に止まりクラス 3 に誤分類される可能性が高いといえる．

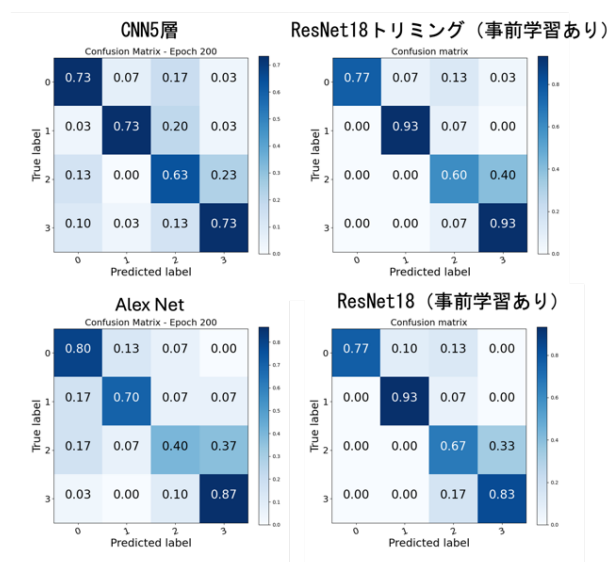


図 4 CNN, ResNet18, AlexNet の混同行列

表 4 CNN, ResNet18, AlexNet の分類結果

	クラス	適合率	再現率	F 値
CNN5 層	3	0.71	0.73	0.72
	2	0.56	0.63	0.59
	1	0.88	0.73	0.80
	0	0.73	0.73	0.73
	平均	0.72	0.71	0.71
ResNet18 事前学習 あり (トリミ ング)	3	0.68	0.93	0.79
	2	0.69	0.60	0.64
	1	0.93	0.93	0.93
	0	1.0	0.77	0.87
	平均	0.83	0.81	0.81
ResNet18 事前学習 あり	3	0.71	0.83	0.77
	2	0.65	0.67	0.66
	1	0.90	0.93	0.92
	0	1.0	0.77	0.87
	平均	0.82	0.80	0.80
AlexNet	3	0.67	0.87	0.75
	2	0.63	0.40	0.49
	1	0.78	0.70	0.74
	0	0.69	0.80	0.74
	平均	0.69	0.69	0.68

今回研究使ったモデル各クラスにおける分類性能評価を表 4 に示す．ResNet18 の方でトリミングあるがどうか，総体的に CNN の 5 層モデルと AlexNet よりも優れた性能を示す．

特にトリミングが 施された ResNet18 は，さらに良好な結果を得て，写真のトリミング処理が分別性能を向上させた事が確認できる．

5. 終わりに

今回の研究では，腐食の程度によって事前に分類された画像データセットを使い，機械学習でそれぞれの特徴を学習・分類した．その結果，事前学習ありのモデルが最も優れた性能を示した．特に，ResNet の残差ブロック技術が学習において大幅にクラス分類の精度に寄与した事が分かった．また，画像のトリミングなどの前処理もモデルの性能向上に役立っていることがわかった．

ただし，画像の前処理を行っても，クラス 2 の分類精度は依然として低く，改善の余地がある．今後は，他の事前学習モデルや異なる画像増強技術を試し，特定クラスの分類精度の低下を改善したい．

参考文献

- 1) 行木靖人, et al. "トンネル照明灯具の腐食・損傷分類における CNN と ResNet モデルの比較検討", 日本情報ディレクトリ学会誌, Journal of Japan Society of Directories, 日本情報ディレクトリ学会, Vol. 22, (2024) pp.34-44.
- 2) Mikołajczyk, Agnieszka, and Michał Grochowski. "Data augmentation for improving deep learning in image classification problem. " 2018 international interdisciplinary PhD workshop (IIPhDW). IEEE, (2018) pp.117-122.
- 3) Akiba, Takuya, et al. "Optuna: A next-generation hyperparameter optimization framework. " Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, (2019) pp.2623-2631.
- 4) He, Kaiming, et al. "Deep residual learning for image recognition. " Proceedings of the IEEE conference on computer vision and pattern recognition, (2016) pp. 770-778.
- 5) Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks. " Advances in neural information processing systems 25, (2012) pp.1-9.