

畳み込みニューラルネットワークを用いた分析フェログラフィ画像解析

Diagnosis of analytical ferrography image using convolutional neural network

福井大・工（正）*本田 知己 福井大・工（院）米道 潤紀 福井大・工（正）今 智彦

トライボテックス（株）（正）川畠 雅彦 （正）安部田 泰 （非）竹内 崇晴

Tomomi Honda*, Junki Yonemichi*, Tomohiko Kon*, Masahiko Kawabata**, Yasushi Abeta**, Takaharu Takeuchi**

*University of Fukui, ** TRIBOTEX Co., Ltd.

1. 緒言

近年、分析フェログラフィの摩耗粒子画像解析に畳み込みニューラルネットワーク(CNN: Convolutional Neural Network)を用いた研究が報告されている^{(1),(2)}。一方、その結果に対してAI(Artificial Intelligence: 人工知能)の判断根拠を可視化し、専門家の分析箇所と同様であるかの検証は行われていない。そこで、本研究ではまず、分析フェログラフィで使用されるフェログラム画像から単一摩耗粒子の切り取りを行い、単一摩耗粒子を教師画像としてCNNによる学習を行った。次に、摩耗粒子の形態を識別するCNNモデルを作成し、その精度について評価した。さらに、解析結果からCNNにおけるAIの判断根拠を可視化するために、Grad-CAM(Gradient-weighted Class Activation Mapping)を用いて判断根拠となる特徴箇所を摩耗粒子画像内に色で表示し、その結果からAIが重要視している摩耗粒子画像内の特徴パラメータについて考察した。

2. 教師画像の作成

分析フェログラフィで使用されたフェログラム画像の中から4つの摩耗粒子の形態(切削・球状・平板状・シビア)が顕著に表れている画像を436枚用意した。436枚のフェログラム画像から単一摩耗粒子を教師画像として切り取る際、「ISO18436-4 準拠 機械設備の状態監視と診断(トライボロジー) カテゴリーII」に記載されている摩耗粒子の形態を参考に単一摩耗粒子の切り取りを行った。切り取りを行った後、それらの粒子画像について再度フェログラフィの分析業務を行っている専門家と議論し、最終的な教師画像を決定した。4つのラベルにおける典型的な単一摩耗粒子の教師画像を図1に示す。教師画像の枚数は切削摩耗粒子80枚、球状摩耗粒子80枚、平板状摩耗粒子80枚、シビア摩耗粒子80枚の計320枚である。

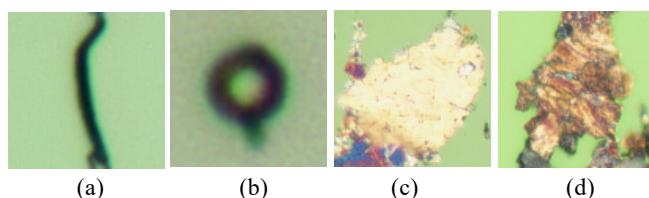


Fig. 1 Wear particle classification.

(a)Cutting Wear Particles, (b)Spherical Particles, (c)Fatigue-Plate Particle, (d)Severe Sliding Wear Particles

3. 解析方法および解析アルゴリズム

3.1 解析方法

ソニーネットワークコミュニケーションズ(株)が提供しているNNC(Neural Network Console)を用いて、単一摩耗粒子の形態を識別するCNNモデルの構築、学習および評価を行った。演算処理にはNVIDIA® T4 GPUを用いた。本研究では図1に示すように摩耗粒子の形態ごとに0-3のラベルをつけてNNCにデータセットした。学習方法はミニバッチ学習を用い、CNNモデルの構造はNNCの構造自動探索を用いて最適化した。その結果、40パターンの構造から学習の繰り返し数であるエポック数100回での学習結果の誤差が最小となったモデル(A)と評価結果の誤差が最小となったモデル(B)の2種類を選定した。

3.2 解析アルゴリズム

3.2.1 学習アルゴリズム

学習アルゴリズムの流れを図2に示す。はじめにデータセットされた全ての教師画像はInputレイヤーで1枚の画像が取り出される。取り出された画像は学習画像の種類を増やすことを目的としてImageAugmentationレイヤーとRandomShiftレイヤーで指定した範囲内の画像の加工が行われる。RandomShiftレイヤーでは、画像を中心から上下左右に1pixelランダムに動かす加工が行われる。画像の加工が完了するとConvolutionレイヤーで1層目の畳み込みが行われる。ここでは400×400pixelの入力画像に対し、4×11pixelのフィルタを20枚使用して1pixelずつずらしながら畳み込みが行われ、1枚の画像から各フィルタの特徴を抽出した397×390pixelの特徴マップが20枚作成される。その後、20枚の特徴マップは活性化関数の1つであるReLUレイヤーで非線形な値に変換される。MaxPoolingレイヤー

一では 2×2 pixel の範囲の最大値を抽出することで、 397×390 pixel の特徴マップが 198×195 pixel まで圧縮される。これにより、この後のネットワーク全体の演算量が削減される。その後、Convolution_2 レイヤーで 2 層目の畳み込みが行われる。Convolution_2 レイヤーでは 198×195 pixel の 20 枚の特徴マップに対し、 3×3 pixel のフィルタを 40 枚使用して 1 pixel ずつずらしながら畳み込みが行われ、20 枚の特徴マップから 196×193 pixel の特徴マップが 40 枚作成される。BatchNormalization レイヤーでは、Convolution レイヤーや Affine レイヤーの後に入力値の正規化を行うことで分類精度の向上および計算の収束を早める効果がある。ここでの入力値の正規化とは、 1×1 pixel 内の数値を 0-1 に変換することを意味する。その後 MaxPooling_2 レイヤーでは、 2×2 pixel の範囲で最大値の抽出が行われ、40 枚の特徴マップは 98×96 pixel まで圧縮される。Tanh レイヤーでは ReLU レイヤーと同様に MaxPooling_2 レイヤーまでの線形関数が非線形関数に変換される。ここまで、2 層の畳み込みによる特徴の抽出が完了し、その後 Affine レイヤーでは、 98×96 pixel の 40 枚の特徴マップが全て結合され、重みとバイアスが加えられて 136 個のニューロンが出力される。BatchNormalization_2 レイヤーでは、136 個のニューロンの正規化が行われ、各ニューロンの数値は 0-1 の範囲で出力される。ReLU_2 レイヤーでは、BatchNormalization_2 レイヤーまでの線形関数が非線形関数に変換される。出力のニューロン数をラベルの数と合わせるために Affine_2 レイヤーでは 136 個のニューロンが全て結合され、重みとバイアスが加えられて 4 個のニューロンが出力される。Softmax レイヤーでは、クラス分類における各確率を 0-1 かつ合計が 1 になるように出力される。AI は誤差を小さくするためにフィルタと重みおよびバイアスの更新を行うが、学習データを全て用いると局所最適点に陥りやすい。そのため、本研究では学習データを分割して学習するミニバッチ勾配降下法を用いて、1 回の学習に教師データを 64 枚ずつ分割し、イテレーション数を 5 として学習を行った。

3.2.2 評価アルゴリズム

指定したエポック時に AI は評価画像の評価を実行する。評価アルゴリズムの流れを図 3 に示す。はじめにデータセットされた全ての評価画像は Input_2 レイヤーで 1 枚の画像が取り出される。その後、学習時に作成したフィルタを使用して畳み込みが行われる。Convolution_3 レイヤーでは、 400×400 pixel の入力画像に対し、 4×11 pixel のフィルタを 20 枚使用して 1 pixel ずつずらしながら畳み込みが行われ、1 枚の画像から各フィルタの特徴を抽出した 397×390 pixel の特徴マップが 20 枚作成される。その後、20 枚の特徴マップは ReLU_3 レイヤーで非線形関数に変換された後、Maxpooling_3 レイヤーで 2×2 pixel の最大値が抽出される。これにより、 397×390 pixel の特徴マップは 198×195 pixel まで圧縮される。Convolution_4 レイヤー

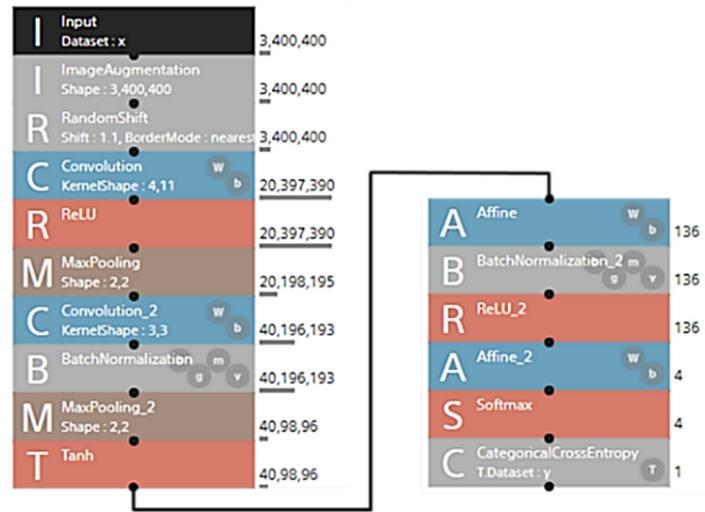


Fig. 2 4 class classification learning algorithm

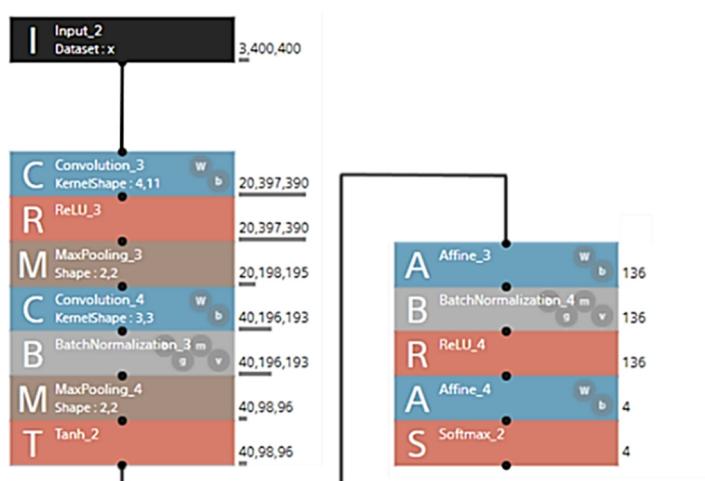


Fig. 3 4 class classification evaluation algorithm

では、 198×195 pixel の 20 枚の特徴マップに対し、 3×3 pixel のフィルタを 40 枚使用して 1 pixel ずつずらしながら畳み込みが行われ、20 枚の特徴マップから 196×193 pixel の 40 枚の特徴マップが作成される。BatchNormalization_3 レイヤーでは、入力値に対する正規化が行われ、 1×1 pixel 内の数値が 0-1 に変換される。MaxPooling_4 レイヤーでは 2×2 pixel の範囲で最大値の抽出が行われ、40 枚の特徴マップは 98×96 pixel まで圧縮される。Tanh_2 レイヤーでは MaxPooling_4 レイヤーまでの線形関数が非線形関数に変換される。ここまで、2 層の畳み込みによる特徴の抽出が完了し、その後 Affine_3 レイヤーでは、 98×96 pixel の 40 枚の特徴マップが全て結合され、重みとバイアスが加えられて 136 個のニューロンが出力される。BatchNormalization_4 レイヤーでは、136 個のニューロンの正規化が行われ、各ニューロンの数値は 0-1 の範囲で出力される。ReLU_4 レイヤーでは、BatchNormalization_4 レイヤーまでの線形関数が非線形関数に変換される。Affine_4 レイヤーでは、出力のニューロン数をモルヒ学の形態ラベルの数と合わせるために 136 個のニューロンが全て結合され、重みとバイアスが加えられて 4 個のニューロンが出力される。Softmax_2

レイヤーでは、クラス分類における各ラベルの確率が 0-1 かつ合計が 1 になるように出力される。これにより、1 枚の画像から 4 種類の摩耗粒子の形態の確率が算出される。

3.2.3 Grad-CAM による判断根拠の可視化

NNC では、付属の拡張機能である Grad-CAM を活用することで判断根拠の可視化を行うことができる。Grad-CAM は、評価アルゴリズムの Convolution_4 レイヤーで畳み込みを行い、正規化、最大値抽出、非線形変換を行った特徴マップに Affine レイヤーで 40 枚の特徴マップを結合する際の重みを掛けてから評価画像に重ね合わせて可視化を行う。そのため、重みの値が高くなると画像内の特徴量が大きくなることを意味する。Grad-CAM では、判断根拠となる箇所が色のグラデーションで表示され、濃い赤の領域は画像内の特徴量が大きいことを示し、濃い青の領域は画像内の特徴量が小さいことを示す。また、画像内の特徴量が増加するにつれて色は濃い青から緑、赤へと変化する。

4. 解析結果および考察

4.1 4 クラス分類

3.1 節で述べたモデル(B)の学習曲線と評価エラーの推移を図 4 に示す。Cost は最適化段階の誤差、Training error と Validation error はそれぞれ教師画像と評価画像での各エポックの最適化完了時点での誤差を表している。エポック数の増加に合わせて重みとバイアスを更新することで、Cost と Training error が減少するように学習していることがわかる。一方で、エポック数を増やすと教師画像に特化した AI となるため、NNC では 5 エポックまでは毎回、その後は 10 エポックごとに評価画像を全て用いた解析を行う。さらに、NNC では評価エラーが最小となるモデルを保存することができ、解析結果は評価エラーが最小となったエポック数のデータを表示する。図 4 に示されるように、本研究ではエポック数 50 での評価結果が解析結果となった。

評価画像は全て単一摩耗粒子画像であり、1 つのラベルで 25 枚、4 つのラベルで計 100 枚使用した。解析結果を表 2 に示す。切削摩耗粒子と球状摩耗粒子は全て正解ラベルの確率が高く、正解ラベルを正しく判定した。平板状摩耗粒子は 25 枚中 14 枚で正解ラベルを正しく判定し、10 枚はシビア摩耗粒子、1 枚は球状摩耗粒子と判定した。シビア摩耗粒子は 25 枚中 21 枚で正解ラベルを正しく判定し、4 枚は平板状摩耗粒子と判定した。以上の結果から、評価画像の形状特徴が明確である切削摩耗粒子と球状摩耗粒子の再現率(Recall)が高いことが示された。一方で、平板状摩耗粒子の再現率は 0.56 であり、シビア摩耗粒子の再現率は 0.84 であった。CNN を用いた摩耗粒子の画像解析において、切削摩耗粒子や球状摩耗粒子の再現率が高く、平板状摩耗粒子やシビア摩耗粒子の再現率が切削摩耗粒子や球状摩耗粒子と比べて低くなることは、先行研究の CNN モデルからも明らかとなっている。例えば Peng らは、フェログラム画像から摩耗粒子の位置を特定する CNN アルゴリズムと摩耗粒子の形態を特定する CNN アルゴリズムを組み合わせた CNN モデル(WP-DRnet)において、切削摩耗粒子と球状摩耗粒子の再現率が 1.00 であったのに対し、平板状摩耗粒子の再現率は 0.72、シビア摩耗粒子の再現率は 0.77 であったことを報告している⁽¹⁾。また Liu らは、CNN モデルの一つである VGG モデルに残差ブロックを組み合わせた CNN モデル(VRCNN)において、切削摩耗粒子の再現率が 0.96、球状摩耗粒子の再現率が 1.00 であったのに対し、平板状摩耗粒子の再現率は 0.90、シビア摩耗粒子の再現率は 0.96 であったことを報告している⁽²⁾。これらは、平板状摩耗粒子やシビア摩耗粒子の表面性状が多様であり、同じ摩耗粒子の形態でも摩耗粒子によってその特徴が表れる箇所や形が微妙に異なることが要因であると考えられる。一方で、摩耗粒子の画像解析において CNN モデルの性能は再現率以外の指標でも評価することができる。例えば、本研究の単一摩耗粒子の形態を識別する CNN モデルでは摩耗粒子画像の解析結果がラベルごとに確率として表示される。そのため、正しく判定した画像においても判定ラベルの確率が高いのか、他のラベルと比較してその確率が僅差だったのかということは、摩耗粒子画像の分類精度を評価する上で重要な指標になると考えられる。そこで、次節では尤度グラフを用いて画像一枚の判定ラベルの確率に注目し、その精度について評価した。

4.2 尤度グラフ

横軸を判定ラベルの確率、縦軸を画像の枚数とした尤度グラフを図 5 に示す。なお、false は判定ラベルが正解ラベルと一致しなかった画像の枚数を示しており、true は判定ラベルが正解ラベルと一致していた画像の枚数を示している。false は全部で 15 枚あり、7 枚が 90% 以上の確率を示している。true は全部で 85 枚あり、61 枚が 90% 以上の確率を示している。以上の結果から、本研究で作成した単一摩耗粒子の形態を識別する CNN モデルにおいて画像一枚の判定ラベルの確率に注目した場合においても、判定結果は特定のラベルで高い値を示すことが明らかとなった。また、true の中に 70% 以上の確率を示した割合は 85 枚中 70 枚であり、正しく判定した画像の 82% を占めている。true の確

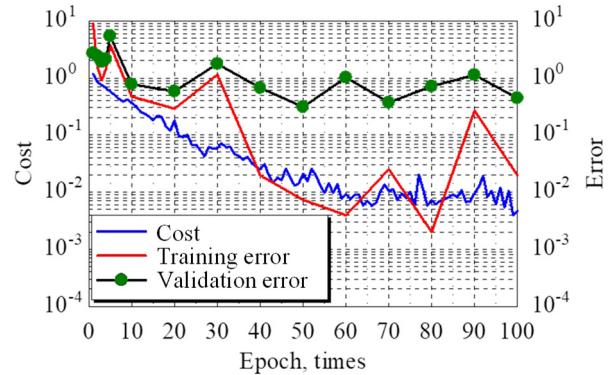


Fig. 4 Learning curve and validation error in (B)

Table 2 Classification accuracy in (B)

	Recall	(a)	(b)	(c)	(d)	Total
(a)	1.00	25	0	0	0	25
(b)	1.00	0	25	0	0	25
(c)	0.56	0	1	14	10	25
(d)	0.84	0	0	4	21	25

率が70%以上を示すことはその他のラベルの確率が30%以下となることを示しており、ラベルごとの確率を明確に区別することができる。この要因として、形状が特徴的な切削摩耗粒子や球状摩耗粒子のラベルを含めているかつラベルの数を4種類に限定していることが挙げられる。すなわち、ラベルの数を増やした場合、摩耗粒子の形態を詳細に識別することが可能となる一方で、他のラベルと比較して確率が僅差となる場合も増加すると予想される。

ラベルの数を増やしていくことを想定した場合、摩耗粒子の画像解析においてAIが摩耗粒子の形態を識別する方法を説明する要求が高まる。一般的に専門家による分析では、摩耗粒子の大きさ、形状、色、表面性状等の特徴を表す情報を総合的に判断して摩耗粒子の形態を識別するため、寄与度が大きい箇所を限定して定量化することは困難である。一方で、教師あり学習や教師なし学習による画像解析では、教師画像の統一性が解析の精度を向上させる上で重要となる。そこで、次節では単一摩耗粒子の形態を識別するCNNモデルにおける画像内の重点箇所をGrad-CAMを用いて可視化することで、摩耗粒子の形態を識別する上でAIが重視している特徴パラメータについて考察した。

4.3 Grad-CAM

評価に使用した画像のGrad-CAMによる解析結果の一例を図6に示す。Grad-CAMでは全ての画像において画像内の明るい部分が赤色となり、暗い部分が青色で表されている。これは摩耗粒子の形態を認識する際に色の明暗が重要であることを示している。例えば切削摩耗粒子は表面が酸化していることが多く、長さに対して幅が狭いため画像全体が黒色となることが多い。球状摩耗粒子は球形であるため、光学顕微鏡では球の周囲の反射光は接眼レンズに戻らず、輪郭が黒くなる。このように、切削摩耗粒子や球状摩耗粒子は画像内に黒色が明確に表れるため、形状を色の明暗から認識しやすい。一方で、平板状摩耗粒子は表面積が広く、摺動面の材質と温度によって摩耗粒子の色が変化するため、摩耗粒子表面の色が一様であることは少ない。例えば、平板状摩耗粒子画像には黄色や黄褐色、薄青色のものが多い。これは摩耗粒子が生成された時の摺動面の温度が関係している。例えば、SUS304ステンレス鋼は400°Cで白色から黄色に変色し、鉄は330°Cで白色から黄褐色に変色する。また、SUJ2高炭素クロム鋼は300°Cで白色から青色になる。このように、平板状摩耗粒子の色は材質と温度により変化する場合がある。本研究で使用した平板状摩耗粒子の教師画像には黄色や黄褐色、薄青色等の比較的明るい色が多いことがわかった。また、シビア摩耗粒子の教師画像には茶色や濃青色のものが多くた。これは摩耗粒子表面が部分的に酸化したことによって鉄の黄褐色やSUJ2の青色が部分的に黒くなつたためであると考えられる。以上の結果から、平板状摩耗粒子やシビア摩耗粒子は表面性状が多様であり、同じ摩耗粒子の形態でも摩耗粒子によってその特徴が表れる箇所や形が微妙に異なるため、AIは摩耗粒子の色をもとに摩耗粒子の形態を識別していることがわかった。

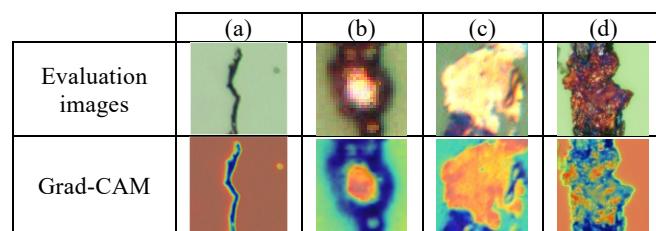


Fig. 6 Grad-CAM results of evaluation images

5. 結言

- (1) 単一摩耗粒子の評価画像100枚中85枚で正解ラベルの確率が他のラベルの確率と比べて最も高く、摩耗粒子画像から摩耗粒子の形態を正しく識別した。さらに、正しく識別した画像85枚中70枚で正解ラベルの確率が70%以上を示し、摩耗粒子の形態を明確に区別することができた。
- (2) CNNによる単一摩耗粒子の画像解析において、Grad-CAMの結果からAIは摩耗粒子画像の色の明暗を重視して識別していることがわかった。

参考文献

- (1) Yeping Peng et al : WP-DRnet: A novel wear particle detection and recognition network for automatic ferrograph image analysis, Tribology International, Vol.151, No.106379 (2020), pp.1-9.
- (2) Liu Z et al : Intelligent classification of online wear particle in lubricating oil using optical direct imaging method and convolutional neural network for rotating machinery, Tribology International, Vol.189, No.109015 (2023), pp.1-12.