



画像分析への AI 適用による 検査計測技術の向上

Hata Nobuyoshi

畑 伸佳*1

Tanakamaru Tempei

田中丸 天兵*3

Kurita Shinobu

栗田 しのぶ*2

Kondo Naomi

近藤 直美*4

当社では検査・計測システムを開発し、そのシステムには画像処理技術を用いてきた。今回さらなる業務の効率化と付加価値向上を目指し、AI 技術を活用した。

Deep Learning の一手法である CNN を用いた手法では、金属製品の SEM 画像から「ポイド」と呼ばれる空隙の有無を約 95% の精度で判定でき、さらにポイドである可能性を可視化したソフトウェアを考案し、作業者による判定を容易にした。pix2pix を利用した手法では、SEM 画像からのポイド部分の抽出に成功したが、ポイドの面積が大きめになる、小さなポイドを見逃すという課題が残った。

本稿では、ポイドの検出・抽出を試行した事例と、今後精度を向上するための検討内容を紹介する。

キーワード： 画像処理、人工知能、AI、Deep Learning、深層学習、CNN、pix2pix、ポイド

1. はじめに

当社では、検査・計測の作業効率化、付加価値向上を目指して検査・計測システムを開発している。またそれらのシステムを開発するにあたり、画像処理技術を適用してきた。近年、検査・計測における画像分析では、より高度な分析が求められるようになり、従来の画像処理技術単独では対応することが難しくなりつつある。このため、画像処理に AI（人工知能 Artificial Intelligence）の技術を取り入れることで、より高度な画像分析をできるようにし、検査・計測作業の自動化・判

断支援の実現を図った。

本稿では、当社でより高度で効率的な画像分析を実現するために採用した AI 技術を紹介する。

2. AI を実現する技術の概要

2.1 AI を実現する Deep Learning の概要

AI を実現するための手法の 1 つである Deep Learning（深層学習）の適用は、2011 年に音声認識の研究分野において、従来の手法では超えられなかった性能の限界を上回り、大幅な性能改善を实

*1：制御システム事業部 産業システム部 製品・サービスグループ

*2：制御システム事業部 産業システム部 デジタルイノベーショングループ

*3：計測事業部 化学・材料部 磯子グループ 主査

*4：研究開発センター 研究管理グループ グループ長

現した⁽¹⁾。また、この技術は、画像処理の研究分野でも2012年の大規模な物体認識コンペティションにおいて、大幅な性能改善を実現した⁽²⁾。これによりDeep Learningを適用したAIの有効性が確認され、音声認識や画像認識にとどまらず、自然言語処理や予測・異常検知などさまざまな分野への適用が進み、GoLeNet⁽³⁾やResNet⁽⁴⁾などさまざまなAIモデルが発表されてきた。

従来の画像処理では、人の判断により入力データからエッジの強さや向きなどの特徴的な情報(特徴量)を抽出して認識するという方法をとっていた。しかしその手法では複雑な画像に対応することは困難であった。それに対し、画像処理にDeep Learningを適用する方法では、大量のデータから共通する認識に必要な特徴量を抽出して「学習」することができるため、複雑な画像に対しても高精度な認識ができるようになる。

当社でもDeep Learningの適用により高度な画像分析を実現する検査計測技術の開発に取り組んできた。

本稿では開発で採用したCNNとpix2pix⁽⁵⁾について説明する。

2.2 CNN

CNN(Convolutional Neural Network)は畳み込み処理を使用したニューラルネットワークの総称であり、画像認識分野や音声認識分野

などで使われる。画像認識分野では何の画像なのかを判別する物体認識、画像から1つまたは複数の物体の位置を特定する物体検出、画像から道路や街路樹、車、人などを分類してラベリングするセグメンテーションなどと幅広く研究・適用されている。CNNは一般的には図1のように入力データ、畳み込み層、プーリング層、全結合層、出力データから構成される。畳み込み層では入力データに対して畳み込み処理(フィルタ処理)を施し特徴マップを作成する。プーリング層では必要な特徴量を抽出して特徴マップを圧縮する。画像認識の場合、画像に写るターゲット物体の位置の変化が認識処理に影響しないようにするために導入する。そして全結合層では特徴マップ全てと結合して計算する。

CNNの一般的な学習方法としては、入力データと正解データ(学習データ)をあらかじめ用意し、最初はネットワークの各パラメータにランダムな値を格納する。次に入力データをネットワークに流して得た出力データと正解データとの誤差を計算する。そして計算した誤差をできる限り小さくするよう、誤差をネットワークの出力側から逆方向に伝播させて各パラメータを更新する。更新したネットワークに再度入力データを流して出力データの計算と誤差を逆伝播させて各パラメータの更新とを繰り返すことで、期待する認識結果を出力するよう「学習」する手法である。

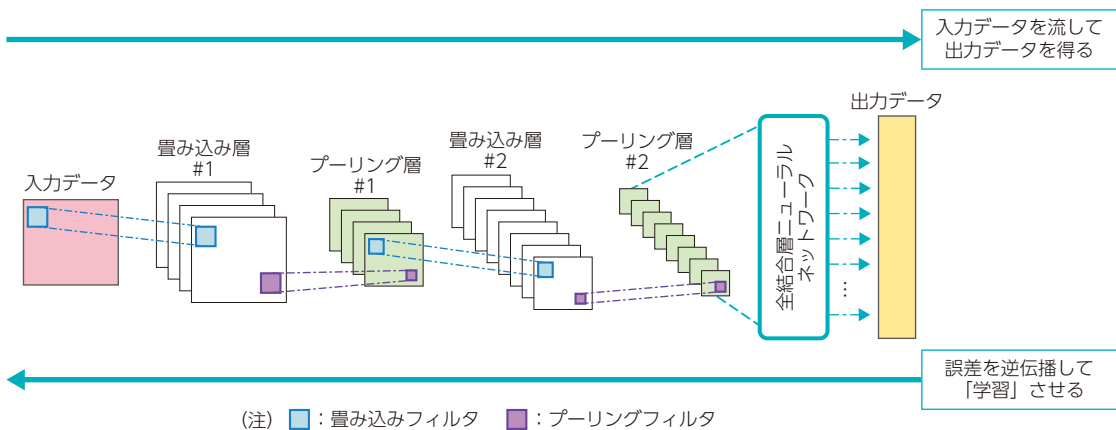


図1 CNNの構造イメージ

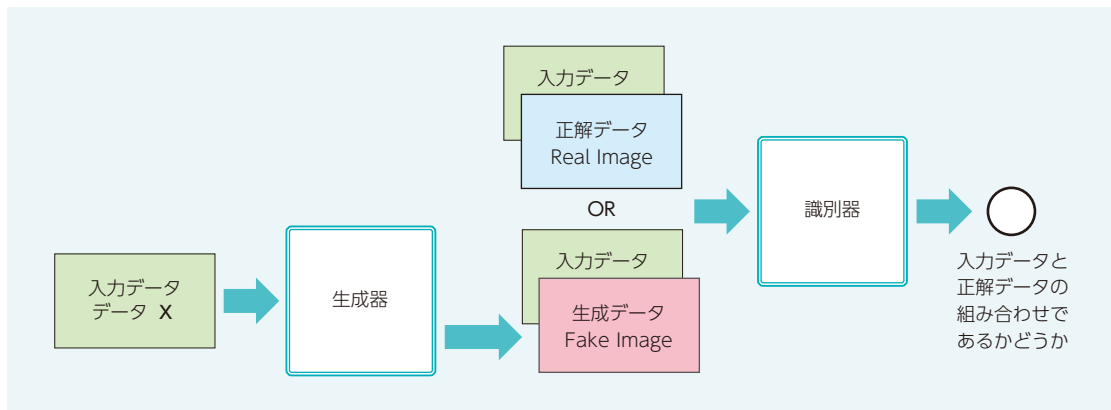


図2 pix2pixの構造イメージ

2.3 pix2pix

pix2pix は画像を与えて新たな画像を生成するための AI 技術の 1 つである。CNN を応用した手法であり、発表された論文では線画画像からカラー画像を生成、モノクロ画像からカラー画像、航空写真からマップ画像、とさまざまな画像を生成できることを示した。

pix2pix は画像をネットワークに入力して期待する画像を出力できるよう「学習」させる方法に特徴があり、図 2 のように 2 つのネットワークから構成する。1 つ目のネットワークは生成器と呼ばれ、入力データから画像データを生成する。2 つ目のネットワークは識別器と呼ばれ、入力したデータが生成器で生成した画像データと正解データどちらなのかを判別する。生成器は識別器を欺く画像データの生成を目的に、識別器は生成器が作った画像データを見破ることを目的に競い合うようにして学習させることで、生成器から正解データのような画像を出力できるようにする。

生成器のネットワークには、図 3 の U-Net⁽⁶⁾ という Encoder-Decoder 構造のネットワークを用いる。Encoder-Decoder 構造の Encoder は、入力データに対して畳み込み層とプーリング層を利用して徐々に特徴マップを小さくして、特徴を集約する。そして Decoder は小さくなった特徴マップを逆畳み込み層により徐々に大きくし

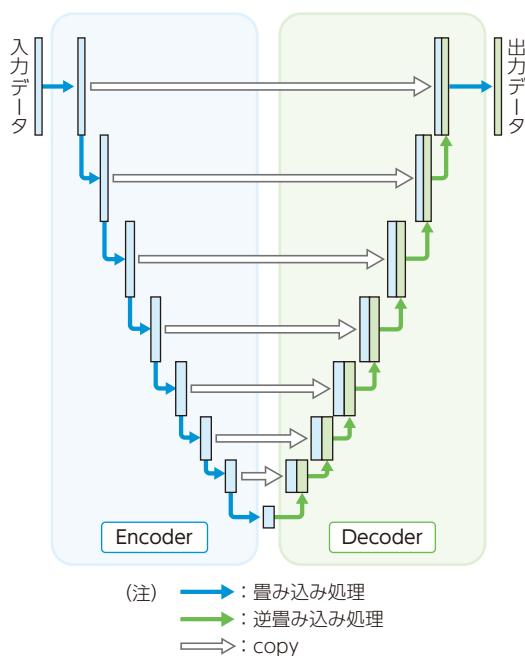


図3 U-Netの構造イメージ

て、集約した特徴から画像を生成する。U-Net は Decoder の入力に前層の出力データのほかに同じ深さの Encoder の出力データも使うという他の Encoder-Decoder 構造ネットワークとは異なる方法をとっている。Decoder の入力に前層からの集約した抽象的な特徴のほかに Encoder からの細かい特徴も用いることで出力データにおいて物体の形状を細かく描画できるようになったことが報告されている⁽⁵⁾。

3. AI 技術の検査計測技術への適用

当社では金属組織を SEM (Scanning Electron Microscope、走査電子顕微鏡) で観察し、「ポイド」と呼ばれる空隙を計測してそれらの個数・面積から金属製品の余寿命を評価している。金属組織をレプリカフィルムに転写し、SEM で撮影したグレースケール画像 (図 4 左) でポイドは白く見えるが、ゴミやレプリカフィルムの破れ、組織など、同じ白色で見える部位があり、単純にしきい値を設けて超過した部分をポイドとして抽出することはできない。またポイド自体が小さく見落とすしやすいため、熟練者が一つ一つ目視してポイド部分を抽出している。多くの SEM 画像に対して一つ一つ目視して抽出するため作業に時間がかかるという課題、作業者によって抽出結果が変わるといった課題があり、画像処理技術のみでは解決が難しいため前章に記載の AI 技術を使うことで解決を目指した。

3.1 CNN によるポイドの有無の判定

CNN は画像認識分野でよく使われており、入力した画像が例えば人や犬・猫などの何に分類されるのかを認識し、最も適切な分類へ仕分け、出力することができる。本件では、入力した SEM

画像がどの分類 (ポイド有り無し) の 2 つの分類に属するのかを認識することで、ポイドの有無を判定できることが期待できるため採用した。

3.1.1 学習データ

AI モデルを構築するにあたり、ポイドが写っているグレースケールの SEM 画像を 876 枚用意した。SEM 画像の画像サイズはそれぞれ 1280 × 960 [pixel] であり、これらを 112 × 112 [pixel] に分割し、その画像を学習データとした。分類した画像は、ポイドが有る画像と無い画像でそれぞれ「1」または「0」とラベリングし、「1」と判定されるデータをポイド有りとして判定させた (正解データ)。なお、分割の際に 112 で割り切れない画像は残りの部分を SEM 画像の背景と同じ「黒」に塗りつぶした。用意したデータの 9 割を学習用、残りの 1 割を評価用に割り当てた。

3.1.2 学習方法

AI モデルの構造は表 1 の通りである。112 × 112 [pixel] のグレースケール画像を入力して、ポイド有り無し の 2 つの分類に仕分け、それぞれの推定値を出力する。最適化アルゴリズムには一般的に使われている Adam⁽⁷⁾ を採用し、バッチサイズ 256 で 16 エポック学習した。

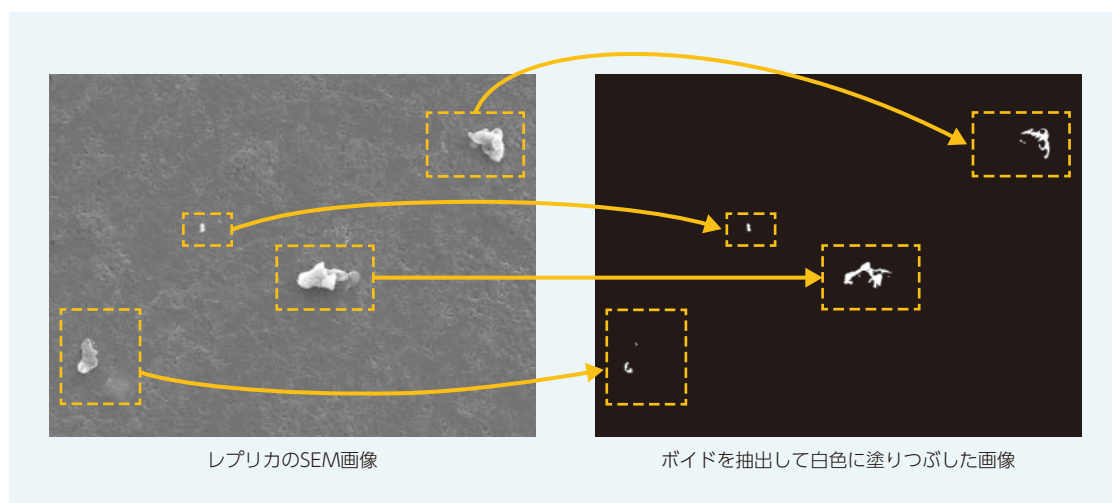
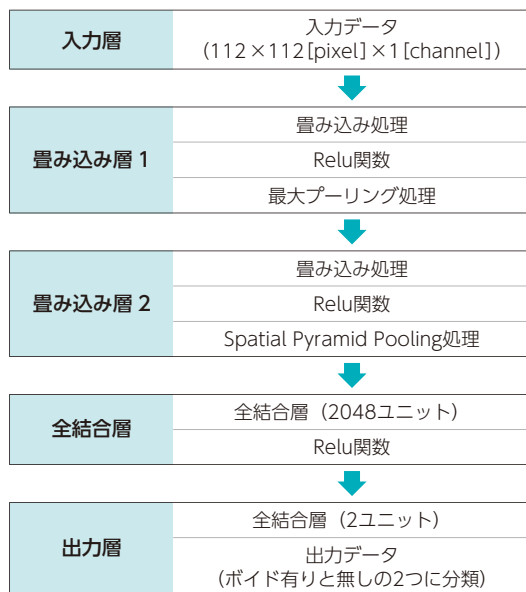


図 4 レプリカフィルムの SEM 画像から確認されるポイド抽出のイメージ

表1 採用した AI モデルの構造



3.1.3 学習結果

学習の結果、テスト用として用意したポイド有りと無しの計 1335 枚の SEM 画像のうち 1274 枚を正しく判定でき、95.4% の判定精度を得られた。評価結果を詳しく確認したところ、ポイドが有るにも関わらず無しと誤判定したケースでは、ポイドが画像データの端に有り、一部が欠けて写っていることが多くみられた。ポイドの見落としを防ぐために当社にて開発したポイド検出ソフトウェアでは、入力された SEM 画像に対して 112 × 112 [pixel] の範囲を 30 [pixel] ずつスライドしながら AI 判定させるという工夫を取り入れた。ポイド検出ソフトウェアでのポイド検出イメージを図 5 に示す。描画した枠線の色が赤色に近いほどポイドの可能性が高くなり、赤色の枠線はポイドの可能性が 100%、オレンジ色の枠線はポイドの可能性が 99% と判定したことを表している。図 5 では赤色の枠線の中にポイド（白色の物体）が見られる。これらの工夫を取り入れた結果、全体的に暗くコ

ントラストが低い SEM 画像でなければ、ポイドの可能性のしきい値を 99% 以上とすることで全てのポイドを漏れなく検出できることを確認した。

3.1.4 課題

SEM 画像の撮影条件などにより SEM 画像の明るさ・コントラストは一定ではなく、コントラストが強すぎたり弱すぎたりすると AI 判定で誤判定する傾向がみられた。どのコントラストであっても適切に判定できるよう、入力する SEM 画像の輝度値を正規化してコントラストを整える方法が考えられる。

また、ポイド検出ソフトウェアではポイドを検出したらそれを囲む矩形の枠線を描画するが、広範囲でポイドを検出すると描画した矩形枠線が複雑に重なりどこにポイドが有るのかわからなくなるという課題がある。直感的にポイドの位置がわかるようにするために、描画する枠線を矩形から多角形線に変更し、かつポイドが有る可能性が高い箇所絞って描画する枠線を減らす方法が考えられる。

今後はポイド検出ソフトウェアに入力した SEM 画像の輝度・コントラストを適切に調整する機能を付けるとともに、適切な輝度・コントラストを数値化して常に適切な SEM 画像を入力することで AI 判定可能となることを目指していく。

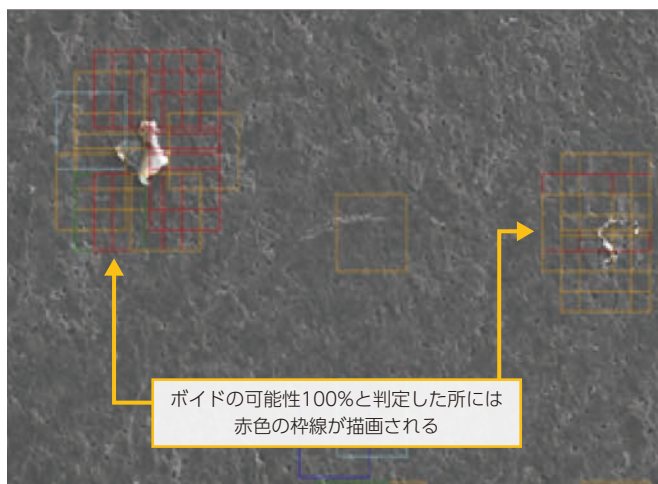


図 5 ポイド検出ソフトウェアを用いて検出したポイド

3.2 pix2pix による画像生成でのボイド抽出

pix2pix という画像を与えて新たな画像を生成する AI 技術を利用し、SEM 画像を与えてボイドのみを抽出した白黒画像を生成する AI モデルの開発を目指した。

3.2.1 学習データ

AI モデルを構築するにあたり、学習データとしてグレースケールの SEM 画像と、熟練者が目視して抽出したボイド部分を白色に塗りつぶした白黒画像の組み合わせを 497 組用意し、このうち 397 組を学習に使うデータ、100 組を AI モデルの精度評価に使うデータとして割り当てた。SEM 画像とボイド白黒画像の画像サイズはそれ

ぞれ 1280 × 960 [pixel] である。学習時は SEM 画像とボイド白黒画像を 256 × 256 [pixel] ごとに分割し、分割した SEM 画像を AI モデルへの入力データ、分割したボイド白黒画像を正解画像として用いた。

3.2.2 学習方法

pix2pix の既報⁽⁵⁾ ではカラー画像を入力してカラー画像を出力するため生成器の入力・出力チャンネル数は 3 に設定されていたが、学習データの SEM 画像とボイド白黒画像はそれぞれ 1 チャンネルのグレースケール画像であるため、生成器の入力・出力チャンネル数を 1 に変更した。AI モデルの各パラメータを更新させるための誤差の求め方、最適化アルゴリズムは pix2pix の既報

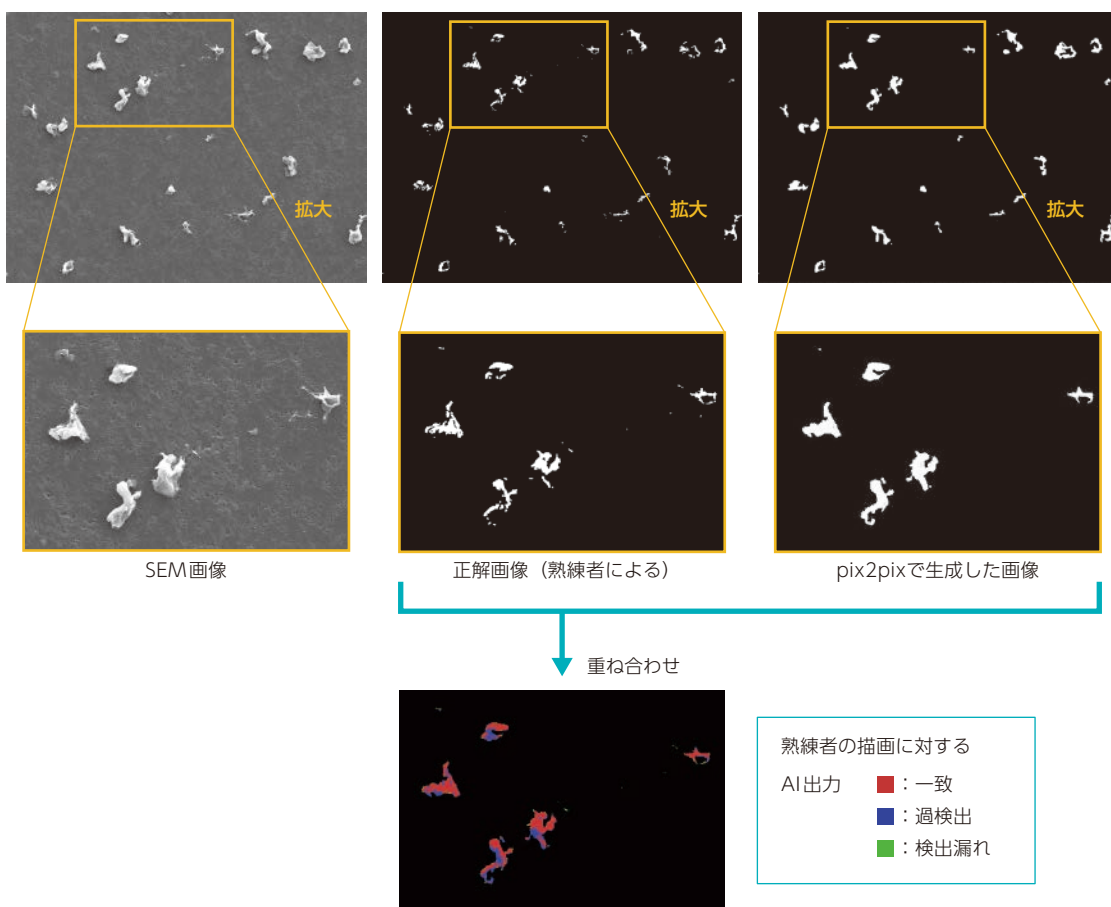


図6 熟練者により生成したボイド抽出画像と AI との比較

掲載の手法をそのまま採用した。学習回数は 40 エポックとした。

3.2.3 学習結果

学習の結果、SEM 画像に写る大半のポイドを抽出できることを確認した (図 6)。しかしポイド周辺の輝度が低いところまでをポイドとして抽出している結果が多くみられた。また、複数のポイドが集まっている部分の間隙を白く塗りつぶして 1 個のポイドとして抽出する箇所もあった。さらに小さく細かいポイドについては抽出できていない箇所も多くみられた。その結果、面積については熟練者が抽出した結果よりポイドとして抽出する部分が大きくなり、面積比で 1.6 倍となった。個数については小さく細かなポイドを抽出できていない、複数のポイドをまとめて 1 個として抽出している箇所があったため、個数比では 0.46 倍と少なくなった。今回、面積比は 1.6 倍、個数は 0.46 倍と出力されたが、ポイドと素地模様の判別は十分にできていることから、入力像の工夫や、学習をより進めることで実用に向けたさらなる精度向上が可能と見込まれる。

3.2.4 課題

pix2pix による画像生成でのポイドの抽出では、ポイドを余分に抽出する課題と小さく細かなポイドを抽出できていないという課題がある。余分な抽出を防いで小さく細かなポイドも抽出できるようにするために、3 点の改善方法が考えられる。

1 つ目は入力する SEM 画像の輝度・コントラストを調整することである。本件で構築した AI には、精度よくポイドを抽出できる輝度・コントラストが存在する。AI 処理に適した輝度・コントラストを数値化することにより、どのような SEM 画像でも画像調整で適切なポイド抽出が可能になると考えられる。

2 つ目はデータを追加しての学習を増強させることである。学習データを増やしてさまざまな形、サイズのポイドをより多く学ばせることで改善

できることが考えられる。

3 つ目は学び方の変更である。pix2pix は競い合いながら学習するという方法をとっているが、「識別器を欺けるように学習する」をやめて正解画像と完全に同じ画像を生成できるよう学び方を変えることで、熟練者が抽出したポイド抽出結果により近い結果を出力できるようになると考えられる。

以上 3 点の改善を試みたうえで、さらに精度向上の必要がある場合、pix2pix で採用した AI モデルの構造をより深くするなど AI モデルの構造を改良する方法がある。近年研究が進み、より AI モデルの層を深くして学ばせることで精度の高い AI モデルを生成できるようになっており、AI モデルの構造の改良によりどのような SEM 画像であっても大小さまざまなポイドを抽出できるようになると考えられる。

pix2pix の AI 処理の中にランダム要素が含まれるという課題がある。乱数シードを固定して毎回同じ結果になるようにしているが、パソコンの環境によっては同じ SEM 画像でも入力するタイミングによっては完全に同じ画像を生成できない可能性がある。ランダム要素を完全に排除して常に同じ結果を出力できるよう、AI モデルの構造を見直す必要があると考える。

今後は最適な輝度・コントラストを数値化して適切な SEM 画像を生成できるよう手順を整えるとともに、学習データの増強や AI 構造の見直しなどして AI のさらなる精度向上を目指していく。

4. まとめ

当社では、これまでに培ってきた画像処理技術を用いて検査・計測の作業効率化、付加価値向上を目的としたシステムの開発をしている。

近年のより高度な画像分析のニーズに対応するため、AI 技術を採用してシステムを研究・開発している。例えば CNN による SEM 画像に写るポイドの有無・位置を判定するシステムでは約

95%の判定精度を得られている。AIの改善や入力するSEM画像の最適な明るさ・コントラストを探るなどして判定精度の改善をして実用化を目指す。pix2pixによる画像生成でのボイド抽出ではSEM画像からボイド部分のみを白く塗りつぶした画像を生成することができたが、面積比で約1.6倍と大きめに抽出してしまう課題がある。熟練作業員の判定内容により近づけたAIを構築できるように、学習データの増強や構造の見直し、入力するSEM画像の最適な明るさ・コントラストを探るなどして精度を向上させて実工事への適用を目指す。

AI技術は日々進歩しており、新しいAI技術への挑戦と、培ってきた画像処理とAIとの連携技術の模索を実践することが重要であり、今後も判定精度の改善と他案件への転用を視野に入れ開発を進める。

参考文献

- (1) G.E. Hinton, L. Deng, D. Yu, G.E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath and B. Kingsbury. Deep Neural Networks for Acoustic Modeling in Speech Recognition. In Signal Processing Magazine, Vol.29, No.6, 2012, pp.82-97
- (2) A. Krizhevsky, I. Sutskever, and G.E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In Advances in Neural Information Processing Systems (NIPS), 2012, pp.1097-1105
- (3) C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanboucke, and A. Rabinovich. Going deeper with convolutions. In Conference on Computer Vision and Pattern Recognition (CVPR), 2014
- (4) K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image Recognition. In Conference on Computer Vision and Pattern Recognition (CVPR), 2016
- (5) P. Isola, J. Zhu, T. Zhou, and A. A. Efros. Image-to-Image Translation with Conditional Adversarial Networks. In Conference on Computer Vision and Pattern Recognition (CVPR), 2016
- (6) O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation, In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015
- (7) D.P. Kingma and J. Ba. Adam: A method for stochastic optimization. ICLR, 2015



制御システム事業部 産業システム部
製品・サービスグループ

畑 伸佳

TEL 045-759-2487
FAX 045-759-2491



制御システム事業部 産業システム部
デジタルイノベーショングループ

栗田 しのぶ

TEL 045-759-2486
FAX 045-759-2491



計測事業部 化学・材料部
磯子グループ 主査

田中丸 天兵

TEL 045-759-2122
FAX 045-759-2119



研究開発センター 研究管理グループ
グループ長

近藤 直美

TEL 045-791-3522
FAX 045-791-3547