



AI を用いた画像計測自動化

Nomura Kyohei
野村 恭兵*1

Kondo Naomi
近藤 直美*4

Tanakamaru Tempei
田中丸 天兵*2

Sato Katsuhiko
佐藤 雄大*5

Hata Nobuyoshi
畑 伸佳*3

Sakakibara Yohei
榎原 洋平*6

当社では、検査・計測の作業効率化および作業者によるばらつきの低減、技術伝承補助といった付加価値向上を目指して AI による各種金属組織特徴抽出システム開発を進めている。

本稿ではそれら取り組みの 1 つとして、火力発電ボイラで用いられているフェライト系耐熱鋼のクリープボイドについて、画像生成 AI である pix2pix を使い、ボイドを含む SEM 画像よりボイドの判定ならびにそのボイド領域の自動計測を試みた。その結果、作成した 4 種類の AI モデルのうち、寿命の初期、中期、後期を分類せず、かつ SEM の元画像を「回転 + 反転」によりデータ拡張して得られたモデルで評価指数 F が最も高い値を示した。一方、コントラストが低い SEM 画像ではボイド検出精度が低下する傾向があることも確認された。コントラストが低い SEM 画像については、画像の明るさ / コントラストを調整することでボイドの検出精度が向上することも確認した。

キーワード： AI、画像自動計測、pix2pix、クリープボイド、電子顕微鏡、SEM 画像、評価指数 F、画像処理、明るさ / コントラスト範囲

1. はじめに

金属材料の機械特性は材料自身を構成する金属組織に依存し、機械製品の設計から製品使用開始後の保守、損傷発生時の原因究明などさまざまな場面で重要な情報として収集・活用されている。それら金属組織情報は一般に金相試験にて取得され、当社でも検査・計測事業における材料

調査サービスの一環⁽¹⁾として、試験室では製品ないしサンプルの切断面から、現地では Suzuki's Universal Micro Printing Method (SUMP 法：各種製品表面の金属組織をフィルム等に転写させて非破壊的に観察する手法。SUMP 法にて金属組織を転写したフィルム等を以降レプリカと称す)により製品表面から金属組織情報を取得するサー

*1：計測事業部 化学・材料部 材料グループ グループ長 工学博士

*2：計測事業部 化学・材料部 礫子グループ 主査

*3：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

*4：研究開発センター 研究管理グループ グループ長

*5：株式会社IHI 技術開発本部 技術基盤センター 材料・構造技術部

*6：株式会社IHI 技術開発本部 技術基盤センター 材料・構造技術部 主幹 工学博士

ビスを提供している。また昨今では、検査・計測の作業効率化および作業者によるばらつきの低減、技術伝承補助などの付加価値向上を目指し、前述した金属組織から AI による材料組織の特徴を抽出するシステム開発にも取り組んでいる⁽²⁾。

一方、火力発電ボイラで用いられるフェライト系耐熱材料では、プラント運転に伴い、特に溶接継手の熱影響部においてクリープボイドと呼ばれる材料劣化が発生・増加し⁽³⁾、その発生量を定量的に把握することが発電プラントの安定運用に重要となる。ボイドの定量把握方法としては、定期検査等にてボイラ配管の表面レプリカを採取し、光学顕微鏡や電子顕微鏡を用いて観察、得られた画像よりボイドの個数密度や面積率等の情報を定量化する手法があり⁽⁴⁾、それらの情報を用いて使用材料の寿命評価が行われている。このようなボイドの観察・評価はこれまで熟練技術者により実施されてきたが、今後少子高齢化がより進んでいくことが見込まれており、ボイドを定量評価するための熟練技術者の確保や技術伝承がより困難な状況となっている。この解決策の一つとして、当社で取り組んでいる AI 技術を用いてボイド定量を行うことができれば、前述した技術伝承・維持に役立てることができる。

本稿では当社で取り組んだ AI を用いたボイド計測の自動化について、その検討結果を紹介する。具体的には、画像生成 AI である pix2pix^{(2)・(5)}を用い、火力発電ボイラで用いられているフェライト系耐熱鋼のクリープボイドについて、走査型電子顕微鏡 (Scanning Electron Microscope : SEM) を用いて取得した SEM 画像よりボイドの判定ならびにそのボイド領域の自動計測を試みた。

2. AI 技術 pix2pix の学習条件および評価方法

pix2pix⁽⁵⁾ は画像を与えて新たな画像を生成する AI 技術の 1 つである。詳細は前稿⁽²⁾を参照のこと。本稿では前稿同様、表面情報を転写

して得たレプリカについて、走査型電子顕微鏡を用いて SEM 画像を取得し、pix2pix に与えた。pix2pix ではボイドのみを抽出した白黒画像を生成させた後、ボイドの正解画像と比較・評価した。

2.1 学習データおよび学習方法

AI モデルを構築するにあたり、使用データとしてグレースケールの SEM 画像 (代表例 : 図 1) とその画像を熟練技術者が目視しボイドとして抽出した部分を白色に塗りつぶした白黒 2 値化画像 (ボイド正解画像) の組み合わせを 48 組用意し (1 組の学習データは SEM 画像とボイド正解画像各 100 枚、48 組のデータ中には材料の寿命消費率が初期～後期までの幅広い範囲のデータを準備)、このうち 46 組を学習データに、2 組を AI モデルの検証データとして割り当てた。SEM 画像およびボイド正解画像のサイズは画像 1 枚あたり 1280 × 960 [pixel] である。学習時は SEM 画像とボイド正解画像を 128 × 96 [pixel] ごとに分割し、分割した各画像を AI モデルへ入力した。学習方法については既稿⁽²⁾をベースとしたが、学習画像数の拡張も行った。また、学習回数は 30 エポックとした。



図 1 ボイド SEM 画像代表例

2.2 AI 出力結果の正誤判定および精度評価

統計解析で使用される評価指数 F 値⁽⁶⁾を用いて検証データより AI が生成した画像と正解画像の比較を行った。AI 出力結果の正誤判定は

表 1 正誤判定の分類

ボイドの判定	検証（正解）画像あり	検証（正解）画像なし
AI がありと判定	TP（ボイドとして検出した - 正解）	FP（過検出）
AI がなしと判定	FN（検出漏れ）	TN（ボイドとして検出しなかった - 正解）

適合率 P：AI が「あり」と判断したもののうち、検証画像も「あり」だった割合

$$P = TP / (TP + FP) \quad \cdots (1)$$

再現率 R：検証画像が「あり」であるもののうち、AI が「あり」と判断した割合

$$R = TP / (TP + FN) \quad \cdots (2)$$

評価指数 F：P と R の調和平均であり、0 ～ 1.0 の値で示される。P と R が 1.0 の場合、F も 1.0 となる

$$F = 2 \{PxR / (P + R)\} \quad \cdots (3)$$

表 1 に示す 4 種類に分類し、評価指数 F 値は上記式 (1) ～ (3) により算出した。

以上を評価指針とし、次章以降、検証結果を述べる。なお、適合率 P はその値が 1.0 に近いほどボイドの過検出が少なく、再現率 R はその値が 1.0 に近いほどボイドの検出漏れが少ないことを意味する。従って、いずれの値も高く、かつ評価指数 F が高いものほど AI として良いモデルが作成されたと判断することができる。

3. pix2pix を用いた AI モデルの検討

3.1 評価精度に及ぼす学習データパターンの影響

46 組の学習データを全て与えたパターンおよび寿命初期のみの 12 組、中期のみの 22 組、後期のみの 12 組をそれぞれ与えたパターンの計

4 種類の学習データパターンにて pix2pix による AI モデルを作成し、No.1 および 2 の 2 組の検証データにて各 AI モデルの評価精度を検証した。各検証データの違いは No.2 の方が No.1 よりコントラストが高くボイドと素地の境界が明確な点である。また、検証データ No.1 および 2 はいずれも寿命の初期となる検証データである。

AI モデルごとの精度検証結果を表 2 に示す。コントラストの低い No.1 では学習データパターンにより適合率 P および再現率 R が変化し、全データ使用または、寿命後期のみの学習データパターンにおいて評価指数 F が高い値を示した。一方でコントラストの高い No.2 では学習データパターンによる適合率 P および再現率 R の変化は僅かであった。双方の結果を踏まえると、学習データパターンは全ての学習データを使用することが望ましいと言える。加えて、コントラストは高い方が検証精度も高くなる傾向があることも確認された。

表 2 各 AI モデルによる精度検証結果一覧

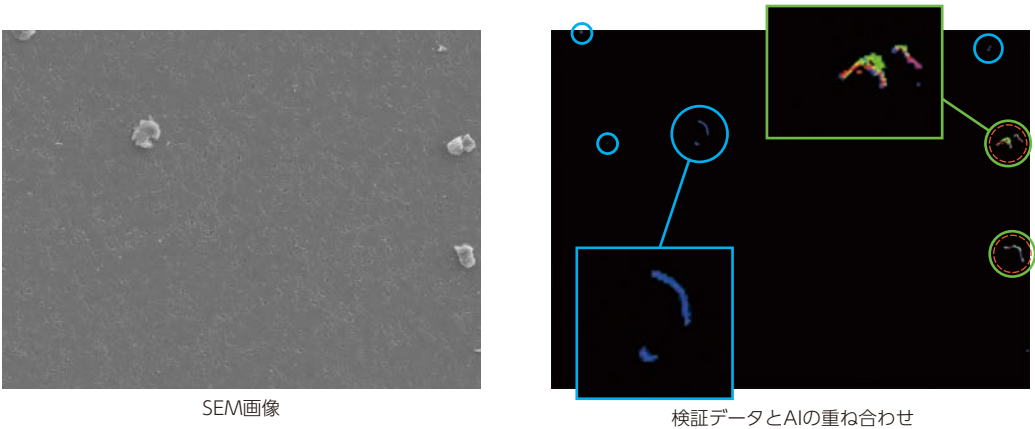
検証データ	学習データパターン	正解値との重ね合わせ結果 [pixel]					適合率	再現率	評価指数 F
		正解ボイド	出力ボイド	一致赤 TP	過検出青 FP	検出漏れ緑 FN	P 過検出 少	R 検出漏れ 少	F
No.1	全データ	110694	49718	41114	8604	69580	0.827	0.371	0.513
	寿命初期	110694	44098	33959	10139	76735	0.770	0.307	0.439
	寿命中期	110694	77209	33362	43847	77332	0.432	0.301	0.355
	寿命後期	110694	101952	63413	38539	47281	0.622	0.573	0.596
No.2	全データ	160374	164169	125536	38633	34838	0.765	0.783	0.774
	寿命初期	160374	171105	123374	47731	37000	0.721	0.769	0.744
	寿命中期	160374	165414	120612	44802	39762	0.729	0.752	0.740
	寿命後期	160374	191809	130198	61611	30176	0.679	0.812	0.739

次に、46 組の全データにて学習した AI モデルにおいてどのような場合に AI がボイドの検出漏れ、ないし過検出するか確認した。正誤判定は、AI で生成した 2 値化画像と正解画像を比較し、正解を赤、検出漏れを緑、過検出を青で描画させた。コントラストの低い No.1 の検証データについて、SEM 画像および AI により生成された 2 値化画像とボイド正解画像の比較画像代表例を図 2

に示す。コントラストの低い No.1 のデータでは、AI は明るさの低い粒子や素地部分を過検出し、また、ボイドの明るさが高い部分の近傍で検出漏れる傾向がみられた。

3.2 評価精度に及ぼす学習データ拡張の影響

3.1 項の検証データ No.1 および 2 のいずれに



	正解画像あり	正解画像なし
AI がありと判定	TP (ボイドとして検出した - 正解)	FP (過検出)
AI がなしと判定	FN (検出漏れ)	TN (ボイドとして検出しなかった - 正解)

図 2 SEM 画像および AI により生成された 2 値化画像とボイド正解画像の比較画像代表例（検証データ No.1）

表 3 学習データのボイド画像拡張による評価精度検証結果一覧

検証データ	AI 学習条件	正解値との重ね合わせ結果 [pixel]					適合率	再現率	評価指数 F
		正解ボイド	出力ボイド	一致赤 TP	過検出青 FP	検出漏れ緑 FN	P 過検出 少	R 検出漏れ 少	F
No.1	水増し なし	有効なモデルを生成できず							
	反転のみ	111898	143670	65709	77961	46189	0.457	0.587	0.514
	回転のみ	111898	71409	56236	15173	55662	0.788	0.503	0.614
	回転 + 傾斜	111898	63458	49797	13661	62101	0.785	0.445	0.568
	回転 + 反転	111898	72579	58398	14181	53500	0.805	0.522	0.633
	傾斜 + 伸長	111898	106112	75290	30822	36608	0.710	0.673	0.691
	回転 + 反転 + 傾斜	111898	81017	60854	20163	51044	0.751	0.544	0.631
	回転 + 反転 + 傾斜 + 伸長	111898	117687	41625	76062	70273	0.354	0.372	0.363
No.2	水増し なし	有効なモデルを生成できず							
	反転のみ	161877	275331	54871	220460	107006	0.199	0.339	0.251
	回転のみ	161877	218407	142135	76272	19742	0.651	0.878	0.748
	回転 + 傾斜	161877	198875	135850	63025	26027	0.683	0.839	0.753
	回転 + 反転	161877	210015	140807	69208	21070	0.670	0.870	0.757
	傾斜 + 伸長	161877	281749	148768	132981	13109	0.528	0.919	0.671
	回転 + 反転 + 傾斜	161877	225431	143132	82299	18745	0.635	0.884	0.739
	回転 + 反転 + 傾斜 + 伸長	161877	257668	133185	124483	28692	0.517	0.823	0.635

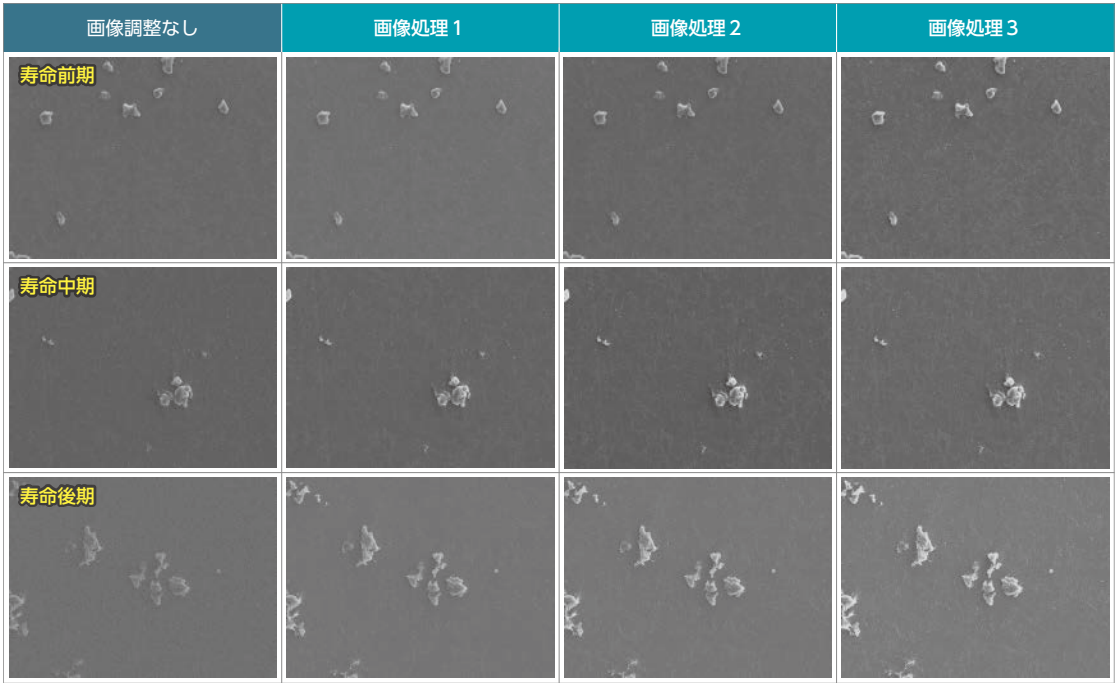
においても学習データが最も多いモデルで評価指数 F の値が高い結果となった。そこで、学習に用いた各データ画像に反転および回転、変形等加え、学習データを拡張することで評価精度が向上するか検証した。学習データのボイド画像拡張による評価精度検証結果を表 3 に示す。No.1 および 2 でそれぞれ傾向に若干の差はあるものの、AI 学習に使用するデータ数を拡張することで適合率 P および再現率 R が向上、評価指数 F も高くなる傾向がみられた。また、いずれもデータ数が多くなりすぎると適合率 P および再現率 R が低下し評価指数 F が低下する、過学習の傾向も確認された。今回の検討範囲では、いずれの検証データでも評価指数 F の値が高く、かつ適合率 P および再現率 R のバランスが良かったモデルは「回転 + 反転」で作成したモデルであった。また、No.1 および 2 の評価指数 F を比較すると、SEM 画像のコントラストが高い No.2 の方が高い F 値を示す傾向は 3.1 項と変わらなかった。以上より、学習データ

拡張は「回転 + 反転」を行った場合に最も評価精度が高くなることが確認された。また、本項で作成された pix2pix の AI モデルでは No.1 のようなコントラストの低い SEM 画像のボイド抽出は不得手であることも確認された。なお、データ拡張をせずに学習した AI モデルにおいて、有効な AI モデルが作成されなかった理由としては、視野内に占めるボイドの比率が少なすぎたことが考えられる。

4. pix2pix を用いた AI モデルの精度向上

4.1 検証データの明るさ / コントラスト (B/C) 調整による評価精度への影響

3 章の検討にて本稿の学習データ量の範囲において、検証データのコントラストが低い場合は学習データ量により評価指数 F が変化し、コントラストが高い検証データでは評価指数 F は学習



画像処理 1：B/C 調整：弱（ボイドの B/C を少し高めた：元画像に最も近い）
画像処理 2：B/C 調整：中（画像処理 1 と 3 の中間）
画像処理 3：B/C 調整：強（ボイドの B/C を最も高めた：元画像から最も遠い）

図 3 各検証データにおける B/C 調整前後の SEM 画像代表例

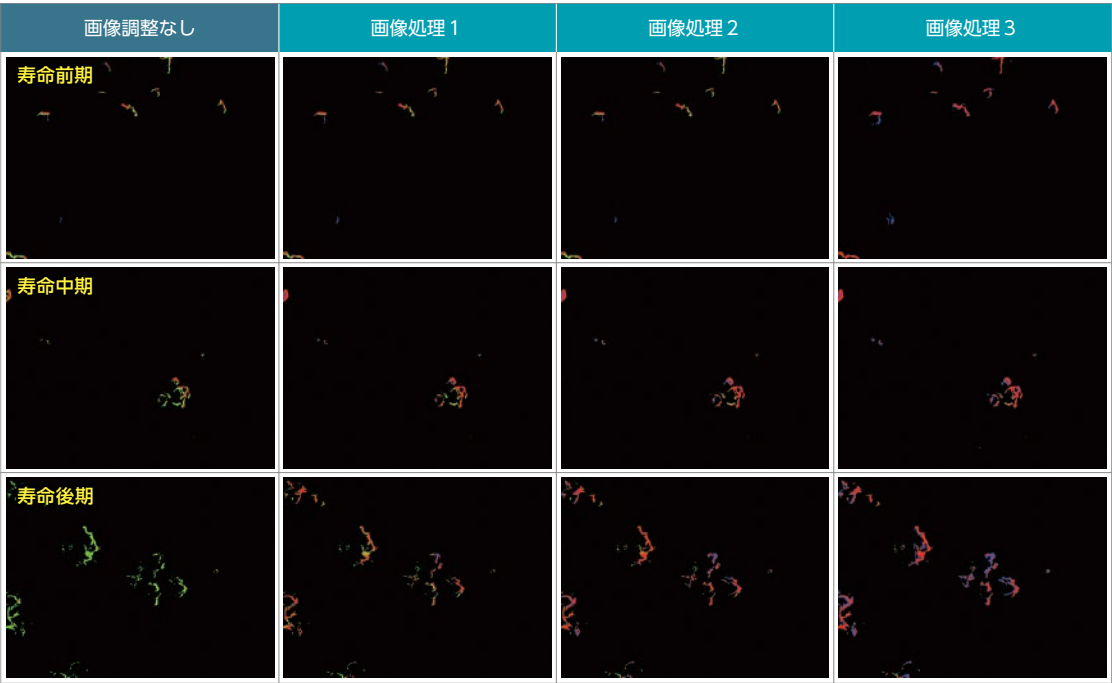
データ量にほとんど影響されなかった。すなわち、3章にて作成したAIモデルでは適切なB/C範囲のデータを与えることでボイド定量における評価精度がより向上する可能性がある。そこで、コントラストの低い検証データを用いてそれらSEM

画像のB/C調整を行い、評価精度の変化を確認した。なお、B/C調整による精度向上確認は、寿命初期のNo.1のSEM画像に加えて、寿命中期および寿命後期の検証データも新たに準備し、実施した。

表4 B/C調整したSEM画像のAIによる評価精度一覧

AIモデル：回転＋反転		正解値との重ね合わせ結果[pixel]					適合率	再現率	評価指数F
検証データ	検証画像の調整	正解ボイド	出力ボイド	一致赤 TP	過検出青 FP	検出漏れ緑 FN	P 過検出 少	R 検出漏れ 少	F
No.1 寿命初期	調整なし	111898	72579	58398	14181	53500	0.805	0.522	0.633
	画像処理1	111898	101857	75978	25879	35920	0.746	0.679	0.711
	画像処理2	111898	87808	68667	19141	43231	0.782	0.614	0.688
	画像処理3	111898	206725	103526	103199	8372	0.501	0.925	0.650
No.3 寿命中期	調整なし	258360	140690	99648	41042	158712	0.708	0.386	0.499
	画像処理1	258360	231670	183459	48211	74901	0.792	0.710	0.749
	画像処理2	258360	279972	213855	66117	44505	0.764	0.828	0.795
	画像処理3	258360	334108	230360	103748	28000	0.689	0.892	0.778
No.4 寿命後期	調整なし	575085	54037	21373	32664	553712	0.396	0.037	0.068
	画像処理1	575085	360433	318182	42251	256903	0.883	0.553	0.680
	画像処理2	575085	567493	440278	127215	134807	0.776	0.766	0.771
	画像処理3	575085	866373	525106	341267	49979	0.606	0.913	0.729

画像処理1：B/C調整：弱（ボイドのB/Cを少し高めた：元画像に最も近い）
画像処理2：B/C調整：中（画像処理1と3の中間）
画像処理3：B/C調整：強（ボイドのB/Cを最も高めた：元画像から最も遠い）



画像処理1：B/C調整：弱（ボイドのB/Cを少し高めた：元画像に最も近い）
画像処理2：B/C調整：中（画像処理1と3の中間）
画像処理3：B/C調整：強（ボイドのB/Cを最も高めた：元画像から最も遠い）

図4 B/C調整したSEM画像を用いてAIで生成した画像とボイド正解画像の比較

各検証データにおける B/C 調整後の SEM 画像代表例を図 3 に、B/C 調整した SEM 画像の AI による評価精度一覧を表 4 に示す。画像調整なしの場合、寿命消費が大きいものほど評価指数 F は低く、特に寿命後期である No.4 の検証データでは F 値は 0.068 となり、ボイドの検出漏れが著しかった。一方、画像調整を行った場合では寿命初期の検証データ No.1 では画像処理の影響がほとんどみられなかったものの、寿命中期以降では評価指数 F は大幅に改善し、特にボイドの検出漏れとなる再現率 R が著しく改善された。

各検証データについて、B/C 調整した SEM 画像を用いて AI で生成した画像とボイド正解画像を比較した画像代表例を図 4 に示す。いずれの検証データにおいても SEM 画像の B/C 調整を行わない場合、緑色で示すボイドの検出漏れが多く確認された。しかしながら、B/C 調整の程度を大きくするとボイドの検出漏れは減少し、赤色で示す正解の部分が増加していることも確認された。また、B/C 調整の程度を過度に大きくすると青色で示す過検出部が増加する傾向があることも確認された。今回検討した B/C 調整においては画像調整 2 の場合に赤色で示す正解の部分が最も多く、正解画像により近い結果が得られた。また、素地の模様についても問題なく処理できていることも確認された。

4.2 AI に適した画像処理範囲の検討

4.1 項において、No.3 および 4 のデータを用いた検証では SEM 画像の B/C 調整により評価指数 F が大きく変化し、いずれにおいても画像処理 2 で評価指数 F が最も高い値となった。また、ボイドの検出漏れおよび過検出も最も抑制されていた。そこで、AI にて高精度にボイド検出可能な SEM 画像の B/C 範囲を明確にすることを目的とし、検証データ No.3 および 4 の SEM 元画像ならびに 4.1 項の画像処理 2 の各 100 枚ずつ、計 400 枚の画像に対してグレースケールの 256 階調と頻度を抽出、階調分布を確認した。検証データ No.3 および 4 の階調分布を図 5 に示す。検証データごとに階調分布の幅に差異はみられたものの、いずれの検証データにおいても元画像の階調分布形状を維持した上で階調幅を広げつつ白飛び・黒飛び（SEM 画像中で輝度の階調が 255 を上回る、または 0 を下回る領域）が抑制された階調分布となっていた。

次に、検証データ No.3 および 4 の画像処理 2 の SEM 画像について、具体的にどの階調に対して AI がボイド検出を行っているか確認した。結果をそれぞれ図 6 および 7 に示す。検証データ No.3（寿命中期）では、素地の明るさの階調は平均 86 階調であることに對し、ボイド部の階調

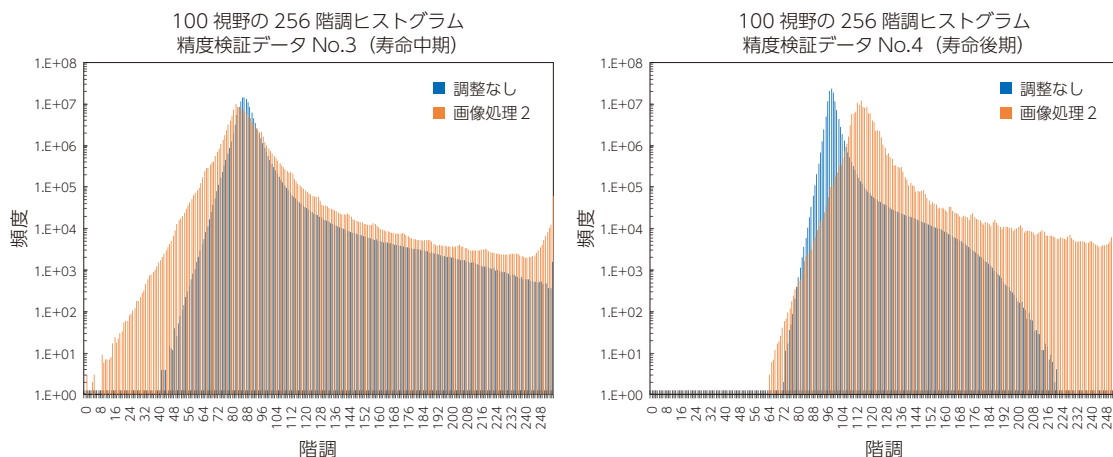


図 5 検証データ No.3 および 4 の画像処理有無による階調分布変化

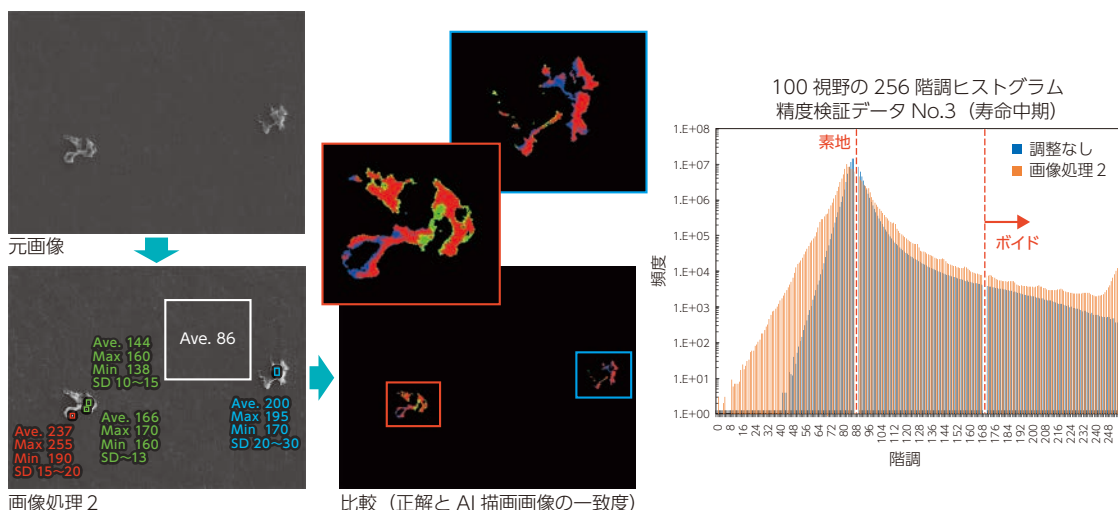


図 6 検証データ No.3 におけるボイド検出評価結果

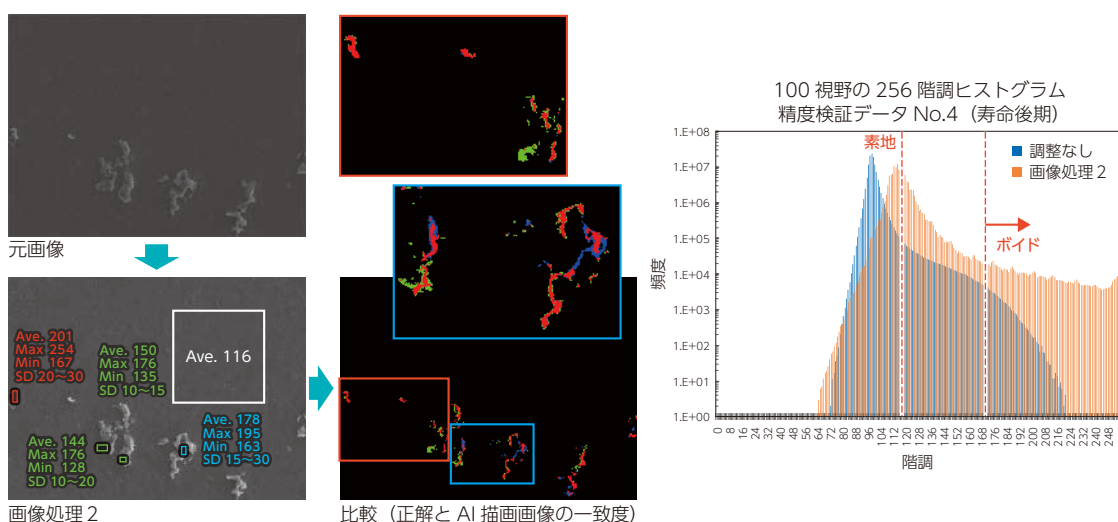


図 7 検証データ No.4 におけるボイド検出評価結果

は 170 以上、かつ微小エリアの階調幅（コントラスト）となる標準偏差がおおむね 15 ～ 30 程度であった。検証データ No.4（寿命後期）では、素地の階調は平均 116 階調であることに對し、ボイド部の階調は 160 以上、かつ微小エリアの標準偏差がおおむね 15 ～ 30 程度であった。すなわち、AI はボイドと素地の明るさの階調差に加えて、階調幅となる標準偏差も踏まえてボイド検出を実施していると考えられる。

以上の結果より、本稿で作成した AI では素地

とボイドの明るさに一定量の差を持たせ、かつボイド部全体を白飛び・黒飛びさせない SEM 画像をインプットデータとすることでボイドの検出精度を向上できることが明らかとなった。

5. まとめと今後の課題

AI を用いてボイドを自動計測することを目的とし、AI モデルの作成およびその精度向上について検討、その結果を紹介した。詳細は以下のとおりである。

(1) pix2pix を用いた AI モデルの検討

48 組（1 組の学習データはグレースケールの SEM 画像とボイド正解画像各 100 枚）の全データ中、46 組を学習データとし、4 種類の AI モデルを作成、2 組の検証データを用いて AI モデルの精度検証を実施した。その結果、最適な AI モデルは 46 組の全学習データを使用した AI モデルであった。

全学習データを使用した AI モデルを用いて学習データを拡張し、評価精度への影響を検証した結果、SEM の元画像を「回転 + 反転」により拡張した AI モデルで最も評価指数 F が向上することが確認された。ただし、本検討で作成された AI モデルではコントラストの低い SEM 画像のボイド抽出が不得手であることも確認された。

(2) pix2pix を用いた AI モデルの精度向上

寿命初期・中期・後期の 3 種類の検証データについて各 SEM 画像の明るさ (B) / コントラスト (C) 調整を行い評価精度への影響を調査した。その結果、今回検討した B/C 調整においては SEM 元画像の B/C が AI 処理に適していない場合でも画像の B/C 調整を行うことでボイド計測の精度を向上可能なことを明らかにした。また、画像調整 2 の場合に正解画像に最も近い結果が得られ、素地の模様を除外してボイド計測可能なことも確認した。

SEM 元画像と画像処理 2 を施した SEM 画

像のグレースケール階調分布を確認した結果、今回作成した AI では素地とボイドの明るさに一定量の差があり、かつボイド部全体を白飛び・黒飛びさせない SEM 画像をインプットデータとすることでボイド定量における評価精度の向上が見込めることを明らかにした。

以上の結果より、今回作成した pix2pix の AI モデルは SEM 画像のボイド計測自動化に有用であることを確認した。今後の課題としては AI を適用させるための SEM 画像インプットデータの明るさ / コントラスト範囲をより明確にすることである。また、今回得られた F 値が実際の画像計測に適用可能な精度であるかは評価基準が別となるため、ユーザーが実運用に求める性能を本 AI が満たすか確認した上で AI 適用を進める必要もある。

参考文献

- (1) IHI 検査計測ホームページ：
<https://www.iic-hq.co.jp/>
- (2) 畑 伸佳 他：画像分析への AI 適用による検査計測技術の向上、IIC REVIEW、No.73、2025/04、pp.40-46
- (3) 渡部 隆 他：9Cr-1Mo-V-Nb 鋼溶接継手の長時間クリープにおける Type IV 破壊と微細組織、鉄と鋼、Vol.90、No.4 (2004)、pp.206-212
- (4) 新田 明人：ボイラの寿命予測、安全工学、Vol.30、No.6 (1991)、pp.465-471
- (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) J. B. Brown: Classifiers and their Metrics Quantified, DOI: 10.1002/minf.201700127



計測事業部 化学・材料部
材料グループ グループ長 工学博士
野村 恭兵

TEL 045-791-3516



計測事業部 化学・材料部
機子グループ 主査
田中丸 天兵

TEL 045-759-2122



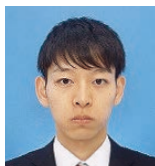
制御システム事業部
産業システム部 製品・サービスグループ
畑 伸佳

TEL 045-759-2487



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



株式会社 IHI
技術開発本部 技術基盤センター
材料・構造技術部
佐藤 雄大

TEL 070-3239-9315



株式会社 IHI
技術開発本部 技術基盤センター
材料・構造技術部 主幹 工学博士
榊原 洋平

TEL 070-3239-6601