少量データによる AI 構築技術及び実応用に関する研究

前島 崇宏*　西本 圭志**

1. はじめに

近年の AI 技術の著しい進展もあり、製造ラインにおける外観検査に AI が搭載された装置が使われるようになってきている。しかし、検査精度が高い AI を構築するには学習用画像を多く集める必要があり労力を要する、NG 品のピッキングまで自動化すると装置が大掛かりな高価なものとなってしまうなどの理由により資金力に乏しい中小企業が実用化するには課題がある。

2. 目的

中小企業にも導入しやすい、作業者にピッキングさせる AI 検査システムを目指し、AI の学習の効率化と AI の判定結果を作業者に示す装置を開発することとした。

3. 研究内容

3.1 AI を構築する際の労力削減を図るため研究開発

3.1.1 アノテーションコストの削減

3.1.1.1 課題

異物や欠陥等の検査に AI を用いた画像認識技術の適用が期待されているが、検査精度が高い AI を構築するためには、膨大な学習データが必要となる。画像の各ピクセルに対して正常品や異物などの分類が可能な手法として、セマンティックセグメンテーションが知られている。セマンティックセグメンテーションはピクセル単位で高精度な検査が行える一方、学習にはピクセル単位でクラス分けされた画像が必要となるため、画像判別や物体検出などの手法と比較し、アノテーションコストが高いという課題がある。

3.1.1.2 目的

本研究では、少量の異物混入画像に、「入手やアノテーションが容易な正常品のみが写った画像（背景や異物が写っていない画像）」と、「異物のみが写った画像（背景や正常品が写っていない画像）」を追加し、学習することで、アノテーションコストを削減しながらセマンティックセグメンテーションを用いた検査精度を向上させることを目的とした。

3.1.1.3 提案手法

学習には、正常品を模擬した M6 ボルト、異物を模擬した M8 ボルトおよび、異物混入画像を模擬した M6 ボルトに M8 ボルトが 1 本混入した画像を用いた。

図 1 に学習に用いた画像とアノテーション画像の例を示す。生産現場の検査に応用することを想定し、評価用画像には、異物混入画像 50 枚（475×475 ピクセル）を用いた。

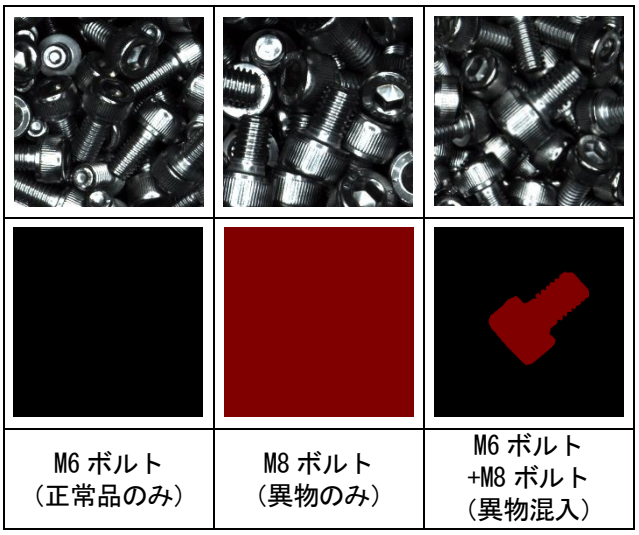


図 1　学習に用いた画像の例（上段：生画像、下段：アノテーション画像）（750×750 ピクセル）

3.1.1.4 学習条件

ADE20K データセットで、学習済みの PSPNet[1] を用いて転移学習を行った。前処理として、ランダムに平行移動、回転（0 度、90 度、180 度、270 度）、反転後、画像中央の 475×475 ピクセルを切り抜き、学習を行った。4,000 エポック学習後、最も mIoU（mean Intersection over Union）が高くなったときの値を求めた。

異物混入画像の枚数が 0 枚、1 枚、5 枚、10 枚、50 枚のときに、正常品のみ画像および異物のみ画像をそれぞれ 50 枚（計 100 枚）、100 枚（計 200 枚）、200 枚（計 400 枚）、400 枚（計 800 枚）、800 枚（計 1,600 枚）追加した場合の mIoU をそれぞれ求めた。

また、提案手法を用いた学習の有効性を評価するため、従来手法である異物混入画像（10 枚、20 枚、30 枚、40 枚、50 枚）を用いた学習についても、同様の評価を行った。

3.1.1.5 研究結果と考察

表 1 に、提案手法を用いて学習した場合の学習条件と mIoU の関係を、表 2 に従来手法を用いて学習した場合の学習条件と mIoU の関係を示す。

表 1　学習条件と mIoU の関係（提案手法）

n=3

異物混入画像の枚数〔枚〕	正常品/異物のみ画像の枚数〔枚〕				
	50/50	100/100	200/200	400/400	800/800
0	0.454	0.454	0.475	0.432	0.458
1	0.768	0.835	0.829	0.826	0.870
5	0.911	0.920	0.918	0.915	0.915
10	0.933	0.934	0.933	0.937	0.939
50	0.953	0.954	0.954	0.955	0.955

* IT・マテリアルグループ　**研究推進グループ

表 2　学習条件と mIoU の関係（従来手法）　n=3

異物混入画像の枚数〔枚〕	10	20	30	40	50
mIoU	0.297	0.902	0.919	0.927	0.936

3.1.1.5.1 従来手法との比較

従来手法である異物混入画像 50 枚で学習した場合、表 2 に示したように、の mIoU は 0.936 であった。異物混入画像 10 枚に正常品のみ画像および異物のみ画像をそれぞれ 800 枚追加して学習した場合、表１に示したように、mIoU は 0.939 であり、従来手法の検査精度を上回った。

さらに、「正常品のみ画像」や「異物のみ画像」は、自動でアノテーション可能なため、人によるアノテーションが必要な画像を５分の１に削減できた。

3.1.1.5.2 正常品/異物のみ画像の枚数と検査精度の関係

3.1.1.5.2.1　異物混入画像を用いない場合

表 1 に示したように、異物混入画像を用いず、「正常品のみ画像」および「異物のみ画像」を用いて学習した場合の mIoU は、学習用画像の枚数にかかわらず、0.5 以下と小さな値となった。

一方、学習用画像に 1 枚でも異物混入画像が含まれている場合、「正常品のみ画像」および「異物のみ画像」を追加し学習することで、mIoU の値が 0.76 以上となり、異物混入画像が含まれない場合と比較して、高い値となった。

異物混入画像を用いず学習した場合、評価用の異物混入画像の多くを正常品のみ画像と誤判別したことが検査精度低下の原因であった。

3.1.1.5.2.2　異物混入画像を用いた場合

表 1 に示したように、異物混入画像 10 枚、正常品のみ画像、異物のみ画像各 50 枚、計 110 枚で学習した場合の mIoU は 0.933 となった。一方、異物混入画像１枚、正常品のみ画像、異物のみ画像各 800 枚、計 1601 枚で学習した場合の mIoU は 0.870 である。学習用画像の枚数が約 15 倍に増加したにもかかわらず、mIoU は低い値となった。「正常品のみ画像」、「異物のみ画像」のアノテーション画像にはボルトの輪郭に関する情報が含まれていないため、異物混入画像と比較して、学習用画像１枚当たりの検査精度向上の効果が低いためと考えられる。

表２に示したように異物混入画像 10 枚を用いて学習した場合、mIoU は 0.297 と非常に低い値となった。その理由は、モデルが学習を行うために必要な最低限の画像枚数に達していなかったためと考えられる。表 1 に示したように、「正常品のみ画像」、「異物のみ画像」を追加すると、mIoU が 0.93 以上に向上しているので、提案手法は、異物混入画像の入手が困難な場合に有効であることが分かる。

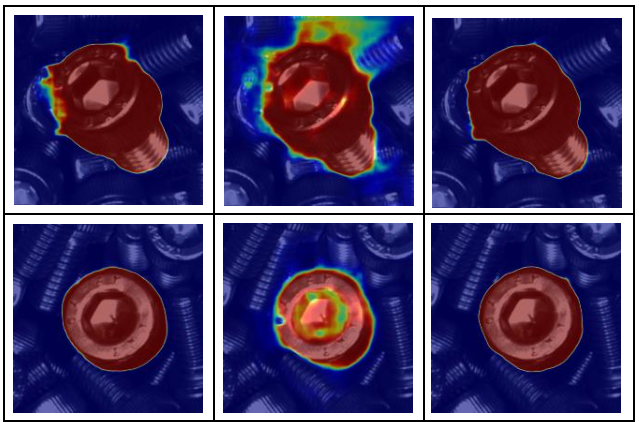
3.1.1.5.2.3　正常品のみ/異物のみ画像を増やした場合

表 1 に示したように、異物混入画像 50 枚に正常品のみ画像および異物のみ画像をそれぞれ 50 枚追加した場合の mIoU は 0.936 から 0.953 と 0.017 向上した。一方、正常品のみ画像および異物のみ画像をさらに追加し学習した場合の mIoU は最大でも 0.955 となり正常品のみ画像および異物のみ画像をさらに追加したことによる検査精度向上の効果は小さくなった。

異物混入画像１枚に正常品のみ画像および異物のみ画像をそれぞれ 50 枚追加した場合の mIoU は 0.768 と低い値となった。一方、正常品のみ画像および異物のみ画像をそれぞれ 800 枚追加した場合の mIoU は 0.870 と 50 枚追加した場合と比較し 0.102 向上した。このことから提案手法は学習用の異物混入画像が少ない場合に特に検査精度向上の効果があると考えられる。

3.1.1.5.3 判別結果の可視化

図 2 に、従来手法及び提案手法を用いて学習した場合の判別結果を示す。「M8 ボルトの確率が高いと判断した個所」を赤色で、「M8 ボルトの確率が低いと判断した個所」を青色で示している。異物混入画像 1 枚を用いた学習では、図１で示した M8 ボルトの側面全体が写った画像を用いたが、図２に示したように M8 ボルトの向きが異なる画像であっても、おおむね判別ができていることがわかる。正常品のみ画像や異物のみ画像を学習することで、これらの判別が可能になったと考えられる。



従来手法 異物混入画像 50 枚	提案手法 異物混入画像 1 枚	提案手法 異物混入画像 10 枚

図 2　判別結果の可視化

3.1.1.6　まとめ

本研究では、少量の異物混入画像に入手やアノテーションが容易な「正常品のみ画像」と「異物のみ画像」を追加し、学習する手法を提案した。従来手法である異物混入画像を用いて学習した場合と比較して、提案手法でのmIoUが向上している。すなわち、提案手法を用いることで、アノテーションコストを軽減し、検査精度を向上させることができた。

3.2 複合現実デバイスを用いた直感的な提示方法についての研究開発

3.2.1　作業支援システムと研究開発した AI の統合

3.2.1.1　課題と目的

これまで、複合デバイス（MR デバイス）を用いて、品質保持のための異物除去作業支援システムを 3 種類（図 3）、開発してきた。この際、カメラにて取得した画像から異物を検出する画像解析 AI は研究開発中のため、画像解析による異物検出アルゴリズムを実装することで、異物検出の対応を行っていた。

図 3　研究開発を行った作業支援システム

画像解析による異物検出アルゴリズムの場合、特定の色や形による異物検出となるため、汎用性に欠けるという課題があった。本年度は、「研究開発を行った画像解析用 AI を、画像解析による異物検出アルゴリズムの代わりに実装する」ことを目指した。

また、研究開発をおこなった 3 種類の作業支援システムは、用いた MR デバイスにより、それぞれの優位点・課題点がある。優位点・課題点をまとめ、各企業や各機関が作業支援システムを導入するにあたり、どのシステムが導入しやすいか、考察を行った。

3.2.1.2　AI 処理と MR 作業支援システム処理の統合

「AI 処理」と「MR デバイスを用いた作業支援システム処理」を統合するにあたり、当初は、1 つの処理に統合することを考えていた。しかし、実際に処理統合を行ってみると、2 つの処理の間には課題があることが判明した。

3.2.1.2.1　処理時間の違いの吸収

AI が、カメラから画像を取得し、画像解析し、異物を検出するまで、約 1 秒程度かかることが判明した。前述の画像解析による異物検出アルゴリズムは、色や形で判断する単純なアルゴリズムであるため、異物を検出して作業指示を表示するまでの時間は、異物の数によるが、50ms～200ms 程度と早い。

一つのプログラムに統合すると、AI の処理が処理実行を占有する時間が長いので、MR の処理がその間動かなくなり、期待される処理実行時間で動作しなくなってしまうことが判明した。具体的には、MR デバイスが作業指示マーカを描画表示する時間に遅れが生じるという問題が生じた。

そこで、「AI 用の処理と MR 用の処理を統合し、1 スレッドで実行する」ことを止め、「AI と MR の処理を独立させ、演算装置のスレッドをそれぞれの処理に割りあて、個別に独立したタイミングで処理する」ように対応した。

3.2.1.2.2　処理間のデータインターフェース追加

AI と MR の処理を独立させたため、同一処理時とは異なり、簡単に AI の解析結果のデータを MR の処理に受け渡すことができなくなった。

そこで、新たに AI の解析結果を受け渡すスレッド間のインターフェースを追加した。このスレッド間インターフェースは、Python の mmap という仕組みを用いた。任意のファイルを用意し、ファイルの内容をメモリ空間にマッピングし、複数のスレッドやプロセスが、マッピングされた内容を共有して操作できる仕組みを提供するライブラリである。

mmap を使用することで、AI の解析結果を MR の処理に受け渡すデータインターフェースを構築した。また、AI の処理が mmap の提供するメモリ空間への書込み直後に MR の処理が読み出しを行った場合の処理時間は、最短約 1msec 程度で収まった。

なお、「AI が画像を解析し異物の位置を特定後、次の画像を取得して画像解析するまでの 1 秒間」は、「画像取得時間情報、異物位置情報と、事前に測定しておいたベルトコンベアの速度ベクトルから、MR 側の処理にて異物の位置を予測し、作業指示マーカを追従させる」作業に割り当てている。

以上により、作業支援システムと研究開発した AI の統合を行うことができた。

3.2.2　3 種類の作業支援システムの比較分析

研究開発を行った 3 種類の作業支援システムは、

・プロジェクションマッピング方式
・スマートグラス方式
・透過型ディスプレイ方式

であり、それぞれ用いた MR のデバイスが異なる。各システムに対し、比較分析を実施して、優位点と課題の抽出を行った。その結果を表 3 に示す。

表３　３種類の作業支援システムの優位点・課題点

方式	優位点	課題点
プロジェクション マッピング	・機材が安価 (10 万円以下) ・作業者が身に着ける機材が無い。	・作業場所を考慮すると、設置場所が限定される。 ・作業指示が、プロジェクタ投稿範囲内に限定される。
スマートグラス	・設置場所を選ばない。 ・作業指示が眼前にあるので、遮られない。 ・自由に動ける。	・グラス着用によるストレス ・機材が高価 ・連続稼働の限度時間 （２時間程度） ・眼精疲労と酔いに似た体感 （初回または長時間の使用時）
透過型 ディスプレイ	・作業対象物と作業者の目の間に表示部がある。 →作業指示が遮られない。 ・作業者が身に着ける機材が無い。	・設置場所が限定される。 ・ディスプレイ表示部の大きさで指示範囲が限定される。 ・機材が高価 ・両眼視差による眼精疲労

3.2.3 導入のしやすさについての考察

各企業・各機関が作業支援システムを導入するにあたり、考慮すると予想される観点を想定し、導入のしやすさについて考察した。

3.2.3.1　導入にあたって考慮すると予想される観点

本研究にて開発した３種類の作業支援システムを導入するにあたり、本研究を認知している製造業数社へ、口頭にてインタビューを行った。その中から回答が重複した、導入コスト、設置位置の自由度、連続作業可能時間、作業者の感じるストレスの４点を、企業が導入にあたり考慮する項目として想定した。

3.2.3.2　まとめ

前述の 3.2.2 にて比較分析を行った結果に照らし合わせることで分析を行い、表４における「①」をプロジェクションマッピング方式、「②」をスマートグラス方式、「③」を透過型ディスプレイ方式として、結果をまとめた。

表４　作業支援システム導入にあたっての優位点

考察ポイント	①	②	③
導入コスト	○	×	×
設置位置の自由度	×	○	×
連続作業時間	○	×	○
作業者のストレス	○	×	×

○:優れている　×:課題あり

作業支援システムの導入コストや連続作業時間、作業者にかかるストレスを考慮した場合、プロジェクションマッピング方式が最も導入しやすい方式であると考える。しかし、作業環境において、特に高さ方向にある程度の余裕がない（プロジェクタの投影範囲を確保できない）場合等はこの限りではなく、他の２つの方式も選択肢に入ってくると考える。

4.　今後について

作業支援システムで作業対象物を見つけるアルゴリズムを従来の画像解析方式から AI による画像解析方式に置き換えた事で、より汎用的な広い分野に応用することが可能となった。更に今後、指示マーカの形状や色、大きさなど、暫定仕様になっている部分の研究を行うことで、より使いやすいシステムを目指す。

また、本研究については、さらに認知度を上げる必要があると考える。今後、企業訪問、展示会やイベント、SNS や動画配信など、様々な手段にて広報を行うことで、研究結果の認知度を向上させ、興味を持っていただき、現場に導入する企業を開拓していく。

なお、本研究は令和２年度から令和６年度まで文部科学省特別電源所在県科学技術振興事業の一環として、実施しているものである。

5.　参考文献等

1) Zhao, H., Shi, J., Qi, X., Wang, X. and Jia, J. : Pyramid Scene Parsing Network, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6230-6239, 2017.