

少量データによる AI 構築技術及び AI の実応用に関する研究（第 2 報）

前島 崇宏* 西本 圭志** 平間 毅** 青木 邦知**

1. はじめに

近年のAI技術の進歩により、人の判断をAIに代替させることが可能になりつつあることから、様々な業界でAI活用の動きが広がっている。

特に異物や欠陥等の目視検査の現場では、検査の判断基準が明確にできない場合等に、AI画像認識技術の適用が期待されている。しかし、一般的にはAIを構築するために膨大な学習データが必要とされ、データの収集にも労力がかかる。そのため、AIの導入促進を図るには構築に要するデータ数の削減や工程内でのAI判定結果の活用法が課題となっている。

2. 目的

AI を応用する際の検査精度向上と学習データ数削減を目指し、少量データによる AI 構築技術について研究を行う。また、AI による検査結果を現場で効果的に活用するため、複合現実デバイスによる作業指示手法について、研究を行う。

本年度は画像合成を利用した学習手法が検査精度とデータ削減効果に与える影響を評価するとともに、スマートグラスと搬送ラインを連動した作業指示システムの構築に必要な技術について、研究開発を行った。

3. 研究内容

3.1 画像合成を利用した学習効果の評価

食品工場などの正常品や異物が個別に入手可能な環境を想定し、画像収集コストやアノテーションコストが高い「異物混入画像」（正常品に異物が混入した画像）の代わりに、「正常品のみ画像」と「異物のみ画像」の合成画像を用いて学習する手法を検討した。

この提案手法の評価において、M6 ボルト（正常品を想定）中に混入した M8 ボルト（異物を想定）を検出するという課題を設定し、検査精度やデータ削減効果に与える影響について検証を行った。

3.1.1 学習・評価用データセットの収集

正常品としてM6ボルト、異物としてM8ボルトを用いて、令和2年度に開発した画像自動収集装置により「正常品のみ画像」及び「異物のみ画像」を収集した。また、「異物混入画像」を手動で取得し、異物の露出具合に応じ「easy」（露出8割以上）、「normal」（露出5割～8割）、「hard」（露出5割以下）の3種類に分別した。図1に収集したデータセットの画像例を示す。



図 1 学習・評価に用いたデータセットの画像例

3.1.2 画像の合成方法

合成方法は mixup¹⁾、CutMix²⁾ と提案手法である正規分布を用いた画像合成の 3 種類とし、合成する画像の割合 λ を変更し学習を行った。提案手法では画像中心の透過率が 100%、画像全体の透過率の平均値が λ となるマスクを正規分布で作成し、画像の合成を行った。

図 2 に各合成方法で用いたマスクと、それにより「正常品のみ画像」と「異物のみ画像」を合成した画像の例を示す。ここで、異物検査では、わずかな異物を含む場合でも適切に異物と判別される必要があることから、合成後の画像クラスは、 λ の値によらず ($\lambda=0$ の場合を含む) 全て異物クラスとして設定した。

また、本研究では画像全体に占める異物画像の割合が検査精度に与える影響を評価するため、「正常品のみ画像」と「異物のみ画像」を合成した場合は、「正常品のみ画像」の割合を $1-\lambda$ 、「異物のみ画像」の割合を λ として画像を合成した。

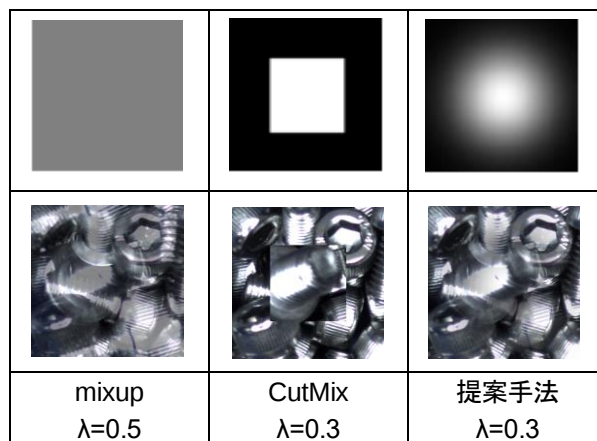


図 2 合成に用いたマスク(上)と合成画像の例(下)

3.1.3 学習・評価条件

正常品と異物の 2 クラス分類を行うため、ニューラルネットワークモデルである ImageNet で学習済みの ResNet18³⁾を用い、「正常品のみ画像」と「異物のみ画像」をランダムに組み合わせて合成した画像により、学習を行った。図 3 に合成画像を利用した学習・評価のフローを示す。

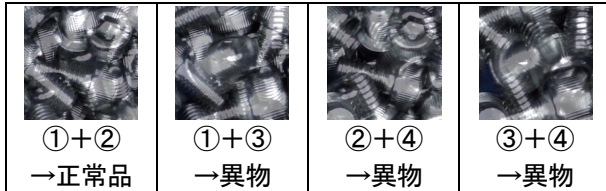
400 エポック学習後、「正常品のみ画像」と「異物混入画像」の ROC 曲線 (Receiver Operating Characteristic curve) から求めた AUC (Area Under the Curve) により検査精度の評価を行った。

また、画像合成を行わない従来手法との比較のため、従来手法として「正常品のみ画像」と「異物混入画像 (easy)」を用いて学習後、同様の評価を行った。ここで、「正常品のみ画像」は画像合成時と同一の画像、「異物混入画像 (easy)」は評価用とは異なる画像を使用した。

学習画像（正常品のみ、異物のみ）



画像をランダムに組み合わせて合成し学習

(mixup ($\lambda=0.5$) で合成した場合)

学習後「正常品のみ画像」、「異物混入画像」を評価



図 3 合成画像を利用した学習・評価のフロー

3.1.4 検査精度の評価

画像合成を利用した場合の検査精度の評価のため、「正常品のみ画像」及び「異物のみ画像」をそれぞれ 1,000枚用いて、学習した。表1に、画像合成方法や λ の違いによるAUCを比較した結果を示す。 λ を変更したときに最も高いAUCが得られたセルを灰色で示している。

表 1 画像合成方法の違いによる AUC の比較 (n=1)

λ	mixup			CutMix			提案手法		
	easy	normal	hard	easy	normal	hard	easy	normal	hard
0	0.947	0.734	0.611	0.947	0.734	0.611	0.947	0.734	0.611
0.1	0.787	0.703	0.604	0.961	0.876	0.730	0.995	0.971	0.922
0.2	0.667	0.655	0.602	0.999	0.967	0.878	1.000	0.991	0.948
0.3	0.778	0.748	0.684	0.999	0.980	0.914	1.000	0.994	0.943
0.4	0.843	0.811	0.729	0.999	0.982	0.886	1.000	0.989	0.912
0.5	0.994	0.953	0.847	1.000	0.974	0.816	1.000	0.979	0.859
0.6	0.949	0.791	0.566	1.000	0.967	0.793	0.998	0.933	0.711
0.7	0.940	0.751	0.552	1.000	0.950	0.738	0.991	0.898	0.686
0.8	0.959	0.769	0.589	0.998	0.913	0.671	0.985	0.856	0.662
0.9	0.986	0.853	0.641	0.985	0.839	0.618	0.979	0.831	0.622
1	0.993	0.884	0.675	0.993	0.884	0.675	0.993	0.884	0.675

画像合成を行わない従来手法（「正常品のみ画像」及び「異物混入画像 (easy)」をそれぞれ 1,000枚用いて学習）の「異物混入画像 (easy, normal, hard)」の AUC はそれぞれ 1.000、0.985、0.863 であった (n=3)。

3.1.5 データ削減効果の評価

画像合成を利用した学習により、従来手法と比較して検査精度を維持したまま、どの程度データが削減できるか検討を行った。

画像合成方法は高い検査精度が得られた条件である提案手法 ($\lambda=0.3$) とし、学習用の「正常品のみ画像」と「異物のみ画像」(従来手法は「異物混入画像 (easy)」) をそれぞれ 100 枚ずつ増やしたときの検査精度を比較した。表 2 に学習画像枚数と AUC (平均値) の関係を

示す。

表 2 学習画像枚数と AUC の関係 (n=3)

画像枚数	従来手法			提案手法 ($\lambda=0.3$)		
	easy	normal	hard	easy	normal	hard
100	1.000	0.939	0.787	0.975	0.916	0.820
200	1.000	0.962	0.813	0.993	0.958	0.882
300	1.000	0.974	0.826	0.998	0.972	0.880
400	1.000	0.976	0.829	0.999	0.982	0.901
500	1.000	0.980	0.846	1.000	0.987	0.923
600	1.000	0.981	0.846	1.000	0.989	0.929
700	1.000	0.982	0.856	1.000	0.991	0.940
800	1.000	0.982	0.855	1.000	0.991	0.935
900	1.000	0.984	0.857	1.000	0.991	0.940
1000	1.000	0.985	0.863	1.000	0.993	0.943

3.1.6 異物位置の可視化

AI による判断結果を MR デバイス等で作業者に指示するためには、異物位置の可視化が必要となる。そこで本研究では、画像全体に対して 28 ピクセルピッチで 224×224 ピクセルの矩形領域を走査、分類し、異物の確率が高いと判断した個所を赤色で、異物の確率が低いと判断した個所を青色で示すこととした。

判別難易度の高い異物検査において最も高い AUC が得られた条件である提案手法 ($\lambda=0.2$) により正常品画像 11.2 万枚 (背景画像を含む) 及び「異物のみ画像」6 万枚を学習したモデルを用いて異物確率分布を可視化した例を図 4 に示す。これにより MR デバイス等で作業者に異物位置を可視化する見通しが得られた。

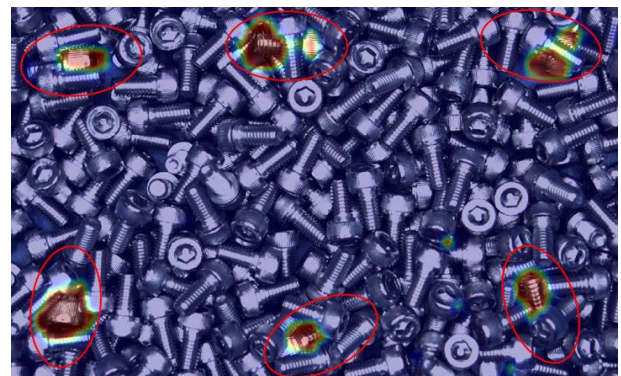


図 4 異物確率分布

3.2 複合現実デバイスによる AI の判断結果の提示

本研究では、AI の判断結果の提示を行うにあたり、複合現実デバイスの一つであるスマートグラスと搬送ラインを連動した作業指示システムの構築に必要な技術の研究開発を行った。

3.2.1 システム構成

作業員が搬送ライン上で異物や欠陥品等の目視検査及びその除去作業を行う現場への適用を想定し、図 5 に示す作業指示システムを検討した。システムは指示対象物の位置を検出する検査用カメラ、作業指示を行うスマートグラス、検査用カメラとスマートグラスの位置姿勢関係を算出する基準画像マーカで構成される。



図 5 作業指示システムの構成

(1) スマートグラスから画像マーカを認識する技術

検査用カメラ及びスマートグラスの初期位置合わせに基準画像マーカを用いるため、スマートグラスから画像マーカを認識する必要がある。

(2) スマートグラスと検査用カメラの位置合わせ技術

検査用カメラで検出した作業対象物の位置をスマートグラスで指示するため、検査用カメラとスマートグラスの位置姿勢関係を算出する必要がある。

3.2.2 スマートグラスの選定

本システムは、工場等で働く作業員がスマートグラスを装着して作業する事を想定しており、スマートグラスに必要な機能として、周囲にある物体と干渉しないよう演算装置一体型であること、作業員がメガネを着用したまま装着可能であること、ハンドトラッキング機能等により外部コントローラー等を持つことなく操作が可能であることが、求められる。

これらの機能を満足するスマートグラスとして、図 6 に示す Microsoft 社の HoloLens2 を採用した。



図 6 HoloLens2 の外観

3.2.3 スマートグラスで画像マーカを認識する技術

HoloLens2 から画像マーカを認識するため、HoloLens2 搭載カメラで撮影した画像から連続的に画像マーカを認識する処理を実装した。

画像マーカには AR マーカの一つである ArUco マーカを使用し、検査用カメラ座標系から見た画像マーカ座標系の 3 次元位置姿勢を求めるため、マーカ認識ライブラリ Vuforia⁴⁾を用いた。認識結果を視覚的に確認するため、認識した画像マーカの 3 次元位置姿勢に合わせて立方体の 3DCG を重畳表示させた結果を図 7 に示す。

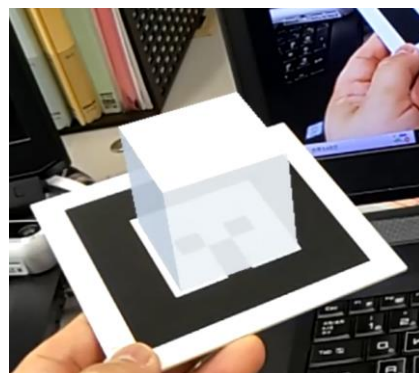


図 7 画像マーカ認識結果の可視化

3.2.4 HoloLens2 と検査用カメラの位置合わせ技術

検査用カメラで検出した作業対象物に HoloLens2 で 3DCG を重畳表示するために、検査用カメラで検出した作業対象物の 3 次元位置姿勢を、HoloLens2 から見た 3 次元位置姿勢に変換する必要がある。また、検査用カメラの座標系は固定されているが、HoloLens2 の座標系は装着者の動きにより動的に変化するため、検査用カメラと HoloLens2 の位置姿勢関係を逐次的に求める必要がある。本研究で開発した手法を以下に示す。

1) 検査用カメラで搬送ラインに固定した画像マーカを撮像し、画像処理により検査用カメラから見た画像マーカの位置姿勢を算出した。ここで、画像マーカ認識ライブラリ ar_track_alver⁵⁾を用いた。

2) HoloLens2 搭載カメラで搬送ラインに固定した画像マーカを撮像し、3.2.3 と同様の処理で HoloLens2 搭載カメラから見た画像マーカの位置姿勢関係を算出した。以上から、画像マーカを介して検査用カメラ及び画像撮影時における HoloLens2 の位置姿勢関係を算出する。

3) 画像撮影時における HoloLens2 の位置姿勢関係と HoloLens2 内部に有する周辺環境の 3 次元マップを連動させることで、HoloLens2 から見た画像マーカの位置姿勢をトラッキングし、検査用カメラと HoloLens2 の位置姿勢関係を逐次的に出力する処理を実装した。本出力結果及び令和 2 年度の研究で求めた検査用カメラ座標系における作業対象物の 3 次元位置姿勢をもとに、HoloLens2 座標系から見た作業対象物の 3 次元位置姿勢を算出することで、作業対象物上に 3DCG を重畳表示する。

以上の一連の処理を実装し、令和 2 年度の研究と同様の環境で 3DCG 表示を行った結果、作業者の動きによらず適切な位置に指示できることを確認した。

3.2.5 3DCG の重ね合わせ誤差に関する評価

HoloLens2 の表示位置精度を評価するため、方眼紙に取り付けた画像マーカの中心位置から 5cm ずつ ±20cm の範囲に円形の 3DCG を表示させ、方眼紙上の位置誤差を測定した。図 8 に実験時の様子を示す。また、画像マーカ垂直方向については定規を使用することで位置誤差を測定した。その結果、画像マーカの中心位

置から±20cm の範囲において、3DCG の表示位置精度誤差が±0.2cm に収まることを確認できた。

±0.2cm の誤差は、実用上問題にならない精度であると考えられるが、HoloLens2 の装着具合や視覚の個人差で表示位置がずれる可能性がある。そこで、HoloLens2 上に表示した仮想ボタンを操作することで 3DCG の表示位置を 3 軸方向にオフセット調整する機能を追加実装した。これにより、作業員自身の操作で 3DCG の表示位置を微調整することが可能となる。

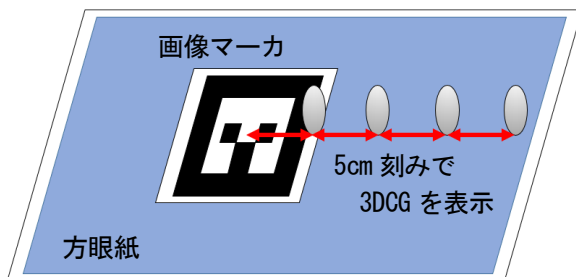


図 8 表示誤差測定の様子

4. まとめと考察

4.1 画像合成を利用した学習効果の評価

表 1 によると、AUC の最大値は判別難易度によらず、mixup、CutMix、提案手法の順で大きくなる傾向があった。また、提案手法では、評価画像の難易度が増すにつれて、AUC を最大にする λ が 0.4、0.3、0.2 と小さくなる傾向がみられた。画像合成方法や λ により AUC に違いが生じたが、これは評価画像と合成画像の類似具合を反映した結果であると考えられる。

表 2 によると、従来手法で学習用画像枚数が各 1,000 枚のとき、easy の AUC が 1.000、normal の AUC が 0.985、hard の AUC が 0.863 であったのに対し、提案手法では学習用画像枚数が各 500 枚のとき easy の AUC は同等精度を維持しつつ、normal の AUC が 0.987、hard の AUC が 0.923 と従来手法を上回る精度であった。また、画像合成を用いた手法では学習用画像枚数により検査精度が大きく変化する傾向がみられたが、これは画像合成の組み合わせ数が枚数の二乗に比例して増加するためと考えられる。

4.2 複合現実デバイスによる AI の判断結果の提示

スマートグラスとして HoloLens2 を使用し、搬送ラインと連動した作業指示システムの構築に必要な要素技術の研究開発を行った。従来、HoloLens2 を装着した作業員が画像マーカの撮像範囲外に移動した場合に、画像マーカの位置姿勢情報や 3DCG の重畳表示が失われるという課題があった。これに対して、本研究では、作業員が移動した場合にも 3DCG の重畳表示を継続できたことから、より作業員の移動自由度が向上し、実作業現場での利便性が向上すると考えられる。

また、HoloLens2 の表示位置精度を評価した結果、±2mm 程度の誤差であったが、オフセット調整機能を併用することにより、さらに精度の高い作業指示につ

ながると考えられる。

5. 今後の課題

5.1 画像合成を利用した学習効果の評価

本研究で用いた「異物のみ画像」は画像全体に異物が映り込んでいるため、確実に異物が混入した画像が合成可能である。しかし、実際の食品工場で手に入る画像の中には、正常品の一部が変色しているために異物と判別する必要がある画像や、背景が多く映り込む異物画像がある。これらの異物画像の一部を正常品画像と合成させた場合、異物ではない領域の画像と正常品画像を合成する可能性がある。このとき、異物が全く含まれない合成画像を異物画像として学習することとなり、検査精度低下の要因になると考えられる。今後は、そのような検査対象に対してもアノテーションコストを軽減しつつ、検査精度を向上させる手法を検討していく必要がある。

5.2 複合現実デバイスによる AI の判断結果の提示

本研究では単独のスマートグラスを対象としたが、実現場では複数台の連携が必要な場面も想定される。今後は、搬送ラインの上流と下流に作業員がおり、異物除去作業を行う場合を想定し、上流の作業員が異物除去を行った際に、上流の作業員の視界のみならず、下流の作業員の視界からも作業指示が消えるようなシステムの研究を行うために、複合現実デバイスの複数台連携に必要な技術の研究開発にも取り組む予定である。

また、本研究で開発したユーザーインターフェースは試作的なデザインである。実用化に向けては、スマートグラスの 3DCG 表示視野角外にある作業対象物の位置を作業員に効率的に指示する手法の研究開発や、スマートグラスの視認性評価が必要である。

さらに、スマートグラスを長時間使用することで、作業員に身体的な負荷やストレスを与えることが懸念される。今後は、スマートグラス着用時における作業員の疲労度やストレス具合を定量化する試みも必要かと考える。

6. 参考文献等

- 1) Hongyi Zhang, et al., “mixup: Beyond Empirical Risk Minimization”, ICLR2018.
- 2) Sangdoo Yun, et al., “CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features”, ICCV 2019.
- 3) He, et al., “Deep residual learning for image recognition.” In proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016. 2).
- 4) Vuforia, <https://developer.vuforia.com/>
- 5) ar_track_alver, http://wiki.ros.org/ar_track_alvar.