

## FIT2023 情報科学技術フォーラム での研究発表

高 木 裕 也

Yuya TAKAGI

情報メディア学専攻修士課程 2年

### 1. はじめに

2023年9月6日から8日まで大阪公立大学中百舌鳥キャンパスで開催された第22回情報科学技術フォーラム（FIT2023）に参加した。私は、8日の画像認識・メディア理解5のセッションで、“Vision Transformer を用いた鍛造部品の不良品検出”というタイトルで研究発表を行った。

### 2. 研究内容

工業製品において外観検査は安全性や信頼性などの観点から必要な工程である。しかし、現在、外観検査の多くは人間による目視で行われている。そのため、検査員の習熟度や疲労などによる検査精度のバラツキが生まれることで、品質を一定に保てないことや、人的ミスによる欠陥品の見落としが起こる。また、人手不足や検査員が検査できる数に限りがあるといった様々な課題が挙げられる。そこで、画像検査による外観検査の自動化によってこれらの課題を解決することを目的として Vision Transformer（ViT）を用いた鍛造部品の不良品検出について研究を行った。また、鍛造部品の製造で用いられるプレスの金型が摩耗することで、製造時期によって表面テクスチャが変化することが考えられる。そのため、画像検査機の再学習が求められるが、再学習にスパンが短いとコストがかかるためロバストなモデルが求められる。そこで、近年、画像タスクにおいて優秀な成績を上げている ViT と、従来用いられてきた畳み込みニューラルネットワーク（CNN）を用いてロバスト性について比較を行った。

### 3. データセット

本研究では、高橋金属（株）提供の鍛造部品の底面部と側面部が写っている画像をデータセットとする。側面部について傷や欠け、打痕が見られる不良品画像と、それらが見られない、または無視できる程度の大きさである良品画像でラベル付けされている。用いるデータセットは大きく分けて3つあり、それぞれ2022年に撮影されたデータ、2021年に撮影されたデータ、2020年に撮影されたデータがある。2022年のデータは、学習用データ、検証用データ、評価用データに分割されている。不良品画像の例を図1に示す。

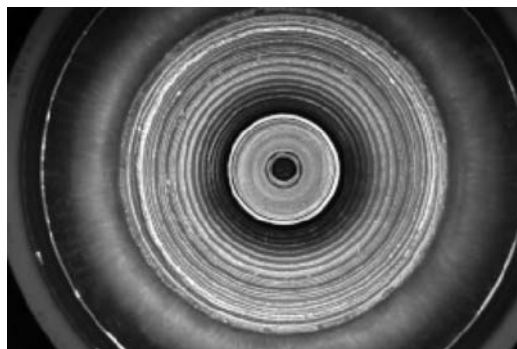


図1 サンプルデータ

2021年のデータと、2020年のデータは、学習データとは異なる時期に撮影されたデータに対してのロバスト性を評価するために用いる。

### 4. 前処理

データセットの不良品画像は側面部についての欠陥だけなので、原画像から側面部のみが写った画像に変換する。原画像を鍛造部品の中心を原点として極座標変換を行い、横軸を角度、縦軸を動径とした178×256画素にリサイズし、上部50画素を除去することで、128×256画素の画像にする。その後、グレースケール化を行う。変換後の不良品画像の例を図2に示す。

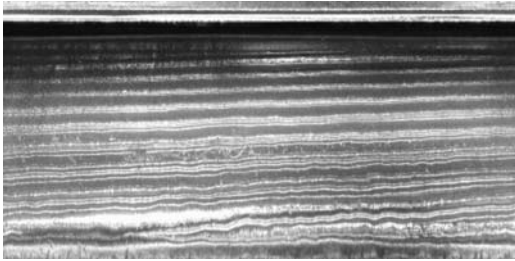


図 2 前処理後の画像

## 5. 識別実験結果

作成した学習済みモデルを用いて、2022 年の評価用データ、2021 年のデータ、2020 年のデータについて、良品画像および不良品画像に対して識別を行った。それぞれのデータに対する AUC の値を表 1 に示す。

表 1 ViT と CNN の AUC

|     | 2022 年 | 2021 年 | 2020 年 |
|-----|--------|--------|--------|
| ViT | 1.0    | 0.99   | 0.86   |
| CNN | 1.0    | 0.64   | 0.66   |

表 1 より、ViT と CNN とともに学習データと同じ時期に撮影された 2022 年評価用データは AUC が 1.0 であり、高い精度に識別ができています。2021 年のデータの関しては、ViT の AUC は 0.99、CNN の AUC はそれぞれ 0.64 であり、ViT は高い値となっている一方で、CNN は大きく値が低下しています。2020 年のデータについては、ViT の AUC が 0.86 である。また、CNN の AUC は 0.66 であり、ViT と CNN どちらも AUC の値が低下しているが、ViT の方が CNN よりも AUC が高い。

ViT と CNN について、それぞれのデータに対する識別の ROC 曲線を図 3 に示す。

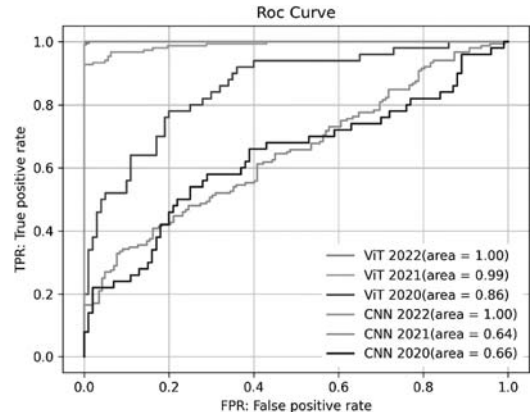


図 3 ROC 曲線

画像の誤識別率を 0 に抑えた上での良品画像の誤識別率が、学習データとは異なる撮影時期のデータだと ViT と CNN どちらも増加している。しかし、ViT は CNN よりも良品画像の誤識別率は低く、2021 年のデータに対しては差が大きい。

## 6. おわりに

本研究では鍛造部品の側面部について ViT と CNN を用いて、不良品画像および良品画像について識別を行った。学習データと同時期に撮影された評価用データに対する識別結果としては ViT、CNN とともに正確な識別が行えた。学習データと異なる撮影時期の鍛造部品のデータについては識別精度が低下してしまうが、ViT は CNN に比べて精度の低下が抑えられており、よりロバストなモデルが作成できた。

発表はスムーズに行うことができ、かつ、質問にも丁寧に回答することができた。また、同じ分野の研究発表を聞くことで、知識を深めることができ、研究のアプローチを自分の研究にも応用できる可能性もあったため、今後の研究活動において有益でした。