

ものづくり現場の生産性向上のための AI 技術の活用に関する研究開発

生駒晃大*、渡辺博己*、松原早苗*

A study on the use of artificial intelligence to improve productivity in manufacturing

IKOMA Akihiro*, WATANABE Hiroki* and MATSUBARA Sanae*

本研究では、AI 技術を活用したものづくり現場における作業の自動化や効率化を支援するための技術開発に取り組んでいる。今年度は、製品の外観検査を対象に、AI 技術による高精度な異常検知を低コストで実現するための手法について検討を行った。また、事前学習済みのモデルを活用した異常検知手法に関して、モデルの性能をより向上させるための処理の実装と評価を行ったので、その結果について報告する。

1. はじめに

製造業をはじめとするものづくりの現場では、熟練技術者の高齢化などに伴う人材不足の問題や、要求される製品品質の高度化、生産コストのさらなる削減などを背景に、業務改善や生産性向上に関する取り組みが年々重要度を増している。そのため、生産性低下の要因の一つとなっている人手によるアナログな作業や繰り返し作業を自動化・効率化するため、ロボットなどのデジタル技術の導入が積極的に行われている。特に、従来技術と比較して高い認識性能や汎用性を持つ人工知能（AI）技術にも注目が集まっており、製造現場で収集される様々なデータのより高度な活用が期待されている。

このように、AI 技術のものづくり現場への導入に期待が寄せられる一方、AI 技術の利用には、大量のデータの収集や動作環境の構築、精度向上のための学習やチューニングに多くの時間や費用が必要となる場合がある。そのため、人手や資金などの経営資源の制約が大きい中小企業にとっては、導入が困難となる場合も多い。

そこで本研究では、ものづくり現場の生産性向上を目的とし、AI 技術の現場実装を支援するための技術開発に取り組んでいる。今年度は、製造工程における外観検査を対象に、AI 技術を用いた高精度な手法をできるだけ低コストで実現するため、深層学習（Deep Learning）による異常検知モデルの検討と開発に取り組んだので、その結果を報告する。

2. 深層学習による異常検知モデルの構築

2.1 正常データのみによるモデルの学習

深層学習を用いた外観検査手法として、良品と不良品の画像データをそれぞれ大量に収集し、良・不良を分類可能なモデルを構築することが考えられる（図1(a)）。しかし、実際の製造現場では、不良品のデータを大量に確保することは困難な場合が多く、データ収集に時間を要することや、収集したデータのバラツキが精度に悪影

響を与えるなどの課題がある。そこで本研究では、良品から得られる正常データのみで構築可能な異常検知モデルについて検討を行った。

正常データのみで学習可能な異常検知モデルとして、Autoencoder（AE）の Encoder-Decoder 構造による再構成誤差を利用した手法¹⁾や、Generative Adversarial Network（GAN）の Generator と Discriminator による敵対的学習を利用した手法²⁾などが考案されている（図1(b)）。これらの手法では、正常データの学習により、データが持つ潜在的な特徴やその分布を十分に表現可能な生成モデルを構築することで、学習していない異常なデータを正確に再現できないことを異常検知に利用している。しかし、表現能力の高い複雑なモデルを構築するほど、学習時間の増加や、学習が不安定になるなどの問題がある。そのため、最適なモデル構造やパラメータを探索するために、モデルの学習と検証を何度も繰り返すなどの手間が発生してしまう。

2.2 学習済みモデルの活用

このようなモデルの構築、学習に費やす時間を削減するため、ImageNet などの膨大な数のデータを用いて事前に学習されたモデル（以下、学習済みモデル）を異常検知に活用する手法が登場している。

学習済みモデルは、それ単体で異常検知を行うことはできないが、大量のデータを用いた学習により、入力されたデータから有用な特徴量を抽出するための機構がモデル内に獲得されている。そのため、学習済みモデルを特徴抽出器として使用し、抽出した特徴量に対して異常度を算出するための処理を加えることで異常検知を行う。

学習済みモデルでの特徴抽出と、統計的な手法による異常度の計算を組み合わせる手法として SPADE³⁾がある。本手法は、判別対象データの特徴量と、正常データ全ての特徴量とのユークリッド距離をそれぞれ算出し、判別データ近傍にある正常データとの距離の和を異常度としている。また、PaDiM⁴⁾では、正常データ全体の特徴量から平均と分散を算出し、判別対象データの特徴量とのマハラノビス距離を異常度としている（図1(c)）。

* 情報技術部

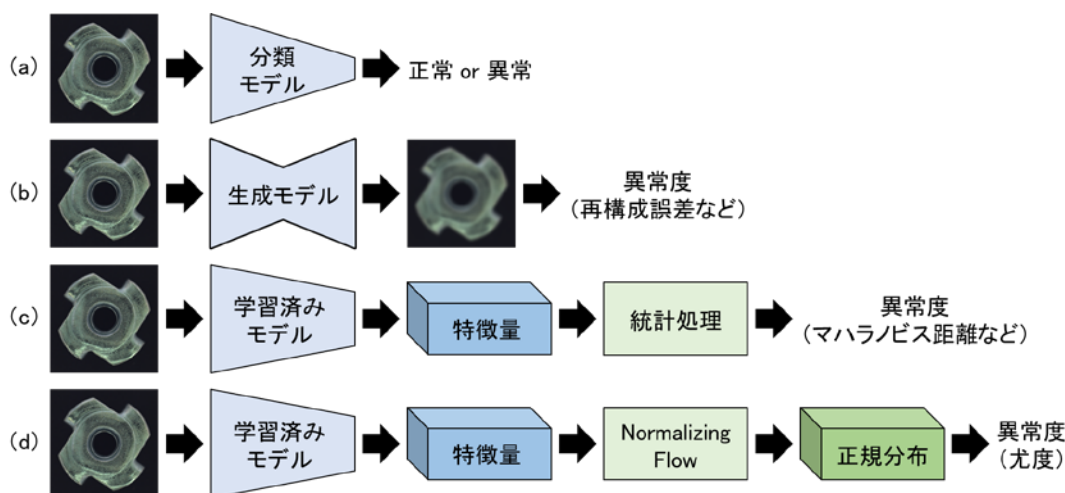


図1 深層学習を用いた外観検査手法の比較

学習済みモデルを用いたこれらの手法では、判別対象となるデータでの学習を行わないにも関わらず、異常検知が可能となっているため、現場実装を行う上では取り扱いやすく、有効な手段であると考えられる。しかし、学習済みモデルの出力をそのまま使用するため、選択したモデルによっては特徴量サイズが大きくなり、統計処理や推論処理に時間がかかってしまうことや、画像中の物体の位置やスケールなどの、入力データの変化に対するロバスト性が問題とされている。

2. 3 Normalizing Flow の導入

学習済みモデルを活用しつつ、前述の課題を解決し、異常検知の精度をより高めるための手段として、本研究では、学習済みモデルで抽出した特徴量に対して、Normalizing Flow⁵⁾による変数変換の処理を導入した（図1(d)）。

学習済みモデルで抽出した特徴量は、データを表現するのに有用な特徴となっている一方、データ全体の特徴分布がどのようなになっているのかは考慮されていない。Normalizing Flow は、このような有用だが複雑で解析が困難な分布を、より解析に適した単純な分布へと変換することが可能な手法であり、正規分布などの扱いやすい分布を変換先とする場合が多い。なお、Normalizing Flow による変換では、変換の前後でデータの次元数などの形状は変化しない。

ここで、変換対象となる分布を $p_Y(y)$ 、目的とする分布を $p_Z(z)$ 、変換に使用する関数を f_{NF} とすると、 $p_Y(y)$ は式(1)の通り表現できる。なお、式中の $\det \partial f_{NF} / \partial y$ は、ヤコビ行列の行列式（ヤコビアン）である。

$$p_Y(y) = p_Z(f_{NF}(y)) \left| \det \frac{\partial f_{NF}}{\partial y} \right| \quad (1)$$

式(1)の関数 f_{NF} にはニューラルネットワークを使用することもでき、目的とする分布 $p_Z(z)$ に合わせ、 $p_Y(y)$ の負の対数尤度を最小化するように学習を行うことで、ネットワークのパラメータを最適化することが可能である。しかし、扱うデータが高次元になるほど、式(1)のヤコ

ビアン⁶⁾の計算コストが増大するため、パラメータ更新に必要な勾配計算を効率的に行うための工夫が必要となる。そこで、本研究では、式(2)で示す Affine Coupling⁶⁾による変換を適用した。

$$\begin{aligned} z_1 &= y_1 \\ z_2 &= \exp(s(y_1)) \odot y_2 + t(y_1) \end{aligned} \quad (2)$$

式(2)では、変換対象となる y を次元方向に二等分したものを y_1 、 y_2 とし、それぞれに異なる処理を適用して求めた z_1 、 z_2 を結合したものを変換結果 z としている。 y_2 の変換に使用する s と t は、スケールと平行移動の変換を表す関数であり、 y_1 を入力して得た出力をもとに y_2 の変換を行っている。また、一度の変換では、 y_1 がそのまま z_1 として出力されるなど、十分な変換を表現できないため、変換結果 z の次元方向での順番を入れ替え、再度 y として同様の変換を複数回繰り返すことで、複雑な変換を実現している。

Normalizing Flow での変換先となる分布 $p_Z(z)$ を多変量標準正規分布 $z \sim N(0, I)$ とした場合、 $p_Y(y)$ の負の対数尤度の最小化を目的とした誤差関数 L は、式(3)のようになる。なお、Affine Coupling におけるヤコビアンは、式(2)における $\exp(s(y_1))$ の各要素の積として効率的に計算可能となっている。

$$\begin{aligned} L(y) &= -\log p_Y(y) \\ &= -\log p_Z(f_{NF}(y)) - \log \left| \det \frac{\partial f_{NF}}{\partial y} \right| \\ &= \frac{1}{2} \|z\|_2^2 - \log \left| \det \frac{\partial f_{NF}}{\partial y} \right| \end{aligned} \quad (3)$$

式(3)に示す誤差関数を用いて、正常データのみでNormalizing Flow による変換を学習することで、変換後の特徴量 z の正規分布に対する尤度を異常度として使用することができる。正常データの特徴量は、分布に沿って尤度の高い位置に変換される一方、学習していない異常データの特徴量は、分布の中心から外れた位置に変換されることで尤度が低くなるため、正常と異常の判別が可能となる。

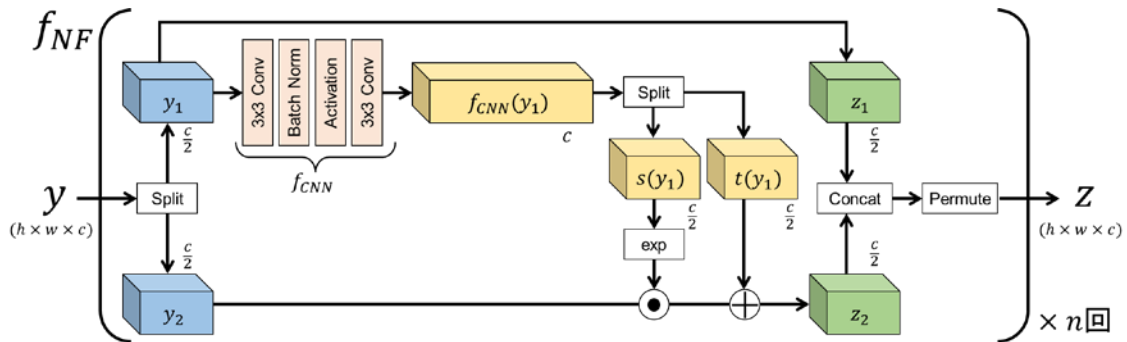


図2 特微量の変換に使用する Normalizing Flow のモデル構造

3. 実験

3.1 データセット

本実験では、外観検査用の実画像データセットとして、異常検知手法の精度検証に用いられる MVTEC AD⁷⁾を使用する。MVTEC AD には、15 種類の製品の画像が含まれており、ネジや錠剤などのオブジェクト系が 10 種類、カーペットやタイルなどのテクスチャ系が 5 種類で構成されている。学習用データには良品画像のみが含まれており、枚数は製品ごとに異なるが、それぞれ 200 枚程度と、深層学習で扱うには少量のデータとなっている。また、評価用データには良品と不良品の両方の画像が含まれており、全ての製品の評価用データを合計すると、良品画像が 467 枚、不良品画像が 1258 枚となっている。

データへの前処理として、画像サイズを 256×256 にリサイズし、画像中心から 244×244 のサイズになるようにトリミングを行う。また、本実験では、入力データの変化に対する異常検知モデルのロバスト性を評価するため、評価用データに対して、回転、平行移動、拡大・縮小の 3 つの変換を適用したデータについても評価を行った。なお、1 枚の画像に対して各変換を適用した画像を 2 枚生成することで、データの総数は元の評価用データの 6 倍となっている。

3.2 実験条件

特徴抽出に使用する学習済みモデルとして、本実験では、ImageNet で事前に学習された EfficientNet⁹⁾を使用する。EfficientNet には、モデル内のブロック構造の深さに合わせ、B0 から B7 までの 8 種類のモデルが存在し、数字が大きくなるほどモデルサイズが大きくなる。EfficientNet を用いた異常検知に関する先行研究⁹⁾の結果を参考に、本実験では EfficientNet-B5 を特徴抽出器として選択し、入力側から数えて 7 番目のブロックの出力値を特微量として使用した。入力画像サイズが 244×244 の場合、出力される特微量の形状は、 $7 \times 7 \times 304$ (高さ $h \times$ 幅 $w \times$ チャンネル数 c) となる。

特微量の変換を行う Normalizing Flow の構造として、式(2)による Affine Coupling での変換に従い、図 2 で示すモデルを構築した。変換に使用する s と t の関数には、畳み込みニューラルネットワーク f_{CNN} を使用し、ネットワークの出力を 2 つに分割することで、 s と t で別々のネ

表1 MVTEC AD の評価用データに対する ROC-AUC (元の評価用データ / 回転等の変換後の評価用データ)

Category	PaDiM ⁴⁾	NF	NF+DA
Objects	bottle	1.000 / 0.832	1.000 / 0.989
	cable	0.992 / 0.974	0.992 / 0.977
	capsule	0.943 / 0.671	0.971 / 0.714
	hazelnut	0.981 / 0.980	0.996 / 0.995
	metal nut	0.994 / 0.790	0.976 / 0.889
	pill	0.966 / 0.631	0.941 / 0.802
	screw	0.874 / 0.632	0.940 / 0.719
	toothbrush	0.989 / 0.778	0.986 / 0.843
	transistor	1.000 / 0.881	0.995 / 0.903
	zipper	0.978 / 0.788	0.997 / 0.881
Textures	carpet	1.000 / 0.998	1.000 / 0.998
	grid	0.985 / 0.987	0.977 / 0.985
	leather	1.000 / 1.000	1.000 / 1.000
	tile	0.987 / 0.993	0.988 / 0.992
	wood	0.986 / 0.982	0.992 / 0.990
Avg.	0.978 / 0.861	0.983 / 0.912	0.981 / 0.975

ットワークを構築するよりも計算コストを抑えている。また、 f_{CNN} は 2 つの畳み込み (Conv) と、バッチ正規化 (Batch Norm)、活性化関数 (Activation) というシンプルな構造となっているため、モデルの構築や学習における負担は少ない。なお、活性化関数には EfficientNet で使用されているのと同じ Swish 関数を採用し、変換処理の繰り返し回数 n は 4 回とした。

学習時のハイパーパラメータとして、バッチサイズ 8、エポック数 100、最適化アルゴリズムには重み減衰値 $1e-5$ の Adam を採用し、学習率は $[1e-5, 1e-3]$ の区間での Warmup と Cosine 減衰によるスケジューリングを行った。

3.3 実験結果

変換適用前後での 2 つの評価用データに対して、構築したモデルで異常検知を行い、ROC 曲線の AUC (ROC-AUC) を算出した結果を表 1 に示す。ROC 曲線は、異常検出率 (True Positive Rate) と誤検出率 (False Positive Rate) の関係をグラフ化したものであり、AUC はその曲線の下部分の面積を表している。ROC-AUC が大きい

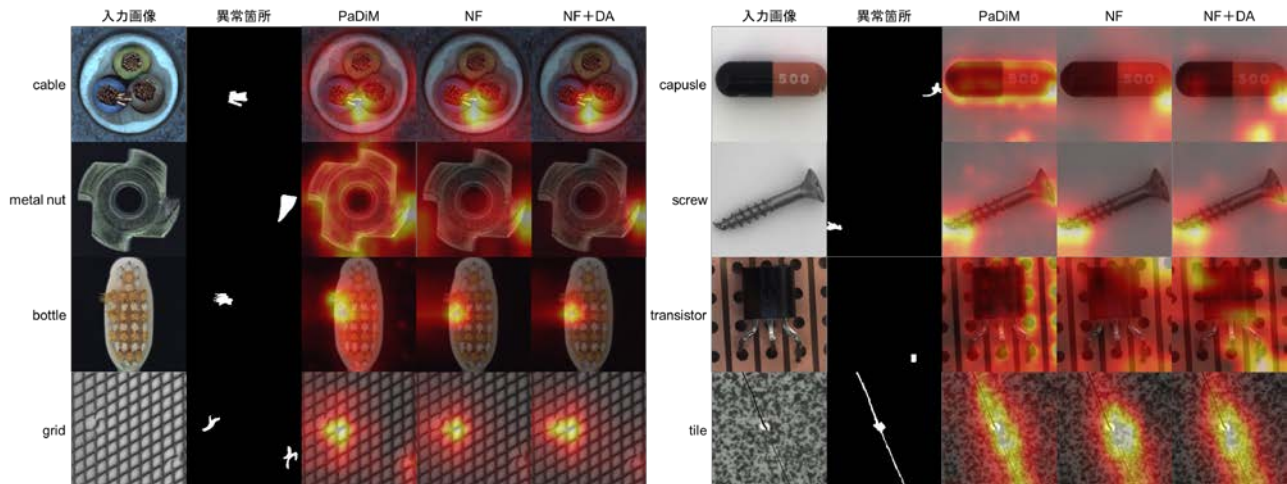


図3 各手法における異常度のヒートマップ画像

ほど、正常と異常の判別性能が高いことを示しており、完全に判別可能な場合には1となる。

表1のNFが、Normalizing Flowを用いた異常検知モデル（以下、NFモデル）の結果である。また、NFモデルと同様に、学習済みのEfficientNet-B5を特徴抽出器として使用したPaDiMの結果も併せて示す。

変換前のデータセットに対しては、どちらの手法も高い判別性能を示しており、正常データと学習済みモデルを活用することで、高精度な異常検知を低コストに実現可能なことが確認できる。一方、変換後のデータセットに対しては、どちらの手法も判別性能は変換前より低下する結果となった。しかし、NFモデルでは、数値の低下をより抑えられているため、ロバスト性の向上が確認できる。また、オブジェクト系に比べテクスチャ系の製品に対する性能は変換前後であまり違いが見られないが、これは、テクスチャ系の画像が画像全体で類似パターンの繰り返しになっていることから、今回適用した位置やスケールの変換の影響を受けないためだと考えられる。

表1のNF+DAは、NFモデルの性能をより向上させるため、学習に使用する正常データに対してデータ拡張（Data Augmentation）を適用した結果である。データ拡張は学習時の過学習防止や、汎化性能を向上させるための手法であり、学習データを追加で収集することなくデータ数を増殖できるため、今回のような正常データの枚数が限られているような場合でも有効である。なお、今回適用した拡張処理は、評価用データに対して適用したのと同様に回転、平行移動、拡大・縮小となっている。

3つのモデルについて、ROC-AUCの平均値を比較すると、変換後のデータに対してはNF+DAが0.975と最も良い結果となり、データ拡張によるさらなる性能の向上が確認できた。また、変換前のデータに対しては、3つのモデルで大きな差が見られないため、画像の撮影段階で製品の位置決めを正確に行える場合や、テクスチャへの異常検知の場合には、PaDiMなどの統計的アプローチでも十分な性能を確保できると考えられる。

図3は、各モデルにおける異常度の分布を可視化したヒートマップ画像である。画像中の異常箇所付近の異常度が大きくなっていることを確認することで、モデルが異常箇所に注目して正しく判定を行っているかどうかの定性的な評価に利用することが可能である。モデルごとの異常度の分布を比較すると、NFモデルやデータ拡張を適用したモデルの方がより異常箇所に注視していることが確認できる。

4. まとめ

本研究では、製造工程における外観検査を対象に、AI技術を用いた異常検知手法の検討と開発を行った。

収集が容易な正常データのみを用いて構築可能な手法や、大規模データで事前に学習されたモデルを活用した手法に着目することで、高精度な異常検知モデルを低コストで実現する方法について検討を行った。

また、Normalizing Flowによる特徴量変換やデータ拡張を組み合わせた異常検知モデルを構築することで、入力データの位置やスケールの変化に対するロバスト性の向上を実現した。

今後は、本研究で得られた知見を活用し、実際の製造現場に適用した場合の有効性や課題を検証するとともに、異常検知モデルの精度や汎用性をさらに向上させるための手法に関しても研究を進めていきたい。

【参考文献】

- 1) P. Bergmann, et al., arXiv:1807.02011, 2018
- 2) S. Akcay, et al., Proc of ACCV, pp. 622-637, 2018
- 3) N. Cohen, et al., arXiv:2005.02357, 2020
- 4) T. Defard, et al., Proc of ICPR, pp. 475-489, 2018
- 5) D. J. Rezende, et al., Proc of ICML, pp. 1530-1538, 2015
- 6) L. Dinh, et al., arXiv:1605.08803, 2017
- 7) P. Bergmann, et al., Proc of CVPR, pp. 9592-9600, 2019
- 8) M. Tan, et al., Proc of ICML, pp. 6105-6114, 2019
- 9) O. Rippel, et al., Proc of ICPR, pp. 6726-6733, 2019