

機械学習を用いた画像外観検査の高速化技術の開発

Method for accelerating machine learning in visual inspection

藤本 真奈* 大島 亮太* 村瀬 博典*
 Mana Fujimoto Ryota Oshima Hironori Murase

Abstract

We present a new semi-supervised anomaly detection technique for image visual inspection in the manufacturing industry. This technique is based on Patch Distribution Modeling (PaDiM)¹⁾ methodology with the dual objectives of enhancing performance and accelerating processing speed. Specifically, we have adopted an approach to enhance the processes of feature extraction and dimensionality reduction. We confirmed that the proposed method can lead to significantly reduction in processing time when compared to the PaDiM approach while maintaining comparable performance levels. This makes it possible to apply the proposed method to image visual inspections in the manufacturing industry, with the expectation of enhancing the efficiency and quality of such inspections.

1. はじめに

部品や製品の品質を維持・保証するために、製造現場では外観検査を実施している。外観検査は人による目視検査が現在も主流であるが、近年、高速かつ一貫した判定が可能な画像による自動検査が普及している。

画像外観検査は現在ルールによる判定(以下、ルール判定)が一般的である。ルール判定はあらかじめ人間が設定したルールに従って判定する方法である (Fig. 1)。処理内容をフローチャートで示すことが可能なため、結果に至るまでの過程が解釈しやすいというメリットがあるが、ルールが複雑な

パターンの場合にはフローチャートに書き起こすことが難しいため、より高度な自動検査のために機械学習 (Machine Learning) を用いた画像外観検査が注目されている。ルール判定は結果の解釈が容易である強み、機械学習を用いた判定(以下、AI判定)は複雑なパターンに対応可能である強みをもつため、それぞれの強みを生かす画像外観検査自動化システムの構築を進めている。

2. 目的

当社では残材検知²⁾³⁾、鋼材超音波検査の波形分類⁴⁾などで機械学習を製造現場へ適用している。それらは教師あり学習(学習データに答えをつけて学習する手法)であり、予測精度は高いが、外観検査へ適用する場合は不良品を網羅する必要がある。しかし外観検査の場合は、そもそも不良品の発生率が低く種類も多様なため、学習に必要な不良品を網羅するには多大な時間と労力を要する。そのため外観検査に教師あり学習は不向きのため、外観検査用の新たな機械学習手法が必須となる。そこで不良品画像が学習に不要で未知の不良の検出が可能な半教師あり異常検知技術を用いた様々な研究がなされている。

Rule	Black stains → NG		
True Label	OK	NG	NG
Rule-based judgement	OK	NG	OK

Fig. 1. Example of rule-based method.

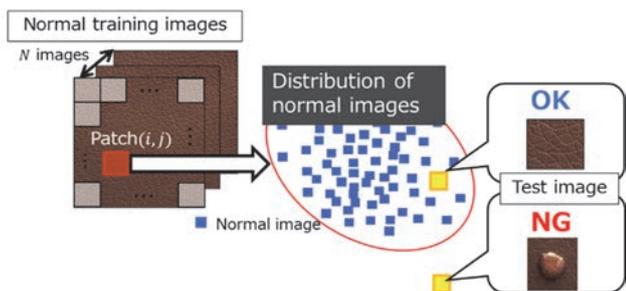


Fig. 2. Concept of anomaly detection.

例えば、PaDiM(Patch Distribution Modeling)¹⁾は画像内の位置(i, j)ごとに良品画像の特徴の分布を学習し、学習した正常の分布と比較することで異常を検出する方法である。(Fig. 2)

PaDiM は画像データセットにおける異常検知で最も優れた性能を示した技術(2020年11月時点)であるが、実際に現場へ適用するにはルール判定と同程度の処理時間が求められる。前述のとおり、画像外観検査はルール判定とAI判定のそれぞれの強みを生かしたシステムを目指しているため、ルール判定とAI判定を同時に実行し、二つの判定結果から最終的な判定をする仕組みを考えている (Fig. 3)。したがってAI判定はルール判定とほぼ同じ処理時間で結果を判定する必要がある。

また画像外観検査で使用される画像は、高精細な画像を採用していることが多いため、より処理の高速化が求められる場合も考えられる。

本開発では、機械学習による半教師あり異常検知技術を現場へ適用するために必要な性能と速度の両立を可能にし、既存手法よりもさらに高速化を目指した新たな方法を提案し、その有効性を実証する。

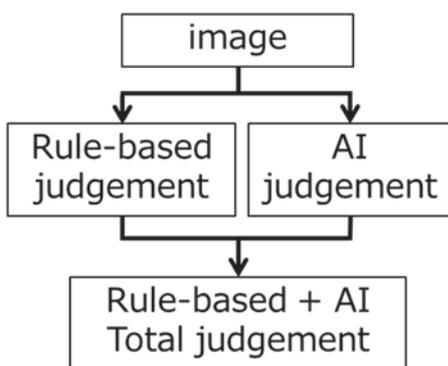


Fig. 3. Flowchart of image-based visual inspection.

3. 開発目標

本開発では、性能を評価するために、公開データセットであるMVTec AD⁹⁾を用いた。MVTec ADは、全15種類の工業用製品を対象とした5354枚の高解像度画像による異常検知用のデータセットである。画像外観検査は様々なカテゴリに対応する必要があるため、多種多様な状況を含むベンチマークデータセットであるMVTec ADデータセットを採用した。

性能は、実用レベルの目安として、AUROC(Area Under the ROC Curve):0.94以上とした。AUROCは、ROC曲線(Receiver Operating Characteristics Curve)の下の面積のことであり、機械学習モデルの性能指標として多く利用されている (Fig. 4)。ROC曲線とは、縦軸に真陽性率(TPR:True Positive Rate) (式(1))、横軸に偽陽性率(FPR:False Positive Rate) (式(2))を取ってプロットした場合の曲線である。AUROCは1.0に近いほど良いモデルとされている。

処理時間は最初の適用予定先のルール判定が0.7s/1製品のため、0.7s以下とした。

Table 1. Confusion matrix.

		Predicted	
		Negative	Positive
Actual	Negative	True Negative (TN)	False Positive (FP)
	Positive	False Negative (FN)	True Positive (TP)

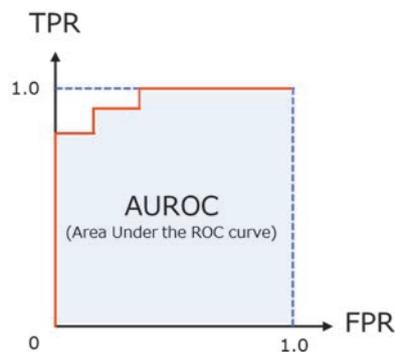


Fig. 4. ROC curve.

$$TPR = \frac{TP}{TP + FN} \quad (式(1))$$

$$FPR = \frac{FP}{TN + FP} \quad (式(2))$$

4. 開発課題

本開発では、半教師あり異常検知技術の一つである PaDiM¹⁾を採用した。画像内の位置 (i, j) ごとに良品か否かを判定する手法のため、被写体の位置がずれにくい画像には微細な欠陥も見逃さないという特徴をもっており、外観検査には適した手法である。しかしこの手法をそのまま使用すると開発目標(処理時間)に未達であることが判明した。(Table 2)

性能を維持しつつ処理時間を削減するため、以下の開発課題を提案する。

Table 2. Development goals.

	Target value	Current value
AUROC (Pixel level)	0.94	0.943
Processing time[s]	0.70	0.84

(1) 解像度

PaDiM では、事前に大規模データセットで学習済みのモデルによって様々な画像の情報(以下、特徴量)を取り出す技術を採用している(Fig. 5)。

従来手法では、Fig. 5 のように、大規模な自然画像データセットである ImageNet⁶⁾で事前学習した Wide ResNet50-2⁷⁾ というモデルを使用して、階層の異なる三カ所からそれぞれ特徴量を取り出している。簡易～複雑な特徴量まで取り出すことが可能というメリットがあるが、低解像度な特徴マップの解像度を1 番高解像度な特徴マップ (Fig. 5 の場合 56×56) に合わせる必要がある (Fig. 5 の場合 Upsampling)。

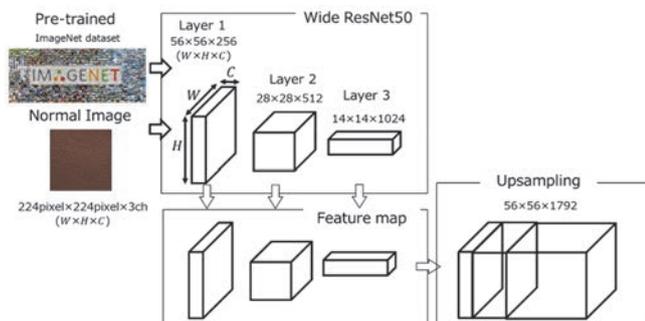


Fig. 5. Feature extraction of normal image (for training).

PaDiM は画像内の位置 $(i, j) \in [1, W] \times [1, H]$ ごとに判定を実施する。その局所的な領域をパッチと呼び、パッチ数 $(W \times H)$ が多いほど処理時間がかかる。(Fig. 6)

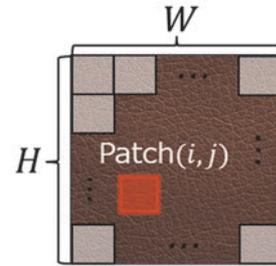


Fig. 6. Image of patches.

(2) 特徴選択方法

使用する特徴量が多いとデータ量が大きくなりメモリ使用量の過大等の理由から次元削減を実施する(Fig. 7)。使用する特徴量を少なくすると速度面で有利であるが、異常検知に有効な特徴マップを削ってしまう恐れがある。そこで次元削減をするにあたり、いかに正常と異常の識別可能な特徴を選別できるかが重要となる。PaDiM では、ランダムに特徴を選択しているが、性能面、速度面で同レベルもしくは更に良い特徴の選定方法が必要となる。

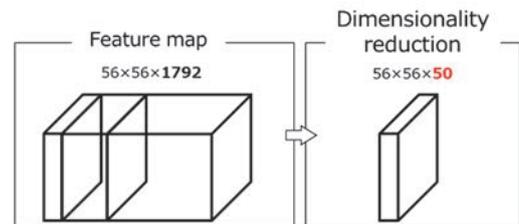


Fig. 7. Dimensionality reduction.

5. 課題解決の方法

5. 1 解像度の改善案

4. (1) の問題を解決する案として、Wide ResNet50-2 の2 ブロック目最終層の特徴量のみを取得する方法を提示する(Fig. 8)。

この方法は特徴の種類が限定されるが、高解像度に調節する必要がないため、パッチ数 $(W \times H)$ が少なくて済む。その分判定回数が減るため、処理時間を短縮できる。

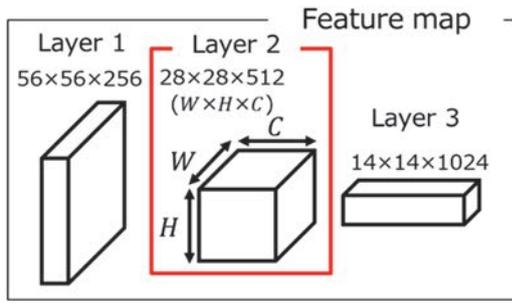


Fig. 8. Feature extraction method.

特徴量を抜き取る層を2ブロック目の最終層にした理由は、正常と異常の見分けが可能な特徴を捉えることができ、かつ位置情報が低下しないからである。取り出す特徴の層が浅い(Layer1)ほど簡易な特徴(ex. エッジ、色など)を、逆に深い(Layer3)ほどより複雑な特徴(ex. 物体の形状、構造)を捉えることができる。しかし Layer3 は位置情報が低下してしまうため、今回はLayer2を採用している。

5. 2 特徴選択方法の改善案

4. (2) の問題を解決する案として、特徴毎に N 枚の正常画像(学習用)の平均値を算出し、平均値の小さい特徴を選択する方法を提示する。(Fig. 9)

とあるチャンネル $C (C = 1, 2, \dots, d)$ の特徴量を $x_{c(ij)}$ とする。チャンネルは、画像データの各ピクセルに含まれる情報の種類を示す。特徴量 $x_{c(ij)}$ における正常画像 (N 枚) の平均値算出方法は以下の式(3)になる。

$$\mu_{c(ij)} = \frac{1}{N} \sum_{k=1}^N x_{c(ij)}^k \quad (3)$$

$\mu_{c(ij)}$ はチャンネル数 d 分得られるため、 d 個の $\mu_{c(ij)}$ から小さいチャンネルを選択する。

特徴マップの値は複数のカーネルを通して得られる値である。カーネルは、画像の中の重要な部分を強調する役割があり、カーネルに強く反応する特徴は大きな値として抽出さ

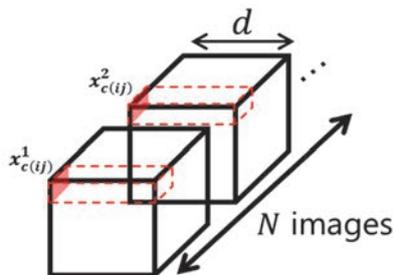


Fig. 9. Method to select features with small mean values per feature channel.

れる。異常検知においては、正常の潜在空間を定義し、異常との違いを明確にする必要がある。つまり正常データには反応しない特徴量を見つけることができれば、異常データには反応する特徴量を見つけることができるのではないかと考えた。そこで特徴量のチャンネル毎に N 枚の正常画像(学習用)の平均値を算出することで、正常データには反応しないチャンネルを特定する手法を考えた。

6. 実験

6. 1 実験方法

前述(4. (1))の通り、特徴量を取り出すために ImageNet で事前学習された Wide ResNet50-2 モデルを使用した。また、MVTec AD データセット画像サイズは 700×700 pixel ~ 1024×1024 pixel の範囲内であるが、全て 256×256 pixel にリサイズし、画像中央の 224×224 pixel をセンタークロップして使用した。チャンネル数は PaDiM では 550 次元だったが、速度面を考慮し、今回は 50 次元で統一した。

5. 1 で示した解像度の改善案、5. 2 で示した特徴選択方法の改善案それぞれ条件を変えて実験を行った。

6. 2 結果

6. 2. 1 性能評価

各データセットカテゴリにおいて、性能指標である AUROC 値を算出した結果を Table 3 に示す。今回はピクセルレベルにおける AUROC 値を採用している。

Table 3. Comparison of AUROC values between previous and proposed methods.

method		① previous	② 5.1	③ 5.2	④ proposed
object	bottle	0.973	0.978	0.966	0.979
	cable	0.872	0.893	0.926	0.944
	capsule	0.969	0.976	0.976	0.977
	hazelnut	0.946	0.948	0.955	0.968
	metal_nut	0.949	0.912	0.933	0.926
	pill	0.955	0.889	0.910	0.905
	screw	0.948	0.952	0.855	0.963
	toothbrush	0.983	0.982	0.978	0.985
	transistor	0.939	0.931	0.976	0.946
	zipper	0.978	0.969	0.967	0.968
	average	0.951	0.943	0.944	0.956
texture	carpet	0.962	0.950	0.970	0.982
	grid	0.908	0.774	0.945	0.855
	leather	0.988	0.978	0.980	0.989
	tile	0.863	0.778	0.906	0.940
	wood	0.909	0.898	0.892	0.918
	average	0.926	0.876	0.939	0.937
average all classes		0.943	0.921	0.942	0.950

5. 1の比較をすると(①、②)、AUROC値は②解像度の改善案を適用すると①従来手法よりも低下する。これは①従来手法のほうが様々な解像度の特徴量を使用できるからと考えられる。また5. 2の比較をすると(①、③)、AUROC値はほぼ同等の値となっている。つまり、提案した特徴選択方法を使用することで、高解像度な特徴マップを使用せず従来手法と同等のAUROC値が得られることがわかった。

AUROC値は④提案手法(5. 1、5. 2の方法を採用)において、0.950(pixel level)と目標値である0.940(pixel level)を達成できている。

6. 2. 2 処理時間

処理時間をまとめた表をTable 4に示す。CPU(Intel® Xeon® W-2225 CPU 4.10GHz)を使用し、それぞれの処理時間を測定した結果を示す。

Table 4より、5. 1の比較をすると(①、②)、処理時間は約1/2以下まで短くなっている。従来手法のほうが高解像度の特徴マップが得られるため、パッチ数($W \times H$)が増える(①: $W=H=56$ 、②: $W=H=28$)。各パッチで異常度の計算を実施するため、パッチ数が多いほど異常度計算処理数が増加するため、処理時間も増加すると考えられる。

また5. 2の比較をすると(①、③)、従来手法のほうが約0.4s処理時間が短くなっている。従来手法は全てのパッチで同じチャンネルをランダムに選択しているのに対し、提案手法

はパッチごとにチャンネルを選択しているため多少処理時間が増加したと考えられる。

処理時間は④提案手法(5. 1、5. 2の方法を採用)において、全体で0.417sと目標を達成できている。

よって、性能と処理時間を考慮すると、提案手法が従来手法と比較して同等の性能を持ちつつ処理時間を約1/2に削減できることを確認できた。

7. おわりに

本開発では、製造業の画像外観検査において、既存のPaDiMを改良し、更なる高速化を目指した新たな手法を提案した。特に、解像度と特徴選択方法において、既存手法の課題を改善するアプローチを採用した。

解像度の改善案として、モデルの最終層の特徴量のみを取得することで、パッチ数を削減し、異常度計算処理数を減らした。また特徴選択方法の改善案として、特徴毎に正常画像(学習用)の平均値を算出し、平均値の小さい特徴を選択する方法で性能の向上を図った。

実験結果から、提案手法が従来手法と比較して、同等の性能を持ちつつ、処理時間を約1/2に削減できることが確認できた。この技術は様々な画像に適用でき、製造業の更なる検査の効率化と品質向上に貢献することが期待できる。

参考文献

- 1) T. Defard, et al. "PaDiM: A Patch Distribution Modeling Framework for Anomaly Detection and Localization," ICPR, 2021.
- 2) 村瀬博典, 大島亮太, 甲村啓伍, 愛知製鋼技報, 35 (2019), 1, 10-15.
- 3) 岡田大樹, 幸村健太, 村瀬博典, 梶野晶敬, 愛知製鋼技報, 38 (2022), 1, 7-13.
- 4) 大島亮太, 村瀬博典, 澤清和, 愛知製鋼技報, 39 (2023), 1, 22-26.
- 5) P. Bergmann, et al. "MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection," CVPR, (2019), pp. 9592-9600.
- 6) J. Deng, et al. "ImageNet: A large-scale hierarchical image database," CVPR, (2009).
- 7) S. Zagoruyko, et al. "Wide Residual Networks," BMVC, 2016.

Table 4. Comparison of Processing time between previous and proposed methods.

method		① previous	② 5.1	③ 5.2	④ proposed
object	bottle	0.863	0.319	1.345	0.422
	cable	0.912	0.402	1.245	0.413
	capsule	0.841	0.377	1.232	0.414
	hazelnut	0.851	0.355	1.223	0.415
	metal_nut	0.845	0.350	1.229	0.415
	pill	0.848	0.349	1.239	0.416
	screw	0.842	0.349	1.243	0.416
	toothbrush	0.813	0.347	1.247	0.418
	transistor	0.795	0.347	1.247	0.419
	zipper	0.784	0.346	1.246	0.419
	average	0.840	0.354	1.250	0.417
texture	carpet	0.841	0.365	1.234	0.415
	grid	0.841	0.357	1.229	0.416
	leather	0.843	0.352	1.216	0.415
	tile	0.836	0.348	1.247	0.417
	wood	0.790	0.347	1.248	0.419
		average	0.830	0.354	1.235
average all classes		0.836	0.354	1.245	0.417