

# 正常画像を用いた異常画像検知手法の検討

山野井祐介<sup>1</sup> 原潤一<sup>2</sup> 渡辺裕<sup>1,2</sup>

**概要**：ディープラーニングによる画像認識は多量のデータを集めることができる分野においては活用が見込まれるが、データ収集が困難な異常検知には活用することが難しい。そこで、正常なデータを中心に学習する手法や、他のデータセットに対して学習済みのモデルを転用する手法が注目されている。その代表的な異常検知手法が SPADE, PaDiM である。事前の実験から、これらの手法は AUROC では高い評価を示しているが、AUPR の評価は低いことがわかった。そこで、SPADE, PaDiM の AUPR での評価の向上のため、「特徴点マッチングによる対象物検出(提案手法 1)」、「特徴量抽出するレイヤーの変更(提案手法 2)」を提案する。評価実験では、一部のデータセットで AUPR の向上を示し、その結果から AUPR 向上に有効なレイヤーを確認する。

## A Study of Anomaly Detection Method Using Normal Images

YUSUKE YAMANOI<sup>†1</sup> JUNICHI HARA<sup>†2</sup> HIROSHI WATANABE<sup>†1,2</sup>

### 1. まえがき

ディープラーニングによる画像認識は、認識対象の学習時に多量の画像データを必要とする。そのため、データ収集に困難を伴う異常検知には活用することが難しい。そこで、正常なデータを中心に学習する手法や、異常検知用ではない他のデータセットに対して学習済みのモデルを転用する手法が注目されている。その代表的な異常検知手法が SPADE[1], PaDiM[2]の手法である。これらの手法での性能評価として AUROC が指標として使用されている。AUROC は陰性、陽性に対する判別能のバランスに着目した評価である。両手法の AUROC は画像単位、画素単位どちらにおいても高いため、異常検知の基本構成として期待できる。しかし、事前の実験から陽性を陽性と判断できているかに着目した評価である AUPR の評価は低いことがわかった。そこで、SPADE, PaDiM の AUPR での評価の向上のため、「特徴点マッチングによる対象物検出(提案手法 1)」、「特徴量抽出するレイヤーの変更(提案手法 2)」を提案する。

### 2. 関連技術

#### 2.1 正常データを活用した異常検知手法

##### 2.1.1 SPADE

ImageNet[3]を学習した Wide ResNet[4]を特徴量抽出器として特徴量抽出を行い、正常な画像の特徴量との最近傍特徴量の距離を異常スコアとして異常検知している。

##### 2.1.2 PaDiM

ImageNet を学習した Wide ResNet を特徴量抽出器として特徴量抽出を行い、正常な画像の特徴量の平均と共分散を算出し、それをを用いて正常な画像の特徴量とのマハラノビス距離[5]を測り、それを異常スコアとして異常検知している。

### 2.2 性能評価指標

#### 2.2.1 AUROC

AUROC は、機械学習の分類タスクの性能評価指標の一つである。テスト結果の偽陽性率に対する再現率の変化をグラフ化し、その線より下の部分の面積で評価を行う。値は 0~1 の範囲をとり、値が 1 に近いほど判別能が高いことを示す。陽性を陽性と判別できているか、陰性を陰性と判別できているかの評価であるため、陰性、陽性に対する判別能のバランスに着目した評価である。そのため、扱うデータが不均衡でない場合に有用である。

#### 2.2.2 AUPR

AUPR は、機械学習の分類タスクの性能評価指標の一つである。テスト結果の再現率に対する適合率の変化をグラフ化し、その線より下の部分の面積で評価を行う。値は 0~1 の範囲をとり、値が 1 に近いほど判別能が高いことを示す。陽性を陽性と判断できているかに着目した評価であるため、陰性が大半であるような偏りのあるデータセットでも陽性を評価することができる。

### 3. 予備実験

#### 3.1 概要

本研究の主要技術である SPADE と PaDiM について AUPR での評価を確認する予備実験を行う。これらの従来手法は画素単位のクラス分類の性能評価指標として AUROC を採用している。しかし、正常に対して異常の数が少ないデータセットの場合、AUROC では異常に対する判別能を正しく評価できない。そこで、異常に対する判別能を評価するため、従来手法を AUPR で評価する。さらに、SPADE と PaDiM で実際に推定した異常スコアを可視化することで問題点の洗い出しを行う。

1 早稲田大学大学院基幹理工学研究所

2 早稲田大学国際情報通信センター

### 3.2 従来異常検知手法の AUROC と AUPR

以下の手順で予備実験を行った。異常検知手法である SPADE と PaDiM を工業検査に特化した異常検知手法のベンチマーク用データセットである MVTecAD[6][7]の正常画像のみ使って学習し、その後、異常画像の画素ごとのクラス分類を AUROC, AUPR の二つの評価指標で評価した。なお、学習及び評価は MVTecAD の 15 クラスそれぞれで独立して行った。両手法における AUROC と AUPR の 15 クラスの平均値を表 1 に示す。SPADE と PaDiM の AUPR による評価は、AUROC による評価よりも大幅に低いことが確認された。

表 1 SPADE と PaDiM の 15 クラスの平均評価値

モデル	AUROC	AUPR
SPADE	0.968	0.536
PaDiM	0.969	0.491

### 3.3 従来異常検知手法の結果の可視化

SPADE と PaDiM で実際に推定した異常スコアを可視化した。AUPR が SPADE, PaDiM とともに 15 クラスの中で最低であった screw クラスの画像の一例を図 1 に示す。異常検知の対象物であるネジだけでなく背景に異常の反応が出ていることがわかる。従来手法では、画像全体の特徴量を抽出し異常判定を行うため、本来注目すべきでない背景が評価に影響していることがわかった。

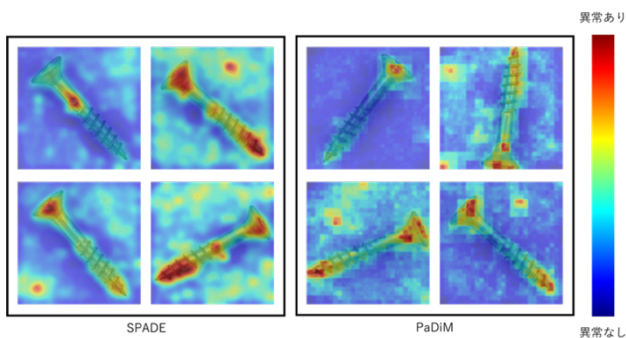


図 1 screw クラスの異常スコアマップ

## 4. 提案手法

### 4.1 提案手法 1：特徴点マッチングによる対象物検出

3.3 節で示したように、SPADE 及び PaDiM では、画像内の背景に異常の反応が出ていることがわかる。本来、背景は異常検知する必要のない範囲である。そこで、画像から異常検出対象物だけの範囲を抽出する方法を提案する。以下に、処理手順を示す。

- (1) データセットに適した特徴点抽出手法の選定
- (2) 特徴点マッチングによる対象物マスクの作成
- (3) 従来異常検知手法への組み込み

まず、対象物の種類によって有している特徴が異なるため、対象物ごとに、代表的な特徴点抽出手法で特徴点マッチングのテストを行う。各クラスの任意の正常画像と、その画像以外の正常画像との特徴点マッチングを行い、マッチングした特徴点の周囲を対象物の範囲とする。その後、対象物がマスク内に収まっているか、対象物の形を捉えているかを、目視によって評価し、満たしている場合、その特徴点抽出手法を採用する。データセットの各対象物に対して、有効であった特徴点抽出手法を用いて、テスト画像と正常画像との特徴点マッチングを行う。特徴点抽出手法の選定時と同じく、マッチングした点の周辺を対象物の範囲としマスクを作成する。SPADE と PaDiM の出力である画素ごとの異常スコアに対し、このマスクを任意の比率で掛け合わせたものを、提案手法 1 の出力とし、AUPR で評価を行う。図 2 に提案手法 1 の過程を示す。

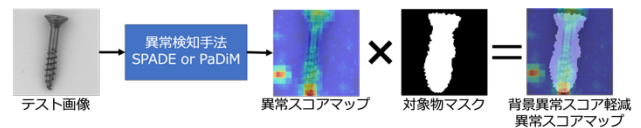


図 2 提案手法 1 の過程

### 4.2 提案手法 2：特徴量抽出するレイヤーの変更

従来異常検知手法である SPADE と PaDiM は、ImageNet を学習済みの Wide ResNet を特徴量抽出器として用いている。両手法共に、全体四つのレイヤーのうち、前半の三つのレイヤーのみを特徴量抽出器として使用し、AUROC で評価を行なっている。AUPR での評価は行ってないため、従来手法のレイヤーの選択が最適であるとは限らない。そこで、特徴量を抽出するレイヤーの変更を提案する。SPADE, PaDiM の概略図を図 3 に示す。SPADE, PaDiM ではレイヤー1~3 を使用していたが、提案手法ではレイヤー 4 を含めた上で、様々な組み合わせで実験を行う。特徴量を抽出するレイヤー以外の処理は既存の SPADE, PaDiM に従う。

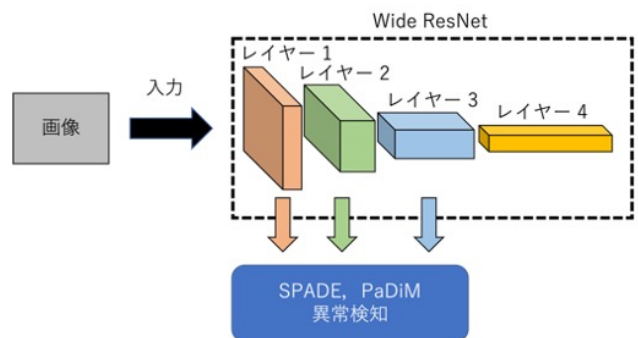


図 3 従来異常検知手法

## 5. 評価実験

### 5.1 データセット

工業検査に特化した異常検知手法のベンチマーク用データセットである MVTECAD を用いる。15 種類の異なるオブジェクトやテクスチャ画像からなり、各カテゴリは欠陥のない画像と、様々な種類の欠陥がある画像を含んでいる。各クラス平均正常画像数は約 250 枚、平均異常画像数は約 80 枚である。また、画像中の異常画素は約 5% であり、正常画素と異常画素の数に偏りがある。

### 5.2 提案手法 1 : 特徴点マッチングによる対象物検出

MVTECAD データセットの背景が存在する 7 クラスに適切な特徴点抽出手法を見つける。代表的な特徴点抽出手法である, ORB, AKAZE, BRISK によって特徴点の抽出を行う。正常画像との特徴点マッチングを行い、マッチングした特徴点の周囲を対象物の範囲とする。ここで、評価は 0, 1, 2 の 3 段階、目視で行った。0 は「対象物を全く捉えていない」、1 は「対象物を一部捉えている」、2 は「対象物を捉えている」を表す。その結果を表 2 に示す。screw クラスは画像内の外縁にコーナーが多く存在しているためその形を捉えることができた。toothbrush クラスは画像内の外縁は丸みを帯びているものの、歯ブラシの毛先部分が特徴点として抽出できたため、形を捉えることができた。他のクラスは物体表面の凹凸や画像内にエッジやコーナーが少ないため、抽出できる特徴点も少なく、マッチングした特徴点では対象物をカバーできなかった。

表 2 クラスと特徴点抽出手法の相性

クラス	ORB	AKAZE	BRISK
bottle	0	0	0
cable	1	0	0
pill	0	0	0
hazelnut	0	0	0
toothbrush	0	2	2
capsule	0	1	0
screw	2	2	1

表 2 から AKAZE による特徴点抽出が toothbrush, screw クラスに有効であることが確認されたため、この二つのクラスに対して、AKAZE による特徴点マッチングを行い、作成したマスク外の異常スコア倍率を 0, 0.2, 0.4, 0.6, 0.8, 1.0 にした結果の AUPR を求めた。マスク外倍率が 1.0 のとき、対象物マスクを追加していないのと同様である。その結果を図 4 に示す。SPADE, PaDiM とともにマスク外倍率が 0.6~0.8 のとき、マスク外倍率が 1.0 (従来手法) のときよりも toothbrush クラスの AUPR が向上している。よって、背景に誤って振られた異常スコアを軽減することができて

いることが確認できた。screw クラスについては、AUPR の向上は確認できなかった。これは、対象物マスクが完全に対象物を捉えきれていないためである。特にネジの先端が欠けている場合、特徴点マッチングではその欠けた部分を対象物の範囲にすることができなかった。

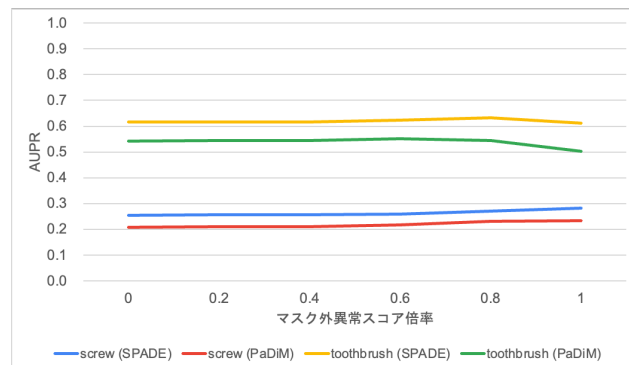


図 4 対象物マスク導入結果

### 5.3 提案手法 2 : 特徴量抽出するレイヤーの変更

従来の異常検知手法である SPADE と PaDiM が特徴量抽出器として用いている ImageNet を学習済みの Wide ResNet の特徴量抽出レイヤーを変更する。なお、SPADE, PaDiM とともに特徴量抽出するレイヤーの変更以外は従来通りの手法である。レイヤー 1, 2, 3 から特徴量を抽出した場合は「1,2,3」と表記する。SPADE, PaDiM とともに「1,2,3」が従来手法で採用されていたレイヤーである。その結果を図 5 に示す。図 5 の SPADE と PaDiM の AUPR の相関係数は 0.70 であり、正の相関が認められた。このことから SPADE, PaDiM の両手法において有効なレイヤーは相関しているといえる。また、レイヤー 1, 4 で大きく AUPR が低下している。レイヤー 4 の有無による AUPR の違いを表 3 に示す。この表から、レイヤー 4 が AUPR を下げていることがわかる。特徴量抽出器の Wide ResNet は ImageNet を学習済みであるためレイヤーが後半になる程、ImageNet の分類問題に最適化されている。そのため、最終レイヤーであるレイヤー 4 から抽出される特徴量は MVTECAD での異常分類に有効ではなかったと考えられる。さらに、クラスごとの AUPR では、2 種類のレイヤーを組み合わせたときが最高値であることが 60% 以上のクラスで確認できた。よって、AUPR で評価する際は、少ないレイヤーである分計算量も削減できるため、レイヤーは 2 種類で十分であることがわかった。

表 3 レイヤー 4 の有無による AUPR

レイヤー 4 の有無	AUPR	
	SPADE	PaDiM
あり	0.497	0.420
なし	0.531	0.485

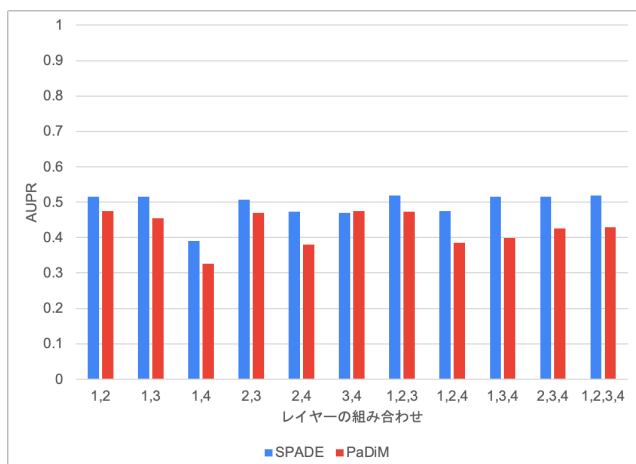


図 5 微量抽出レイヤーによる AUPR の変化

## 6. むすび

本研究では、正常なデータを用いた異常検知手法の AUPR 向上のための手法を提案した。提案手法 1 の特徴点マッチングによる対象物検出では、異常検知エリアを対象物に限定した手法を提案した。MVtecAD データセットの screw クラスと toothbrush クラスに対し対象物のマスクが作成できた。対象物マスクの従来異常検知手法への組み込みでは toothbrush クラスの AUPR の向上が確認できた。提案手法 2 の特徴量抽出するレイヤーの変更では、様々なレイヤーの組み合わせによる実験で異常検知での AUPR 向上に有効なレイヤーがわかった。さらに、60%以上のクラスでは、2 種類のレイヤーで 3 種類するとき以上の AUPR が確認できた。

提案手法 1 は、特徴点が検出でき、特徴点マッチングによって対象物をカバーできる必要があるため、限られたデータにしか適用できなかつた。よって、凹凸が少ない滑らかな物体においても画像中の対象物の範囲を特定する手法が必要である。この場合、正常データの対象物のマスクを作成し、機械学習モデルで学習する方法が考えられる。さらに、本手法では欠損した部分がマスクの対象外になってしまったため、欠損した部分も対象物として捉える手法も必要である。提案手法 2 では、データによって適切なレイヤーの組み合わせが異なつた。そのため、データの特徴によって動的にレイヤーを決定することが課題である。

## 参考文献

- [1] Cohen, N. and Hoshen, Y.. Sub-Image Anomaly Detection with Deep Pyramid Correspondences. arXiv preprint. arXiv:2005.02357. 2020.
- [2] Defard, T., Setkov, A., Loesch, A. and Audigier, R.. PaDiM: a Patch Distribution Modeling Framework for Anomaly Detection and Localization. arXiv preprint. arXiv:2011.08785. 2020.
- [3] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L.. ImageNet: A Large-Scale Hierarchical Image Database. CVPR. June 2009.
- [4] Zagoruyko, S. and Komodakis, N.. Wide Residual Network. arXiv preprint. arXiv:1605.07146. 2016.

- [5] Mahalanobis, P.. On the generalized distance in statistics. National Institute of Science of India. 1936.
- [6] Bergmann, P., Fauser, M., Sattlegger, D., and Steger, C.. Mvtec ad-a comprehensive real-world dataset for unsupervised anomaly detection. CVPR. June 2019.
- [7] Bergmann, P., Batzner, K., Fauser, M., Sattlegger, D., and Steger, C.. The MVtec Anomaly Detection Dataset: A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. International Journal of Computer Vision. Jan. 2021.