

2021 年度

早稲田大学大学院基幹理工学研究科情報理工・情報通信専攻 修士論文

正常画像を用いた異常画像検知手法の検討

A Study of Anomaly Detection Method

Using Normal Images

山野井 祐介

(5120F088-1)

提出日：2021.01.24

指導教員：渡辺 裕 印

研究指導名：オーディオビジュアル情報処理研究

目次

第 1 章	序論	1
1.1	研究の背景.....	1
1.2	本研究の目的.....	1
1.3	本論文の構成.....	2
第 2 章	関連技術	3
2.1	まえがき.....	3
2.2	CNN.....	3
2.2.1	CNN とは.....	3
2.2.2	ResNet.....	3
2.2.3	Wide ResNet.....	4
2.3	正常データを活用した異常検知手法.....	4
2.3.1	SPADE.....	4
2.3.2	PaDiM.....	5
2.4	性能評価指標.....	7
2.4.1	AUROC.....	7
2.4.2	AUPR.....	7
2.5	特徴点抽出手法.....	8
2.6	むすび.....	8
第 3 章	予備実験	9
3.1	まえがき.....	9
3.2	データセット.....	9
3.3	従来異常検知手法の AUROC と AUPR.....	9
3.3.1	実験結果.....	9
3.3.2	考察.....	11
3.4	従来異常検知手法の結果の可視化.....	13
3.4.1	実験結果.....	13
3.4.2	考察.....	13
3.5	むすび.....	13
第 4 章	提案手法	14
4.1	まえがき.....	14

4.2	提案手法 1：特徴点マッチングによる対象物検出	14
4.2.1	概要.....	14
4.2.2	特徴点抽出手法の選定	14
4.2.3	異常検知手法への組み込み	15
4.3	提案手法 2：特徴量抽出するレイヤーの変更.....	16
4.3.1	概要.....	16
4.3.2	特徴量抽出するレイヤーの変更.....	16
4.4	むすび	16
第 5 章	評価実験と考察.....	18
5.1	まえがき	18
5.2	提案手法 1：特徴点マッチングによる対象物検出	18
5.2.1	クラスごとの適切な特徴点抽出手法の選定	18
5.2.2	対象物マスク+従来異常検知手法	18
5.2.3	考察.....	19
5.3	提案手法 2：特徴量抽出するレイヤーの変更.....	21
5.3.1	特徴量抽出するレイヤーの変更.....	21
5.3.2	考察.....	24
5.4	むすび	25
第 6 章	結論と今後の課題.....	26
6.1	結論.....	26
6.2	今後の課題.....	26
謝辞	27
参考文献	28
図一覧	29
表一覧	30
研究業績	31

第1章 序論

1.1 研究の背景

近年、ディープラーニングによる画像認識の精度が飛躍的に向上している。2015年には、大規模画像データセットである ImageNet[1]を用いた ILSVRC 画像認識コンテストで、人間の認識精度よりも優れた ResNet[2]というモデルが登場し、話題となった。このモデルは従来の手法よりもネットワークの層を深くすることで、学習できるパラメータの数を増加させた。その結果、より多くの画像を学習することができ、認識精度をあげた。実際、ResNet では 128 万枚の画像を学習に使っている。

ResNet のような大規模なモデルは、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、多量のデータ収集が困難な分野、具体的には、部品の製造過程で行われる工業検査、橋や道路のインフラ点検、X 線検査や CT 検査の医用画像診断などには活用することができない。このような分野の検査・診断を行える熟練した人材は少なく、雇用コストも高い。また、人間による評価はばらつきが生じてしまうため、一定の品質レベルを保つことも難しい。そのため、より安価に正確で安定した製造・診断を行うために、画像認識による異常検知は非常に効果的である。しかし、こういった分野では、正常な状態のデータは比較的容易に集めることができるが、異常なデータは稀である可能性が高く、多量に集めることは困難である。

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。中でも、ImageNet を学習済みのモデルは様々な種類の画像を多量に学習していることから、汎用的な特徴量抽出器として用いることができ、これを転用するケースが増えてきている。しかし、関連のないデータセットを学習しているためその特徴量はそれぞれの異常検知に最適化されていない。その解決策として、SPADE[3]や PaDiM[4]と呼ばれる、学習済みモデルを用いた異常検知への適応手法が近年登場している。これらの手法は、ImageNet を学習済みのモデルを用いて正常画像から特徴量を抽出し、それらの特徴量と、同じくして得られたテスト画像の特徴量と比較することで異常検知を行っている。その正解率は画像単位、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。

1.2 本研究の目的

SPADE や PaDiM などの学習済みモデルの異常検知への適応手法の性能改善を試みる。異常検知の性能は、AUROC で評価されている。この AUROC は 2.4.1 項で説明するように、正常画素と異常画素の両方を評価するため、画像中の異常画素が少ないような不均衡データの評価には適していない。そこで、予備実験として従来手法を不均衡デ

一タの評価に適した AUPR により評価する。AUPR は 2.4.2 項で説明するように、異常画素に着目した評価である。その結果を踏まえ、特徴点マッチングによる画像中の対象物の範囲を特定する手法や、学習済みモデルから特徴量抽出するレイヤーを変更する手法を提案する。その結果、従来の異常検知手法の AUPR を向上させることで、取りこぼしのない安定した検査・診断手法としての実用化を目指す。

1.3 本論文の構成

本論文の構成を以下に示す。

第 1 章は本章であり、本研究における背景、目的について述べる。

第 2 章では、正常データを活用した異常検知手法とその性能評価指標について述べる。

第 3 章では、従来手法を他の性能評価指標で評価した予備実験について述べる。

第 4 章では、本研究で提案する手法について述べる。

第 5 章では、提案手法を評価する実験について述べ、提案手法の有効性について述べる。

第 6 章では、本研究における結論と今後の課題について述べる。

第2章 関連技術

2.1 まえがき

本章では本研究で利用する関連技術として、CNN, 正常データを活用した異常検知手法, その性能評価指標について述べる. 提案手法で用いる CNN として, ResNet とその後継の Wide ResNet[5]について述べる. 正常データを活用した異常検知手法として, SPADE, PaDiM について述べる. これらの手法は ImageNet を学習済みのモデルを活用することで, 正常データのみを用いて異常検知を行う手法である. また, 性能評価指標である AUROC と AUPR について述べる. 従来手法は AUROC で評価されているが, 第3章の予備実験では AUPR での評価を行う.

2.2 CNN

2.2.1 CNN とは

CNN(Convolutional Neural Network)はディープラーニングのためのネットワークアーキテクチャである. ニューラルネットワークに畳み込み層とプーリング層を追加したものであり, 画像認識に広く用いられている[6].

2.2.2 ResNet

ResNet(Residual Network)は, He らが考案したニューラルネットワークのモデルである[2]. 従来よりも深いネットワークの学習を容易にする残差学習フレームワークを提案している. **Shortcut connection** を含んだ残差ブロックの概略を図 1 に示す. ここで, x は層の入力ベクトル, $F(x)$ は学習すべき特徴量抽出器を表している. この仕組みによって, 逆伝搬時の勾配が直接的に伝搬するようになり, 深いネットワークの際, 逆伝搬時に勾配が消失する問題が解決された. その結果, 層のより深い CNN を学習することが可能となり, 従来手法よりも良い認識精度を出している.

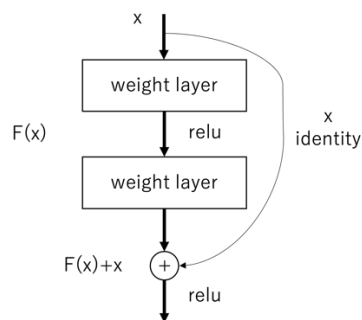


図 1 **Shortcut connection** を含んだ残差ブロック

2.2.3 Wide ResNet

Wide ResNet(Wide Residual Network)は, Zagoruyko らが提案したニューラルネットワークのモデルである[5]. ResNet のような深い残差ネットワークは, 数千層までスケールアップしても性能が向上することが示されている. しかし, 非常に深い残差ネットワークの学習では, 特徴の再利用率が低下し, その結果, これらのネットワークの学習には非常に時間がかかる. そこで, 残差ネットワークの深さを浅くし, 畳み込みのフィルタ数を増やした新しいアーキテクチャを提案している.

2.3 正常データを活用した異常検知手法

2.3.1 SPADE

正常データを活用した異常検知手法として, Cohen らが提案した SPADE がある[3]. Bergman らの分析[7]によると, ImageNet を学習させた ResNet の特徴は, 異常検知のデータセットを学習させた自己教師付きの特徴よりも, 異常検知に有用であることが示されている. そこで, ImageNet で事前に学習した ResNet を特徴量抽出器として異常検知に使用している. 学習段階では, 正常な画像を学習済みのモデルに入力し, 特定の3層から特徴量を抽出のうえ結合し, 保存している. SPADE の学習過程を図 2 に示す.

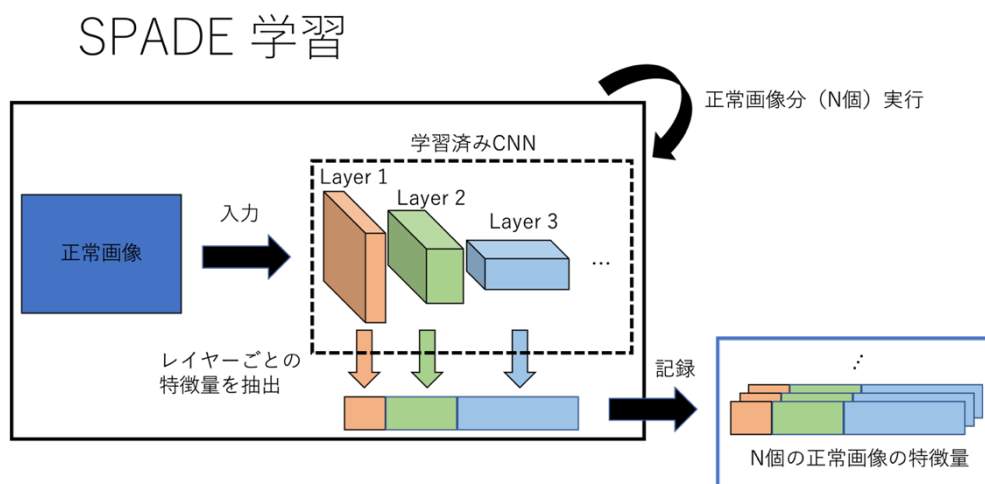


図 2 SPADE の学習過程

ここで, 特徴量抽出器 F に画像 x_i を入力したとき得られる特徴を f_i とする. f_i を式(1)に示す.

$$f_i = F(x_i) \quad (1)$$

その後、テスト画像に対しても特徴量を抽出し、正常画像の特徴量とのユークリッド距離を算出することで、 K 個の最近傍正常画像を検索している。SPADE のテスト過程を図 3 に示す。

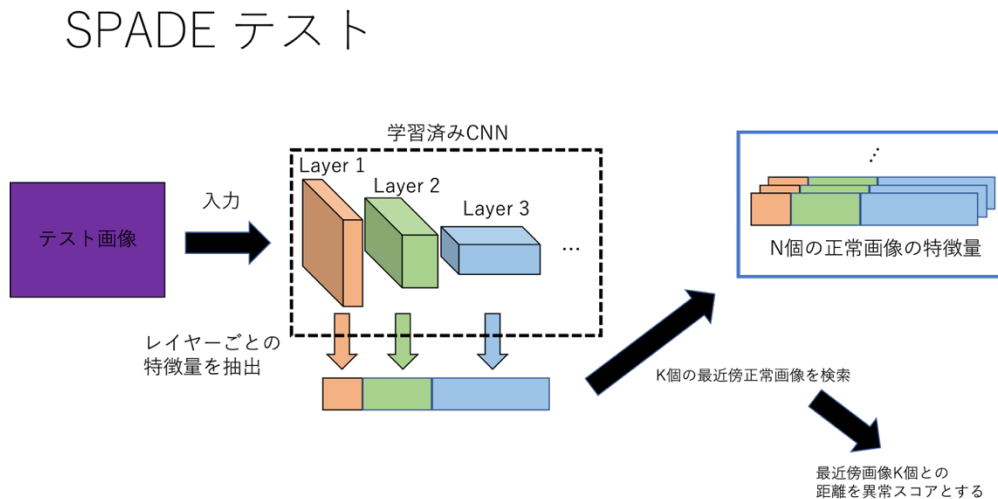


図 3 SPADE のテスト過程

K 個の最近傍正常画像特徴量とのユークリッド距離の平均を取ることで、画像単位での異常検知を行なっている。テスト画像を y ，その K 個の最近傍正常画像特徴量を $N_k(f_y)$ とすると、距離 $d(y)$ は式(2)で求めることができる。

$$d(y) = \frac{1}{K} \sum_{f \in N_k(f_y)} \|f - f_y\|^2 \quad (2)$$

次に、画素単位での異常検知についてである。画素 p における特徴量抽出器の抽出した特徴量を $F(x_i, p)$ と示す。画素 p におけるテスト画像の特徴量と、画素 p における K 個の最近傍正常画像の特徴量との平均距離 $d(y, p)$ を取ることで画素単位での異常検知を行なっている。

$$d(y, p) = \frac{1}{K} \sum_{f \in N_k(F(y, p))} \|f - F(y, p)\|^2 \quad (3)$$

2.3.2 PaDiM

正常データを活用した異常検知手法として、Defard らが提案した PaDiM がある[4].

PaDiM の特徴量抽出は SPADE と同じく、ImageNet を学習させた ResNet によって特徴量抽出を行なっている。正常な画像を学習済みモデルに入力し、特定の 3 層から特徴量を抽出、その後、計算量削減のためランダムに特徴量を削減している。N 枚の正常な学習画像の位置(i,j)における、特徴量 x_{ij} の集合 X_{ij} を式(1)に示す。

$$X_{ij} = \{x_{ij}^k, k \in [1, N]\} \quad (4)$$

ここで、 X_{ij} は多変量正規分布 $N(\mu_{ij}, \Sigma_{ij})$ によって生成されると仮定し、そのパラメータである平均 μ_{ij} 、共分散 Σ_{ij} を算出している。PaDiM の学習過程を図 4 に示す。

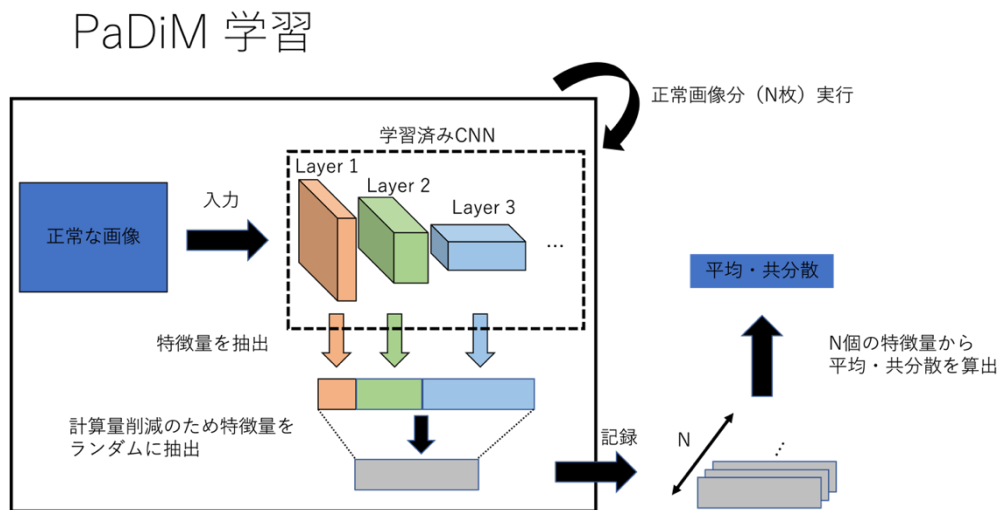


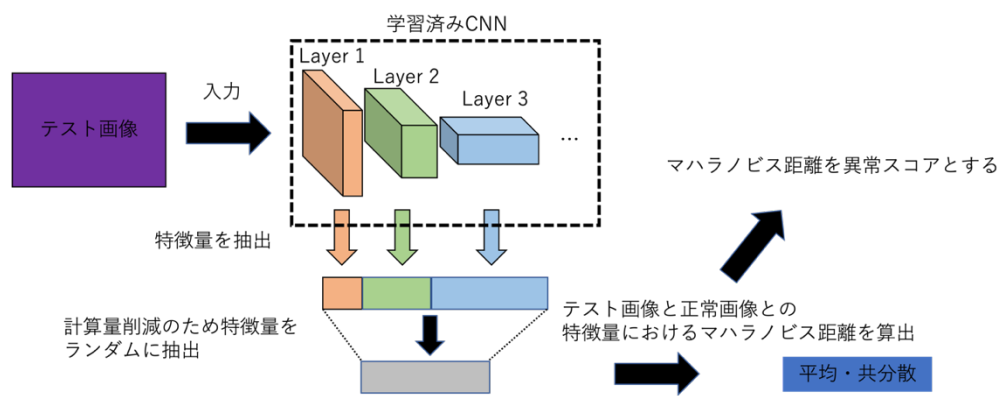
図 4 PaDiM の学習過程

テスト画像においても正常画像と同じように特徴量抽出し、先ほど算出した正常画像の特徴量の平均と共分散を用いて、正常画像とのマハラノビス距離 $M(x_{ij})$ [8] を算出している。マハラノビス距離 $M(x_{ij})$ を式(5)に示す。

$$M(x_{ij}) = \sqrt{(x_{ij} - \mu_{ij})^T \sum_{ij}^{-1} (x_{ij} - \mu_{ij})} \quad (5)$$

この距離を異常スコアとし、異常検知を行なっている。PaDiM のテスト過程を図 5 に示す。

PaDiM テスト



2.4 性能評価指標

2.4.1 AUROC

Area Under Receiver Operator Characteristics curve (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。閾値を設定せずに分類タスクを評価することが可能である。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。図 6 に ROC 曲線の例を示す。青い線が ROC 曲線、黄色で塗られた箇所が AUROC で評価に使う面積である。

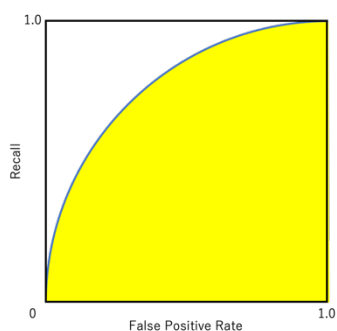


図 6 ROC 曲線の例

2.4.2 AUPR

Area Under the Precision-Recall curve (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。閾値を設定せずに分類タスクを評価することが可能である。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。図7にPR曲線の例を示す。青い線がPR曲線、黄色で塗られた箇所がAUROCで評価に使う面積である。

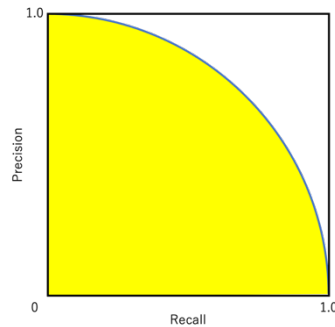


図7 PR曲線の例

2.5 特徴点抽出手法

特徴点抽出とは、画像中の特徴的な点を抽出することである。手法によって異なるが、特徴点はコーナーな輝度の変化などによって定義される。Open Source Computer Vision Library (OpenCV)[9]に実装されている代表的な特徴点抽出手法がORB, AKAZE, BRIEFである。なお、OpenCVは、コンピュータビジョン分野の代表的なアルゴリズムや画像を扱う基本処理をライブラリー化したものである。これらの特徴点抽出手法は、画像を回転や拡大縮小した際の影響を受けづらくロバストな手法である。

2.6 むすび

本章では本研究で用いるCNN、正常データを活用した異常検知手法、機械学習の分類タスクの性能評価指標、特徴点抽出手法について述べた。

第3章 予備実験

3.1 まえがき

本章では，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．

3.2 データセット

データセットには，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[10][11]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．図 8 に MVTecAD の画像例を示す．なお，オブジェクト画像は異常検知には関係のない背景を含み，テクスチャ画像は画像全体が対象物である．

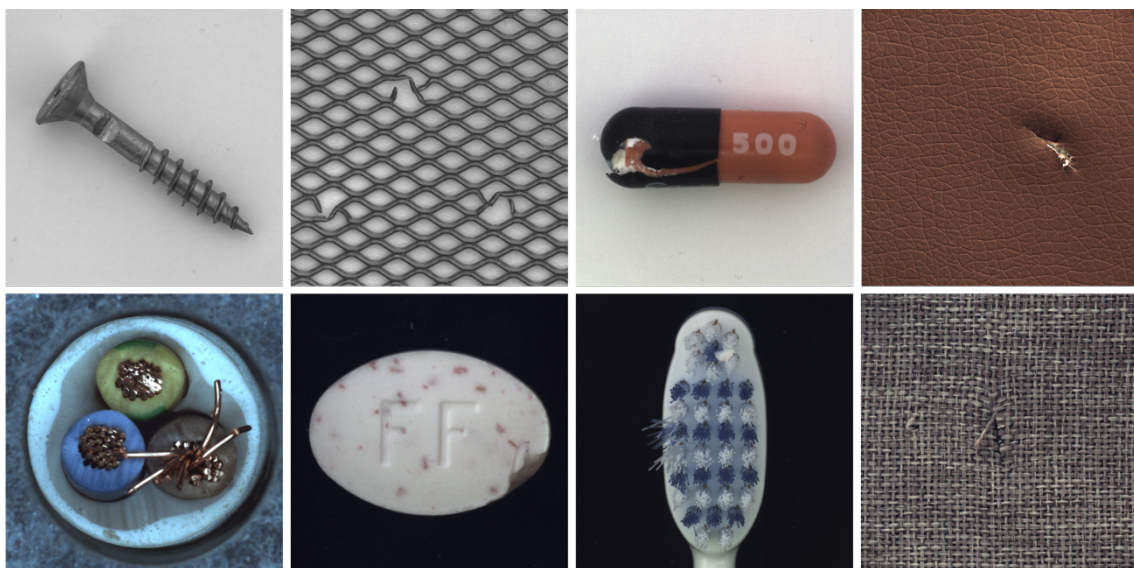


図 8 MVTecAD の画像例

3.3 従来異常検知手法の AUROC と AUPR

3.3.1 実験結果

予備実験は以下の手順で行った．異常検知手法である SPADE と PaDiM を MVTecAD 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。MVTecAD の 15 クラスそれぞれで評価を行い、SPADE の結果を表 1 に、PaDiM の結果を表 2 に、両手法における AUROC と AUPR の 15 クラスの平均値を表 3 に示す。

表 1 SPADE の AUROC と AUPR

クラス	AUROC	AUPR
bottle	0.977	0.768
cable	0.943	0.385
capsule	0.986	0.495
carpet	0.990	0.631
grid	0.961	0.336
hazelnut	0.981	0.604
leather	0.992	0.486
metal nut	0.961	0.801
pill	0.935	0.455
screw	0.989	0.281
tile	0.933	0.469
toothbrush	0.989	0.611
transistor	0.963	0.691
wood	0.944	0.461
zipper	0.982	0.559

表 2 PaDiM の AUROC と AUPR

クラス	AUROC	AUPR
bottle	0.978	0.727
cable	0.959	0.455
capsule	0.984	0.449
carpet	0.988	0.600
grid	0.972	0.283
hazelnut	0.976	0.507
leather	0.987	0.360
metal nut	0.958	0.744
pill	0.932	0.415
screw	0.981	0.216
tile	0.950	0.551
toothbrush	0.985	0.491
transistor	0.969	0.716
wood	0.935	0.420
zipper	0.974	0.427

表 3 SPADE と PaDiM の 15 クラスの平均値

モデル	AUROC	AUPR
SPADE	0.968	0.536
PaDiM	0.969	0.491

3.3.2 考察

表 3 の 15 クラスの平均値の結果から、SPADE, PaDiM とともに AUROC は約 0.96 であるが、AUPR は約 0.5 前後であることがわかる。これは、2.4 節で述べた性能評価指標の特徴の違いに起因し、データセットとして用いた MVTecAD の画像中の正常画素と異常画素の割合に偏りがあるためだと考えられる。そこで、MVTecAD のクラスごとの異常画素の割合と AUPR を表 4 に示す。なお、異常画素率による昇順で並べる。さらに、それをグラフにしたものを図 9 に示す。表 4 から手法ごとの異常画素率と AUPR の相関係数を算出すると、SPADE は 0.62, PaDiM は 0.80 である。この結果から異常画素率と AUPR には正の相関があるといえる。よって、異常画素が正常画素に対して少ないことが、AUPR の評価が低くなった要因の一つといえる。

表 4 MVTecAD のクラスごとの異常画素率と AUPR

クラス	異常画素率[%]	SPADE の AUPR	PaDiM の AUPR
screw	0.34	0.281	0.216
leather	0.87	0.486	0.360
grid	0.95	0.336	0.283
capsule	1.1	0.495	0.449
carpet	2.1	0.631	0.600
toothbrush	2.1	0.611	0.491
zipper	2.6	0.559	0.427
hazelnut	3.4	0.604	0.507
pill	4.0	0.455	0.415
cable	4.7	0.385	0.455
wood	5.1	0.461	0.420
bottle	7.6	0.768	0.727
tile	9.8	0.469	0.551
transistor	12	0.691	0.716
metal nut	14.5	0.801	0.744

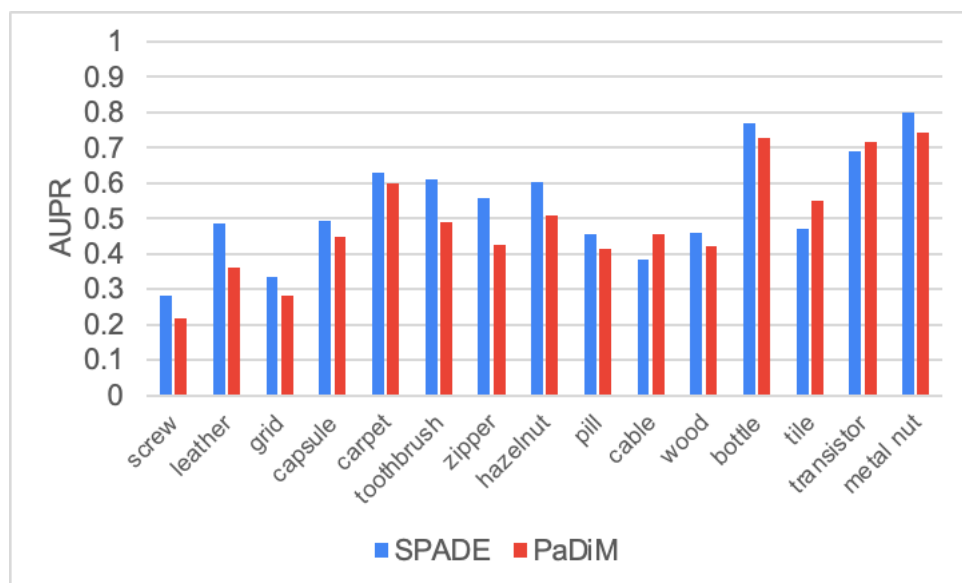


図 9 MVTecAD のクラスごとの AUPR

3.4 従来異常検知手法の結果の可視化

3.4.1 実験結果

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 10 に示す。

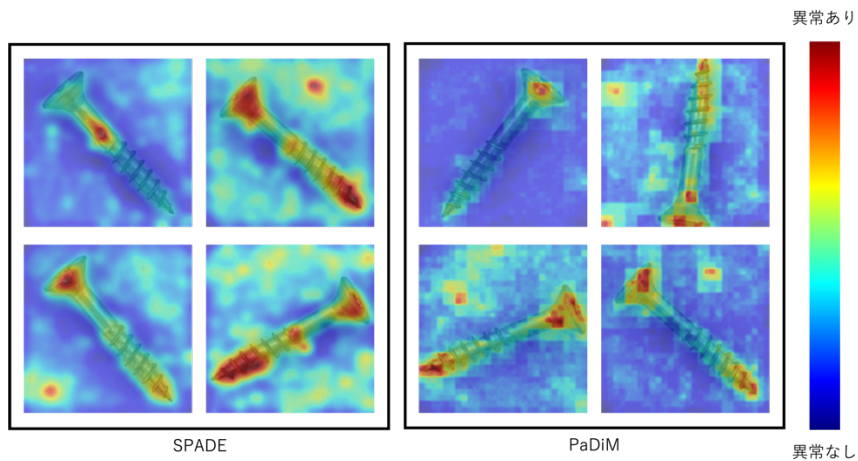


図 10 screw クラスの異常スコアマップ

3.4.2 考察

表 4 から screw クラスは従来手法の AUPR とともに MVTecAD データセットの中で最低である。そのため、考察の対象とする。screw クラスは、画像中の対象物も小さく、白い背景が画像の大半を占める。図 10 の異常スコアマップの結果から、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ものがあることがわかる。その原因の一つが画像に含まれるノイズや、白い背景に付着した塵や埃である。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の違いが結果に影響してしまっている。

3.5 むすび

本章では、SPADE と PaDiM に関する AUPR での評価を確認する予備実験を行った。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。対象画像を調査することにより背景を誤って異常と分類していることがわかった。

第4章 提案手法

4.1 まえがき

本章では、3.4 節で示した背景を誤って異常と分類する問題点を解決するため、特徴点マッチングによる対象物検出（提案手法 1）と、Wide ResNet の使用レイヤーを AUPR に最適化するため、Wide ResNet の特徴量抽出するレイヤーの変更（提案手法 2）を提案する。提案手法 1, 2 ともに従来の異常検知手法である SPADE, PaDiM を改変したものである。提案手法 1 では、対象物の範囲を推定することで、その背景の誤認識を防ぎ。提案手法 2 では、SPADE, PaDiM で特徴量抽出器として用いている Wide ResNet の使用レイヤーを AUPR に最適化し、AUPR を向上させる提案を行う。

4.2 提案手法 1：特徴点マッチングによる対象物検出

4.2.1 概要

3.4 節の従来異常検知手法の結果の可視化から、画像内の背景に異常の反応が出ている。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のみを抽出する方法を提案する。これにより、SPADE や PaDiM での異常検知で、背景に高い異常スコアがつけられてしまった場合でも、その影響を軽減することができる。

4.2.2 特徴点抽出手法の選定

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。図 11 に示すように、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。

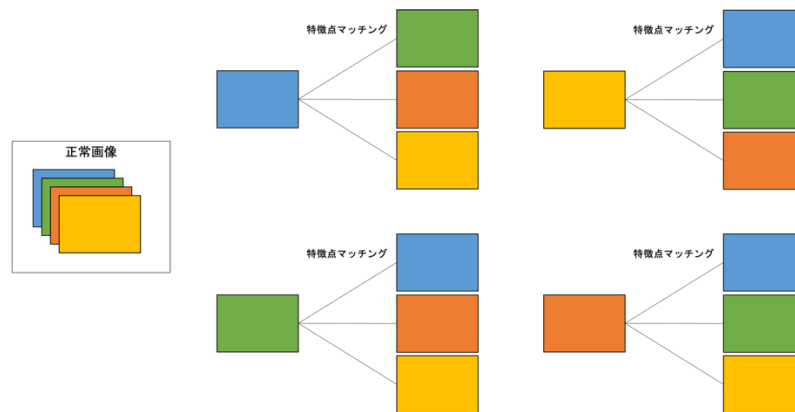


図 11 正常画像のみの特徴点マッチング

4.2.3 異常検知手法への組み込み

データセットの各対象物に対して，4.2.2 項の実験で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．その後，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．従来の異常検知手法（SPADE, 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，提案手法の出力とし，AUPR, AUROC の評価を行う．提案手法 1 の概略図を図 12 に示す．

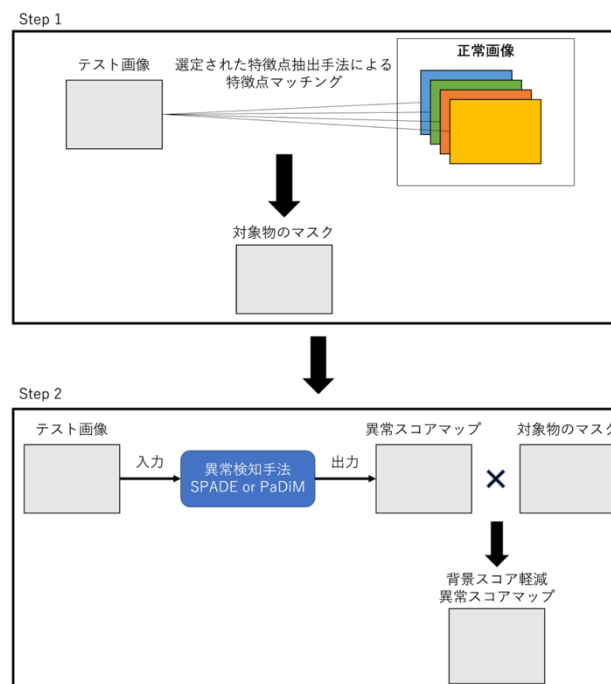


図 12 提案手法 1 の概略

4.3 提案手法 2：特徴量抽出するレイヤーの変更

4.3.1 概要

従来の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 4 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。しかし、本研究で用いるデータセット MVTEC-AD は第 3 章で述べたように AUPR で評価すべきであるため、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。

4.3.2 特徴量抽出するレイヤーの変更

SPADE, PaDiM で特徴量抽出器として用いられている Wide ResNet の概略図を図 13 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の箇所以外の処理は SPADE, PaDiM に従う。

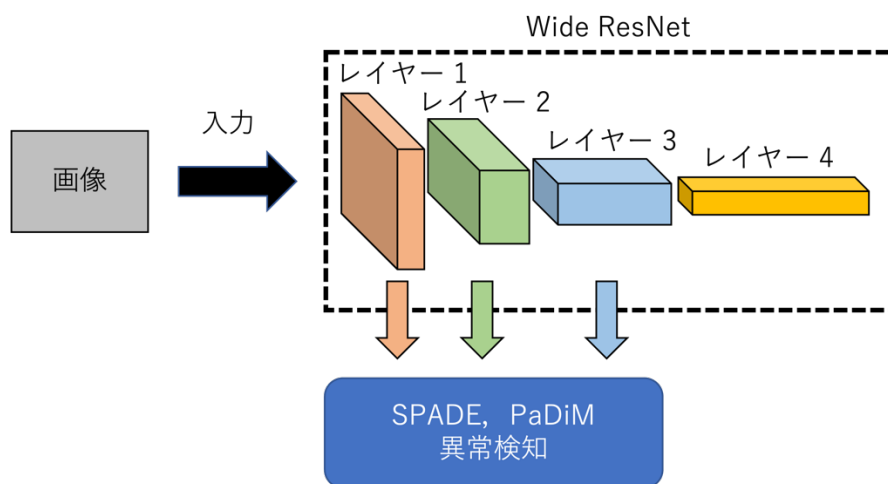


図 13 Wide ResNet のレイヤー構成

4.4 むすび

本章では、提案手法の概要について述べた。

提案手法の一つ目が特徴点マッチングによる対象物検出である。3.4 節の従来異常検知手法の結果の可視化から、画像内の背景に異常の反応が出ていた。本来、背景は異常検知する必要のない範囲である。画像から異常検出対象物のみの範囲を抽出する方法を提案した。

提案手法二つ目が Wide ResNet の特徴量抽出するレイヤーの変更である。従来の異常

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おり, 両手法共に Wide ResNet の特徴量抽出レイヤーは AUROC での評価をもとに選択されている. しかし, 本研究で用いるデータセット MVTecAD は第 3 章で述べたように AUPR で評価すべきであるため, 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した.

第5章 評価実験と考察

5.1 まえがき

本章では，第4章で提案した手法の評価実験とその結果の考察を行う．提案手法1では，クラスごとに適切な特徴点抽出手法を探し，その特徴点抽出手法を用いて対象物のマスクを作成する．その後，SPADE, PaDiM にそのマスクを組み込んだ上で，評価実験，考察を行う．提案手法2では，SPADE, PaDiM での特徴量抽出する Wide ResNet のレイヤーの変更を行い，評価実験，考察を行う．データセットは，3.2 節に記載した MVTecAD を用いる．

5.2 提案手法1：特徴点マッチングによる対象物検出

5.2.1 クラスごとの適切な特徴点抽出手法の選定

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．2.5 節で述べた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．実験及び評価はそれぞれのクラスで独立して行った．各クラスの任意の正常な画像と，その画像以外の全ての正常画像との特徴点マッチングを行う．その結果を表5に示す．ここで，評価は0~2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．

表5 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

5.2.2 対象物マスク+従来異常検知手法

表5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う．特徴マッチングによって作成したマスク外の異常スコアを0, 0.2, 0.4, 0.6, 0.8, 1.0倍

した結果の AUPR, AUROC を求めた. なお, ImageNet を学習済みの Wide ResNet を特徴量抽出器として用いた. SPADE の結果を表 6 に, PaDiM の結果を表 7 に示す. 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.

表 6 対象物マスクを導入した SPADE

マスク外 倍率	AUPR		AUROC	
	toothbrush	screw	toothbrush	screw
0	0.616	0.255	0.962	0.918
0.2	0.617	0.256	0.976	0.961
0.4	0.617	0.256	0.979	0.965
0.6	0.623	0.258	0.986	0.978
0.8	0.634	0.270	0.989	0.986
1.0	0.611	0.281	0.989	0.989

表 7 対象物マスクを導入した PaDiM

マスク外 倍率	AUPR		AUROC	
	toothbrush	screw	toothbrush	screw
0	0.542	0.209	0.961	0.913
0.2	0.543	0.209	0.976	0.956
0.4	0.545	0.211	0.981	0.968
0.6	0.552	0.216	0.987	0.978
0.8	0.543	0.231	0.988	0.982
1.0	0.502	0.235	0.987	0.981

5.2.3 考察

まず, 5.2.1 項の特徴点抽出手法の選出についてである. 表 5 に示したデータセットの種類と特徴点抽出手法の相性の結果から, 「対象物を捉えている」(評価 2) は, toothbrush と screw クラスのみだった. ここで screw と toothbrush クラスの特徴マッチングの結果を図 14 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られている. 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.

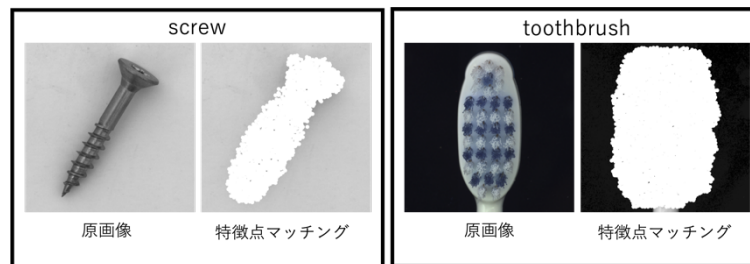


図 14 screw と toothbrush の特徴点マッチング結果

一方, 図 15 に示す pill や capsule, hazelnut クラスは画像内の外縁が丸みを帯びている, かつ物体表面の凹凸が少ないため画像内にエッジやコーナーが少ない. そのため, 抽出できる特徴点が少なく, 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. さらに, hazelnut クラスの表面の模様には個体差があり, 特徴点マッチングには不向きであった.

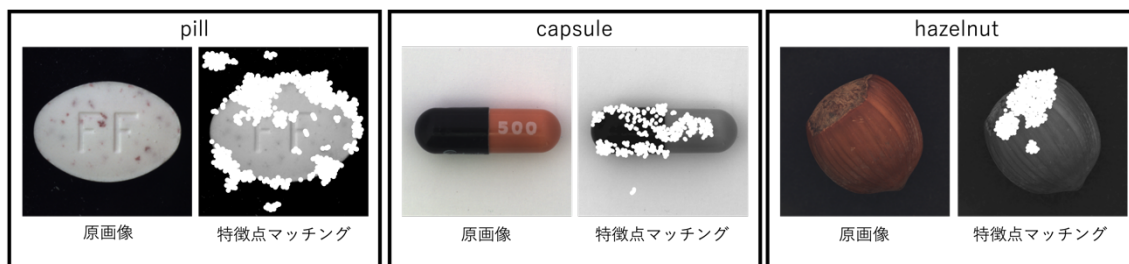


図 15 pill, capsule, hazelnut クラスの特徴点マッチング結果

次に, 5.2.2 項の対象物マスクを異常検知手法に組み込んだ結果についてである. 表 6, 表 7 から 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) としよりも toothbrush クラスの AUPR が向上している. 特に, AUROC がほぼ変化させずに AUPR のみが向上している点がこの手法の成果である. 本来, 作成した対象物マスクが完全に対象物を捉えていれば, マスク外倍率が 0, つまりマスクの中だけを評価したときの結果が一番高い AUPR のはずである. しかし, マスク外倍率が 0.6~0.8 のときに高い AUPR となった. 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. 正常画像との特徴点マッチングで対象物マスクを作成したため, 異常画像の特徴とマッチングできない部分があった. screw クラスについては, AUPR, AUROC の向上は確認できなかった. これも, 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 16 にその一例を示す.

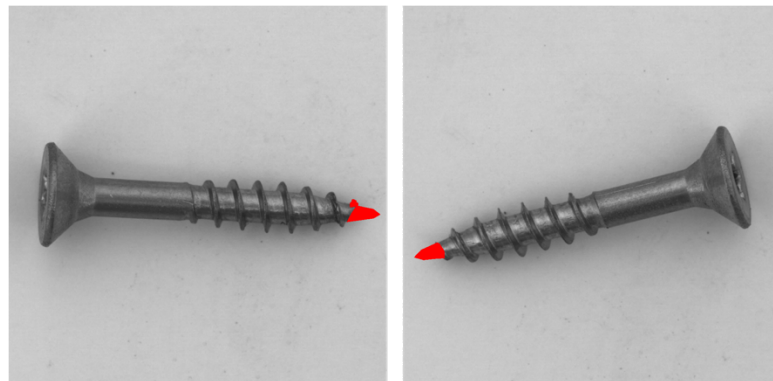


図 16 先端が欠けている screw (赤色が欠損箇所)

5.3 提案手法 2 : 特徴量抽出するレイヤーの変更

5.3.1 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 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 とともに抽出するレイヤーの変更以外は従来通りの手法である。SPADE の結果を表 8, 表 9 に, PaDiM の結果を表 10, 表 11 に示す。MVTecAD の 15 クラスの平均 AUPR を SPADE, PaDiM それぞれ表 12, 表 13 に示す。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。

表 8 特徴量抽出レイヤーによる SPADE の AUPR (クラスごと) その 1

レイヤー	bottle	cable	capsule	carpet	grid	hazelnut	leather	metal nut
1,2	0.751	0.370	0.472	0.626	0.305	0.630	0.463	0.808
1,3	0.751	0.370	0.472	0.626	0.305	0.630	0.463	0.808
1,4	0.627	0.156	0.415	0.409	0.087	0.615	0.492	0.681
2,3	0.763	0.486	0.464	0.607	0.311	0.547	0.452	0.778
2,4	0.778	0.371	0.499	0.581	0.227	0.505	0.505	0.763
3,4	0.705	0.511	0.381	0.582	0.249	0.516	0.402	0.754
1,2,3	0.768	0.385	0.495	0.631	0.336	0.604	0.486	0.801
1,2,4	0.756	0.260	0.505	0.599	0.231	0.584	0.549	0.780
1,3,4	0.751	0.370	0.472	0.626	0.305	0.630	0.463	0.808
2,3,4	0.751	0.370	0.472	0.626	0.305	0.630	0.463	0.808
1,2,3,4	0.768	0.385	0.495	0.631	0.336	0.604	0.486	0.801

表 9 特徴量抽出レイヤーによる SPADE の AUPR (クラスごと) その 2

レイヤー	pill	screw	tile	toothbrush	transistor	wood	zipper
1,2	0.518	0.261	0.494	0.620	0.719	0.436	0.502
1,3	0.518	0.261	0.494	0.620	0.719	0.436	0.502
1,4	0.397	0.104	0.267	0.556	0.331	0.455	0.489
2,3	0.434	0.269	0.471	0.595	0.736	0.433	0.507
2,4	0.337	0.253	0.349	0.571	0.610	0.459	0.591
3,4	0.451	0.144	0.500	0.570	0.765	0.380	0.381
1,2,3	0.455	0.281	0.469	0.611	0.691	0.461	0.559
1,2,4	0.373	0.227	0.345	0.586	0.526	0.489	0.604
1,3,4	0.518	0.261	0.494	0.620	0.719	0.436	0.502
2,3,4	0.518	0.261	0.494	0.620	0.719	0.436	0.502
1,2,3,4	0.455	0.281	0.469	0.611	0.691	0.461	0.559

表 10 特徴量抽出レイヤーによる PaDiM の AUPR (クラスごと) その 1

レイヤー	bottle	cable	capsule	carpet	grid	hazelnut	leather	metal nut
1,2	0.759	0.314	0.465	0.570	0.365	0.511	0.540	0.728
1,3	0.686	0.491	0.395	0.522	0.234	0.537	0.360	0.713
1,4	0.530	0.313	0.406	0.277	0.088	0.400	0.142	0.584
2,3	0.714	0.498	0.425	0.562	0.258	0.483	0.411	0.737
2,4	0.639	0.353	0.427	0.338	0.146	0.439	0.161	0.734
3,4	0.759	0.314	0.465	0.570	0.365	0.511	0.540	0.728
1,2,3	0.727	0.455	0.449	0.600	0.283	0.507	0.360	0.744
1,2,4	0.681	0.307	0.450	0.322	0.136	0.462	0.159	0.772
1,3,4	0.603	0.404	0.373	0.360	0.217	0.449	0.199	0.720
2,3,4	0.656	0.394	0.392	0.422	0.202	0.527	0.226	0.747
1,2,3,4	0.685	0.464	0.448	0.374	0.196	0.473	0.221	0.805

表 11 特徴量抽出レイヤーによる PaDiM の AUPR (クラスごと) その 2

レイヤー	pill	screw	tile	toothbrush	transistor	wood	zipper
1,2	0.424	0.219	0.427	0.496	0.568	0.507	0.509
1,3	0.470	0.158	0.514	0.482	0.723	0.411	0.360
1,4	0.563	0.041	0.418	0.316	0.513	0.224	0.284
2,3	0.428	0.207	0.509	0.528	0.711	0.415	0.411
2,4	0.540	0.109	0.438	0.462	0.549	0.255	0.368
3,4	0.424	0.219	0.427	0.496	0.568	0.507	0.509
1,2,3	0.415	0.216	0.551	0.491	0.716	0.420	0.427
1,2,4	0.501	0.130	0.429	0.504	0.559	0.272	0.400
1,3,4	0.488	0.114	0.471	0.516	0.652	0.282	0.343
2,3,4	0.494	0.154	0.514	0.516	0.685	0.311	0.379
1,2,3,4	0.468	0.143	0.482	0.554	0.665	0.313	0.414

表 12 特徴量抽出レイヤーによる SPADE の AUPR (15 クラス平均)

レイヤー	平均 AUPR
12	0.532
13	0.532
14	0.406
23	0.523
24	0.493
34	0.486
123	0.536
124	0.494
134	0.532
234	0.532
1234	0.536

表 13 特徴量抽出レイヤーによる PaDiM の AUPR (15 クラス平均)

レイヤー	平均 AUPR
12	0.493
13	0.470
14	0.340
23	0.487
24	0.397
34	0.493
123	0.491
124	0.406
134	0.413
234	0.441
1234	0.447

5.3.2 考察

表 12, 表 13 にまとめた MVTecAD の 15 クラスの平均 AUPR をグラフにまとめたものを図 17 に示す. SPADE と PaDiM 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. 「レイヤー1,4」で大きく AUPR が低下している. 「レイヤー1,4」は Wide ResNet の最初と最後のレイヤーとなる. レイヤー4 の有無による AUPR の違いを表 14 に示す. 表 14 からレイヤー4 が AUPR を下げていることがわかる. 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.

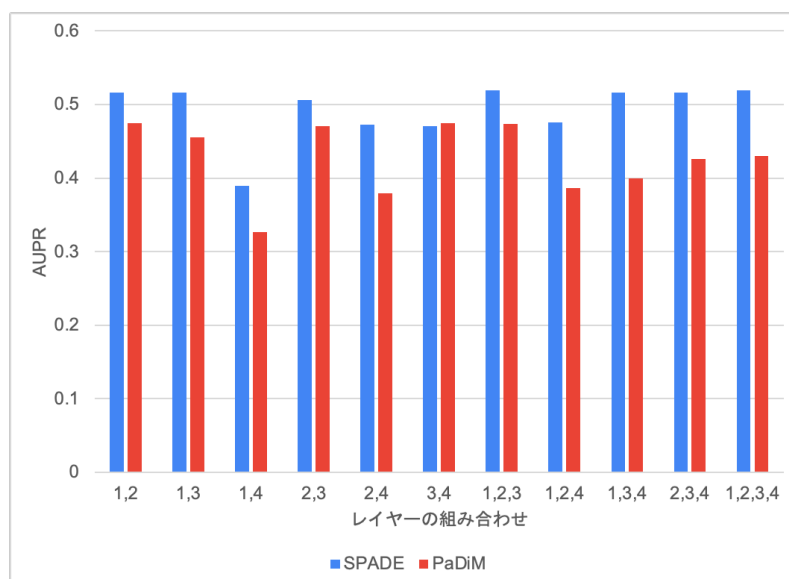


図 17 特徴量抽出レイヤーによる AUPR の変化

表 14 レイヤー4の有無と AUPR

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

表 8, 表 9, 表 10, 表 11 に示したクラスごとの AUPR から, 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多い. クラスごとに適切なレイヤーの組み合わせは異なるが, AUPR で評価する際は, レイヤーは 2 種類で良いとわかった. また, 少ないレイヤーであればその分計算量も削減できることも利点である.

5.4 むすび

本章では, 提案手法の有効性を確認するための評価実験の結果を示し, それについての考察を述べた. 提案手法 1 の特徴点マッチングによる対象物検出では, screw クラスと toothbrush クラスで有効であり, 従来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.

第6章 結論と今後の課題

6.1 結論

本研究では、異常検知手法である SPADE と PaDiM に対して、AUPR の向上のため二つの提案を行った。

一つ目は、特徴点マッチングによって対象物のマスクを作成し、SPADE, PaDiM で推定した異常スコアマップに掛け合わせる手法である。第3章の予備実験で示した SPADE, PaDiM が背景を誤って異常と分類する問題点を解決するため、画像から異常検出対象物のみの範囲を抽出する方法を提案、実験した。その結果、MVTecAD の toothbrush クラスで AUPR の向上を確認した。

二つ目は、SPADE, PaDiM で特徴量抽出器として用いる Wide ResNet の抽出するレイヤーを変更する手法である。クラス別では、従来手法よりも高い AUPR も確認された。さらに、異常検知に有効でないレイヤーや、AUPR を下げることなく計算量を減少させることができることがわかった。

以上より、特徴点マッチングによる対象物マスクの導入と特徴量抽出レイヤーの変更という提案手法により、従来の異常検知手法 SPADE, PaDiM の性能を改善できることがわかった。

6.2 今後の課題

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った。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った。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できることが望ましい。

謝辞

本研究のテーマや方向性，研究の進め方，細かい知識や技術など多くのご指導いただきました渡辺裕教授と早稲田大学国際情報通信研究センター招聘研究員の原潤一博士に深く感謝申し上げます。

また，日頃から問題解決の相談に乗っていただき，快適な研究環境を作ってくださいました研究室の皆様に熱く御礼申し上げます。

最後に，常に支えてくれた家族に心より感謝します。

参考文献

- [1]. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei: "ImageNet: A Large-Scale Hierarchical Image Database", in CVPR, June 2009.
- [2]. K. He, X. Zhang, S. Ren, and J. Sun: "Deep residual learning for image recognition", In CVPR, June 2016.
- [3]. N. Cohen and Y. Hoshen: "Sub-Image Anomaly Detection with Deep Pyramid Correspondences", arXiv preprint, arXiv:2005.02357, 2020.
- [4]. T. Defard, A. Setkov, A. Loesch, and R. Audigier: "PaDiM: a Patch Distribution Modeling Framework for Anomaly Detection and Localization", arXiv preprint, arXiv:2011.08785, 2020.
- [5]. S. Zagoruyko and N. Komodakis: "Wide Residual Network", arXiv preprint, arXiv:1605.07146, 2016.
- [6]. 斎藤康毅: "ゼロから作る Deep Learning—Python で学ぶディープラーニングの理論と実装", 株式会社オライリー・ジャパン, pp.205-239, 2016.
- [7]. L. Bergman and Y. Hoshen: "Classification-based anomaly detection for general data", In ICLR, Apr. 2020.
- [8]. P. Mahalanobis: "On the generalized distance in statistics", in National Institute of Science of India, 1936.
- [9]. OpenCV, 2021, <https://opencv.org/>
- [10]. P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger: "Mvttec ad-a comprehensive real-world dataset for unsupervised anomaly detection," in CVPR, June 2019.
- [11]. P. Bergmann, K. Batzner, M. Fauser, D. Sattlegger, and C. Steger: "The MVTEC Anomaly Detection Dataset: A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection", in International Journal of Computer Vision, Jan. 2021.

図一覧

図 1	Shortcut connection を含んだ残差ブロック	3
図 2	SPADE の学習過程.....	4
図 3	SPADE のテスト過程.....	5
図 4	PaDiM の学習過程.....	6
図 5	PaDiM のテスト過程.....	7
図 6	ROC 曲線の例	7
図 7	PR 曲線の例.....	8
図 8	MVTecAD の画像例	9
図 9	MVTecAD のクラスごとの AUPR.....	12
図 10	screw クラスの異常スコアマップ	13
図 11	正常画像のみの特徴点マッチング	15
図 12	提案手法 1 の概略	15
図 13	Wide ResNet のレイヤー構成	16
図 14	screw と toothbrush の特徴点マッチング結果.....	20
図 15	pill, capsule, hazelnut クラスの特徴点マッチング結果.....	20
図 16	先端が欠けている screw (赤色が欠損箇所)	21
図 17	特徴量抽出レイヤーによる AUPR の変化.....	25

表一覧

表 1	SPADE の AUROC と AUPR.....	10
表 2	PaDiM の AUROC と AUPR.....	11
表 3	SPADE と PaDiM の 15 クラスの平均値.....	11
表 4	MVTecAD のクラスごとの異常画素率と AUPR.....	12
表 5	クラスと特徴点抽出手法の相性.....	18
表 6	対象物マスクを導入した SPADE.....	19
表 7	対象物マスクを導入した PaDiM.....	19
表 8	特徴量抽出レイヤーによる SPADE の AUPR (クラスごと) その 1.....	21
表 9	特徴量抽出レイヤーによる SPADE の AUPR (クラスごと) その 2.....	22
表 10	特徴量抽出レイヤーによる PaDiM の AUPR (クラスごと) その 1.....	22
表 11	特徴量抽出レイヤーによる PaDiM の AUPR (クラスごと) その 2.....	23
表 12	特徴量抽出レイヤーによる SPADE の AUPR (15 クラス平均).....	23
表 13	特徴量抽出レイヤーによる PaDiM の AUPR (15 クラス平均).....	24
表 14	レイヤー4 の有無と AUPR.....	25

研究業績

- [1]. 山野井祐介, 原潤一, 渡辺裕: “正常画像を用いた異常画像検知手法の検討”, 情報処理学会オーディオビジュアル複合情報処理研究発表会, 2022. (2022年2月掲載予定)