

卒業論文概要書

Summary of Bachelor's Thesis

Date of submission: 02/01/2025 (MM/DD/YYYY)

学科名 Department	情報通信	氏名 Name	井田 陽向	指導 教員 Advisor	渡辺 裕 印
研究指導名 Research guidance	オーディオビジュアル 情報処理研究	学籍番号 Student ID number	1w213016-4		
研究題目 Title	FractalDB を基盤とした画像再構成手法による異常検知モデルの改善 Improving Anomaly Detection Models through the Adaptation of Image Reconstruction Methods Based on FractalDB				

1. まえがき

異常検知は工業製品の品質管理において重要な手法であり、多くの最先端手法は大規模な自然画像データセットである ImageNet で事前学習された深層特徴抽出器に依存している。しかし、ImageNet は著作権やプライバシーに関する制約から非商用利用に限られる。この課題に対し、数式から自動生成されるフラクタル画像[1]を利用するアプローチが提案されている。フラクタル画像は著作権やプライバシーの問題を回避可能である一方、既存研究[2]において自然画像に比べて特徴抽出性能が低いという課題がある。

本研究では、画像再構成手法を活用し、MVTec データセット[3]を学習することで特徴抽出性能を向上させる手法を提案する。この手法により、商用利用における制約の回避と異常検知精度の向上を両立できることを示す。

2. 関連研究

2.1 FractalDB

FractalDB[1]は、1000 カテゴリで構成され、各カテゴリには 1000 枚のフラクタル画像が含まれる。これらのフラクタル画像は、Iterated Function System (IFS)を用いて数式からレンダリングされ、

IFS のパラメータに基づいてカテゴリが分類される。

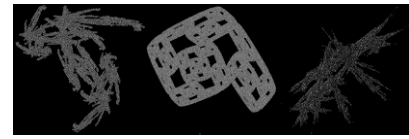


図 1 FractalDB[1] の画像例

2.2 PaDiM

PaDiM[4]は、異常検知手法の一つである。ResNet などの学習済みモデルに正常画像を入力して特徴抽出し、複数の層の特徴をチャンネル方向に結合することで特徴ベクトルを作成する。その後、パッチごとに共分散行列と平均を計算し、正常分布を定義する。推論時には、テスト画像に対して同様の手順で特徴ベクトルを計算し、正常画像分布との乖離度を基に異常の有無を検出する。

3. 従来研究とその問題点

従来研究[2]は、FractalDB および ImageNet で事前学習した異常検知モデルの性能を初めて系統的に分析したものであり、WideResNet50 を事前学習済みの特徴抽出器として使用し精度比較を行っている。さらに、FractalDB で事前学習した WideResNet を異常検知タスクに適用する際、低次元の特徴量がより効果的であると指摘し、高次元

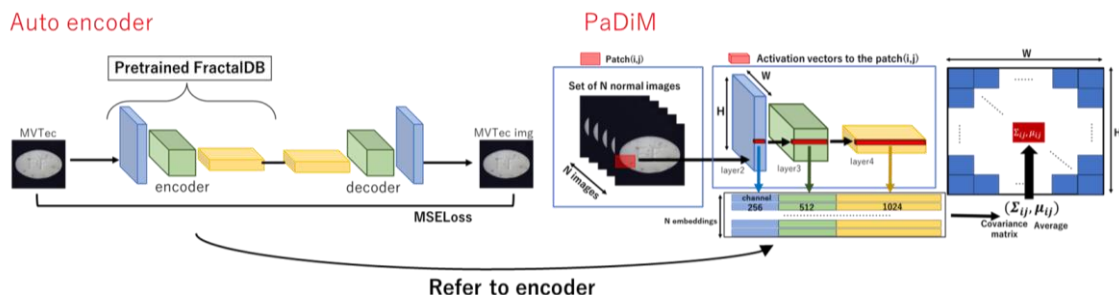


図 2 FractalDB を基盤とした画像再構成手法による異常検知モデル

表 1 Mvtec データセットの各カテゴリに対する PaDiM の精度比較 (img_roc, pix_roc)

	Layer	Amount	method	Bottle	Capusle	Metalnut	Grid	Haelnut	Tile	Average
従来手法	2,3,4	1792	ALL	0.990/0.960	0.795/0.953	0.813/0.919	0.836 /0.908	<u>0.953</u> /0.983	<u>0.930</u> /0.829	0.879/0.920
	2,3,4	550	Random	0.990/0.960	0.798/0.952	0.813/0.916	0.813/0.887	0.956 /0.982	0.916/ <u>0.833</u>	0.873/0.916
	2,3	550	Random	0.991/0.971	0.815/ <u>0.966</u>	0.765/0.931	0.825/ <u>0.926</u>	0.956 / 0.985	0.935 /0.797	0.866/ 0.932
提案手法	2,3,4	1792	All	1.000 / <u>0.974</u>	0.881 / <u>0.966</u>	<u>0.864</u> / <u>0.945</u>	<u>0.826</u> /0.867	0.945/0.965	0.877/ 0.837	<u>0.882</u> /0.924
	2,3,4	550	Random	1.000 /0.973	<u>0.874</u> /0.963	0.873 /0.941	0.812/0.869	0.941/0.962	0.868/0.831	0.878/0.917
	2,3	550	Random	<u>0.998</u> / 0.984	0.864/ 0.970	0.851/ 0.949	0.782/ 0.919	0.937/0.971	0.881/0.719	0.887 / <u>0.931</u>

の特徴量を用いずに実験を実施している。しかし、この研究は性能比較を主な目的としており、事前学習対象の変更に残っているため、FractalDBを活用したモデルの精度向上に関しては、さらなる改良の余地が大きく残されている。

4. 提案手法

本研究では、画像再構成手法を用いて FractalDB で事前学習した異常検知モデルの特徴抽出性能を改善する手法を提案する。図 2 に示す提案手法では、MVTec データセットの正常画像を入力として、AutoEncoder を構築し、Encoder 部分に FractalDB で事前学習したモデルを使用する。平均二乗誤差を損失関数として学習後、Encoder 部分のみを PaDiM に転用することで、正常画像の正確な特徴量分布を策定し、異常画像の特徴抽出性能が向上する。

5. 実験

5.1 AutoEncoder の出力結果

MvtTec データセットを用いて AutoEncoder モデルの学習を行った結果を図 3 に示す。図 3 より、入出力の差分が少ないことから、画像再構成手法の有効性が定性的に分かる。

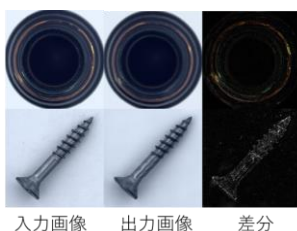


図 3 AutoEncoder の出力結果例

5.2 PaDiM の精度比較

従来手法と提案手法による PaDiM の異常検知結果を表 1 に示す。評価指標として、画像レベル、ピクセルレベルの AUROC を使用する。表 1 では、15 カテゴリのうち最も精度改善された 3 カテゴリと、改善しなかった 3 カテゴリを記載している。表 1 より、全カテゴリの平均精度において、提案手法が従来手法を上回ることを示した。

6. 結論

本研究では、FractalDB を基盤とし、画像再構成手法を使用した異常検知モデルを提案した。本提案により、商用利用における制約の回避と異常検知精度の向上の両立を示した。

参考文献

- [1] H. Kataoka et al., “Pre-training without Natural Images,” 15th Asian Conference on Computer Vision (ACCV), Nov. 2024.
- [2] C. I. Ugwu, S. Casarin et al., “Fractal as pre-training Datasets for Anomaly Detection and Localization,” 2024 IEEE CVPR workshop on fair, data-efficient, and Trusted Computer Vision (TCV2024), Jun. 2024.
- [3] T. Lin et al., “MVTec AD - A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9584-9592, Jun. 2019.
- [4] T. Defard et al., “PaDiM: a Patch Distribution Modeling Framework for Anomaly Detection and Localization,” 25th International Conference on Pattern Recognition (ICPR), Jan. 2021.

2024 年度 卒業論文

FractalDB を基盤とした画像再構成手法による
異常検知モデルの改善

Improving Anomaly Detection Models through the Adaptation of
Image Reconstruction Methods Based on FractalDB

指導教員 渡辺 裕 教授

提出日：2025 年 2 月 1 日

早稲田大学 基幹理工学部 情報通信学科

1W213016 - 4

井田 陽向

目次

第1章	序論.....	3
1.1	研究背景	3
1.2	関連研究と問題点, および研究目的	3
1.3	本論文の構成	4
第2章	関連技術.....	5
2.1	まえがき	5
2.2	FlactalDB	5
2.3	PaDiM	5
2.4	異常検知タスクへのフラクタル画像の適応	6
2.5	むすび	7
第3章	提案手法.....	8
3.1	まえがき	8
3.2	提案手法	8
3.3	むすび	9
第4章	実験・実験結果.....	10
4.1	まえがき	10
4.2	データセット	10
4.3	評価方法	10
4.4	実験結果	10
4.4.1	AutoEncoder の出力結果	10
4.4.2	PaDiM の精度比較	11
4.5	考察	12
4.6	むすび	13
第5章	結論と今後の課題.....	14
5.1	結論	14
5.2	今後の課題	14
	謝辞.....	15

参考文献.....	16
圖一覽.....	17
表一覽.....	18

第1章 序論

1.1 研究背景

異常検知[1]は、正常なパターンから逸脱した異常なパターンを検出する技術であり、産業製品の品質管理において、不良品の検出や不良個所の特定に必要な不可欠な手法である。また、他にも医療疾患スクリーニング[2]やビデオ監視[3]などの多くの分野において注目されている。現在、多くの最先端手法は、高水準な異常検知精度を誇っているが、そのほとんどが大規模な自然画像データセットである ImageNet[4]で事前学習された特徴抽出器に依存している。ImageNet は公開されている画像の著作権を有しておらず、またプライバシーなどの配慮に関しても懸念されており、その利用は非商用目的に限定されている。

以上のことから、異常検知タスクを商用利用する際に、著作権やプライバシーの制約を回避しながら学習する仕組みが重要視されている。

1.2 関連研究と問題点、および研究目的

著作権やプライバシーの課題を回避する手段として、事前学習において自然画像の代わりにフラクタル画像[5]を使用する方法が提案されている。フラクタル画像は、数式によって生成された複雑な幾何構造で構成されており、倫理的・偏見的な懸念がなく、誰でもオープンソースで画像を作成できるため、上記の問題に対する解決策として注目されている。また、フラクタル画像を事前学習したモデルは追加学習を行うことで様々な認識タスクに応用でき、ImageNet と同等の性能を有している。しかし、教師無し学習が一般的な異常検知タスクにおいて、フラクタル画像を事前学習に用いた手法[6]では ImageNet を用いた場合に比べて精度が大きく劣っており、実社会での商用利用においては、いまだ実用性が低いという現状がある。

本研究では、この課題に対処するため、MVTec データセット[7]を用いた画像再構成手法の学習により特徴抽出性能を向上させる手法を提案する。異常検知対象である産業用データセットは商用利用時に独自に準備することが一般的であり、著作権やプライバシーの制約に縛られない。この手法により、商用利用における制約の回避と異常検知精度の改善を両立できることを示す。

1.3 本論文の構成

本論文の構成を以下に示す.

- 第1章 本章であり, 本研究の背景, 関連研究と問題点および研究目的について述べる.
- 第2章 本研究で用いる異常検知手法および関連技術について述べる.
- 第3章 本研究の提案手法について述べる.
- 第4章 本研究における実験の方法, 結果および考察について述べる.
- 第5章 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

本章では自動生成画像の手法として用いられている FractalDB[5]について述べる．また，本研究で使用した異常検知手法の一つである PaDiM[8]について述べる．加えて，PaDiM にフラクタル画像を適応する手法[6]についても述べる．

2.2 FractalDB

FractalDB は，1000 カテゴリで構成され，各カテゴリには 1000 枚のフラクタル画像が含まれる．フラクタル画像の一例をエラー! 参照元が見つかりません。に示す．これらのフラクタル画像は，Iterated Function System (IFS)を用いて数式からレンダリングされ，IFS のパラメータに基づいてカテゴリが分類される．

生成されたフラクタル画像は，パラメータを少し変えるだけで大きく異なるため，データの多様性を確保することが出来る．また，FractalDB は自然画像や合成画像と酷似していない．

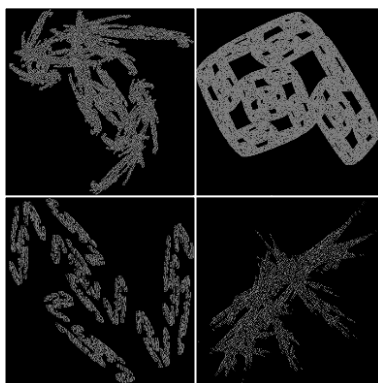


図 2.1 フラクタル画像の例 9

2.3 PaDiM

PaDiM は教師無し学習の異常検知手法の一種であり，正常画像と異常画像の特徴分布の違いを活用して異常を検知するアルゴリズムである．学習時のモデル構造をエラー! 参照元が見つかりません。に，推論時のモデル構造を図 2.3 に示す．PaDiM は画像単位で異常を判定する手法（Anomaly Detection）と，ピクセル単位で異常を判定する手法（Anomaly Localization）に大別され，両方を主目的として研究されている．

学習時は，正常画像のみを用いる．ResNet などの学習済みモデルを使用して特徴を抽出した後，複数の層の特徴量をチャンネル方向に結合し，特徴ベクトルを作成する．その後パッチごとに特徴ベクトルを基に共分散行列と平均を求め，正常画像の分布を定める．

推論時は、テスト用画像を学習時と同様に特徴ベクトルに変換し、正常画像分布との乖離度合いによって異常画像を検知している．この乖離度合いはマハラノビス距離[10]と呼ばれ、距離が離れているほど異常である可能性が高いと判断される．最終的に、すべてのパッチの距離情報を基に、異常スコアマップを生成し、異常個所を特定する．

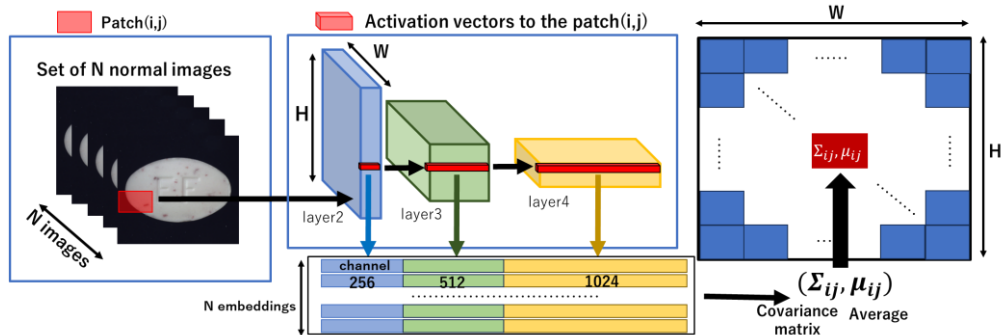


図 2.2 PaDiM における学習時のモデル概要

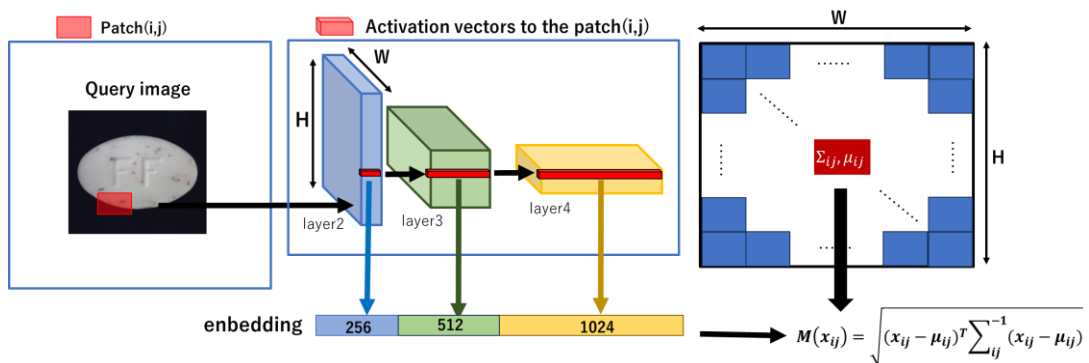


図 2.3 PaDiM における推論時のモデル概要

2.4 異常検知タスクへのフラクタル画像の適応

フラクタル画像で事前学習したモデルは分類タスクなどの他にも、異常検知タスクに対しても利活用が研究されている．この分野の従来研究[6]は、FractalDB および ImageNet で事前学習した異常検知モデルの性能を初めて系統的に分析したものであり、WideResNet50[11]を事前学習済みの特徴抽出器として使用し精度比較を行っている．従来研究のモデル概要を図 2.4 に示す．さらに、FractalDB で事前学習した WideResNet50 を異常検知タスクに適用する際、低次元の特徴量がより効果的であると指摘し、高次元の特徴量を用いない方法と比較実験を行っている．しかし、この研究は性能比較を主な

目的としており，提案は事前学習モデルの変更に留まっているため，FractalDB を活用したモデルの精度改善に関しては，さらなる改良の余地が大きく残されている．

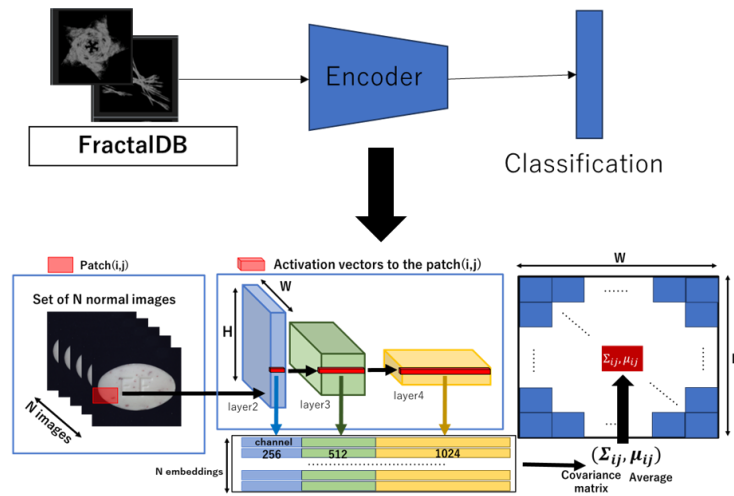


図 2.4 従来手法のモデル概要

2.5 むすび

本章では，自動生成画像として生成されるフラクタル画像について述べた．また，異常検知の一種である PaDiM についても述べた．加えて、フラクタル画像を PaDiM に適応した手法についても述べた．

第3章 提案手法

3.1 まえがき

本章では，画像再構成手法を用いて FractalDB で事前学習した異常検知モデルの特徴抽出性能を改善する手法を提案する．

3.2 提案手法

提案手法の全体的な概要を図 3.1 に示す．本研究では，異常検知性能を向上させるため，MVTec データセットの正常画像のみを入力とした AutoEncoder を構築した．フラクタル画像における先行研究[5]では，Cifar10 などのデータセットをファインチューニングすることで ImageNet と同等の精度を誇る分類モデルを実現したが，異常検知タスクにおける先行研究[6]では追加学習が成されてない．そのため，画像再構成を用いた追加学習を提案する．Encoder 部分に FractalDB で事前学習したモデルを使用する．学習においては，入力画像と出力画像の平均二乗誤差(MSE)を損失関数として設定し，Encoder 部分のみを PaDiM に転用する．それにより，正常画像の正確な特徴量分布の策定や，推論時の異常画像の特徴抽出性能が向上する．

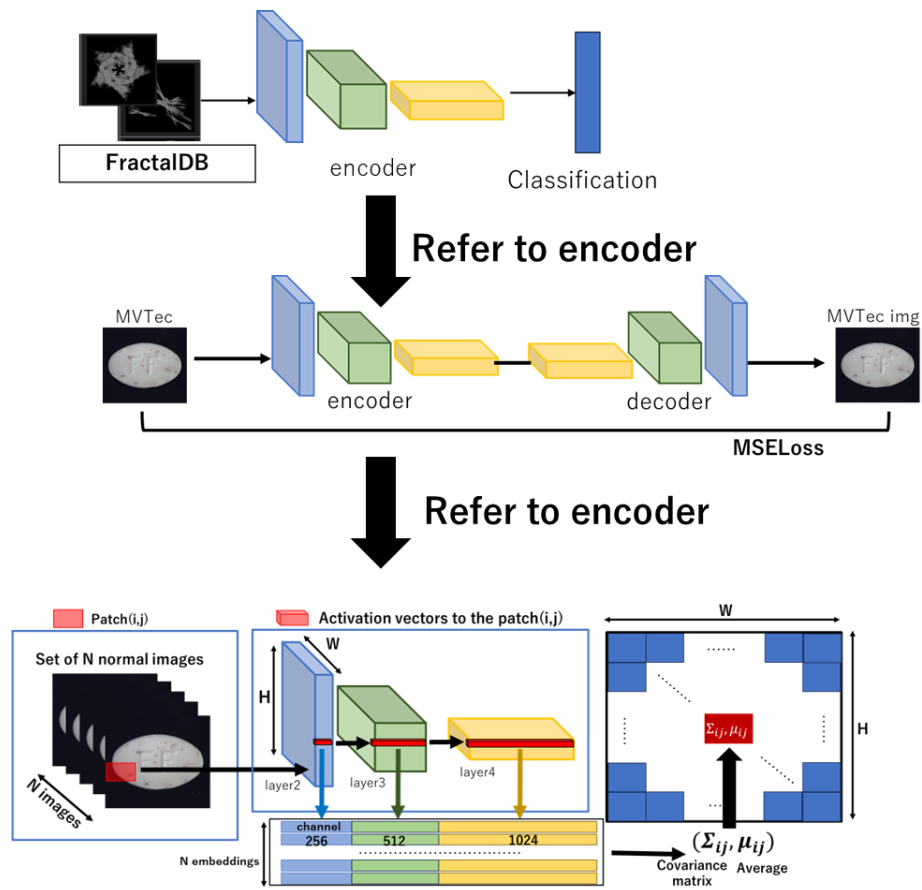


図 3.1 FractalDB を基盤とした画像再構成手法による異常検知モデル

3.3 むすび

本章では, 本研究の提案手法である画像再構成手法を用いた異常検知モデルの精度向上について述べた.

第4章 実験・実験結果

4.1 まえがき

本章では、本研究で用いたデータセット、実験方法、実験結果、考察について述べる。具体的には、まず本研究の評価時の入力として用いたデータセット、評価方法、提案手法の性能について述べる。その後、実際の処理についての定性的な評価と、結果に対する考察を述べる。

4.2 データセット

オートエンコーダーの入力画像として、MVTec データセットを使用した。この素材は、<https://www.mvtec.com/company/research/datasets/mvtec-ad> から入手できる。MVTec データセットは、工業製品に焦点を当てた異常検知手法のベンチマークのためのデータセットである。5000 を超える高解像度の画像が含まれており、15 種類のカテゴリに分類されている。各カテゴリは正常画像と、様々なパターンの異常画像で構成されている。

4.3 評価方法

本研究では、評価指標として画像単位およびピクセル単位の AUROC (Area Under the Receiver Operating Characteristic curve) を使用している。AUROC は、ROC 曲線の下面積を示す指標である。ROC 曲線とは正常画像のうち正常と予測された割合を縦軸、異常画像のうち異常と予測された割合を横軸としプロットした場合の曲線を表現したものである。そのため、モデルの性能は 0 から 1 の間で定量的に評価され、値が大きいほど分類精度が良いことを示す。

画像単位の AUROC (`img_roc`) は、画像を正常画像と異常画像に分類する能力を示し、モデル全体的な分類性能を評価するために用いる。ピクセル単位の AUROC (`pix_roc`) は、正常領域と異常領域を正確に識別する能力を示し、画像内の詳細な異常検知性能を評価するために用いる。

4.4 実験結果

4.4.1 AutoEncoder の出力結果

MVTec データセットを用いて AutoEncoder モデルの学習を行った結果を図 4.1 に示す。図 4.1 は、左から入力画像、出力画像、差分を表している。それぞれのカテゴリにおいて、入出力画像の差分が小さいことから、画像再構成手法が正常画像の特徴を適切

に学習し，高い再構成性能を発揮していることがわかる．

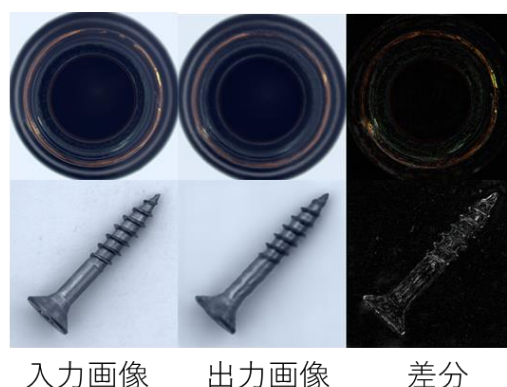


図 4.1 MVTec データセットに対する AutoEncoder モデルの出力結果

4.4.2 PaDiM の精度比較

従来手法と提案手法による PaDiM の異常検知結果の定量的評価を表 4.1 に，定性的な評価を図 4.2 に示す．共分散行列と平均値を算出するために用いる特徴量は従来手法を参考に，以下の 3 パターンで実験した．また，特徴量をランダムに取得する方法では，各 5 回実験し，精度を平均したものを最終的な定量的結果とした．

- ALL:

2 層目から 4 層目の全ての特徴量 (1792 チャンネル) を選択

- Random 550:

2 層目から 4 層目の全ての特徴量 (1792 チャンネル) からランダムに 550 チャンネル選択

- Layer 2,3:

2 層目と 3 層目の全ての特徴量 (768 チャンネル) からランダムに 550 チャンネル選択

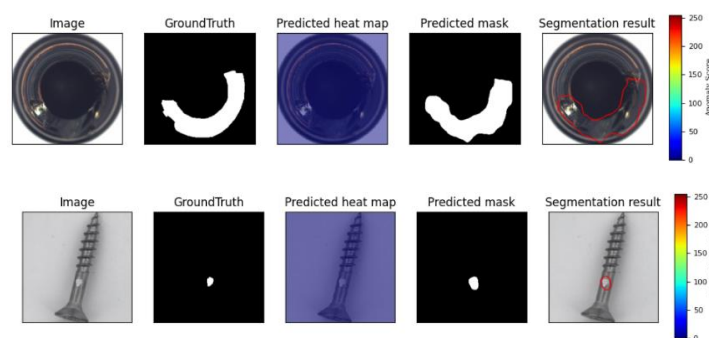


図 4.2 PaDiM の定性的評価

表 4.1 PaDiM の定量的評価 (img_roc, pix_roc)

image/pixel	bottle	cable	capsule	carpet	grid	hazelnut	leather	metal_nut
従手法								
All	0.990 / 0.960	<u>0.883</u> / 0.934	0.795 / 0.953	<u>0.827</u> / 0.923	0.836 / 0.908	<u>0.953</u> / <u>0.983</u>	1.000 / 0.980	0.813 / 0.919
Random550	0.990 / 0.960	0.887 / <u>0.933</u>	0.798 / 0.952	0.814 / 0.917	0.813 / 0.887	0.956 / 0.982	1.000 / 0.978	0.813 / 0.916
Layer 2,3	0.991 / 0.971	0.816 / 0.930	0.815 / <u>0.966</u>	0.825 / 0.951	0.825 / 0.926	0.956 / 0.985	1.000 / <u>0.992</u>	0.765 / 0.931
提案手法								
All	1.000 / <u>0.974</u>	0.806 / 0.929	0.881 / <u>0.966</u>	0.815 / 0.899	<u>0.826</u> / 0.867	0.945 / 0.965	0.993 / 0.951	<u>0.864</u> / <u>0.945</u>
Random550	1.000 / 0.973	0.802 / 0.927	<u>0.874</u> / 0.963	0.811 / 0.894	0.717 / 0.847	0.941 / 0.962	1.000 / 0.886	0.873 / 0.941
Layer 2,3	<u>0.998</u> / 0.984	0.827 / 0.932	0.864 / 0.970	0.961 / <u>0.941</u>	0.782 / <u>0.919</u>	0.937 / 0.971	<u>0.999</u> / 0.994	0.851 / 0.949

image/pixel	pill	screw	tile	toothbrush	transistor	wood	zipper	average
従手法								
All	0.701 / 0.743	<u>0.645</u> / <u>0.975</u>	<u>0.930</u> / 0.829	1.000 / <u>0.964</u>	0.941 / 0.969	<u>0.963</u> / 0.799	0.911 / 0.960	0.879 / 0.920
Random550	0.706 / 0.740	0.609 / 0.968	0.916 / <u>0.833</u>	<u>0.994</u> / 0.961	0.943 / 0.969	<u>0.963</u> / 0.798	0.902 / 0.952	0.873 / 0.916
Layer 2,3	0.679 / 0.804	0.594 / 0.971	0.935 / 0.797	1.000 / 0.971	0.932 / 0.954	0.957 / <u>0.847</u>	0.905 / 0.977	0.866 / 0.932
提案手法								
All	<u>0.717</u> / <u>0.875</u>	0.665 / 0.976	0.877 / 0.837	1.000 / 0.901	0.968 / 0.979	0.952 / 0.835	<u>0.936</u> / 0.968	<u>0.883</u> / 0.924
Random550	<u>0.717</u> / 0.869	0.622 / 0.969	0.868 / 0.831	1.000 / 0.898	0.968 / <u>0.978</u>	0.953 / 0.835	0.932 / 0.966	0.872 / 0.916
Layer 2,3	0.728 / 0.882	0.621 / 0.970	0.881 / 0.719	<u>0.994</u> / 0.939	<u>0.950</u> / 0.951	0.970 / 0.876	0.949 / <u>0.971</u>	0.887 / <u>0.931</u>

4.5 考察

従来手法と提案手法それぞれで、同じ特徴量の取り方をしている手法を比べると、平均精度がほぼ全ての項目で改善されていることが表 4.1 から読み取れる。定量的評価のうち、img_roc については全カテゴリの平均値で、従来手法の 87.9% に比べて提案手法では 88.3% と 0.4 ポイント改善している。また Layer2 と 3 での検出結果では、2.1 ポイントの改善となっている。これより、画像再構成手法の導入により、フラクタル画像で事前学習されたモデルの一定の追加学習が特徴抽出性能を改善したといえる。一方で、詳細に比較すると、提案手法より従来手法の方が精度を上回るカテゴリも散見される。これは、MVTec データセットが、傷やシミなどの局所的異常や、位置のずれや欠損などの論理的異常など、多様な種類の異常画像で構成されており、カテゴリ毎にその性質が異なるのが要因の一つであると考えられる。

また、表 4.1 より、提案手法の中で、Layer2, 3 の特徴量のみを用いる手法が全体的に最も精度を改善していることから、従来研究で指摘されている通り、浅い層の特徴量の方が異常検知タスクの特徴抽出性能の改善に有効であることが分かる。

4.6 むすび

本章では、本研究で用いたデータセットである MVTEC と提案手法の評価方法について述べた。また、画像再構成手法の学習結果や、それが教師無し異常検知タスクである PaDiM の精度に与える影響について評価を行った。提案手法では、MVTEC を用いた画像再構成手法による追加学習が、PaDiM の精度改善を達成でき、有効性を示すことが出来た。

第5章 結論と今後の課題

5.1 結論

本研究では, 異常検知タスクにおいて, 大規模自然画像データセットである ImgeNet の代替としてフラクタル画像を事前学習に用いる従来の手法に対し, 画像再構成手法による追加学習手法を提案した. 提案手法ではフラクタル画像で事前学習し, MVTec データセットを用いて AutoEncoder モデルを追加学習することで特徴抽出性能の改善を図った. 結果として, 画像単位, ピクセル単位の異常検知精度を示す `img_roc`, `pix_roc` を改善することが出来た. したがって, 本提案により, 異常検知タスクにおける商用利用の制約回避と, 異常検知精度の改善が両立できることを示した.

5.2 今後の課題

提案手法では, MVTec データセット全体での平均精度において精度改善を達成できた一方, カテゴリ毎の精度において, 従来手法を下回るカテゴリも複数見られた. そのため, カテゴリ毎の性質に合わせ, 最適な学習などを行う必要があると考える.

また, PaDiM で用いる特徴量は WideResNet50 の 2, 3, 4 層を選択し比較を行った. 「浅い層の特徴量の方が特徴抽出性能の改善に有効である」という従来研究や本研究での考察に基づいて WideResNet50 の 1 層目も使用して比較実験を行い, 提案手法の有効性を確認する必要がある.

謝辞

本論文の執筆に当たり、丁寧かつ素晴らしいご指導をくださり、快適な研究環境を与えてくださった渡辺裕教授に心より感謝いたします。

また、日頃から貴重な意見をくださり、充実した研究室環境を提供してくださった渡辺研究室の皆様に感謝いたします。

最後に、私をここまで育ててくださり、常に心を支えてくださり、生活を支えてくださっている家族に感謝いたします。

参考文献

- [1] X. Zhang, S. Li, X. Li, P. Huang, J. Shan, and T. Chen, “DeSTSeg: Segmentation Guided Denoising Student-Teacher for Anomaly Detection,” 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3914-3923, Jun. 2023.
- [2] H. Watanabe, R. Togo, T. Ogawa, and M. Haseyama, “Bone Metastatic Tumor Detection based on AnoGAN Using CT Images,” 2019 IEEE Global Conference on Life Sciences and Technologies (LifeTech), pp. 235-236, Mar. 2019.
- [3] G. Pang, C. Yan, C. Shen, A. v. d. Hengel, and X. Bai, “Self-Trained Deep Ordinal Regression for End-to-End Video Anomaly Detection,” 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 12170-12179, Jun. 2020.
- [4] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 248-255, Jun. 2009.
- [5] H. Kataoka, K. Okayasu, A. Matsumoto, E. Yamagata, R. Yamada, N. Inoue, A. Nakamura, and Y. Satoh, “Pre-training without Natural Images,” 15th Asian Conference on Computer Vision (ACCV), Nov. 2024.
- [6] C. I. Ugwu, S. Casarin, and O. Lanz, “Fractal as pre-training Datasets for Anomaly Detection and Localization,” 2024 IEEE CVPR workshop on fair, data-efficient, and Trusted Computer Vision (TCV2024), Jun. 2024.
- [7] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, “MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9584-9592, Jun. 2019.
- [8] T. Defard, A. Setkov, A. Loesch, and R. Audigier, “PaDiM: a Patch Distribution Modeling Framework for Anomaly Detection and Localization,” 25th International Conference on Pattern Recognition (ICPR), Jan. 2021.
- [9] H. Kataoka, K. Okayasu, A. Matsumoto, E. Yamagata, R. Yamada, N. Inoue, A. Nakamura, and Y. Satoh, “Pre-training without Natural Images,” International Journal of Computer Vision (IJCV), Mar. 2021.
- [10] 飯田孝久, 福島崇博, 篠崎 信雄, “2 値変量が混在する場合のマハラノビス距離による異常検出,” 応用統計学 (日本応用統計学会) vol37, no.2, pp. 55-76, Sep. 2008.
- [11] S. Zagoruyko, and N. Komodakis, “Wide Residual Networks,” In British Machine Vision Conference (BMVC), Sep. 2016.

図一覧

図 2.1 フラクタル画像	エラー! ブックマークが定義されていません。
図 2.2 PaDiM における学習時のモデル概要	エラー! ブックマークが定義されてい ません。
図 2.3 PaDiM における推論時のモデル概要	6
図 2.4 従来手法のモデル概要	7
図 3.1 FractalDB を基盤とした画像再構成手法による異常検知モデル	9
図 4.1 MVTEC データセットにおける出力結果	11
図 4.2 PaDiM の定性的評価	エラー! ブックマークが定義されていません。

表一覧

表 4.1 PaDiM の定量的評価 (img_roc, pix_roc)	12
---	----