

# VisA データセットにおける異常検知手法の性能評価

## Performance Evaluation of Anomaly Detection on the VisA Dataset

山本 一騎<sup>†</sup>      中西 知嘉子<sup>‡</sup>  
Kazuki Yamamoto      Chikako Nakanishi

### 1. はじめに

近年、製造業における画像認識 AI を用いた異常検知技術が注目を集めている。異常検知は正常なデータから特徴を学習し、それと異なる特徴を異常であると判断する技術である。この技術は、製造業における不良品の検出や Web サイトでの不正アクセス調査に利用されており、コンピュータを活用することで人件費の削減やヒューマンエラーの防止が期待できる。

これまで異常検知の研究においては、MVTec 社が提供する MVTec-AD データセットが主に使用されてきた。しかし、近年では Amazon Web Services (AWS) が新たに提供する Visual Anomaly (VisA) データセットが注目されている。MVTec-AD は、傷や穴、異物混入といったシンプルで明確な異常に焦点を当てており、影が存在しないなど実際の現場とは差異があるデータであった。しかし VisA は工業的応用を考慮し、傷、汚れ、変色といったより現実的な表面異常に対応している。

そこで、本研究では「重心最短距離削減手法(MRMG)[1]」と、MVTec-AD において 2022 年に SOTA を達成している PatchCore[2]を活用し、現実的な異常に対応できるかを検証する。

### 2. 使用データと手法

#### 2.1. 使用するデータセット

AWS から提供されている VisA[3]を使用する。このデータセットは工業製品における現実的な表面異常の検出を目的として構築されたデータセットであり、12 カテゴリで構成され、各カテゴリはキャンドルやカプセル、プリント基板などの異なる工業製品に対応している。

#### 2.2. MRMG[1]

MRMG は特徴抽出器の中間層の 1 層から得られる特徴量に対し、局所外れ値因子法(LOF)を用いて異常を検知する手法である。

##### 2.2.1. 特徴抽出方法

MRMG では、異常検知のために ResNet50 を特徴抽出器として使用し、全 177 層のうち、カテゴリごとに精度を評価し、最も精度が高くなる層を選択して特徴を抽出する。この方法により、データの種類ごとに最適な特徴を取得し、異常検知の精度向上を図る。本手法は、PatchCore のように画像をパッチ単位に分割してパッチごとに特徴を抽出する方法とは異なり、画像全体を対象としてグローバルな特徴ベクトルを抽出し、それに基づいて異常判定を行う。

##### 2.2.2. 異常検出方法(LOF[4])

MRMG では、異常検知手法として局所外れ値因子法(LOF)を採用している。LOF は、あるデータ点の局所的な密度を周囲のデータ点と比較することで、その点が異常であるかを判断する手法である。

具体的には、学習済みの正常データから得られた特徴量と、判定対象となる特徴量を比較し、周囲と比べて局所密

度が著しく低いデータ点を異常とみなす。

ただし、LOF の性質上、学習データの特徴量が過度に密集している場合、正常なデータであっても異常と誤判定を起こすことがある。

その課題を解消するために MRMG では重心最短距離削減手法を導入している。

##### 2.2.3. 特徴量削減(重心最短距離削減手法)

MRMG では、特徴量の密集による LOF の誤判定を防ぐため、重心最短距離削減手法を導入している。本手法は、学習データの分布を均一化し、異常検知精度の向上を図ることを目的としている。

具体的な処理の流れは以下の通りである：

1. 学習データの特徴量を特徴空間に配置し、その分布に基づいてクラスタの重心を算出する。
2. データ間の距離が最小となるペアを探索し、その中で重心に最も近いデータ点を削除する。
3. この処理を反復的に行い、学習データを 10 個ずつ削減する。
4. 各削減ステップにおいて LOF による精度を算出し、最も高い精度の特徴数を記録する。

この削減手法により、特徴空間の密度が均一化され、LOF による誤判定の抑制につながり、データの削減により、推論時の計算負荷が軽減される。

#### 2.3. PatchCore[2]

PatchCore は、画像の局所領域 (Patch) ごとに特徴量を抽出し、特徴ベクトルを貪欲法で削除することで高速化を実現しながら近傍法で異常を判断する手法である。特徴抽出には ResNet50 を使い、その 2 層の中間層を用いてパッチごとの特徴量を抽出する。

##### 2.3.1. 異常検出方法(近傍法)

PatchCore では、学習段階で正常画像から抽出されたパッチ特徴をメモリバンクとして保持する。推論時には、テスト画像から抽出された各パッチの特徴に対して、メモリバンク内の正常パッチとの距離を計算し、最も近い正常パッチ (最近傍) の距離に基づいて異常度を算出する。このとき、距離が大きいほど、正常からの逸脱が大きいと判断され、異常と判定される。

##### 2.3.2. 特徴量削減(貪欲法)

PatchCore では、特徴量の数を削減するために貪欲法を用いている。まず特徴ベクトル群からランダムに 1 つの特徴ベクトルを選択して集合に加え、集合に属する点から最も遠い点を集合に加えていく処理を繰り返していく。

この処理により、特徴量の次元と数を大幅に削減でき、精度を大きく損なわずに効率的な推論が可能となる。

### 3. 評価方法

本研究では、処理時間の計測は画像 1 枚を判定するのにかかる平均時間とする。精度の評価には ROC 曲線・AUC を用いる。

## 4. 評価と結果

PatchCore と MRMG の AUC スコアおよび処理時間の比較結果を表 4.1 に示す。

表 4.1 から、本研究で提案した MRMG は、多くのカテゴリにおいて PatchCore と比較して同等またはそれ以上の異常検知性能を示していることが分かる。特に「capsules」および「macaroni2」においては、PatchCore を大きく上回る精度を達成した。一方で、「pcb3」および「pipe\_fryum」では、PatchCore がわずかに高い AUC を示す結果となった。また、「macaroni2」は両手法において AUC が 0.8 を下回っており、検出が難しいカテゴリであることがわかった。

また、処理時間では MRMG が PatchCore よりも大幅に短い推論時間を実現している。全カテゴリの平均実行時間では MRMG は PatchCore の 3 分の 1 以下であり、特に candle, macaroni2, pcb1 や pcb3 では 10 倍以上の差が見られた。ただし、pcb2 では、両手法間での実行時間にほとんど差は見られなかった。

表 4.1 異常検知手法の処理時間と精度

|            | MRMG  |       | PatchCore |       |
|------------|-------|-------|-----------|-------|
|            | auc   | 秒     | auc       | 秒     |
| candle     | 0.922 | 0.044 | 0.929     | 0.616 |
| capsules   | 0.861 | 0.126 | 0.692     | 0.393 |
| cashew     | 0.979 | 0.109 | 0.949     | 0.319 |
| chewinggum | 0.963 | 0.035 | 0.957     | 0.328 |
| fryum      | 0.945 | 0.034 | 0.906     | 0.326 |
| macaroni1  | 0.810 | 0.218 | 0.783     | 0.623 |
| macaroni2  | 0.715 | 0.059 | 0.620     | 0.616 |
| pcb1       | 0.879 | 0.029 | 0.846     | 0.622 |
| pcb2       | 0.979 | 0.638 | 0.932     | 0.606 |
| pcb3       | 0.822 | 0.044 | 0.901     | 0.627 |
| pcb4       | 0.991 | 0.107 | 0.994     | 0.620 |
| pipe_fryum | 0.884 | 0.078 | 0.983     | 0.340 |

## 5. 考察

### 5.1. 精度に関する考察

両手法ともに精度が悪かった「macaroni2」は画像内に同じ物体が複数個存在し、それらの位置が画像間で変化するという特徴を持つ。また、これと類似した性質を持つ「capsules」においては、PatchCore の性能が大きく劣り、MRMG が優れた結果を示した。

これらの結果から、物体の位置が一定でないカテゴリに対しては、Patch 単位で処理を行う PatchCore は不利となりやすく、画像全体で処理を行う MRMG の方が位置変化に対するロバスト性が高いことが分かる。一方で、「pcb3」と「pipe\_fryum」では、PatchCore が MRMG を大きく上回る精度であった。これらのカテゴリは、画像内に占める背景の割合が高く、異常領域が小さいという特徴を持っていた。このような状況では、背景の影響を受けにくい Patch 単位の処理が有効に働き、PatchCore の特性とマッチして高精度を実現したと考えられる。

### 5.2. 処理時間に関する考察

実行時間の比較においては、ほぼすべてのカテゴリにおいて MRMG の方が大幅に高速であるという傾向が明確に確認された。

PatchCore は、テスト画像のバッチごとに、保持済み特徴量との最近傍探索を行う必要があるため、推論時の計算コストが非常に高くなる。一方、MRMG は、画像全体を対象として中間層 1 層からグローバルな特徴ベクトルを抽出し、それに対して LOF (局所外れ値因子法) を適用すること、正常データの特徴ベクトル群に対して「重心最短距離削減手法」を適用し、特徴数を大幅に削減していることで、推論処理の計算量を大幅に削減されていると考えられる。

しかし、pcb2 は両手法の実行時間に差がほとんど見られなかった。MRMG における削減処理で削除される特徴ベクトル数が少なかったため、結果として LOF の密度推定に時間を要した可能性があると考えられる。

異常の結果から、MRMG は PatchCore と同等あるいはそれ以上の AUC を維持しながら、推論時間を大幅に短縮できる手法であることが確認された。この特性は、実時間性が求められる応用や、計算資源が限られたエッジデバイス上での運用において特に有効であるといえる。

## 6. 今後の課題

「macaroni2」における精度差に着目すると、MRMG と PatchCore の性能に顕著な差が見られたことから、本カテゴリのデータ特性に起因する両手法の強み・弱みをより詳細に分析する必要があると考えられる。

また、「pcb3」や「pipe\_fryum」のような異常領域が小さく、背景が支配的なカテゴリに対しては、MRMG の性能が十分でなかった。今後は、これらのカテゴリに特化した、精度向上に向けた手法改良を進める。

## 7. 結論

本研究により、MRMG は VisA に対して高速かつ高精度に適應できると分かった。今後は、精度の悪い「macaroni2」について、原因を調査し、精度向上の手法を検討しようと考えている。

### 参考文献

- [1] 北野 侑弥, 中西 知嘉子, “エッジ AI 向けの異常検知の高精度, 高速化の検討”, 電子通信学会総合大会, (2023)
- [2] PatchCore, Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Scholkopf, Thomas Brox, Peter Gehler Towards Total Recall in Industrial Anomaly Detection CVPR, (2022)
- [3] Visual Anomaly(VisA) <<https://registry.opendata.aws/visa/>>
- [4] Markus M. Breunig, Hans-Peter Kriegel, Raymond T Ng, Jörg Sander “Identifying Density-Based Local Outliers“(2000)
- [5] 【ROC 曲線と AUC】機械学習の評価指標についての基礎講座 <[https://www.randpy.tokyo/entry/roc\\_auc/](https://www.randpy.tokyo/entry/roc_auc/)> (2022)

† 大阪工業大学 情報科学研究科 情報科学専攻  
Graduate School of Information Science and Technology  
Osaka Institute of Technology  
‡ 大阪工業大学 情報科学部 情報知能学科  
Department of Information and Computer Science Osaka  
Institute of Technology