

PaDiM: 異常検出と局所化のためのパッチ分布モデリングフレームワーク

Thomas Defard, Aleksandr Setkov, Angelique Loesch, Romaric Audigier

Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France

thomas.defard@imt-atlantique.net, {aleksandr.setkov, angelique.loesch, romaric.audigier}@cea.fr

要約—私たちは、1クラス学習設定において画像内の異常を同時に検出・局所化する新たなフレームワーク「パッチ分布モデリング (PaDiM)」を提案します。PaDiMは、パッチ埋め込みに事前学習済みの畳み込み神経ネットワーク (CNN) を利用し、正常クラスの確率的表現を得るために多変量ガウス分布を採用しています。また、CNNの異なるセマンティックレベル間の相関を活かすことで、異常の局所化を向上させます。PaDiMは、MVTec ADおよびSTCデータセットにおいて、異常検出と局所化の両方で現在の最先端手法を凌駕します。現実の視覚的産業検査に合わせるため、評価プロトコルを拡張し、非一致データセットにおける異常局所化アルゴリズムの性能を評価します。PaDiMの最先端性能と低複雑度は、多くの産業応用における有望な候補となります。

I. INTRODUCTION

人間は、均一な自然画像の集合の中から異質なまたは予期しないパターンを検出することができます。このタスクは異常検出または新規性検出と呼ばれ、視覚的産業検査を含む多くの応用分野があります。しかし、製造ラインにおける異常は極めて稀なイベントであり、手動での検出は手間がかかります。したがって、異常検出の自動化は、注意力の低下を回避し、人間のオペレーター作業を容易にすることで、継続的な品質管理を可能にします。本論文では、異常検出に焦点を当て、特に産業検査の文脈における異常局所化に重点を置きます。コンピュータビジョンにおいて、異常検出は画像に異常スコアを付与する作業です。異常局在化はより複雑なタスクであり、各ピクセルまたはピクセルのパッチに異常スコアを付与し、異常マップを出力します。これにより、異常局在化はより正確で解釈可能な結果を生成します。当手法でMVTec 異常検出 (MVTec AD) データセット [1] の画像から異常を局在化した際の異常マップの例は図1に示されています。

異常検出は、正常クラスと異常クラス間の二値分類です。しかし、異常例が不足している場合が多く、さらに異常は予期しないパターンを示すため、このタスクで完全な監督学習モデルを訓練することはできません。したがって、異常検出モデルは通常、単一クラス学習設定で推定されます。つまり、訓練データセットには正常クラスの画像のみが含まれ、訓練中に異常例は利用できません。テスト時、正常な訓練データセットと異なる例は異常として分類されます。

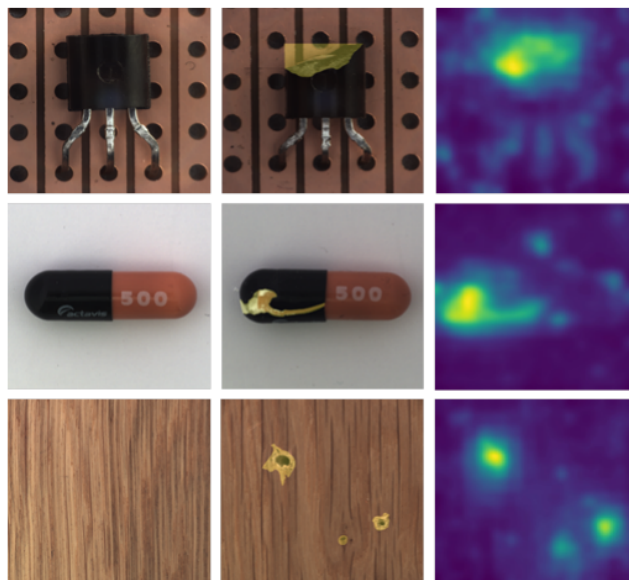


図1. MVTec AD [1] からの画像サンプル。左列: トランジスタ、カプセル、ウッドクラスの正常画像。中央列: 同じクラスの画像で、黄色でハイライトされた ground truth 異常。右列: 当社の PaDiM モデルで得られた異常ヒートマップ。黄色の領域は検出された異常に対応し、青色の領域は正常領域を示します。カラー表示が推奨されます。

最近、異常局所化と異常検出タスクを1クラス学習設定で組み合わせる複数の方法が提案されています [2] - [5]。しかし、これらの方法は、深層神経ネットワークのトレーニング [3]、[6] を必要とし、煩雑な場合があるか、またはテスト時にトレーニングデータセット全体にK-近傍法 (K-NN) アルゴリズム [7] を適用する [4]、[5] ものです。KNNアルゴリズムの線形複雑さは、トレーニングデータセットのサイズが増加するにつれ、時間と空間の複雑さを増加させます。これらの2つのスケーラビリティ問題は、異常局所化アルゴリズムの産業分野での展開を妨げる可能性があります。

上記の問題を緩和するため、私たちは新しい異常検出と局所化アプローチであるPaDiM (パッチ分布モデリング) を提案します。これは事前訓練された畳み込み神経ネットワーク (CNN) を埋め込み抽出に利用し、以下の2つの特性を持っています:

- 各パッチの位置は多変量ガウス分布で記述されます;
- PaDiM は、異なる領域間の相関関係を考慮します。

PaDiM: a Patch Distribution Modeling Framework for Anomaly Detection and Localization

Thomas Defard, Aleksandr Setkov, Angelique Loesch, Romaric Audigier

Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France

thomas.defard@imt-atlantique.net, {aleksandr.setkov, angelique.loesch, romaric.audigier}@cea.fr

Abstract—We present a new framework for Patch Distribution Modeling, PaDiM, to concurrently detect and localize anomalies in images in a one-class learning setting. PaDiM makes use of a pretrained convolutional neural network (CNN) for patch embedding, and of multivariate Gaussian distributions to get a probabilistic representation of the normal class. It also exploits correlations between the different semantic levels of CNN to better localize anomalies. PaDiM outperforms current state-of-the-art approaches for both anomaly detection and localization on the MVTec AD and STC datasets. To match real-world visual industrial inspection, we extend the evaluation protocol to assess performance of anomaly localization algorithms on non-aligned dataset. The state-of-the-art performance and low complexity of PaDiM make it a good candidate for many industrial applications.

I. INTRODUCTION

Humans are able to detect heterogeneous or unexpected patterns in a set of homogeneous natural images. This task is known as anomaly or novelty detection and has a large number of applications, among which visual industrial inspections. However, anomalies are very rare events on manufacturing lines and cumbersome to detect manually. Therefore, anomaly detection automation would enable a constant quality control by avoiding reduced attention span and facilitating human operator work. In this paper, we focus on anomaly detection and, in particular, on anomaly localization, mainly in an industrial inspection context. In computer vision, anomaly detection consists in giving an anomaly score to images. Anomaly localization is a more complex task which assigns each pixel, or each patch of pixels, an anomaly score to output an anomaly map. Thus, anomaly localization yields more precise and interpretable results. Examples of anomaly maps produced by our method to localize anomalies in images from the MVTec Anomaly Detection (MVTec AD) dataset [1] are displayed in Figure 1.

Anomaly detection is a binary classification between the normal and the anomalous classes. However, it is not possible to train a model with full supervision for this task because we frequently lack anomalous examples, and, what is more, anomalies can have unexpected patterns. Hence, anomaly detection models are often estimated in a one-class learning setting, *i.e.*, when the training dataset contains only images from the normal class and anomalous examples are not available during the training. At test time, examples that differ from the normal training dataset are classified as anomalous.

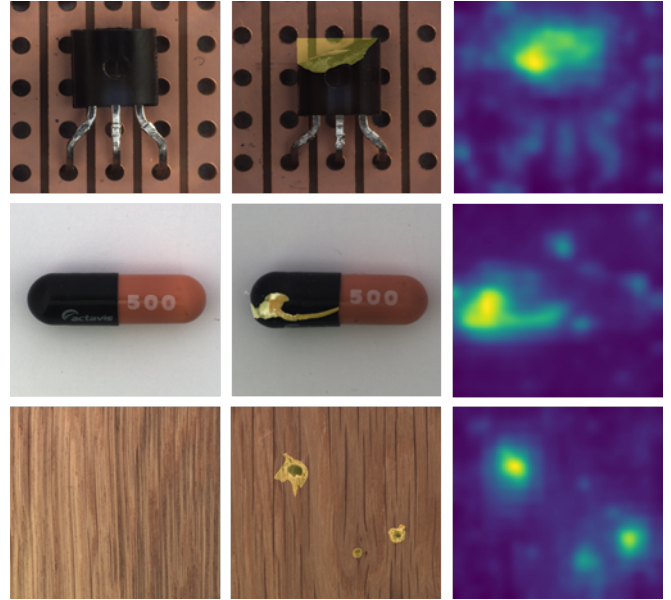


Fig. 1. Image samples from the MVTec AD [1]. *Left column*: normal images of Transistor, Capsule and Wood classes. *Middle column*: images of the same classes with the ground truth anomalies highlighted in yellow. *Right column*: anomaly heatmaps obtained by our PaDiM model. Yellow areas correspond to the detected anomalies, whereas the blue areas indicate the normality zones. Best viewed in color.

Recently, several methods have been proposed to combine anomaly localization and detection tasks in a one-class learning setting [2]–[5]. However, either they require deep neural network training [3], [6] which might be cumbersome, or they use a K-nearest-neighbor (K-NN) algorithm [7] on the entire training dataset at test time [4], [5]. The linear complexity of the KNN algorithm increases the time and space complexity as the size of the training dataset grows. These two scalability issues may hinder the deployment of anomaly localization algorithms in industrial context.

To mitigate the aforementioned issues, we propose a new anomaly detection and localization approach, named PaDiM for Patch Distribution Modeling. It makes use of a pretrained convolutional neural network (CNN) for embedding extraction and has the two following properties:

- Each patch position is described by a multivariate Gaussian distribution;
- PaDiM takes into account the correlations between dif-