

特徴空間は、カーネル手法および深層学習手法 [26] の両方における特徴学習の基盤となっています。別の手法のセットは自己教師付き学習に基づいています。GolanとEl-Yaniv [11] は、入力データに幾何学的変換を施し、その変換を認識しようとするネットワークを訓練するRotNetベース [10] のアプローチを提案しました。彼らは、訓練された分類器が新しい正常な画像にはよく汎化するが、異常な画像には汎化しないというアイデアを利用し、異常検出の基準として使用できるようにしています。Hendrycksら [18] は、アーキテクチャと訓練手順を改善し、優れた性能を実現しました。BergmanとHoshen [4] は、この研究をSVDD型の基準と組み合わせ、非画像データに拡張しました。最近、Bergmanら [3] は、このような自己教師付き学習手法で学習された特徴量が、汎用的なImageNetベースの特徴量抽出器と競合しないことを示しました。kNN (または効率的な近似) に基づく単純な手法が、このような自己教師付き学習手法を大幅に上回りました。

サブ画像手法：これまで説明した手法は、全体的な画像を正常か異常か分類するタスクに取り組み、ほとんどの技術は画像に特化していませんでした。異常を含む特定のピクセルをセグメント化するタスクは画像に特化しており、深層学習コミュニティからほとんど注目されていませんでした。ナポレターノラ [24] は、小さな重なり合うパッチから深層特徴を抽出し、次元削減された特徴量に対してK-meansベースの分類器を使用しました。Bergmannら [5] は、MVTechデータセットでADGANとオートエンコーダーのアプローチを評価し、互いに補完的な強みを確認しました。最近、Venkataramanら [31] は、複数の手法 (GAN損失 [13]、GRADCAM [29]) を組み合わせた注意誘導型VAEアプローチを採用しました。Bergmannら [6] は、事前訓練されたImageNet深層特徴量を用いた学生-教師ベースのオートエンコーダーアプローチを採用しましたが、依然として高コストな訓練段階が必要でした。本研究では、従来の方法よりも精度が高く、高速で安定しており、専用のトレーニング段階を必要としない新たなサブ画像アラインメント手法を提案します。サブ画像異常検出の研究を支援するため、このタスクの評価用に高品質なデータセットが導入されています。例えば：MVTech [5] は、製品のイメージの一部 (へこみや欠落部分など) に欠陥を含む部分を検出する産業用故障検出を模擬したデータセットです。ShanghaiTech Campusデータセット [23]：監視環境を模擬したデータセットで、カメラが混雑したキャンパスを監視し、喧嘩などの異常な物体や活動を検出する目的です。Hendrycksら [17] も、道路の危険物などの異常を含む新たなデータセットを提案しています。私たちは、最も広く使用されているMVTechとShanghaiTech Campus (STC) の2つのデータセットで本研究を評価しました。

3 対応関係に基づくサブ画像異常検出

私たちはサブ画像異常検出のための方法を提案します。私たちの方法は以下の構成要素からなります：i) 画像特徴抽出 ii) K近傍法による正常画像の検索 iii) ディープ特徴ピラミッド対応に基づくピクセルアラインメント。

the feature space used giving rise to both kernel methods as well as deep methods [26] for learning features. Another set of methods is based on self-supervised learning. Golan and El-Yaniv [11] proposed a RotNet-based [10] approach, which performs geometric transformations on the input data and trains a network that attempts to recognize the transformation used. They use the idea that the trained classifier will generalize well to new normal images but not to anomalous images allowing it to be used as an anomaly detection criterion. Hendrycks et al. [18] improved the architecture and training procedure achieving strong performance. Bergman and Hoshen [4] combined this work with an SVDD type criterion and extended it to non-image data. Very recently Bergman et al. [3] showed that the features learned using such self-supervised methods are not competitive with generic ImageNet-based feature extractors. A simple method based on kNN (or efficient approximations) significantly outperformed such self-supervised methods.

Sub-image methods: The methods previously described tackled the task of classifying a whole image as normal or anomalous, and most of the techniques were not specific to images. The task of segmenting the particular pixels containing anomalies is special to images and has achieved far less attention from the deep learning community. Napoletano et al. [24] extracted deep features from small overlapping patches, and used a K-means based classifier over dimensionality reduced features. Bergmann et al. [5] evaluated both a ADGAN and autoencoder approaches on MVTech finding complementary strengths. More recently, Venkataramanan et al. [31] used an attention-guided VAE approach combining multiple methods (GAN loss [13], GRADCAM [29]). Bergmann et al. [6] used a student-teacher based autoencoder approach employing pre-trained ImageNet deep features (still requiring an expensive training stage). In this work, we present a novel sub-image alignment approach which is more accurate, faster, more stable than previous methods and does not require a dedicated training stage. To support research on sub-image anomaly detection, high quality datasets for evaluating this task have been introduced, such as: MVTech [5] - a dataset simulating an industrial fault detection where the objective is to detect parts of images of products that contain faults such as dents or missing parts. The ShanghaiTech Campus dataset [23] - a dataset simulating a surveillance setting where cameras observe a busy campus and the objective is to detect anomalous objects and activities such as fights. Hendrycks et al. [17] also proposed a new dataset containing anomalies such as road hazards. We evaluate our work on the two most used datasets, MVTech and ShanghaiTech Campus (STC).

3 Correspondence-based Sub-Image Anomaly Detection

We present our method for sub-image anomaly detection. Our method consists of several parts: i) image feature extraction ii) K nearest neighbor normal image retrieval iii) pixel alignment with deep feature pyramid correspondences.