

検索に用いられるデータセットのサイズが、データセットが非常に大規模または高次元の場合に問題となる可能性があります。当社のアプローチは、このような複雑さを軽減するように設計されています。まず、グローバルにプールされた特徴量（2048次元のベクトル）に基づいて、初期の画像レベル異常分類を計算します。このようなkNN計算は、中規模のデータセットでは非常に高速に実行可能であり、大規模データセットには異なる高速化技術（例：KD Trees）を適用できます。異常セグメンテーション段階では、画像レベルkNNよりも大幅に遅いピクセルレベルkNN計算が必要です。しかし、当社の方法は異常画像のK近傍のみを対象にサブ画像kNN検索を制限し、計算時間を大幅に短縮します。ほとんどの画像が正常であると仮定するため、異常セグメンテーションの次の段階が必要な画像はごく一部に限定されます。さらに、異常セグメンテーション段階は説明可能性と人間オペレーターとの信頼構築のために必要ですが、多くの場合時間的制約が厳しくないため、計算時間に対する要件が緩やかです。したがって、当手法は複雑さと実行時間の両面から実践的な展開に非常に適しています。

事前学習済み vs. 自己学習済み特徴量：従来のサブ画像異常検出手法は、自己学習済み特徴量または自己学習済みと事前学習済み画像特徴量の組み合わせを使用してきました。この文脈での自己学習アプローチは、通常、オートエンコーダーを訓練し、その再構築誤差を異常検出に利用します。他のアプローチでは、事前学習済みと自己学習済み手法の組み合わせを使用しています。例えば、知覚損失を使用する方法や [6] で提案された事前学習済みエンコーダーを使用する方法などです。私たちの数値結果は、私たちの手法がこれらのアプローチを大幅に上回っていることを示しています。この研究で扱ったように、正常データのためのトレーニングセットにおける限られた監督と小さなデータセットサイズを考慮すると、非常に深い事前学習済みネットワークを凌駕するのは困難だと考えています。したがって、私たちは事前学習済み特徴を使用し、それらを改変しようとはしません。私たちの手法が達成した優れた結果は、このアプローチの有効性を示しています。私たちは、今後の研究ではこの特定のタスク向けに事前学習済み深層特徴の微調整手法に焦点を当て、私たちの手法を超える改善が期待できると考えています。ただし、私たちのアプローチの展開の容易さと汎用性は、多くの実践的な設定において適切な選択となるでしょう。

## 6 Conclusion

私たちは、画像内の異常を検出・セグメント化する新たなアライメントベースの手法を発表しました。私たちの手法は、事前学習済み深層特徴量から抽出されたピクセルレベルの特徴ピラミッドのK近傍に依存しています。私たちの手法は、高精度と合理的な計算複雑性を実現するように設計された2つの段階から構成されています。私たちの方法は、2つの現実的なサブ画像異常検出データセットにおいて、現在の最先端手法を凌駕する性能を示しました。同時に、はるかにシンプルな構造となっています。私たちの方法の展開の容易さは、実践家にとって良い候補となるはずです。

the size of the dataset used for search which can be an issue when the dataset is very large or of high dimensionality. Our approach is designed to mitigate the complexity issues. First, we compute the initial image-level anomaly classification on global-pooled features which are 2048 dimensional vectors. Such kNN computation can be achieved very quickly for moderate sized datasets and different speedup techniques (e.g. KDTrees) can be used for large scale datasets. The anomaly segmentation stage requires pixel-level kNN computation which is significantly slower than image-level kNN. However, our method limits the sub-image kNN search to only the K nearest neighbors of the anomalous image significantly limiting computation time. We assume that the vast majority of images are normal, therefore only a small fraction of images require the next stage of anomaly segmentation. Additionally, the anomaly segmentation stage is required for explainability and trust building with the human operators, but in many cases it is not time-critical therefore putting a laxer requirement on computation time. Our method is therefore quite suitable for practical deployment from a complexity and runtime perspective.

**Pre-trained vs. learned features:** Previous sub-image anomaly detection methods have either used self-learned features or a combination of self-learned and pre-trained images features. Self-learned approaches in this context, typically train an autoencoder and use its reconstruction error for anomaly detection. Other approaches have used a combination of pre-trained and self-learned methods e.g. methods that use perceptual losses and [6] which uses a pre-trained encoder. Our numerical results have shown that our method significantly outperforms such approaches. We believe that given the limited supervision and small dataset size in normal-only training set as tackled in this work, it is rather hard to beat very deep pre-trained networks. We therefore use pre-trained features and do not attempt to modify them. The strong results achieved by our method attest to the effectiveness of this approach. We believe that future work should focus on methods for finetuning the deep pre-trained features for this particular task and expect it to improve over our method. That notwithstanding the ease of deployment and generality of our approach should make it a good choice in many practical settings.

## 6 Conclusion

We presented a novel alignment-based method for detecting and segmenting anomalies inside images. Our method relies on K nearest neighbors of pixel-level feature pyramids extracted by pre-trained deep features. Our method consists of two stages, which are designed to achieve high accuracy and reasonable computational complexity. Our method was shown to outperform the strongest current methods on two realistic sub-image anomaly detection datasets, while being much simpler. The ease of deployment enjoyed by our method should make it a good candidate for practitioners.