

図6. PatchCoreは競合手法に比べて著しく高いサンプル効率を示し、名目上の訓練データのほんの一部で以前の最先端手法と同等の性能を達成しています。比較のため、PaDiMとSPADEはWideResNet50で再実装されています。

表6. mSTC [32, 52]での異常セグメンテーションとMTD [26]での異常検出を、[42]で報告された結果と比較。

mSTC	CAVGA-R _u [52]	SPADE [10]	PaDiM [14]	PatchCore-10
Pixelwise AUROC [%]	85	89.9	91.2	91.8
MTD	GANomaly [2]	1-NN [35]	DifferNet [42]	PatchCore-10
AUROC [%]	76.6	80.0	97.7	97.9

§C.1. 図に示すように、名目上のトレーニングデータの5分の1のみを使用しても、PatchCoreは以前の最先端性能と匹敵する性能を発揮できます。さらに、[42]で実施された16ショット実験と比較すると、PatchCoreは既に事前学習された特徴量の上に正規化フローモデルを適応させる彼らのアプローチよりも優れた性能を示します。[10]で提案された画像レベル記憶アプローチと比較すると、名目上のショットの5分の1のみで、局所化と検出性能が同等であることが確認されました。

4.6. 他のベンチマークでの評価

PatchCoreは、2つの追加の異常検出性能ベンチマークで評価されています：上海科技大学キャンパスデータセット（STC）[32]と磁気タイル欠陥データセット（MTD）[26]。STCの評価は§4.1で説明されている通り、[52]、[14]、[10]に従っています。STC動画データのサブサンプリング版（mSTC）における無監督異常局所化性能を報告します。画像は256×256にリサイズされています[14]。検出コンテキストがImageNetで利用可能な自然画像データに近い場合、階層レベル3と4の深いネットワーク特徴マップを使用しますが、PatchCoreのハイパーパラメーター調整は行いません。表6（上）の結果は、最先端の異常局所化性能を示しており、PatchCoreのこのようなドメインへの良好な転移性を示唆しています。最後に、MTD（磁気タイル欠陥画像のさまざまなサイズを含む）を検証します。ここでは、空間的に剛体なアプローチ（例：PaDiM）を直接適用できません。ここで、名義データは異常サンプルで遭遇するものと同様の高い変動性を既に示しています[42]。[42]で提案されたプロトコルに従い、画像レベルの異常検出性能を測定し、性能が一致する（甚至いは

slightly outperform) that of [42] (Table 6, bottom).

5. Conclusion

本論文では、テスト時に異常なデータを検出・セグメント化するために、名目上の例のみを活用する冷スタート異常検出のためのPatchCoreアルゴリズムを紹介しました。PatchCoreは、ImageNet事前学習ネットワークから抽出された局所的に意識した名目上のパッチレベルの特徴表現を含むメモリバンクを活用し、テスト時に名目上の文脈を最大限保持する一方で、コアセットサブサンプリングにより実行時間を最小化するというバランスを実現しています。その結果、産業用異常検出ベンチマークにおいて低計算コストを実現する最先端の冷スタート画像異常検出・局所化システムが得られました。MVTecでは、関連する小規模トレーニングセット環境において最高のサンプル効率で、画像異常検出AUROCが99%を超える性能を達成しました。

広範な影響。自動化された産業用異常検出はコンピュータビジョンにおける最も成功した応用の一つであるため、PatchCoreによる改善は、この分野の実務家にとって注目値する可能性があります。当研究は産業用異常検出に特化しているため、社会的影響は限定的です。また、根本的なアプローチはより議論の余地のある分野の検出システムに活用される可能性があります。当改善がそのようなシステムの社会的応用を変えるほど重大なものではないと考えています。

制限事項。PatchCoreは、問題領域に特化した適応を必要とせず産業用異常検出において高い有効性を示しますが、適用範囲は利用する事前学習済み特徴量の転移可能性に制限されます。この問題は、PatchCoreの有効性と利用する特徴量の適応を組み合わせることで解決可能です。この興味深い拡張は今後の研究課題として残します。

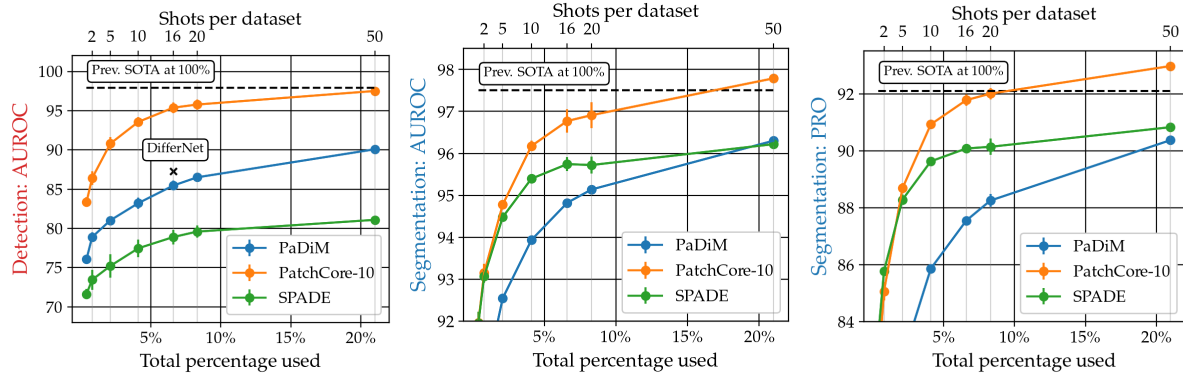


Figure 6. *PatchCore* shows notably higher sample-efficiency than competitors, matching the previous state-of-the-art with a fraction of nominal training data. Note that PaDiM and SPADE were reimplemented with WideResNet50 for comparability.

Table 6. Anomaly Segmentation on mSTC [32, 52] and anomaly detection on MTD [26] compared to results reported in [42].

mSTC	CAVGA-R _n [52]	SPADE [10]	PaDiM [14]	<i>PatchCore-10</i>
Pixelwise AUROC [%]	85	89.9	91.2	91.8
MTD	GANomaly [2]	1-NN [35]	DifferNet [42]	<i>PatchCore-10</i>
AUROC [%]	76.6	80.0	97.7	97.9

§C.1. As shown, using only one fifth of nominal training data, *PatchCore* can still match previous state-of-the-art performance. In addition, comparing to the 16-shot experiments performed in [42], we find *PatchCore* to outperform their approach which adapts a normalizing flows model on top of already pretrained features. Compared to image-level memory approaches in [10], we find matching localization and detection performance with only 5/1 nominal shots.

4.6. Evaluation on other benchmarks

We benchmark *PatchCore* on two additional anomaly detection performance benchmarks: The ShanghaiTech Campus dataset (STC) [32] and the Magnetic Tile Defects dataset (MTD) [26]. Evaluation for STC as described in §4.1 follows [52], [14] and [10]. We report unsupervised anomaly localization performance on a subsampled version of the STC video data (mSTC), with images resized to 256×256 [14]. As the detection context is much closer to natural image data available in ImageNet, we make use of deeper network feature maps at hierarchy levels 3 and 4, but otherwise do not perform any hyperparameter tuning for *PatchCore*. The results in Table 6 (top) show state-of-the-art anomaly localization performance which suggests good transferability of *PatchCore* to such domains. Finally, we examine MTD, containing magnetic tile defect images of varying sizes on which spatially rigid approaches like PaDiM cannot be applied directly. Here, nominal data already exhibits high variability similar to those encountered in anomalous samples [42]. We follow the protocol proposed in [42] to measure image-level anomaly detection performance and find performance to match (and even

slightly outperform) that of [42] (Table 6, bottom).

5. Conclusion

This paper introduced the *PatchCore* algorithm for cold-start anomaly detection, in which knowledge of only nominal examples has to be leveraged to detect and segment anomalous data at test-time. *PatchCore* strikes a balance between retaining a maximum amount of nominal context at test-time through the usage of memory banks comprising locally aware, nominal patch-level feature representations extracted from ImageNet pretrained networks, and minimal runtime through coreset subsampling. The result is a state-of-the-art cold-start image anomaly detection and localization system with low computational cost on industrial anomaly detection benchmarks. On MVTec, we achieve an image anomaly detection AUROC over 99% with highest sample efficiency in relevant small training set regimes.

Broader Impact. As automated industrial anomaly detection is one of the most successful applications of Computer Vision, the improvements gained through *PatchCore* can be of notable interest for practitioners in this domain. As our work focuses specifically on industrial anomaly detection, negative societal impact is limited. And while the fundamental approach can potentially be leveraged for detection systems in more controversial domains, we don’t believe that our improvements are significant enough to change societal application of such systems.

Limitations. While *PatchCore* shows high effectiveness for industrial anomaly detection without the need to specifically adapt to the problem domain at hand, applicability is generally limited by the transferability of the pretrained features leveraged. This can be addressed by merging the effectiveness of *PatchCore* with adaptation of the utilized features. We leave this interesting extension to future work.