the abstract features required in an industrial environment. In addition, nominal context usable by these methods at test time is effectively limited by the small number of extractable high-level feature representations.

In this paper, we present *PatchCore* as an effective remedy by (1) maximizing nominal information available at test time, (2) reducing biases towards ImageNet classes and (3) retaining high inference speeds. Relying on the fact that an image can be already classified as anomalous as soon as a single patch is anomalous [14, 56], *PatchCore* achieves this by utilizing locally aggregated, mid-level features patches. The usage of mid-level network patch features allows *PatchCore* to operate with minimal bias towards ImageNet classes on a high resolution, while a feature aggregation over a local neighbourhood ensures retention of sufficient spatial context. This results in an extensive memory bank allowing *PatchCore* to optimally leverage available nominal context at test time. Finally, for practical applicability, *PatchCore* additionally introduces greedy coreset subsampling [1] for nominal feature banks as a key element to both reduce redundancy in the extracted, patch-level memory bank as well as significantly bringing down storage memory and inference time, making *PatchCore* very attractive for realistic industrial use cases.

Thorough experiments on the diverse MVTec AD [5] as well as the specialized Magnetic Tile Defects (MTD) [26] industrial anomaly detection benchmarks showcase the power of *PatchCore* for industrial anomaly detection. It achieves state-of-the-art image-level detection scores on MVTec AD and MTD, with nearly perfect scores on MVTec AD (up to AUROC 99.6%), reducing detection error of previous methods by **more than half**, as well as state-of-the-art industrial anomaly localization performance. *PatchCore* achieves this while retaining fast inference times without requiring training on the dataset at hand. This makes *PatchCore* very attractive for practical use in industrial anomaly detection. In addition, further experiments showcase the high sample efficiency of *PatchCore*, matching existing anomaly detection methods in performance while using only a fraction of the nominal training data.

2. Related Works

Most anomaly detection models rely on the ability to learn representations inherent to the nominal data. This can be achieved for example through the usage of autoencoding models [44]. To encourage better estimation of the nominal feature distribution, extensions based on Gaussian mixture models [60], generative adversarial training objectives [2,39,43], invariance towards predefined physical augmentations [25], robustness of hidden features to reintroduction of reconstructions [29], prototypical memory banks [21], attention-guidance [52], structural objectives [7, 59] or constrained representation spaces [38] have been proposed. Other unsupervised representation learning methods can similarly be utilised, such as via GANs [13], learning to predict predefined geometric transformations [20] or via normalizing flows [42]. Given respective nominal representations and novel test representations, anomaly detection can then be a simple matter of reconstruction errors [44], distances to k nearest neighbours [18] or finetuning of a one-class classification model such as OC-SVMs [46] or SVDD [50, 56] on top of these features. For the majority of these approaches, anomaly localization comes naturally based on pixel-wise reconstruction errors, saliency-based approaches such as GradCAM [47] or XRAI [28] can be used for anomaly segmentation [42, 45, 52] as well.

Industrial Anomaly Detection. While literature on general anomaly detection through learned nominal representations is vast, industrial image data comes with its own challenges [5], for which recent works starting with [4] have shown state-of-the-art detection performance using models pretrained on large external natural image datasets such as ImageNet [16] without any adaptation to the data at hand. This has given rise to other industrial anomaly detection methods reliant on better reuse of pretrained features such as SPADE [10], which utilizes memory banks comprising various feature hierarchies for finegrained, kNN-based [18] anomaly segmentation and image-level anomaly detection. Similarly, [14] recently proposed PaDiM, which utilizes a locally constrained bag-of-features approach [8], estimating patch-level feature distribution moments (mean and covariance) for patch-level Mahalanobis distance measures [33]. This approach is similar to [40] studied on full images. To better account for the distribution shift between natural pretraining and industrial image data, subsequent adaptation can be done, e.g. via student-teacher knowledge distillation [24] such as in [6, 45] or normalizing flows [17, 30] trained on top of pretrained network features [42].

The specific components used in *PatchCore* are most related to SPADE and PaDiM. SPADE makes use of a memory-bank of nominal features extracted from a pretrained backbone network with separate approaches for image- and pixel-level anomaly detection. *PatchCore* similarly uses a memory bank, however with neighbourhood-aware patch-level features critical to achieve higher performance, as more nominal context is retained and a better fitting inductive bias is incorporated. In addition, the memory bank is coreset-subsampled to ensure low inference cost at higher performance. Coresets have seen longstanding usage in fundamental kNN and kMeans approaches [22] or mixture models [19] by finding subsets that best approximate the structure of some available set and allow for approximate solution finding with notably reduced cost [1, 9]. More recently, coreset-based methods have also found their way into Deep Learning approaches, e.g for network pruning [34], active learning [48] and increasing effective data