very strong external feature extractors, at minimal cost iv) it makes few assumptions on the images e.g. images can be rotation invariant, and of arbitrary size v) it is robust to anomalies in the training set i.e. it can handle the unsupervised case (when coupled with our two-stage approach) vi) it is plug and play, does not have a training stage.

Another contribution of our paper is presenting a novel adaptation of kNN to image group anomaly detection, a task that received scant attention from the deep learning community.

Although using kNN for anomaly detection is not a new method, it is not often used or compared against by most recent image anomaly detection works. Our aim is to bring awareness to this simple but highly effective and general image anomaly detection method. We believe that every new work should compare to this simple method due to its simplicity, robustness, low sample complexity and generality.

2. Previous Work

Pre-deep learning methods: The two classical paradigms for anomaly detection are: reconstruction-based and distribution-based. Reconstruction-based methods use the training set to learn a set of basis functions, which represent the normal data in an effective way. At test time, they attempt to reconstruct a new sample using the learned basis functions. The method assumes that normal data will be reconstructed well, while anomalous data will not. By thresholding the reconstruction cost, the sample is classified as normal or anomalous. Choices of different basis functions include: sparse combinations of other samples (e.g. kNN) (Eskin et al., 2002), principal components (Jolliffe, 2011; Candès et al., 2011), K-means (Hartigan & Wong, 1979). Reconstruction metric include Euclidean, L_1 distance or perceptual losses such as SSIM (Wang et al., 2004). The main weaknesses of reconstruction-based methods are i) difficulty of learning discriminative basis functions ii) finding effective similarity measures is non-trivial. Semi-supervised distribution-based approaches, attempt to learn the probability density function (PDF) of the normal data. Given a new sample, its probability is evaluated and is designated as anomalous if the probability is lower than a certain threshold. Such methods include: parametric models e.g. mixture of Gaussians (GMM). Non-parametric methods include Kernel Density Estimation (Latecki et al., 2007) and kNN (Eskin et al., 2002) (which we also consider reconstruction-based) The main weakness of distributional methods is the difficulty of density estimation for high-dimensional data. Another popular approach is one-class SVM (Scholkopf et al., 2000) and related SVDD (Tax & Duin, 2004). SVDD can be seen as fitting the minimal volume sphere that includes at least a certain percentage of the normal data points. As this method is very sensitive to the feature space, kernel methods were used to learn an effective feature space.

Augmenting classical methods with deep networks: The success of deep neural networks has prompted research combining deep learned features to classical methods. PCA methods were extended to deep auto-encoders (Yang et al., 2017), while their reconstruction costs were extended to deep perceptual losses (Zhang et al., 2018). GANs were also used as a basis function for reconstruction in images. One approach (Zong et al., 2018) to improve distributional models is to first learn to embed data in a semantic, low dimensional space and then model its distribution using standard methods e.g. GMM. SVDD was extended by Ruff et al. (2018) to learn deep features as a superior alternative for kernel methods. This method suffers from a "mode collapse" issue, which has been the subject of followup work. The approach investigated in this paper can be seen as belonging to this category, as classical kNN is extended with deep learned features.

Self-supervised Deep Methods: Instead of using supervision for learning deep representations, self-supervised methods train neural networks to solve an auxiliary task for which obtaining data is free or at least very inexpensive. It should be noted that self-supervised representation typically underperform those learned from large supervised datasets such as Imagenet. Auxiliary tasks for learning high-quality image features include: video frame prediction (Mathieu et al., 2016), image colorization (Zhang et al., 2016; Larsson et al., 2016), and puzzle solving (Noroozi & Favaro, 2016). Recently, Gidaris et al. (2018) used a set of image processing transformations (rotation by $0, 90, 180, 270$ degrees around the image axis), and predicted the true image orientation. They used it to learn high-quality image features. Golan & El-Yaniv (2018), have used similar image-processing task prediction for detecting anomalies in images. This method has shown good performance on detecting images from anomalous classes. The performance of this method was improved by Hendrycks et al. (2019), while it was combined with openset classification and extended to tabular data by Bergman & Hoshen (2020). In this work, we show that self-supervised methods underperform simpler kNN-based methods that use strong generic feature extractors on image anomaly detection tasks.

3. Deep Nearest-Neighbors for Image Anomaly Detection

We investigate a simple K nearest-neighbors (kNN) based method for image anomaly detection. We denote this method, Deep Nearest-Neighbors (DN2).