4 Experiments

We perform an extensive evaluation of our method against the state-of-the-art in sub-image anomaly detection.

4.1 MVTec

To simulate anomaly detection in industrial settings, [5] Bregmann et al. presented a dataset consisting of images from 15 different classes. 5 classes consist of textures such as wood or leather. The other 10 classes contain objects (mostly rigid). For each class, the training set is composed of normal images. The test set is composed of normal images as well as images containing different types of anomalies. This dataset therefore follows the standard protocol where no anomalous images are used in training. The anomalies in this dataset are more finegrained than those typically used in the literature e.g. in CIFAR10 evaluation, where anomalous images come from a completely different image category. Instead, anomalies take the form of a slightly scratched hazelnut or a lightly bent cable. As the anomalies are at the sub-image level, i.e. only affect a part of the image, the dataset provides segmentation maps indicating the precise pixel positions of the anomalous regions.

An example of the operation of our method on the MVTec dataset can be observed in Fig. 2. The anomalous object (a hazelnut) contain a scratch. The retrieved nearest neighbor normal image, contains a complete nut without scratches. By search for correspondences between the two images, our method is able to find correspondences for the normal image regions but not for the anomalous region. This results in an accurate detection of the anomalous region of the image. The anomalous images pixels are presented on the right-most image.

We compared our method against several methods that were introduced over the last several months, as well as longer standing baseline such as OCSVM and nearest neighbors. For each setting, we compared against the methods that reported the suitable metric.

We first compare the quality of deep nearest neighbor matching as a means for finding anomalous images. This is computed by the distance between the test image and the nearest neighbor normal images. Larger distances indicate more anomalous images. We compared the ROC area under the curve (ROCAUC) of the first step of our method and other state-of-the-art methods for image level anomaly detection. We report the average ROCAUC across the 15 classes. Please note that the first stage of our method is identical with DN2 [3]. This comparison is important as it verifies if deep nearest neighbors are effective on these dataset. The comparison is presented in Tab. 1. Our method is shown to outperform a range of state-of-the-art methods utilizing a range of self-supervised anomaly detection learning techniques. This gives evidence that deep features trained on ImageNet (which is very different from ImageNet dataset) are very effective even on datasets that are quite different from ImageNet.

We proceed to evaluate our method on the task of pixel-level anomaly detection. The objective here is to segment the particular pixels that contain anoma-