Report: Contrastive Patch Representation Learning for Industrial Anomaly Detection

Members

Ishan Kanodia - 22AE10014(Team Leader)
Divyanshu Bhatt - 22GG10017
Adeeba Alam Ansari -22IM30002
Satya Ravi Sujit - 21IM3FP10
Anumothu Ankit - 21MF10008

Divyanshu Bhatt - 22GG10017
G Sreekumar - 20GG20021

Abstract

Anomaly detection is essential for identifying defects in industrial manufacturing, such as incorrect parts, misaligned components, and damages. Traditional machine learning methods face challenges due to the rare occurrence and unknown nature of defects. ReConPatch, a novel approach, leverages contrastive patch representation learning to improve anomaly detection. By training a linear modulation of patch features extracted from a pre-trained model and using contrastive representation learning, ReConPatch creates a target-oriented and easily separable representation. This method achieves state-of-the-art performance on the MVTec AD and BTAD datasets without the need for carefully designed input augmentation.

Introduction

Anomaly detection plays a critical role in maintaining product quality in industrial manufacturing by identifying defects like incorrect parts, misaligned components, and damages. Machine learning approaches are widely used for this purpose, but they face significant challenges due to the rarity and unpredictability of defects. Traditional methods often rely on one-class classification, where the model learns to distinguish anomalies based on their distance from normal data. Reconstruction-based approaches, such as auto-encoders and generative adversarial networks (GANs), have been employed to detect anomalies by measuring reconstruction errors. However, these methods struggle with the limited variety of data, making it difficult to estimate a reliable nominal distribution from scratch.

Recent advancements have shown that leveraging visual representations pre-trained on natural image datasets can enhance anomaly detection performance. Despite their effectiveness, pre-trained models may not provide sufficiently distinguishable representations for subtle industrial defects due to the distribution shift between natural and industrial images. To address this issue, ReConPatch introduces a method that adapts features from a pre-trained model to better suit the target industrial dataset.

Related Work

Unsupervised machine learning approaches for anomaly detection using neural networks have been extensively studied. Deep Support Vector Data Description (SVDD) and its patch-wise extension, Patch SVDD, enhance localization and enable fine-grained examination by mapping data to a hyperspherical embedding. Reconstruction-based approaches assume that normal data can be accurately reconstructed, while abnormal data cannot, using models like auto-encoders and GANs. These methods calculate an anomaly score based on the reconstruction error.

To mitigate the lack of data variety, several approaches use pre-trained models from rich natural image datasets. These models measure the distance between the input data's representations and their nearest neighbours or compare hierarchical sub-image features to detect anomalies. DifferNet employs a normalising flow to map pre-trained feature distributions to a nominal distribution, helping identify anomalies. PatchCore defines a corset using locally aware patch features and greedy subsampling, while CFA and PaDiM use Gaussian distributions and patch descriptors for anomaly detection.

Methodology

ReConPatch focuses on learning a representation space that groups similar nominal features closely while spreading out different features. The framework consists of training and inference phases, utilising patch-level features extracted from a pre-trained CNN model. Features are aggregated using adaptive average pooling, and two networks are employed for representation learning: one for learning patch-level features and another for calculating pairwise and contextual similarities.

Training Phase

During training, patch-level features are extracted and aggregated from the CNN model. ReConPatch uses contrastive learning with pairwise and contextual similarities as pseudo-labels to train the feature representation network. The pairwise similarity measures the Gaussian kernel similarity between features, while contextual similarity considers the neighbourhood of an embedding vector. By combining these similarities, the method creates pseudo-labels for training the representation network using a relaxed contrastive loss. The similarity calculation network is updated using an exponential moving average (EMA) to ensure stable training.

Inference Phase

In the inference phase, features of a test sample are extracted and compared to the nominal representative features stored in the memory bank. The anomaly score is calculated based on the distance between the test sample's features and the nearest representative feature. This approach allows for effective anomaly detection without the need for extensive handcrafted input augmentation.

Experimental Results

ReConPatch was evaluated on the MVTec AD and BTAD datasets, achieving state-of-the-art performance with accuracy rates of 99.72% and 95.8%, respectively. These results demonstrate the method's effectiveness in detecting anomalies in various industrial settings without requiring carefully designed input augmentation.

Experimental Benchmarking

Notebook to run the code - Google Colab
Accelerator - T4 GPU

We benchmarked the Model on Two datasets - MVTecAD and BTAD

The model is based on Image Label classification wherein the category labels are obtained from the names of the directories. As part of the BTAD dataset, which consisted of images of three industrial components, normal images and defective images were provided. Assuming that there is a ground truth available for defective images, we used it as our starting point.

In total, we trained the model for 25 epochs with training losses at each epoch for the following parameters:

Dataset for training - **901.2882690429688**

It has been found that validating the data for 25 epochs results in the following result: - : **899.5626831054688**

Results:

-> A ratio of **1.0019182498006063** was reported as the ratio of training/validation loss for BTAD dataset.

-> A ratio of **1.0022420649269164** was reported as the ratio of training/Validation loss for MVTec dataset.

For MVTec dataset

```
1/1 [==============================] - 0s 32ms/step - loss: 186895.2344
Validation Loss: 186895.234375

[ ] print(train_loss/val_loss)

1.0022420649269164
```

For BTAD dataset

```
... 3/3 [==============================] - 0s 72ms/step - loss: 899.5627
Validation Loss: 899.5626831054688

print(train_loss/val_loss)
                                                                 Python
... 1.0019182498006063
```

Future Plans - :

We intend to find the AUROC for this model and apply more visualisation metrics.We intend to tune the hyperparameters of the model to improve its performance.

Conclusion

ReConPatch offers a practical and effective solution for industrial anomaly detection by leveraging contrastive patch representation learning. By adapting features from a pre-trained model to better suit the target dataset and using contrastive learning with pseudo-labels, ReConPatch achieves high anomaly detection accuracy. This method addresses the challenges posed by the rarity and unpredictability of defects, providing a robust tool for maintaining product quality in industrial manufacturing.

References

1. Hyun, J., Kim, S., Jeon, G., Kim, S. H., Bae, K., & Kang, B. J. (2024). ReConPatch: Contrastive Patch Representation Learning for Industrial Anomaly Detection. LG AI Research.
2. Ruff, L., Vandermeulen, R. A., Göpfert, J., Deecke, L., Siddiqui, S. A., Binder, A., ... & Kloft, M. (2020). Deep Semi-Supervised Anomaly Detection. arXiv preprint arXiv:2002.10445.
3. Cohen, I., & Papandreou, G. (2020). PatchCore: Local Feature Aggregation for Robust Anomaly Detection. arXiv preprint arXiv:2011.08545.
4. Davy, A., Greben, J., Wiatr, T., & Douak, A. (2021). PaDiM: A Patch Distribution Modeling Framework for Anomaly Detection and Localization. arXiv preprint arXiv:2011.08785.