Anomaly Classification with Anomaly-Focused Patch Selection by Gaussian Distribution

Yuga ONO
Graduate School of Science and Engineering, Ritsumeikan University
1-1-1 Noji-higashi, Kusatsu, Shiga 525-8577 Japan
ri0098ss@ed.ritsumei.ac.jp

Lin MENG
Department of Electronic and Computer Engineering, Ritsumeikan University
1-1-1, Noji-higashi, Kusatsu, Shiga 525-8577 Japan
menglin@fc.ritsumei.ac.jp

***Abstract*—Numerous methods for detecting and localizing anomalies have been proposed, and many have achieved great success. In practical applications, various factors can lead to different types of anomalies, each of which requires specific treatment. Therefore, it is crucial to not only detect anomalies but also to identify their specific types. This paper builds upon an existing anomaly detection method, using it as a foundational model, and extends its application toward anomaly classification. We present a novel approach aimed at efficiently and accurately identifying the type of anomaly by leveraging patch embeddings and the anomaly score obtained during the initial anomaly detection stage. Our paper introduces Anomaly-Focused Patch Selection (AFPS), which is a unique mechanism that helps select more meaningful patches for training a classification model. AFPS demonstrates superior classification accuracy with 75% fewer number of patches compared to the base method, which simply employs random patch selection.**

***Index Terms*—Deep Learning, Image Recognition, Anomaly Detection, Anomaly Classification,**

I. Introduction

Numerous methods for detecting and localizing anomalies have been proposed, and many have achieved great success. In practical applications, various factors can lead to different types of anomalies, each of which requires specific treatment. This paper aims to not only perform anomaly detection but also anomaly classification, which help us identify the specific types of detected anomalies.

We propose a method to efficiently and accurately identify the type of anomaly by effectively using the patch embeddings and anomaly scores obtained at the anomaly detection stage. Our paper introduces Anomaly-Focused Patch Selection (AFPS), which is a unique mechanism that helps select more meaningful patches for training a classification model. Our method consists of two main stages, anomaly detection, and anomaly classification. The first stage, anomaly detection, uses a pre-trained network to extract features from input images and obtain patch embeddings. The second stage, anomaly classification, uses Vision Transformer as a classification model, which receives patch embeddings as training data to classify anomaly images into two anomaly classes. More detailed information about our classification method is shown below.

Patch Distribution Modeling

Feeding a set of normal images into a pre-trained model and extracting features of them, which is used for modeling the patch distribution of normal images.

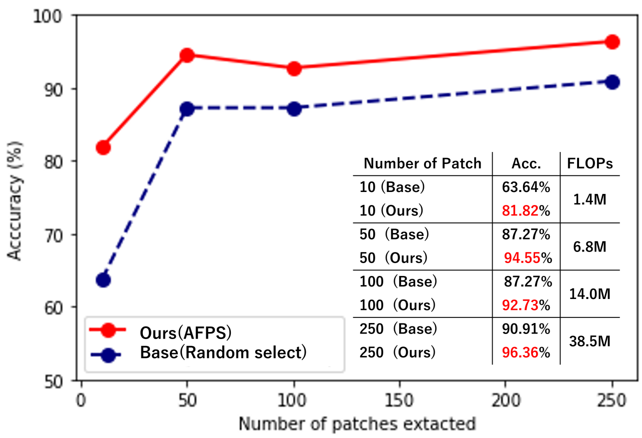

Fig. 1: Classification performances by the number of the patches selected from the original patch embeddings obtained at the anomaly detection stage. Our patch selection method AFPS outperforms the base method, which simply employs random patch selection, regardless of the number of patches selected for training data.

Anomaly Detection

Based on the mean and covariance of the learned distribution, Calculate the anomaly score for each patch in the input image. anomalous area in an image, where its features are different from normal ones, gives a high anomaly score.

Anomaly Classification

We further narrow down the patches for anomaly classification. The patches with high anomaly scores are selected based on anomaly-focused patch selection(AFPS), which allows the classification model to better focus on anomalous areas, leading to an improvement in classification accuracy. The model learns differences between types of anomaly images and classify them into predefined anomaly classes.

II. Related Works

In the past few years, many types of approach based on deep learning for anomaly detection has been made and they brought great success. There are two typical ways for anomaly detection: Reconstruction-based method and Feature extraction-based method.

*1) **Reconstruction-based**:* The reconstruction-based method is widely used for anomaly detection and localization. This method features identifying anomalies by reconstructing input data using a machine learning model, typically autoencoders(AE) [1], variational autoencoders (VAE) [2], [3] or similar architecture. A model is trained to reconstruct input data. Reconstruction error is calculated for each data point by comparing the original input data with the model's reconstructed output. The key idea of this method is that normal data points can be accurately reconstructed by the model, whereas anomalies might bring higher reconstruction errors. During inference, if the reconstructed image differs significantly from the original, it may indicate the presence of an anomaly. GANs [4], [5] can be adapted for anomaly detection by training them to generate only normal data. Anomalies are detected when the generated data significantly deviates from the real data distribution. In addition to CNN based model, Vision Transformer(ViT) [6] has also become a common choice for reconstruction-based anomaly detection [7], [8].

*2) **Embedding Similarity-based**:* This approach uses a deep neural network to extract meaningful features from the normal images for anomaly detection and localization [9], [10]. Since this approach uses a pre-trained model for feature extraction, it does not require parameters update in the model, which is a great advantage of this approach. The anomaly score is calculated based the comparison between the extracted features of the test image and the normal image distribution obtained in the training stage. PaDiM [11] is an outstanding embedding similarity-based method, which has an important role in our anomaly classification method. PaDiM uses Resnet18 [12] or Wide-Resnet50 [13] as a feature extraction model. Features are extracted from three different layers to obtain both local and global contexts. After that, all extracted features are concatenated to get the final embedding vectors.

III. Methodology

Our paper proposes an effective method to perform the anomaly classification task by effectively using features that are used for the anomaly detection phase. In real applications, different matters can cause different types of anomalies., which require an appropriate treatment for each type of anomaly. Therefore, performing anomaly classification as well as anomaly detection can help many types of anomalies to be addressed with better treatments. Fig.5 shows overview of anomaly detection, classification, and Anomaly-Focused Patch Selection. The anomaly detection stage is based on PaDiM, an existing amazing method. This stage is essential to precisely detect the location of the anomaly. It is also instrumental for the classification model to be able to focus on the area of the anomaly and its vicinity at the later anomaly classification stage.

A. Embedding creation for anomaly detection

An existing anomaly detection called PaDiM is used for this phase. Firstly, we obtain the distribution of normal data from some images that have no anomalies. During the inference phase, the model receives an input image with the size of $M \times M$ and gives back extracted features with the size of $(H \times W \times D)$. After this process, those features are analyzed in comparison with the learned distribution of normal data, and the Mahalanobis distance is calculated for each patch. Mahalanobis distance for a single patch $x(i,j)$ is computed with $N(\mu_{ij}, \Sigma_{ij})$ as follows:

$$M(x_{ij}) = \sqrt{(x_{ij} - \mu_{ij})^T \Sigma_{ij}^{-1} (x_{ij} - \mu_{ij})} \tag{1}$$

B. Anomaly-Focused Patch Selection

Fig.3 shows the overview of our patch selection method. Patch embeddings have their own anomaly score for each patch computed based on Mahalanobis distance which indicates how each patch is different from normal data. Patch embeddings consisting of both normal and anomalous include a significant amount of redundant information, which can increase computational costs. We attempt to narrow down the number of patches in order to reduce their redundancy and extract meaningful ones for accurate anomaly classification. We propose an effective patch selection mechanism with Gaussian distribution. They are selected stochastically with Gaussian distribution, whose mean value is set to the patch with the highest anomaly score. This method extracts more anomalous patches located close to the mean and less normal patches located distant from the central area. This effective patch selection mechanism can also result in drastic savings of computational resources. We test our patch selection mechanism with 10, 50, 100 and 250 of selected patches respectively to investigate how the number of patches can affect the training and classification performance. Moreover, to prove the effectiveness of our AFPS, we also test random patch selection which randomly selects patches from the entire set of original patches.

C. classification by ViT

The set of patches selected based on AFPS is then used as training data for the classification model. We utilize a simple Vision Transformer(ViT). The transformer encoder has multiple layers, each containing self-attention mechanisms and feed forward neural networks. Self-attention helps obtain global relationships between tokens. The final token embeddings, after passing through the transformer encoder, are used for classification tasks using fully connected layers. As shown in Fig.4, the dataset has two anomaly classes, Spot and Scratch.

IV. Evaluation

A. Experimental Configurations

1) Implementation Conditions: Experiments in this paper are conducted on Intel(R) Core(TM) i7-13700KF and an

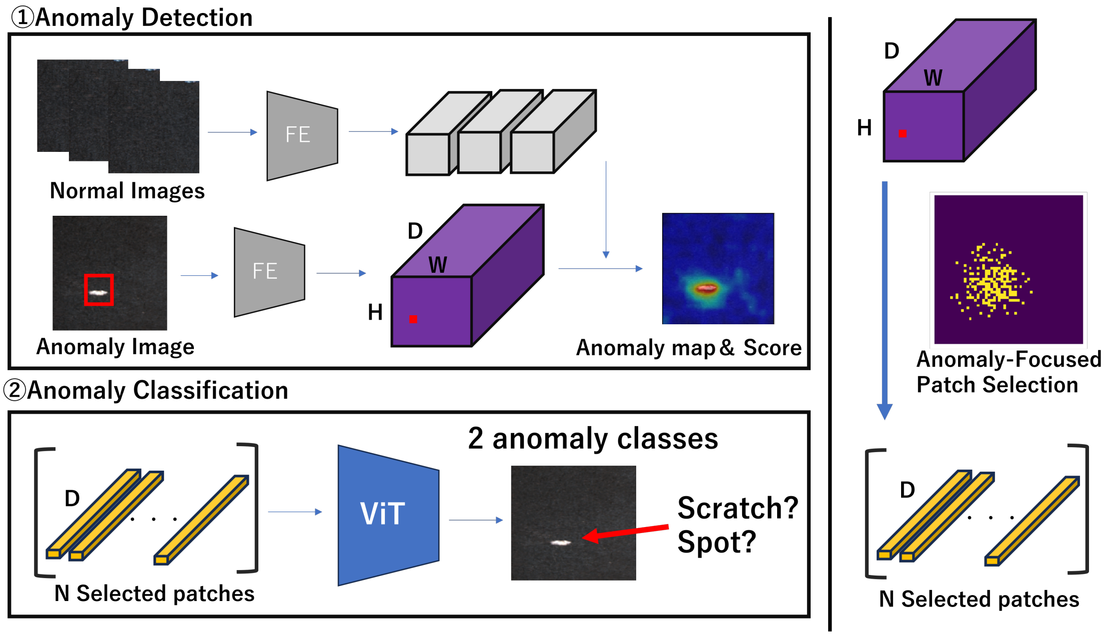

Fig. 2: Overview of our method for anomaly detection and classification. As shown in the left figure, anomaly detection stage is entirely based on PaDiM.The right figure explains how anomaly-focused patch selection(AFPS) works to prepare training data for the classification model.

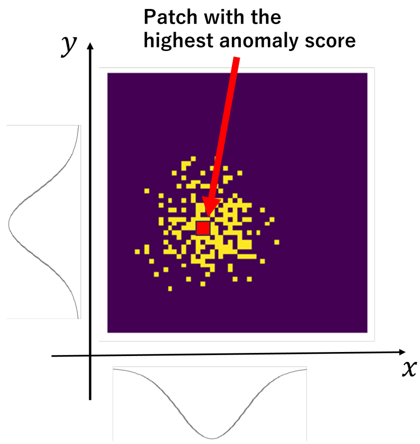

Fig. 3: Yellow dots denote patches that are stochastically selected with Gaussian distribution, whose mean value is set to the patch with the highest anomaly score.

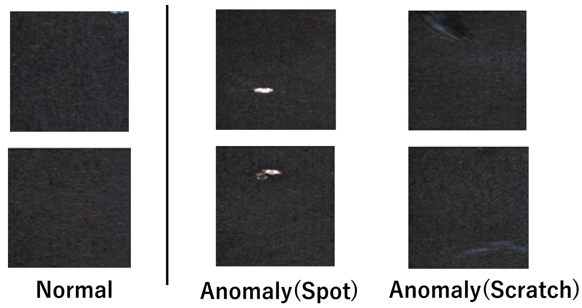

Fig. 4: The set of normal images are used for patch distribution modeling at anomaly detection stage. Anomaly images with two anomaly classes (Spot and Scratch) are used for training and validation data at anomaly classification stage.

NVIDIA GeForce RTX 4090 GPU. The operating system is Ubuntu 22.04.1 LTS, the CUDA version is 11.8, and the GPU acceleration library cuDNN is 8.7.0. PyTorch library is used as the framework for implementing deep learning. In all experiments, the learning rate for the ViT is set to 1e-3 with cosineLRscheduler that drops the learning rate according to pre-defined epochs Adam is used as an optimizer in all experiments.

2) Dataset: Our paper uses Kolektor Surface-Defect Dataset2 (KSDD2) provided by Visual Cognitive Systems Laboratory in University of Ljubljana. [14] The dataset consists of 356 images with visible defects and 2979 images without any anomaly. It has several different types of defects. The original image size is 630 paper uses 1000 normal images to prepare the distribution of normal images and 171 anomaly images with two anomaly classes (scratch and spot) for the test data of the anomaly detection phase. In anomaly classification stage, the set of anomaly images is further divided into train data and test data and used to train and evaluate the model.

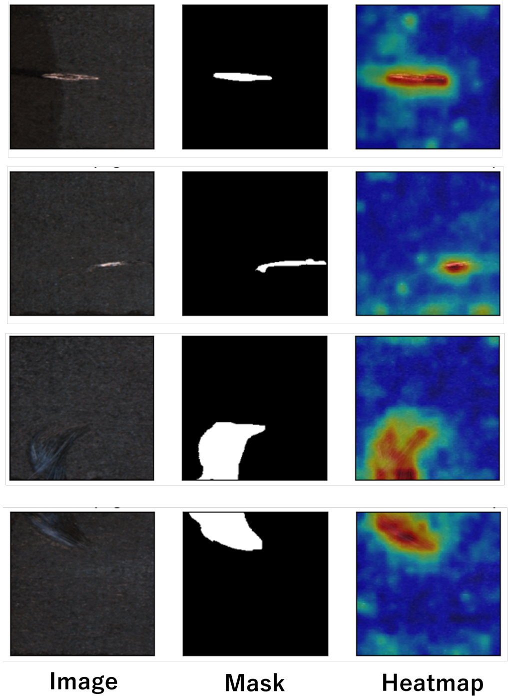

Fig. 5: Detection results on anomaly anomaly detection stage by PaDiM. Both ”Scratch” and ”Spot” are accurately detected and their anomalous areas are highlighted with red.

3) Detail of Models: Pre-trained ResNet-50 [12] is used as feature extraction model in anomaly detection phase. In anomaly classification, Vision Transformer with 8 heads, 4 layers of encoder and decoder blocks is used in all of our experiments. The number of dimensions of input vectors is 64. Thr number of queries varies according to the number of selected patches by AFPS.

4) Evaluation Index: Our patch selection method and classification performance are evaluated with the following metrics, Accuracy, Recall, Precision, F_1 and FLOPs We test our AFPS with 10 50 100, and 250 of selected patches respectively to investigate how the number of patches can affect the training and classification performance. We also perform our AFPS with standard deviation values of 0.1, 0.15, 0.20, 0.25, 0.30.

B. Experimental Results

TableI shows the classification results by the number of patches used as input data to the classification model. The figures in the parentheses

Regardless of the number of selected patches, our AFPS brings better classification performance than the random patch selection. When 250 patches are selected based on AFPS, the model achieves the best accuracy 96.4%, which outperforms the score of the base method by 5%. Also, the results show that in most cases a larger number of patches can yield better evaluation scores.

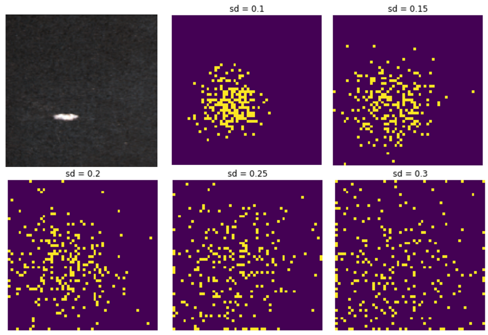

Fig. 6: Simle diagram of anomaly-focused patch selection(AFPS) by a standard deviation value. The location of the patch with the highest anomaly score is set to the mean value. The higher the standard deviation is, the more patches are selected from distant areas.

FLOPS increases in proportion to the number of patches. We further investigate which standard deviation value of the normal distribution can bring better accuracy scores. TableI shows that the standard deviation of 0.15 gives the best accuracy score of 98%.

V. Discussion

A. AFPS vs Random Patch Selection

Comparing the results of Anomaly-Focused Patch Selection (AFPS) with those of the base method (random patch selection), we prove that AFPS surpasses the base method in every evaluation metric and in the effectiveness of the patch selection mechanism. Moreover, using only 50 patches with our method brings the much higher accuracy score than using much more patches of 250 with the base method. This is because our patch selection method strongly focuses on which patch is more significant for anomaly classification. The base method(random patch selection) does not take the location of anomalies into account, which can bring more unnecessary patches for accurate classification. On the other hand, AFPS is designed to extracts more patches around the most anomalous patch and fewer patches distant from the most anomalous patch. This enables the model to concentrate on anomaly classification without any unnecessary information.

B. The effects of standard deviation values

TableI shows that the patches selected with a standard deviation value of 0.15 give the best accuracy score of 98%. As shown in Fig.6, yellow dots are selected patches under the condition of the given standard deviation value. The results show that 0.15 is just the right value to obtain the most effective set of patches for accurate classification. In contrast, the value of 0.1 brings the poor accuracy of 85.5%. This is attributed to the smaller range of patch selection. Patches selected in too small range prevent the ViT model from learning the relationship between the patches.

TABLE I: Main Results

The number of patches	Method	Accuracy	Recall	Precision	F_1	FLOPs
10 (0.3%)	Base	63.4	62.7	63.3	68.8	1.4M
	Ours	**81.2**	**82.0**	**81.2**	**82.7**	1.4M
50 (1.5%)	Base	87.3	87.0	87.3	88.5	6.8M
	Ours	**94.5**	**95.0**	**94.7**	**94.7**	6.8M
100 (3%)	Base	85.5	84.7	86.5	87.5	14.0M
	Ours	**92.7**	**92.4**	**93.2**	**93.6**	14.0M
250 (7.5%)	Base	90.9	90.0	92.85	92.3	38.5M
	Ours	**96.4**	**96.4**	**96.4**	**96.7**	38.5M

TABLE II: Accuracy by the standard deviation value

sd	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**
Accuracy	85.5	**98.2**	92.7	85.5	90.9

VI. Conclusion

We have proposed the effective anomaly classification method based on anomaly-focused patch selection (AFPS). Our method makes use of the output obtained at anomaly detection stage for accurate anomaly classification. We have proved that our AFPS surpasses the base method in every evaluation metric and in the effectiveness of our patch selection mechanism. Our future work is to propose a more efficient patch selection method to parform more accurate anomaly classfication with smaller number of patches and to consolidate two stages of our method into one practical stage for real applications.

References

[1] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Mvtec ad–a comprehensive real-world dataset for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9592–9600, 2019.

[2] Shashanka Venkataramanan, Kuan-Chuan Peng, Rajat Vikram Singh, and Abhijit Mahalanobis. Attention guided anomaly localization in images, 2020.

[3] Adrian Alan Pol, Victor Berger, Gianluca Cerminara, Cecile Germain, and Maurizio Pierini. Anomaly detection with conditional variational autoencoders, 2020.

[4] Thomas Schlegl, Philipp Seeböck, Sebastian M. Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, 2017.

[5] Samet Akcay, Amir Atapour-Abarghouei, and Toby P. Breckon. Ganomaly: Semi-supervised anomaly detection via adversarial training, 2018.

[6] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale, 2021.

[7] Pankaj Mishra, Riccardo Verk, Daniele Fornasier, Claudio Piciarelli, and Gian Luca Foresti. VT-ADL: A vision transformer network for image anomaly detection and localization. In *2021 IEEE 30th International Symposium on Industrial Electronics (ISIE)*. IEEE, jun 2021.

[8] Yunseung Lee and Pilsung Kang. Anovit: Unsupervised anomaly detection and localization with vision transformer-based encoder-decoder, 2022.

[9] Oliver Rippel, Patrick Mertens, and Dorit Merhof. Modeling the distribution of normal data in pre-trained deep features for anomaly detection, 2020.

[10] Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, and Peter Gehler. Towards total recall in industrial anomaly detection, 2022.

[11] Thomas Defard, Aleksandr Setkov, Angelique Loesch, and Romaric Audigier. Padim: a patch distribution modeling framework for anomaly detection and localization, 2020.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[13] Sergey Zagoruyko and Nikos Komodakis. Wide residual networks, 2017.

[14] Jakob Božič, Domen Tabernik, and Danijel Skočaj. Mixed supervision for surface-defect detection: from weakly to fully supervised learning. *Computers in Industry*, 2021.