



# Deep Anomaly Detection via Morphological Transformations <sup>†</sup>

Taehyeon Kim and Yoonsik Choe \*

Department of Electrical & Electronic Engineering, Yonsei University, Seoul 03722, Korea; pyomu@yonsei.ac.kr

\* Correspondence: yschoe@yonsei.ac.kr; Tel.: +82-2-2123-2774

<sup>†</sup> Presented at 1st International Electronic Conference on Applied Sciences, 10–30 November 2020;

Available online: <https://asec2020.sciforum.net/>.

Published: 11 November 2020

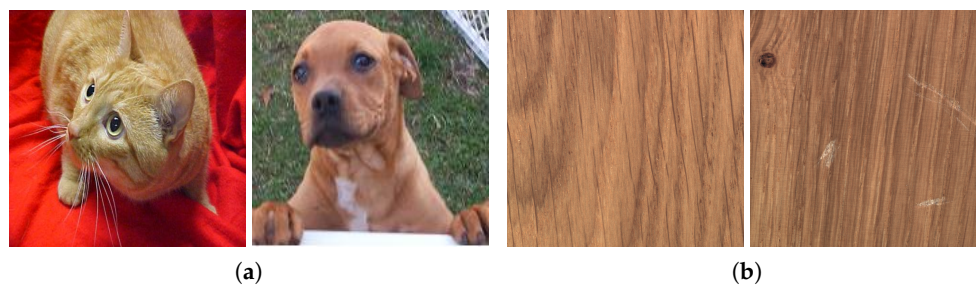


**Abstract:** The goal of deep anomaly detection is to identify abnormal data by utilizing a deep neural network trained by a normal training dataset. In general, industrial visual anomaly detection problems distinguish normal and abnormal data through small morphological differences, such as cracks and stains. Nevertheless, most existing algorithms focus on capturing not morphological features, but semantic features of normal data. Therefore, they yield poor performance on real-world visual inspection, even though they show their superiority in simulations with representative image classification datasets. To solve this problem, we propose a novel deep anomaly detection method that encourages understanding of salient morphological features of normal data. The main idea behind our algorithm is to train a multi-class model to classify between dozens of morphological transformations applied to all the given data. To this end, the proposed algorithm utilizes a self-supervised learning strategy, which makes unsupervised learning straightforward. Additionally, we present a kernel size loss to enhance the proposed neural networks' morphological feature representation power. This objective function is defined as the loss between predicted kernel size and label kernel size via morphologically transformed images with the label kernel. In all experiments on the industrial dataset, the proposed method demonstrates superior performance. For instance, in the MVTec anomaly detection task, our model achieved an area under the receiver operating characteristic (AUROC) value of 72.92%, which is 8.74% higher than the semantic-feature-based deep anomaly detection.

**Keywords:** anomaly detection; self-supervised learning; morphological transformation

## 1. Introduction

Deep anomaly detection means verifying abnormal data via a deep neural network trained by normal instances. It is a significant challenge that has been well studied within various application domains, including video surveillance, disease diagnosis, and visual inspection. In this paper, we tackle the problem of deep anomaly detection in images. The intuition behind most existing methodologies in this problem is training the deep neural network to understand semantically important features of normal data. Hence, most of these studies [1–3] reported their superior results on representative image classification datasets (e.g., MNIST [4] and CIFAR-10 [5]) that are composed of clearly distinguishable classes. However, from the point of view of industrial inspection, these existing methodologies are not useful for solving real-world problems. In a real-world problem, the criterion that discriminates abnormal data from normal data is usually defined as morphological differences, such as cracks, stains, and bends, which cannot be described semantically. For ease of understanding, the visual descriptions of both semantic and morphological differences are shown in Figure 1.



**Figure 1.** The visual descriptions of semantic and morphological differences in images. (a) Semantic difference: Both figures are sampled instances of the cats and dogs [6]. The difference between the “cat” and “dog” classes is called a semantic difference. Generally, a semantic difference involves both semantic and morphological differences. (b) Morphological difference: Both figures are sampled instances from the MVTec dataset [7]. The difference between the “good wood” and “scratched wood” classes is called a morphological difference. A morphological difference does not involve a semantic difference. In other words, instances of both “good wood” and “scratched wood” have the same semantic definition.

In order to utilize morphological features in deep anomaly detection, the proposed method is based on a self-supervised learning algorithm. Self-supervised learning is a form of unsupervised learning where the training data provide the supervision. There is a proxy loss in this learning mechanism that makes the deep neural network achieve the main goal of target application. In other words, by utilizing this training algorithm, the deep neural network can learn what we care about, such as semantic differences or morphological differences. There have been several previous methods in self-supervised-learning-based deep anomaly detection [2,8]. These existing methods focused on training deep neural networks to understand the geometric transformations of normal data, including rotation and translation. In particular, training a deep neural network to classify the rotation degree of normal data is an effective strategy for capturing semantic information of normal data [8]. Obviously, training geometric transformations in self-supervised learning does not help identify abnormal data in the case represented in Figure 1b.

To mitigate this problem, we propose a novel deep anomaly detection algorithm based on self-supervised learning using morphological transformations, including dilation, erosion, and morphological gradient. The proposed method is based on the observation of an industrial anomaly detection problem, which requires a morphological understanding of normal data. Therefore, the proposed method is trained over a self-labeled dataset, which is constructed by the normal instances and their morphologically transformed variants, accomplished by various morphological transformations. At the test procedure, the trained neural network takes input on morphologically transformed test data, and the distribution of softmax activations on trained normal data is useful for detecting abnormal test data. The intuition behind the proposed method is that by training the classifier to discriminate between transformed images, it has to learn valuable morphological features.

In this paper, we performed deep anomaly detection experiments based on the MVTec dataset [7], which was created to measure anomaly detection performance in industrial inspection. There are various industrial defection types (e.g., cracks, stains, bends) per class in this dataset. Additionally, to demonstrate the superior performance of the proposed algorithm in the industrial aspect, we compared it with the latest state-of-the-art deep anomaly detection based on self-supervised learning [2].

In summary, the main contributions of this study are as follows:

- The proposed method achieves superior performance in deep anomaly detection for industrial inspection by training the deep neural network to capture salient morphological features of normal data.

- The proposed algorithm can flexibly adapt to various real-world deep anomaly detection problems by choosing the adequate morphological transformations in image processing technology.
- Because the proposed methodology utilizes self-supervised learning, it has lower computational complexity than other deep anomaly detection methods, such as reconstruction-based algorithms.

## 2. Proposed Method

This section describes the morphological-transformation-based deep anomaly detection algorithm, which is applied to industrial and real-world anomaly detection problems.

### 2.1. Morphological Image Processing

In digital image processing, a mathematical morphology transformation is a mechanism for extracting image components that are useful in representing and describing region shapes, such as boundaries, skeletons, and convex hulls [9]. The proposed deep anomaly detection learns the morphological features through three representative morphological transformations, including erosion, dilation, and morphological gradient, which are described in the following subsections.

#### 2.1.1. Erosion and Dilation

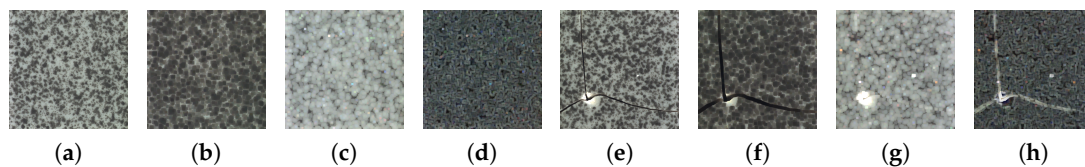
The erosion at any location  $(x, y)$  of image  $i$  by a kernel  $b$  is the minimum value of  $i$  in the region covered by  $b$  when the central point (origin) of  $b$  is at  $(x, y)$ . For instance, if  $b$  is a  $3 \times 3$  kernel, obtaining the erosion at a pixel requires getting the minimum of the nine values of  $i$  included in the  $3 \times 3$  region determined by the kernel when its origin is at that point. In equation form, the erosion is defined as:

$$[i \ominus b](x, y) = \min_{(s,t) \in b} i(x + s, y + t). \quad (1)$$

Likewise, the dilation of  $i$  by  $b$  is designated as the maximum value of  $i$  from all the values of  $i$  contained in the region coincident with  $b$ . That is,

$$[i \oplus b](x, y) = \max_{(s,t) \in b} i(x + s, y + t). \quad (2)$$

Because erosion computes the minimum pixel value of  $i$  in every neighborhood of  $(x, y)$  that is coincident with  $b$ , it is expected that the size of bright features in  $i$  will be reduced, and the size of dark features will be increased. Figure 2b,f shows eroded images of normal and abnormal data in the “tile” class of MVTec, respectively. As mentioned above, from these figures, it can be seen that the area of the dark features is increased in the eroded examples. Similarly, Figure 2c,g shows the results of dilation. The effects are the opposite of those obtained with erosion. The bright features were thickened, and the intensities of the dark features were decreased.



**Figure 2.** Morphologically transformed images in the “tile” class of MVTec [7]: (a) normal image; (b) eroded normal image; (c) dilated normal image; (d) morphological gradient of a normal image; (e) abnormal image; (f) eroded abnormal image; (g) dilated abnormal image; (h) morphological gradient of an abnormal image.

### 2.1.2. Morphological Gradient

To obtain the morphological gradient of an image, dilation and erosion can be used in combination with image subtraction. In this paper, this operation is denoted as follows:

$$i \odot b = (i \oplus b) - (i \ominus b). \quad (3)$$

Because the dilation thickens regions in an image and the erosion shrinks them, the difference between them highlights the boundaries between areas. Therefore, an image in which the edges are emphasized and the homogeneous regions are suppressed has a “derivative-like” (gradient) effect. Figure 2d,h shows morphological gradient images of normal and abnormal data, respectively. Especially in Figure 2h, it can be seen that this morphological transformation emphasizes the cracked area.

### 2.2. Deep Anomaly Detection via Morphological Transformations

The proposed algorithm aims to train the deep neural network with the morphological features of normal data through a self-supervised learning strategy. To achieve this goal, we propose the training of a deep neural network  $F$  to discriminate between the morphological transformation types applied to an image that is given as input. Specifically, we define a set of  $N_1$  discrete morphological transformations,  $N_2$  discrete values for kernel width, and  $N_3$  discrete values for kernel height. In other words, the proposed self-labeled dataset is a multi-class dataset that consists of  $N_1 N_2 N_3$  classes. For clarification, we denote a kernel  $b$  of size  $n_2 \times n_3$  as  $b_{n_2, n_3}$ . Thus, we define a set of  $N_1 N_2 N_3$  discrete morphological transformations as follows:

$$G = \{g(\cdot | n_1, n_2, n_3)\}_{n_1=1, n_2=1, n_3=1}^{N_1, N_2, N_3} \quad (4)$$

where  $g(\cdot | n_1, n_2, n_3)$  denotes application to image  $i$  of the morphological transformation with multi-class label  $\{n_1, n_2, n_3\}$ , which produces the transformed image  $i^{n_1, n_2, n_3} = g(i | n_1, n_2, n_3)$ .

The deep neural network  $F$  takes a transformed image  $i^{n_1^*, n_2^*, n_3^*}$  (where the label  $\{n_1^*, n_2^*, n_3^*\}$  is unknown to  $F$ ) an input. After that, it produces a probability distribution of softmax responses over all possible morphological transformations, which is denoted as follows:

$$F(i^{n_1^*, n_2^*, n_3^*} | \theta) = \{F^{n_1, n_2, n_3}(i^{n_1^*, n_2^*, n_3^*} | \theta)\}_{n_1=1, n_2=1, n_3=1}^{N_1, N_2, N_3} \quad (5)$$

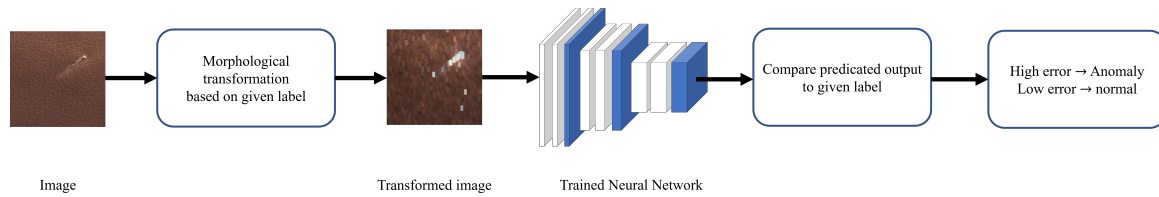
where  $F^{n_1, n_2, n_3}(i^{n_1^*, n_2^*, n_3^*} | \theta)$  is the predicted probability for morphological transformation with  $\{n_1^*, n_2^*, n_3^*\}$  and  $\theta$  denotes the parameters of  $F$ .

Consequently, the proposed objective function is as follows:

$$\min_{\theta} \frac{1}{3T} \sum_{j=1}^T \left( -\frac{1}{N_1} \sum_{n_1=1}^{N_1} \log(F^{n_1}(i^{n_1^*, n_2^*, n_3^*} | \theta)) - \frac{1}{N_2} \sum_{n_2=1}^{N_2} \log(F^{n_2}(i^{n_1^*, n_2^*, n_3^*} | \theta)) - \frac{1}{N_3} \sum_{n_3=1}^{N_3} \log(F^{n_3}(i^{n_1^*, n_2^*, n_3^*} | \theta)) \right), \quad (6)$$

where  $F^{n_1}(i^{n_1^*, n_2^*, n_3^*} | \theta)$ ,  $F^{n_2}(i^{n_1^*, n_2^*, n_3^*} | \theta)$ , and  $F^{n_3}(i^{n_1^*, n_2^*, n_3^*} | \theta)$  denote the predicted probability for  $n_1^*$ ,  $n_2^*$ , and  $n_3^*$ , respectively. Through the above formulation, we force the deep neural network to learn the morphological features of normal images by simultaneously predicting both transformation type and kernel size. Specifically, training to predict kernel size encourages the proposed algorithm to learn useful morphological features in real-world industrial deep anomaly detection. In Figure 3, the overall architecture of the proposed method is presented.





**Figure 3.** The proposed deep anomaly detection aims to distinguish abnormal data using the morphological features of normal data acquired during the training procedure. Therefore, if a given morphologically transformed datum generates a high prediction error, it can be considered abnormal.

### 3. Experimental Results

In this section, the deep anomaly detection experiments were performed to verify the performance in industrial inspection. In addition, to show the superiority of the proposed method over the existing algorithm, which is designed to learn the semantic features of normal data, a performance comparison with [2] is reported. The backbone model of the proposed method is ResNet-18 [10]. In the experimental results, there are three types of the proposed method for verifying the influence of the kernel size learning: type 1:  $n_2 \in \{1, 28, 56\}, n_3 \in \{1, 28, 56\}$ , type 2:  $n_2 \in \{8, 28, 56\}, n_3 \in \{8, 28, 56\}$ , and type 3:  $n_2 \in \{1, 8, 28, 56\}, n_3 \in \{1, 8, 28, 56\}$ . The proposed algorithm was actualized using PyTorch in a GPU implementation [11]. We performed experiments with an RTX 2080Ti 11GB GPU and an Intel i7 CPU.

#### Deep Anomaly Detection on Industrial Dataset

In Table 1, we present the overall experimental results of the proposed method on the representative industrial anomaly dataset, MVTec [7]. From the experimental results, it can be verified that the proposed self-supervised learning designed to capture salient features of normal data achieves superior performance compared to the semantic-feature-based deep anomaly detection. Interestingly, in a performance comparison experiment among the proposed method's three types, although the type 1 case model achieved faster convergence than the other cases, it produced the lowest performance. This observation implies that creating an easily self-labeled dataset in self-supervised learning does not help to lead the deep neural network to where we intended. This phenomenon was proved inductively through the experimental results of the type 3 case. These overall experimental results prove that utilizing morphological image features improves performance in real-world industrial problems. The proposed method can also verify anomalies by inferencing a neural network, which takes a processing time of almost 0.0125 s. In other words, it has low computational complexity.

**Table 1.** Comparison of the area under the receiver operating characteristic (AUROC, %) performance between [2] and the proposed algorithm.

Class	Bottle	Cable	Capsule	Carpet	Grid	Hazelnut	Leather
[2]	83.10	77.81	75.31	38.12	<b>31.47</b>	67.14	64.10
Our type 1	87.86	76.89	<b>77.50</b>	57.22	15.62	68.71	39.67
Our type 2	88.41	77.55	69.92	53.97	29.91	62.29	66.58
Our type 3	<b>95.16</b>	<b>80.34</b>	73.08	<b>57.91</b>	29.99	<b>68.04</b>	<b>82.88</b>
Class	Pill	Screw	Tile	Toothbrush	Transistor	Wood	Average
[2]	62.17	27.73	52.13	82.73	<b>88.25</b>	84.30	64.18
Our type 1	50.60	28.06	84.70	<b>93.33</b>	77.92	85.44	63.17
Our type 2	51.72	46.96	92.71	70.22	84.04	<b>90.96</b>	66.19
Our type 3	<b>57.23</b>	<b>61.86</b>	<b>93.58</b>	91.67	83.29	87.37	<b>72.92</b>

### 4. Conclusions

In this paper, we presented a novel deep anomaly detection that is proper for real-world industrial problems. The proposed algorithm designs self-supervised learning for the morphological feature

representation of normal data. To demonstrate the proposed method's superiority over the existing semantic-feature-learning-based methodology, the experimental results for diverse classes in MVTec were reported. These experimental results show that the proposed algorithm provides an 8.74% higher AUROC performance than the target method with 0.0125 s of processing time. Conclusively, the proposed algorithm achieves high accuracy and low computational complexity simultaneously in real-world industrial anomaly inspection applications. We leave the incorporation of the proposed algorithm with the semantic-feature-based algorithm for future work.

**Author Contributions:** Conceptualization, T.K. and Y.C.; methodology, T.K.; software, T.K.; writing—original draft preparation, T.K.; writing—review and editing, T.K. and Y.C.; supervision, Y.C.; project administration, Y.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** This work was supported by the Technology Development Program (S2798925) funded by the Ministry of SMEs and Startups (MSS, Korea).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Zhou, C.; Paffenroth, R.C. Anomaly detection with robust deep autoencoders. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017.
2. Golan, I.; El-Yaniv, R. Deep anomaly detection using geometric transformations. In Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, QC, Canada, 3–8 December 2018.
3. Gong, D.; Liu, L.; Le V.; Saha, B.; Mansour, M.R.; Venkatesh, S.; Hengel, A.V.D. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019.
4. LeCun, Y.; Cortes, C.; Burges, C.J. *Mnist Handwritten Digit Database*; AT&T Labs: Florham Park, NJ, USA, 2010.
5. Krizhevsky, A.; Hinton, G. Learning Multiple Layers of Features from Tiny Images. Master's Thesis, Department of Computer Science, University of Toronto, Toronto, ON, Canada, 2009.
6. Elson, J.; Douceur, J.R.; Howell, J.; Saul, J. Asirra: A CAPTCHA that exploits interest-aligned manual image categorization. In Proceedings of the ACM Conference on Computer and Communications Security, Alexandria, VA, USA, 29 October–2 November 2007; Volume 7.
7. Bergmann, P.; Fauser, M.; Sattlegger, D.; Steger, C. MVTec AD—A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019.
8. Gidaris, S.; Singh, P.; Komodakis, N. Unsupervised Representation Learning by Predicting Image Rotations. In Proceedings of the IEEE International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
9. Gonzalez, R.C.; Woods, R.E. Morphological Image Processing. In *Digital Image Processing*, 3rd ed.; Prentice Hall: Englewood Cliffs, NJ, USA, 2008; Chapter 9, pp. 649–710.
10. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
11. Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic differentiation in PyTorch. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).