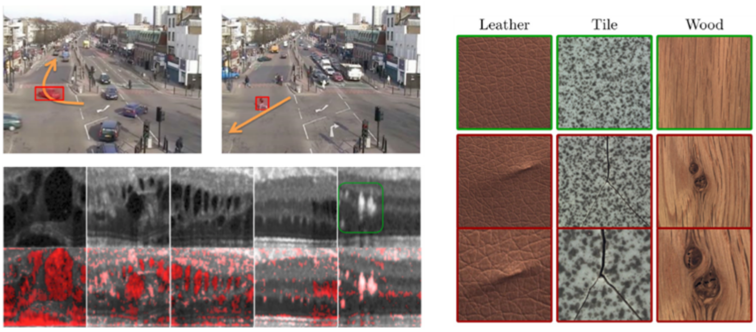
Unsupervised Anomaly Detection in manufacturing domain / Team 3

Cho, Yeong kyu
Song, In gwon
Jang, Do woon
Han, Jae ho

Background & Motivation

Anomaly detection is the technique used to identify unusual patterns that do not conform to expected behavior, called outliers. Recently, deep learning-based anomaly detection algorithms have become increasingly popular by showing better performance than traditional machine learning algorithms. Therefore, deep learning-based anomaly detection has been applied in various fields such as illegal traffic flow,product defect detection, etc.



<Example of deep learning-based anomaly detection>

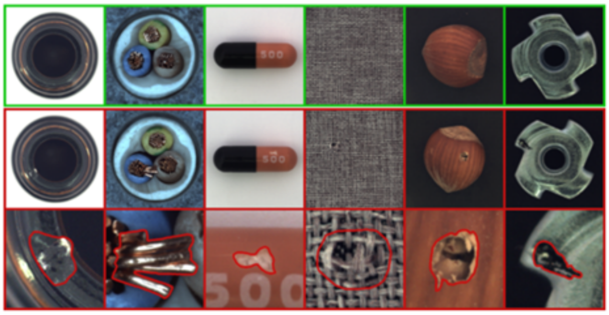
Typically, anomaly detection models can be broadly classified into three categories based on the extent of availability of labels. If labels are used for both normal and abnormal data instances for the training model, then it is supervised anomaly detection. On the other hand, when there are labels for only normal data instances, it is semi-supervised anomaly detection, and if the model detects anomaly based on intrinsic properties of the data instances without labels, it could be called unsupervised anomaly detection.

In general, it shows better performance when labels exist in the dataset during training a model. However, there are some limitations to its practical application in the industry. First, lack of availability of labels for datasets. labels can be created to increase the accuracy of the model. However, creating labels to increase model accuracy is not cost-effective. Second, the number of normal and abnormal data instances has imbalances. In that case, the predictive accuracy of the model is low. Therefore, semi-supervised / unsupervised anomaly detection that is capable of better performance is more preferred than supervised learning for industry fields.

In the manufacturing process, the quality of the product must be verified even if detecting defects is not an easy task. Also, there are difficulties in applying deep learning-based anomaly detection as mentioned above even if it is highly required. In this project, unsupervised anomaly detection based on GAN/Autoencoder will be conducted to find defects in industry data. Moreover, a model that can accurately discriminate abnormal will be developed.

MVtec-ad dataset

In order to reflect the real industry environment, the MVtec dataset will be used for training models. All images were acquired using a high-resolution industrial RGB sensor. MVtec dataset consists of 15 classes made of textures and objects. the training dataset has only normal instances and the test dataset is a mixture of abnormal and normal instances. In this project, it was used to train the model mainly on carper and bottle data.



	Category	# Train	# Test (good)	# Test (defective)	# Defect groups	# Defect regions	Image side length
Textures	Carpet	280	28	89	5	97	1024
	Grid	264	21	57	5	170	1024
	Leather	245	32	92	5	99	1024
	Tile	230	33	84	5	86	840
	Wood	247	19	60	5	168	1024
Objects	Bottle	209	20	63	3	68	900
	Cable	224	58	92	8	151	1024
	Capsule	219	23	109	5	114	1000
	Hazelnut	391	40	70	4	136	1024
	Metal Nut	220	22	93	4	132	700
	Pill	267	26	141	7	245	800
	Screw	320	41	119	5	135	1024
	Toothbrush	60	12	30	1	66	1024
	Transistor	213	60	40	4	44	1024
	Zipper	240	32	119	7	177	1024
	Total	3629	467	1258	73	1888	-

<MVtec dataset images and composition>

Autoencoder based anomaly detection

1. Model explanation

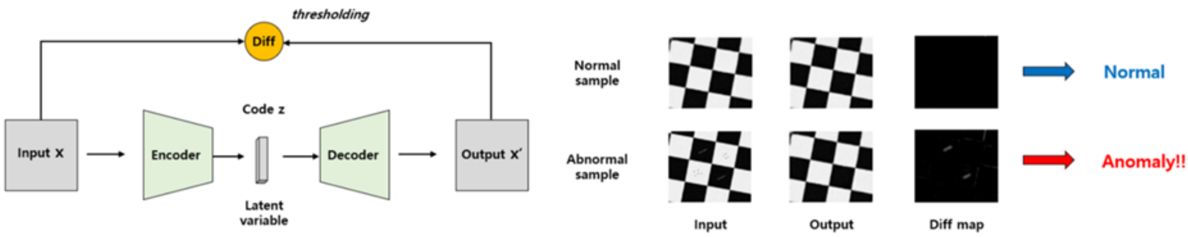
Autoencoder

Autoencoder is unsupervised method for learning feature vectors from raw data x, without any labels. It can be used to generate the reconstructed input. It is constructed by two steps.

(1) Encoder: It represents the input value as a new expression
(2) Decoder: It finds the original input value from the derived expression

CNN Autoencoder is that the encoder and decoder is constructed by CNN which process image data.

Autoencoder based anomaly detection



Autoencoder allows to learn the characteristics of the normal area, which is the main component of the data, without having to label the data. Normal and abnormal images are trained and tested as follows.

(1) In the training process, only normal data is used, and the output of the autoencoder is learned to come out as similar as possible to the input image.

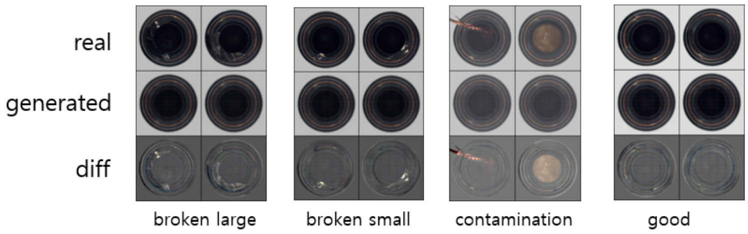
(2) In the test process, when an abnormal sample is added, the autoencoder output is restored like a normal sample. The difference is high between input and output, so the anomaly could be detected.

2. Training Setting

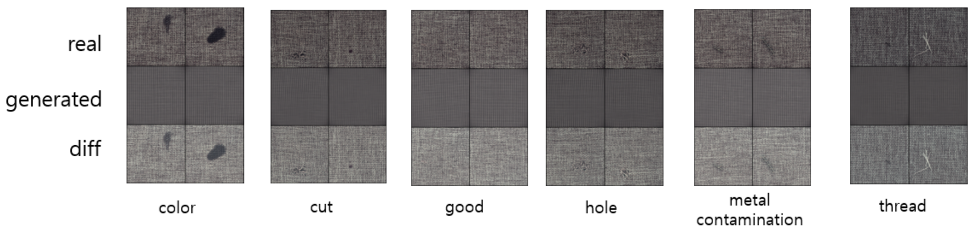
Image size	Augmentation	Normalization	Epoch	Learning rate	Latent dim
256	Random Horizontal Flip	Mean = 0.5 Std = 0.5	300	0.001	1024

3. Result

Autoencoder based anomaly detection model is performed using the difference between real image and generated image. The result of bottle images shows they generate normal data well, and the difference between the real image and the generated image in broken large, broken small, and contamination can be clearly identified. However, in the carpet image cases, they can not generate normal image in many cases, so the model can not perform well. To address this problem, we will consider the GAN based anomaly detection model.

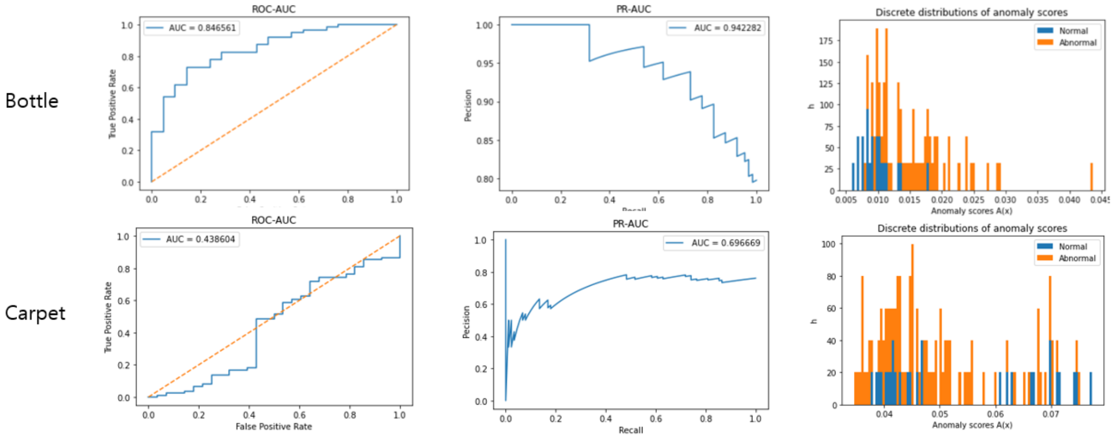


< Result of bottle images >



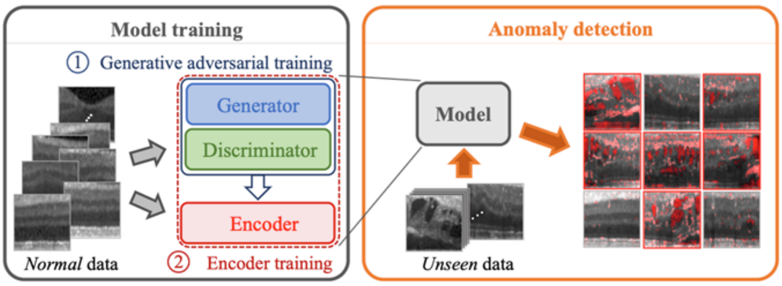
< Result of carpet images >

ROC-AUC and PR-AUC of the bottle class were respectively 0.846 and 0.942, the carpet class respectively showed 0.439 and 0.7. The difference in performance of these two classes was also confirmed in anomaly score distribution.



<ROC-AUC, PR-AUC, Distributions of Anomaly Scores>

GAN based anomaly detection



1. Model introduction

Fast AnoGAN (f-AnoGAN), a generative adversarial network (GAN) based unsupervised learning approach is capable of identifying anomalous images and image segments.

Anomaly detection framework consists of two training steps on normal images:

(1) GAN training, and (2) encoder training based on the trained GAN model.

2. Unsupervised learning of normal anatomical variability (step 1)

The generator learns to generate images of the training distribution capturing normal variability. The discriminator can estimate the fit of generated images to the distribution of training images.

The trained generator and discriminator are utilized with fixed weights for subsequent encoder training and for anomaly scoring

3. Learning a fast mapping from images to encodings in the latent space (step2)

During encoder training, the mapping from real images to latent encodings z is performed by the trainable encoder, while the mapping of z back to image space is performed with the fixed generator.

Additionally calculate image statistics of the real image and the reconstructed image through intermediate feature representation of discriminator.

4. Detection of anomalies

Since the model is only trained on normal images, it only “reconstructs” an image visually similar to the input image and lying on the manifold of normal images.

The anomaly quantification formulation follows directly the specific definition of the loss used for encoder training.

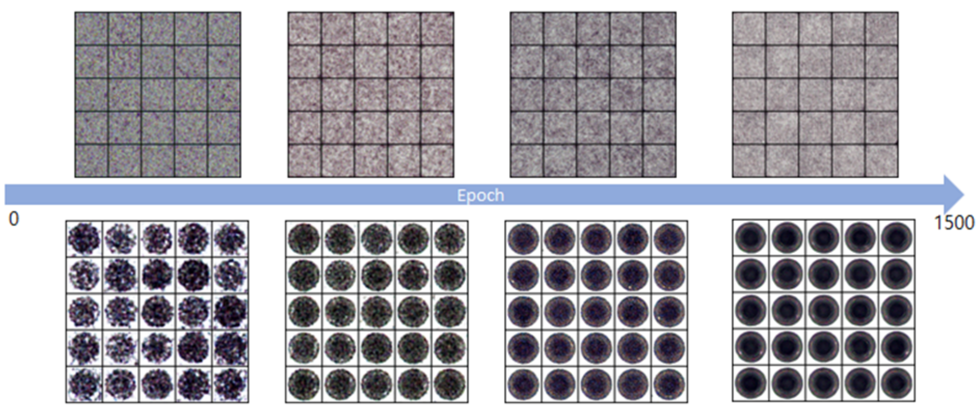
$$\mathcal{A}(\mathbf{x}) = \frac{1}{n} \cdot \|\mathbf{x} - G(E(\mathbf{x}))\|^2 + \frac{\kappa}{n_d} \cdot \|f(\mathbf{x}) - f(G(E(\mathbf{x})))\|^2$$

5. Training Setting

We use following hyperparameter for training f-AnoGAN: resize image to 64, use augmentation randomhorizontalflip and normalize with mean=0.5/std=0.5, epoch for training generator and discriminator is 1500 and for encoder is 200, 32 batch size, 0.0002 learning rate and 100 latent dimension. We update the discriminator every batch iteration and generator every fifth batch iteration.

6. Trained Generator Images Quality

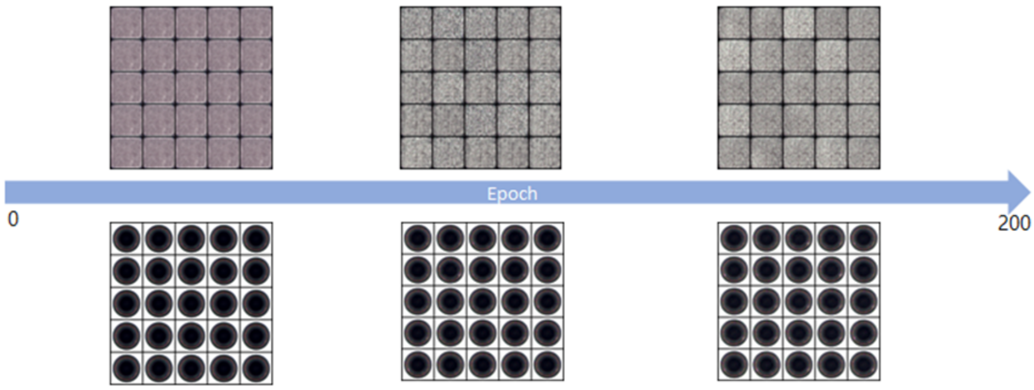
If the GAN is well trained, the possibility that the generator will generate a similar input normal image is high. It can be seen that the bottle image generated by the generator are well made. However, in the carpet class generator did not restore the texture of the original image well.



< Image generated by the generator >

7. Trained Encoder Images Quality

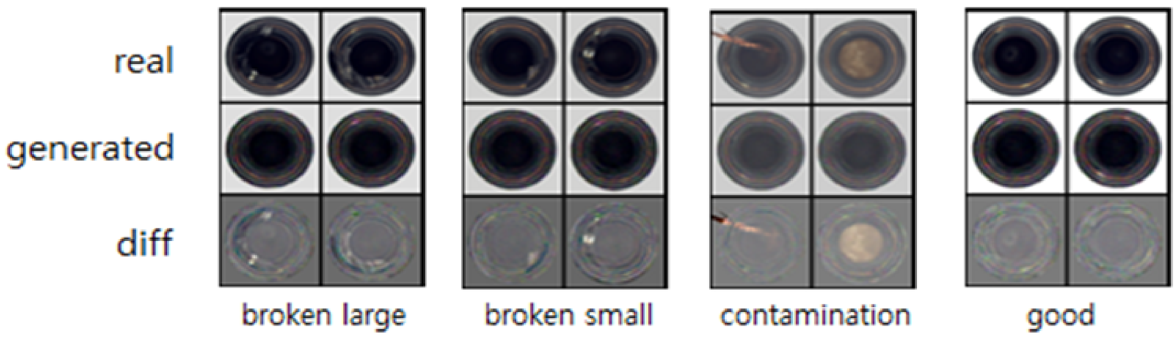
The following process was conducted to ensure that the trained encoder was well learned. First, the latent vector z is extracted from the real image, and a fake image is created based on it. After that, the image is reconstructed by extracting the latent vector z from the generated fake image. In this way, the encoder's quality was being checked for 200 epochs, and the encoder was well trained considering the bottle image reconstructed by the encoder appeared clearly. However, in the carpet class, when you look at the restored image, you can see that it is blurry.



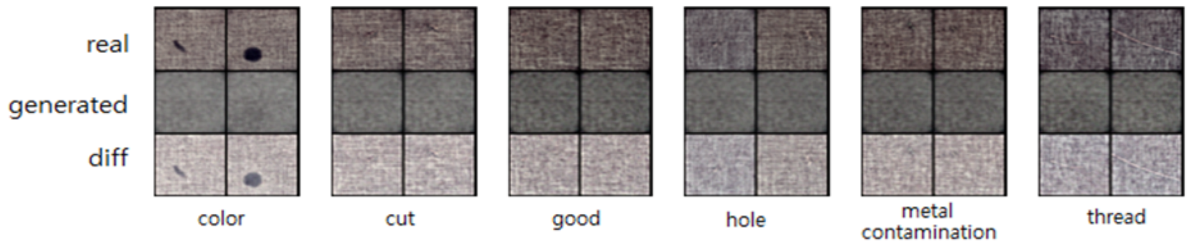
< Image reconstructed by the encoder >

8. Results

Anomaly detection by F-AnoGAN model is performed using the difference between real image and generated image. Looking at the results of the bottle class, the difference between the real image and the generated image in broken large, broken small, and contamination can be clearly identified. Also, there was no difference between the real and fake images in good class. In the results of the carpet class, it is easy to detect the difference by color, but it was difficult to visually confirm the difference in cut, hole, metal contamination, and thread.

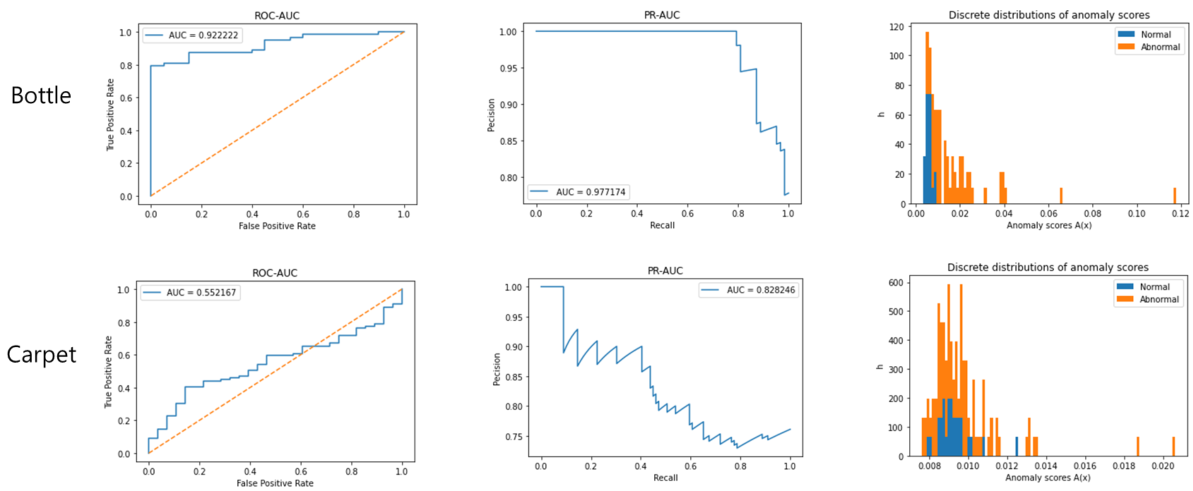


< Result of bottle images >



< Result of carpet images >

ROC-AUC and PR-AUC of the bottle class were respectively 0.922 and 0.977, and the carpet class respectively showed 0.552 and 0.828. The difference in performance of these two classes was also confirmed in anomaly score distribution. In the case of the bottle class, there is a threshold value that distinguishes the normal and the abnormal, but in the carpet class, it is very difficult to define the threshold value.



<ROC-AUC, PR-AUC, Distributions of Anomaly Scores>

Comparison & Limit

Both the Autoencoder and the F-AnoGAN performed well in the anomaly detection for the carpet class. However, the performance of the carpet class was comparatively lower than that of the bottle class. The reason is that both models could not train subtle variations of carpet's textural patterns.

Model	Class	AUROC	AUPR
Autoencoder	Bottle	0.846	0.942
	Carpet	0.439	0.7
f-AnoGAN	Bottle	0.922	0.977
	Carpet	0.552	0.828

Contributions

Although the model we studied has limitations, there are also contributions. The anomaly detection of unsupervised learning can be very useful, because abnormal data is very few compared to normal data in actual work sites. In addition, these models could be useful for manufacturing companies if the threshold value that distinguishes the abnormal things between products is well defined.

Reference

Chalapathy, Raghavendra, and Sanjay Chawla. "Deep learning for anomaly detection: A survey." *arXiv preprint arXiv:1901.03407* (2019).

Xie, Xuemei, et al. "Real-time illegal parking detection system based on deep learning." *Proceedings of the 2017 International Conference on Deep Learning Technologies*. 2017.

Schlegl, Thomas, et al. "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery." *International conference on information processing in medical imaging*. Springer, Cham, 2017.

Bergmann, Paul, et al. "MVTec AD--A comprehensive real-world dataset for unsupervised anomaly detection." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.

Baur, Christoph, et al. "Deep autoencoding models for unsupervised anomaly segmentation in brain MR images." *International MICCAI Brainlesion Workshop*. Springer, Cham, 2018.

Zhai, Shuangfei, et al. "Deep structured energy based models for anomaly detection." *International Conference on Machine Learning*. PMLR, 2016.

Kramer et al, "Nonlinear principal component analysis using autoassociative neural networks," AIChE Journal (1991)

Paul Bergmann et al, "Improving Unsupervised Defect Segmentation by Applying Structural Similarity To Autoencoders " arXiv (2019)

Schlegl, T., Seeböck, P., Waldstein, S. M., Langs, G., & Schmidt-Erfurth, U. (2019). f-AnoGAN:Fast unsupervised anomaly detection with generative adversarial networks. Medical image analysis, 54, 30-44.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.