Unsupervised Anomaly detection in manufacturing domain

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

Index

❖ **Background & Motivation**

- **Anomaly Detection**

❖ **Autoencoder based anomaly detection**

❖ **GAN based anomaly detection**

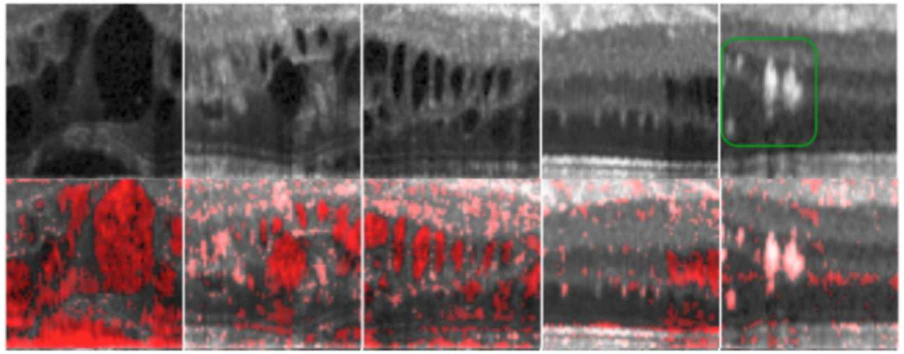
❖ **Comparison**

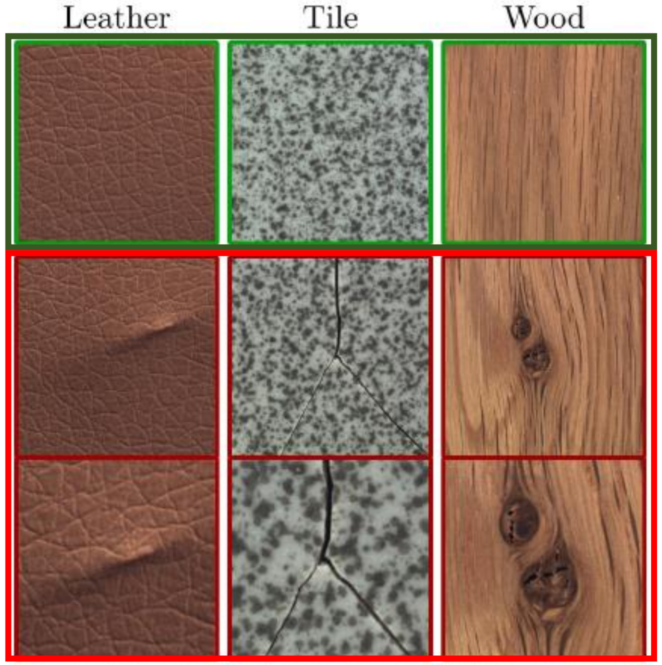
❖ **Limits & Contributions**

Background & motivation

❖ Anomaly Detection

- **Method to identify abnormal** items, events, or observations that differed significatntly from the normal condtion
- Illegal traffic detection, detecting retinal damage, IoT big-data etc.





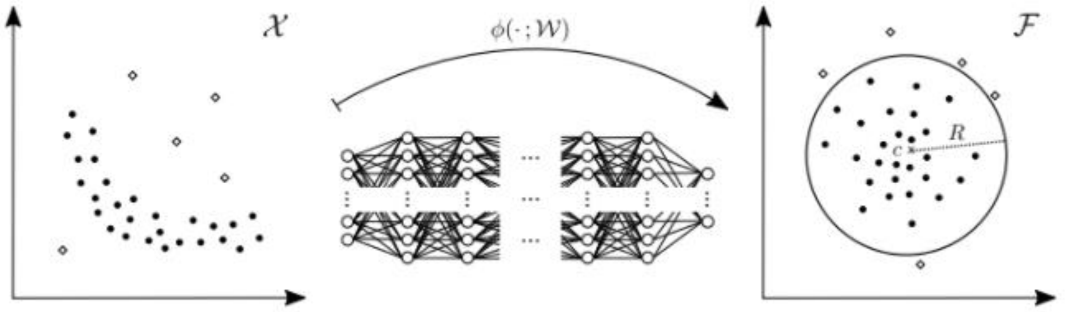
Background & motivation

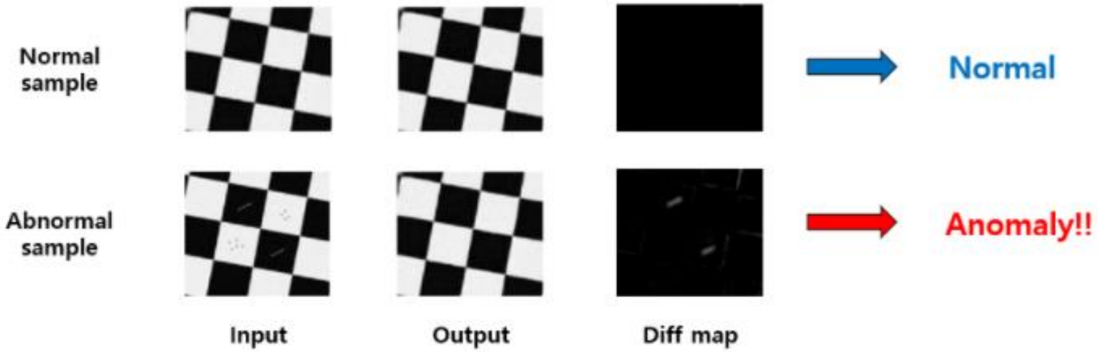
❖ Anomaly Detection

- **Supervised / Semi-supervised / Unsupervised anomaly detection**
- it is classified according to the type of dataset

	Supervised	Semi-supervised	Unsupervised
Training set	O	O	X
Test set	O	X	X

O – with label
X – without label





Background & motivation

❖ Objective

Usually, the dataset from the **industrial field do not have label**

- Accuracy of model will be decrease
- However, creating label one by one is not cost-effective

It is highly required to conduct **unsupervised-learning**

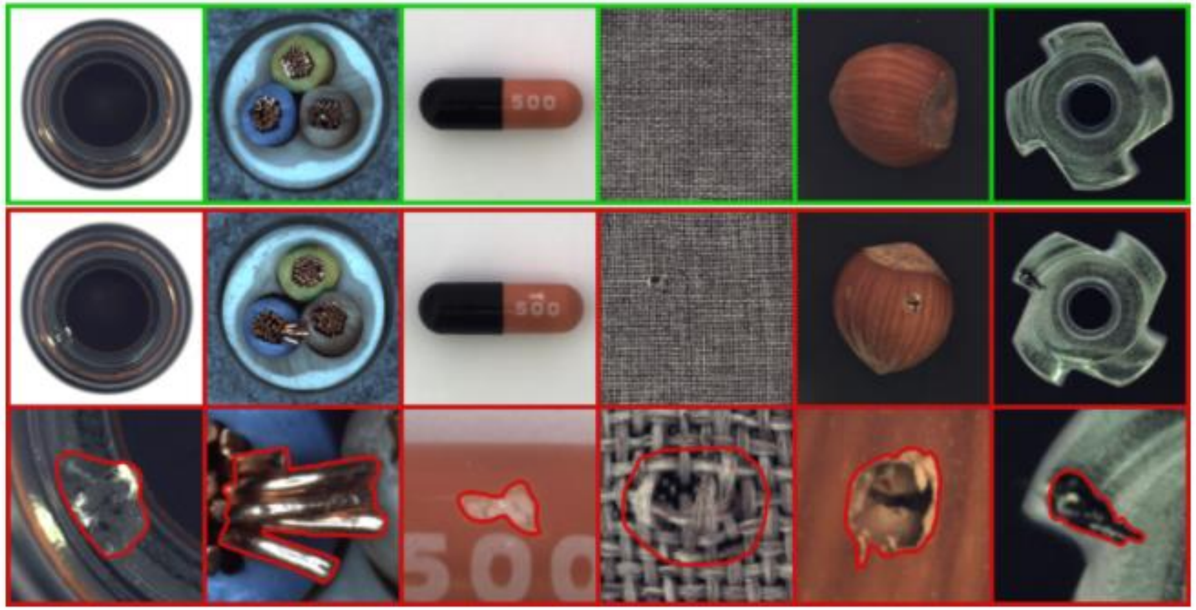
- **Autoencoder, GAN** etc.

Develop a model that can accurately discriminate abnormal in an unsupervised environment

Data description

❖ Mvtec-AD dataset

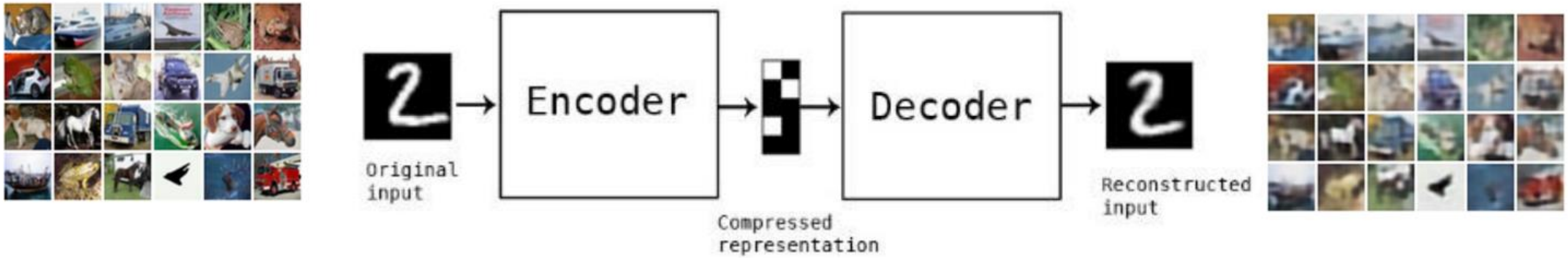
- 15 classes made of textures and objects
- The training data consists only of normal data, the test dataset is a mixture of normal and abnormal data



	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	-

Autoencoder based anomaly detection

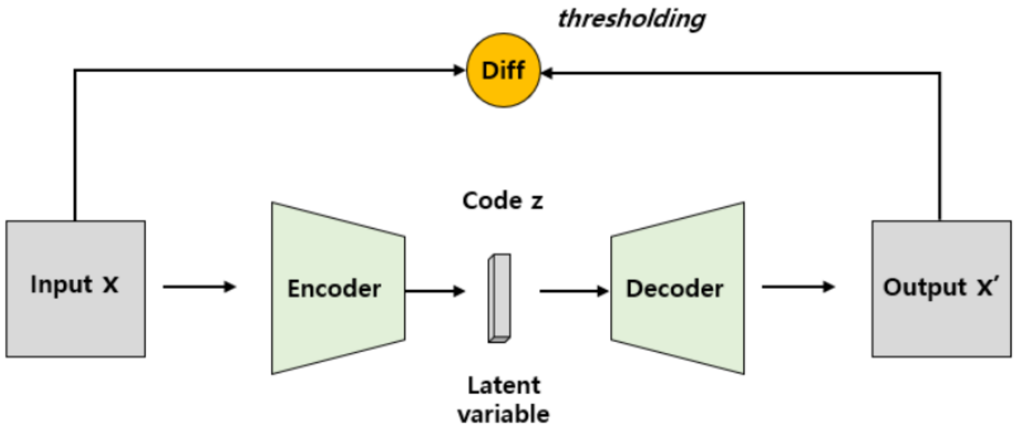
❖ Autoencoder

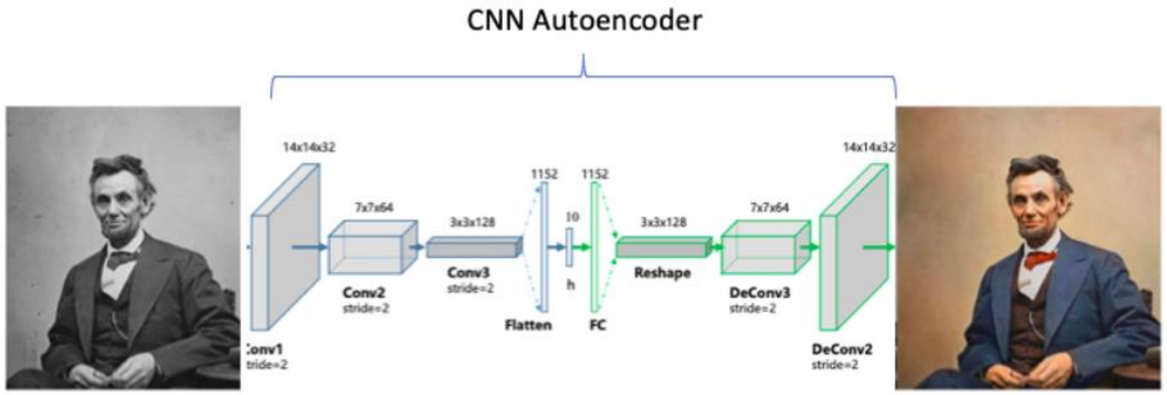


- Autoencoder is **unsupervised method** for learning feature vectors from raw data x, without any labels.
- It is constructed by **encoder** and **decoder**, which encoder represent the input value as a new expression and decoder finds the original input value from the derived expression.
- It can be used to **generate the reconstructed input**.

Autoencoder based anomaly detection

❖ Autoencoder based anomaly detection





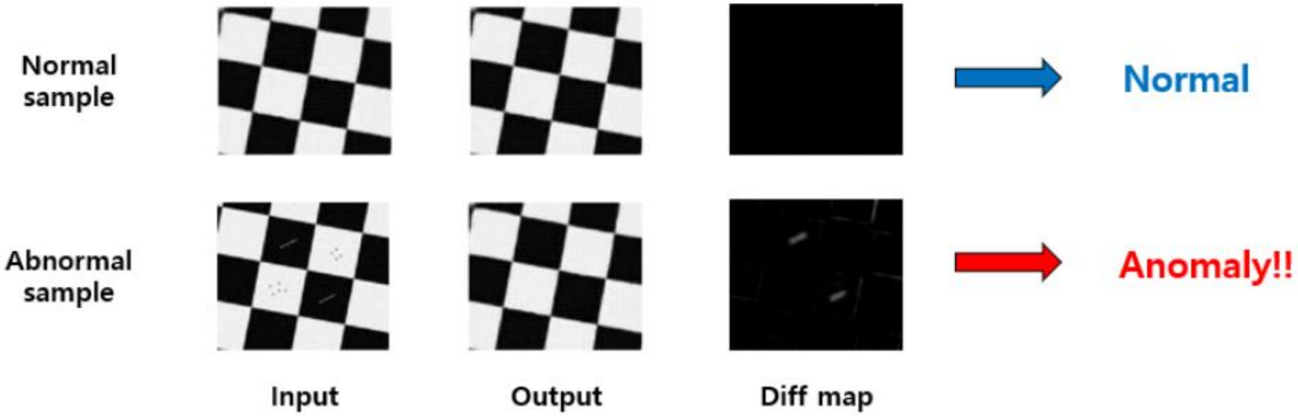
- Autoencoder allows to **learn the characteristics of the normal area**, which is the main component of the data, without having to label the data.

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

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

Autoencoder based anomaly detection

❖ Autoencoder based anomaly detection



- As the figure, if a normal sample is put in the learned autoencoder**, the difference between input and output is low**.
- However, if an abnormal sample is added, **the autoencoder is restored like a normal sample**.
- In abnormal data, **the difference is high between input and output**, so the abnormal sample could be detected.

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

Autoencoder based anomaly detection

❖ Training Setting

- Hyperparameter

image size: 256

preprocessing: Augmentation – RandomHorizontalFlip, Normalize - (0.5, 0.5, 0.5), (0.5, 0.5, 0.5)

epoch: 300

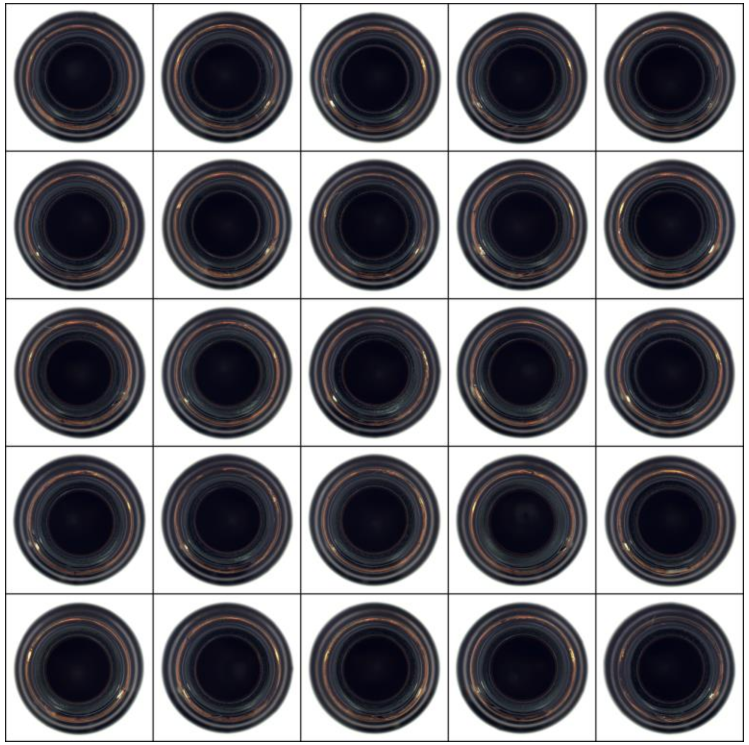
batch size: 70

lr: 0.001

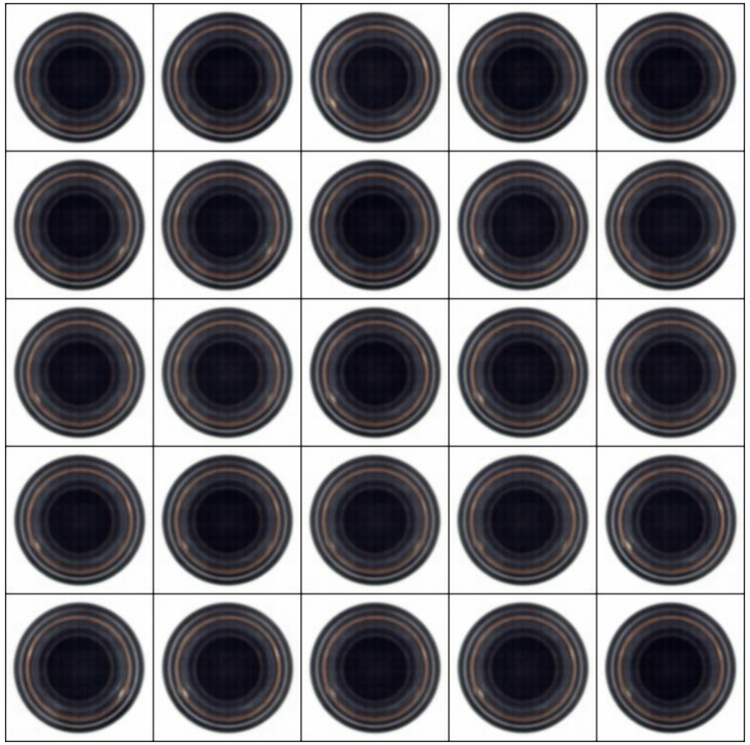
latent dim of z: 1024

Autoencoder based anomaly detection

❖ Trained autoencoder images quality - bottle



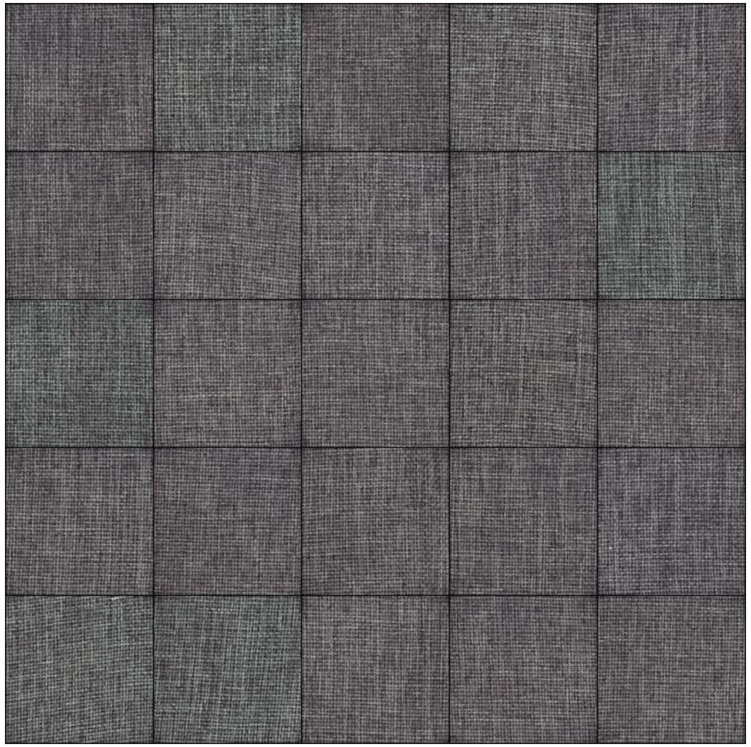
original image



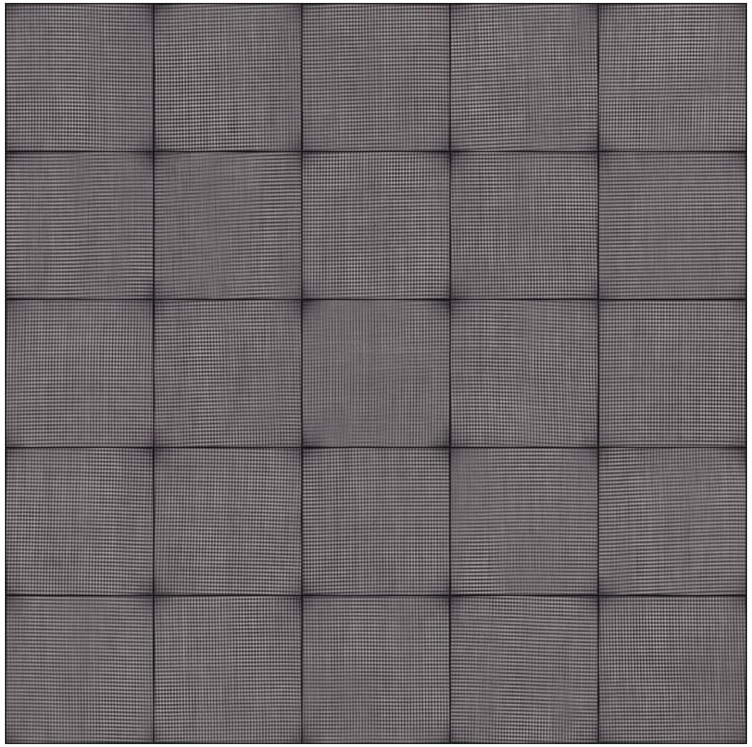
generated image

Autoencoder based anomaly detection

❖ Trained autoencoder images quality - carpet



original image



generated image

Autoencoder based anomaly detection

❖ Evaluation metric

- **ROC AUC (Area Under ROC curve)**

It is a chart that visualizes the tradeoff between **true positive rate (TPR)** and **false positive rate (FPR).**
Basically, for every threshold, we calculate TPR and FPR and plot it on one chart.
The higher TPR and the lower FPR is for each threshold the better and so classifiers that have curves that are more top-left-side are better.
In order to get one number that tells us how good our curve is, we can **calculate the Area Under the ROC Curve**.

- **PR-AUC (Area Under the Precision-Recall Curve)**

It is **a curve that combines precision (PPV) and Recall (TPR) in a single visualization**.

For every threshold, you calculate PPV and TPR and plot it.

The higher on y-axis your curve is the better your model performance.

Calculate the Area Under the Precision-Recall Curve to get one number that describes model performance.

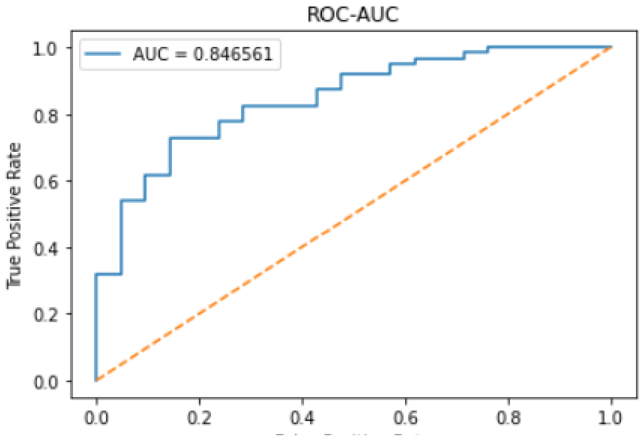
$$(Precision) = \frac{TP}{TP + FP}$$

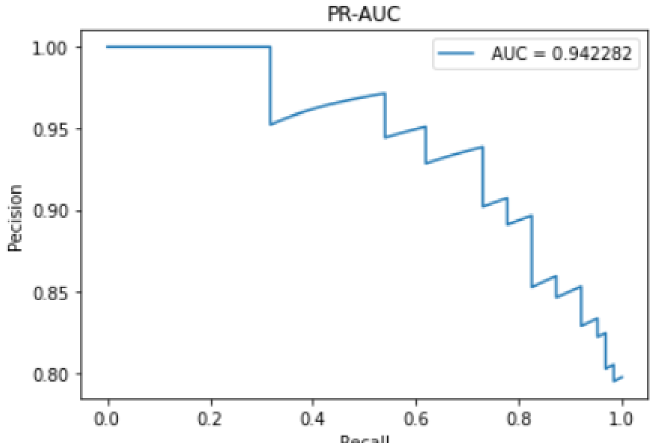
$$(Recall) = \frac{TP}{TP + FN}$$

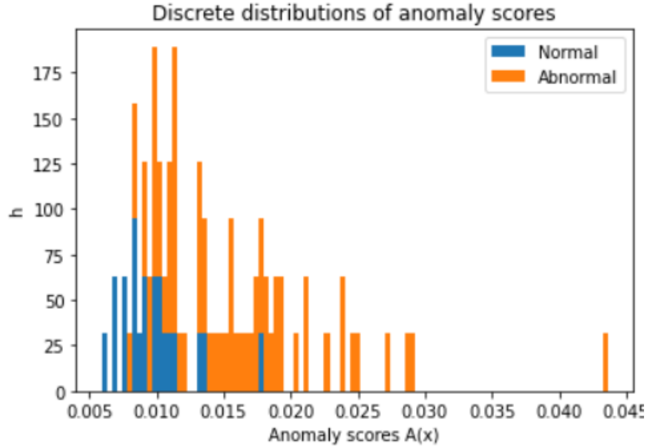
Autoencoder based anomaly detection

❖ Results

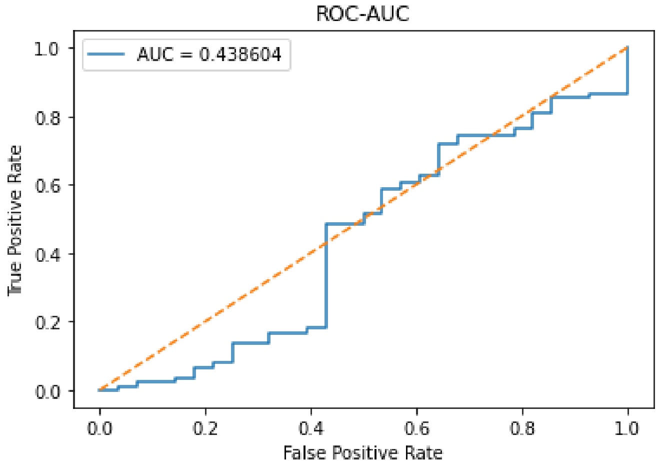
Bottle

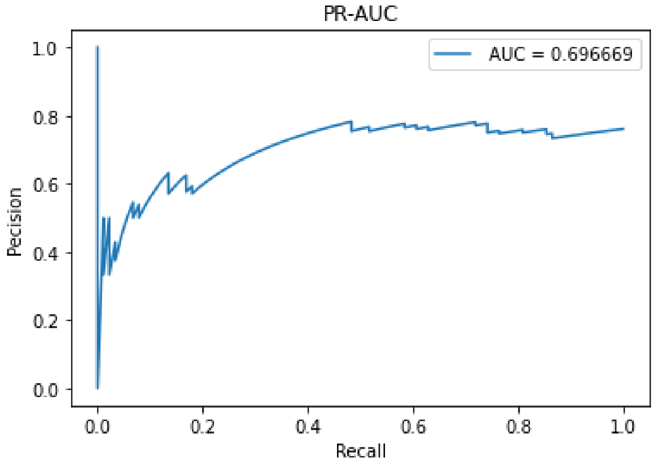


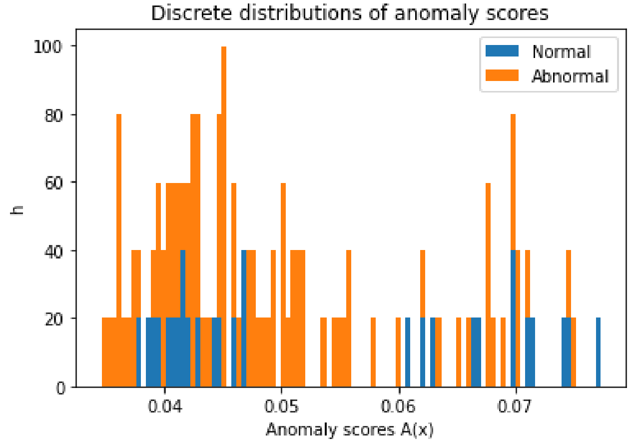




Carpet







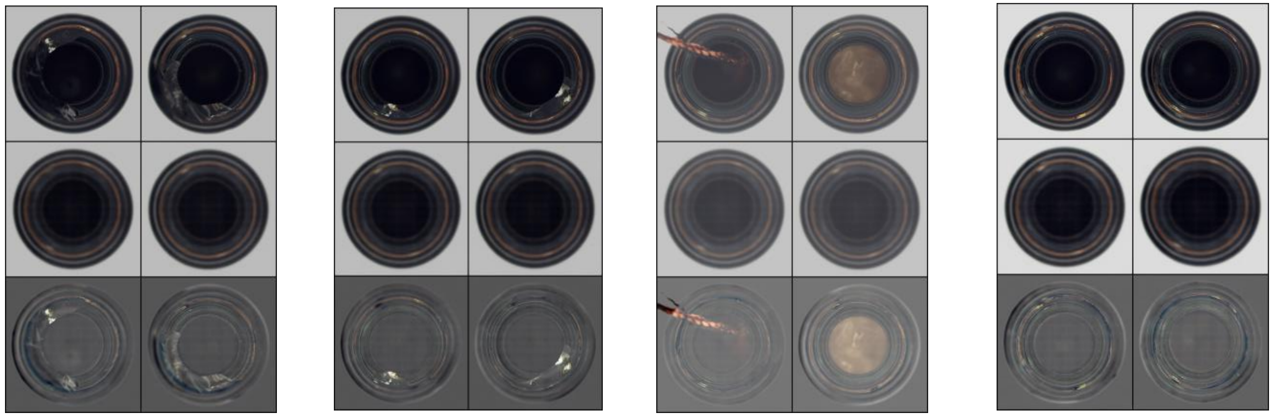
Autoencoder based anomaly detection

❖ Results

real

generated

diff

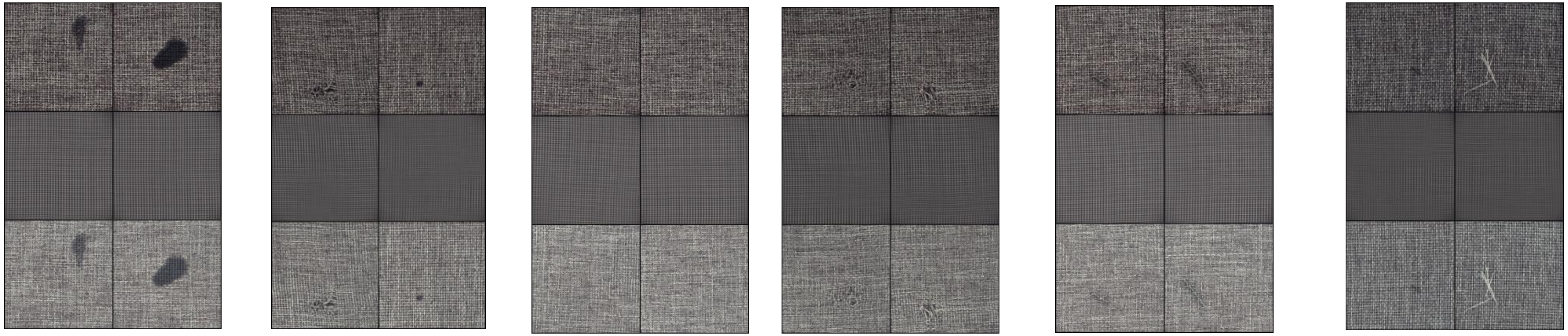


broken large | broken small | contamination | good

real

generated

diff

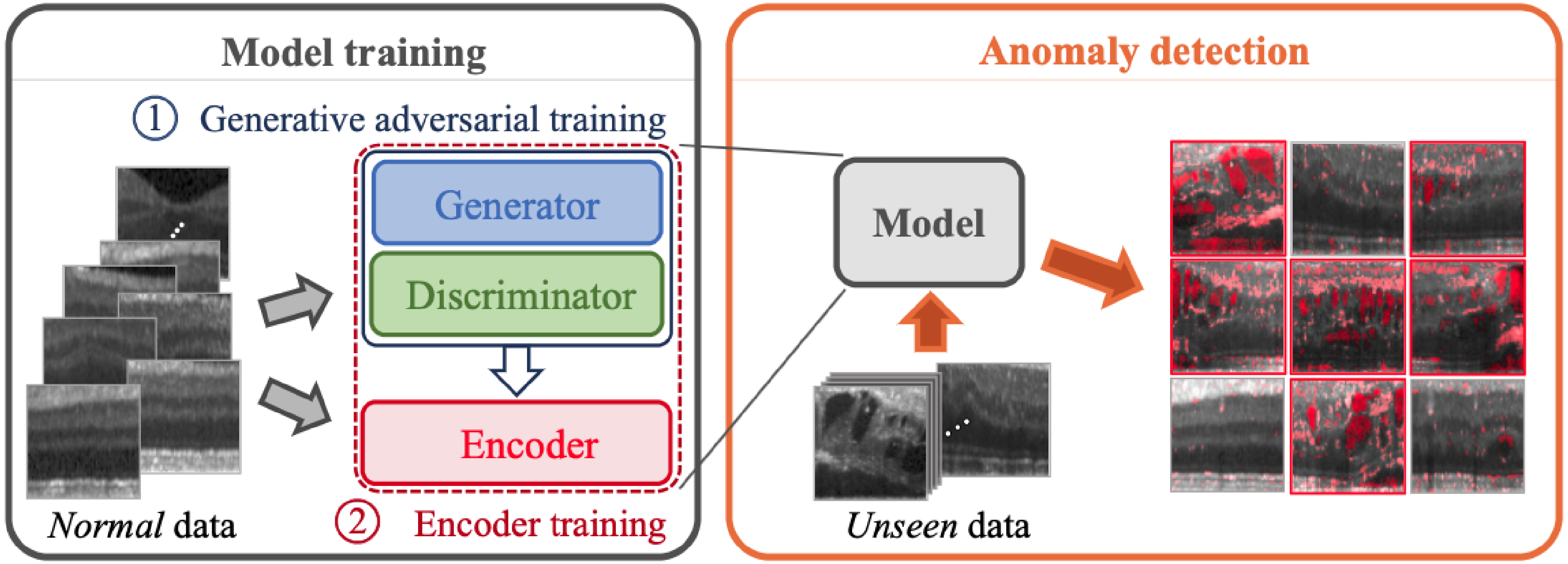


color | cut | good | hole | metal contamination | thread

GAN based anomaly detection

❖ f-AnoGAN

Anomaly detection framework



- GAN based unsupervised learning approach capable of identifying anomalous images and image segments, that can serve as imaging biomarker candidates

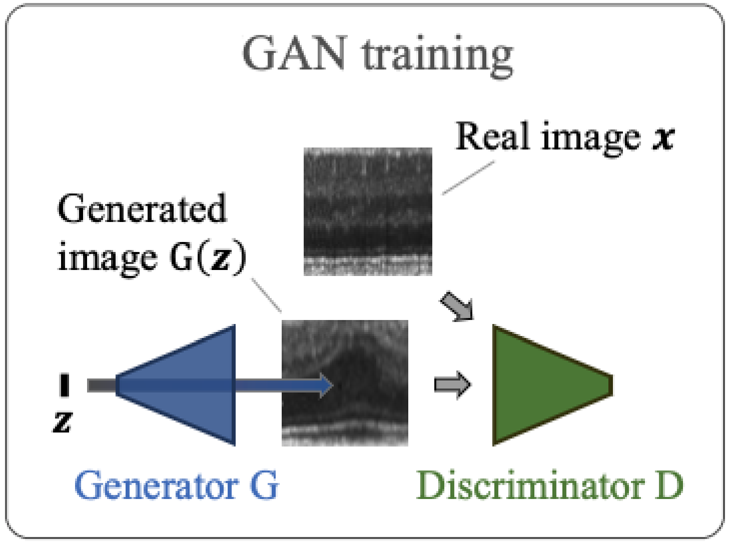
- Both steps of model training, generative adversarial training and encoder training are performed

Schlegl, Thomas, et al. "f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks." *Medical image analysis* 54 (2019)

GAN based anomaly detection

❖ Unsupervised learning of normal anatomical variability (step 1)

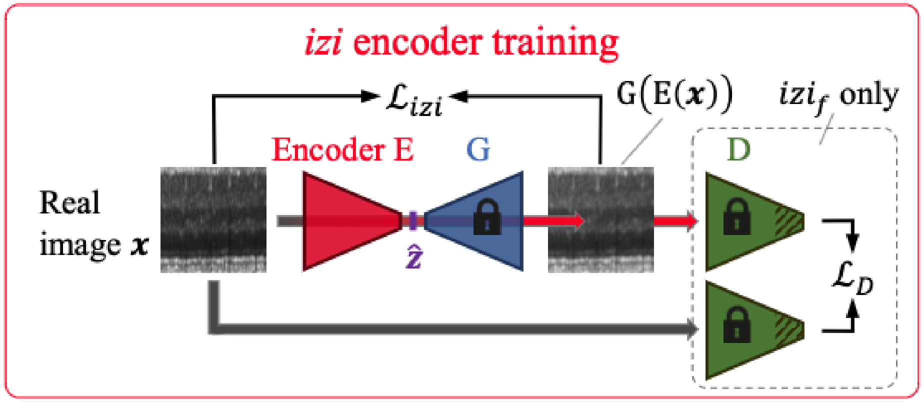
- The generator learns to generate images of the training distribution capturing normal variability
- The discriminator can estimate the fit of generated images to the distribution of training images.
- The trained generator and discriminator are utilized with fixed weights for subsequent encoder training and for anomaly scoring



GAN based anomaly detection

❖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 G.
- Additionally, calculate image statistics of the real image and the reconstructed image through intermediate feature representation of discriminator

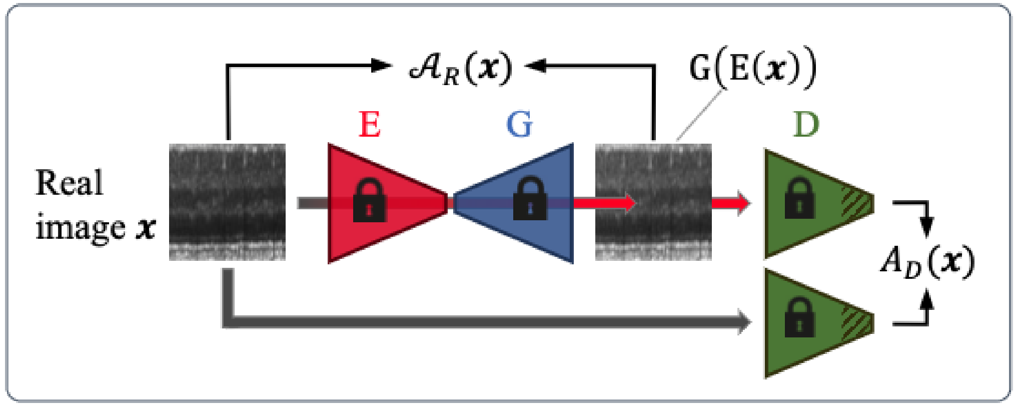


Total loss: $\mathcal{L}_{izi_f}(\mathrm{x}) = \frac{1}{n} \cdot \|\mathrm{x} - G(E(\mathrm{x}))\|^2 + \frac{\kappa}{n_d} \cdot \|f(\mathrm{x}) - f(G(E(\mathrm{x})))\|^2$

GAN based anomaly detection

❖ 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



Anomaly score for new image $$\mathcal{A}(\mathbf{x}) = \underbrace{\frac{1}{n} \cdot \|\mathbf{x} - G(E(\mathbf{x}))\|^2}_{\mathcal{A}_R(\mathbf{x})} + \underbrace{\frac{\kappa}{n_d} \cdot \|f(\mathbf{x}) - f(G(E(\mathbf{x})))\|^2}_{\mathcal{A}_D(\mathbf{x})}$$

pixel-level anomaly localization $$\dot{\mathcal{A}}_R(\mathbf{x}) = |\mathbf{x} - G(E(\mathbf{x}))|$$

GAN based anomaly detection

❖ Training Setting

- Hyperparameter

  - img size: 64

  - preprocessing: Augmentation – RandomHorizontalFlip, Normalize - (0.5, 0.5, 0.5), (0.5, 0.5, 0.5)

  - epoch: 1500 (step1), 200 (step2)

  - batch size: 32

  - lr: 0.0002

  - latent dim of z: 100

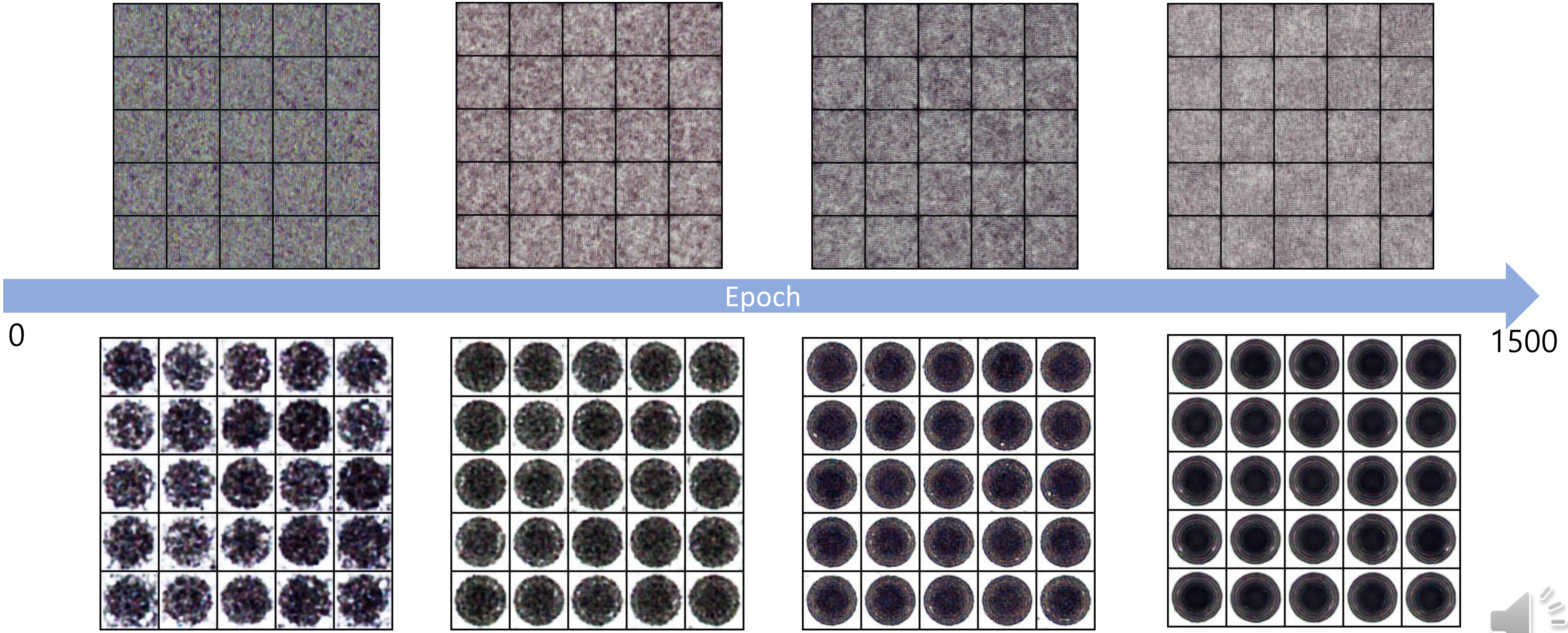
- parameter update

  - discriminator=every batch iteration

  - generator= every fifth batch iteration

GAN based anomaly detection

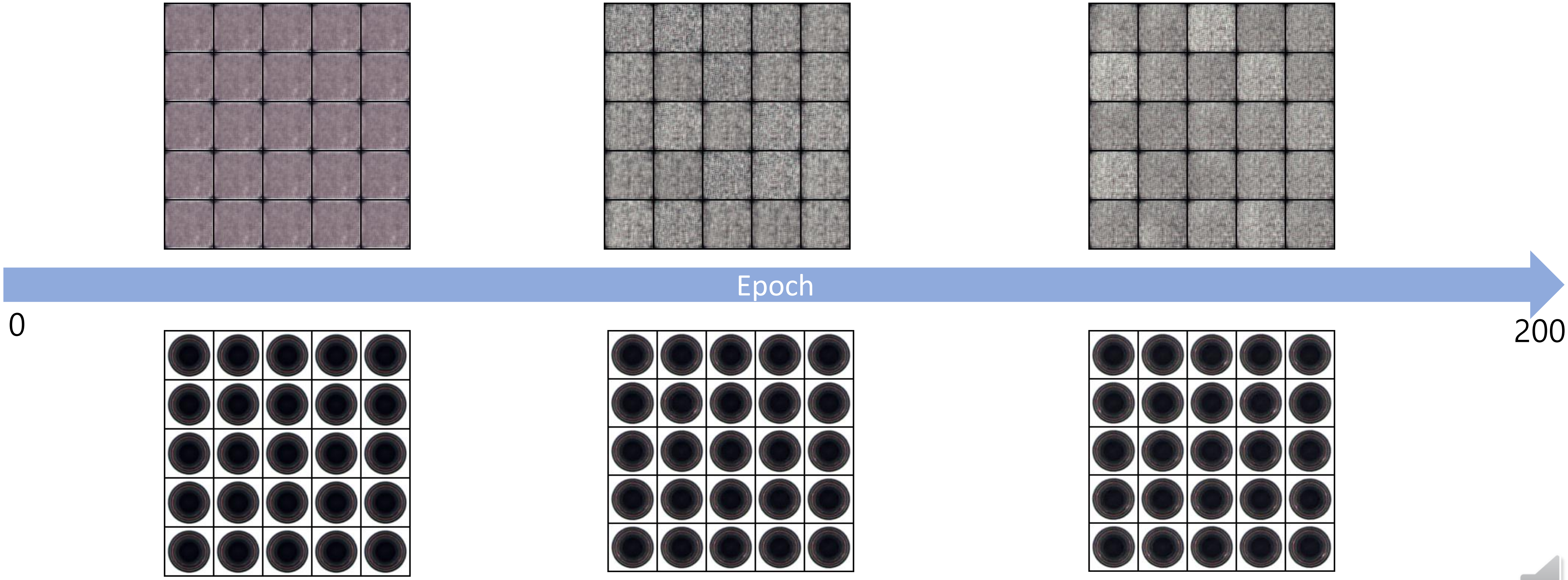
❖ Trained Generator images quality



GAN based anomaly detection

❖ Trained encoder images quality

- Real Image -> Latent vector z -> Fake Image -> Fake z -> Reconfiguration Image



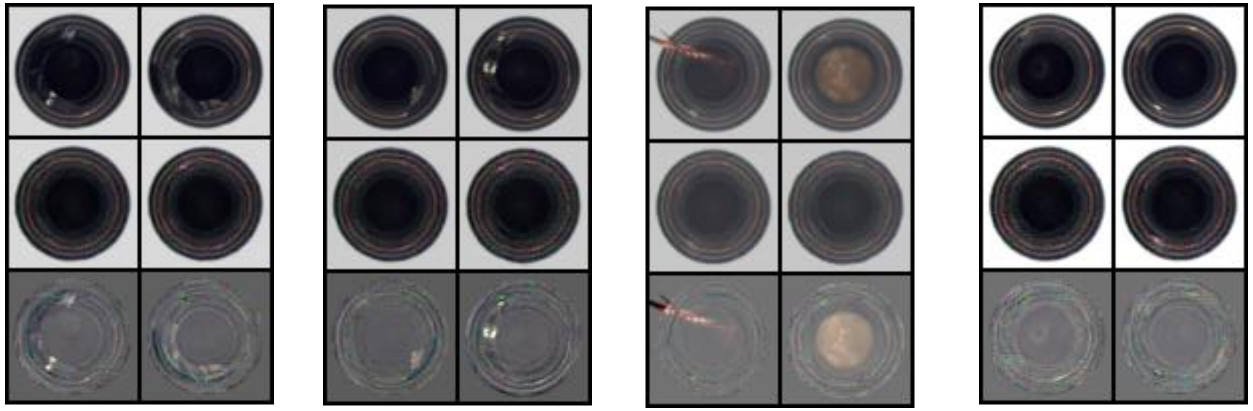
GAN based anomaly detection

❖ Results

real

generated

diff

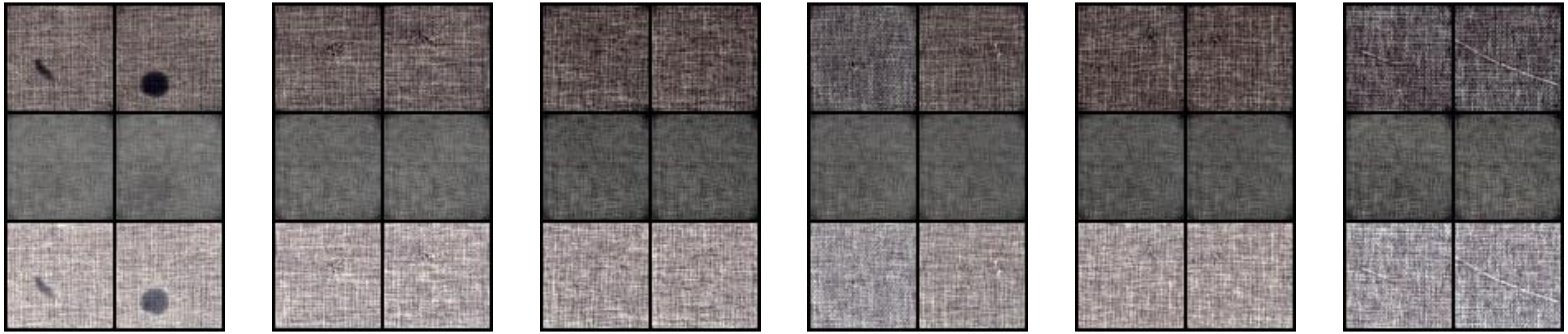


broken large | broken small | contamination | good

real

generated

diff

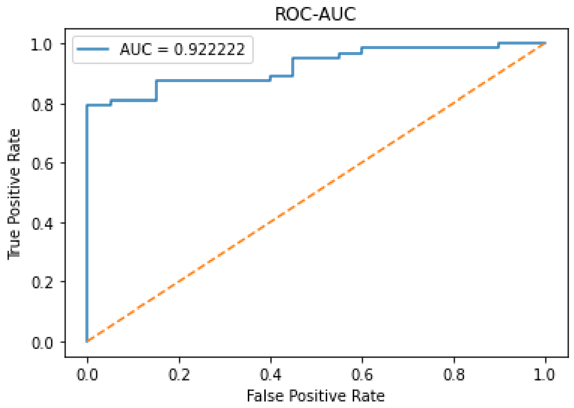


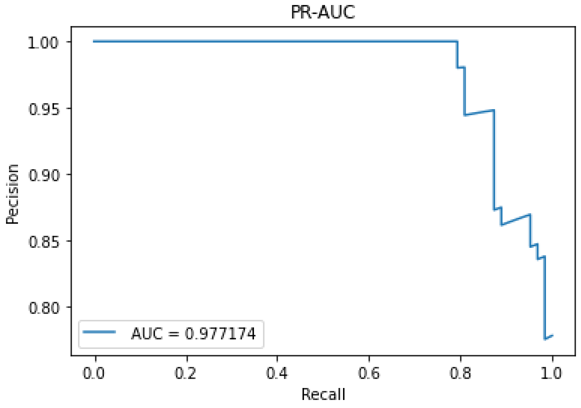
color | cut | good | hole | metal contamination | thread

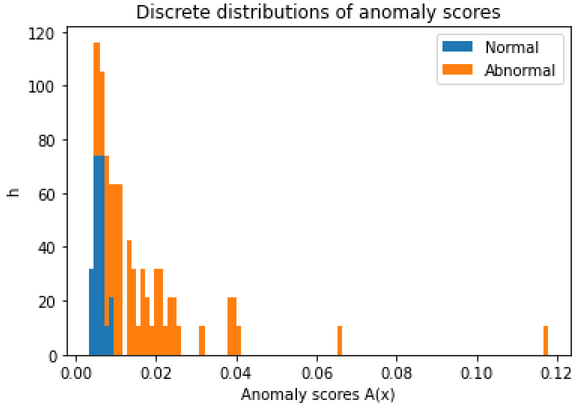
GAN based anomaly detection

❖ Results

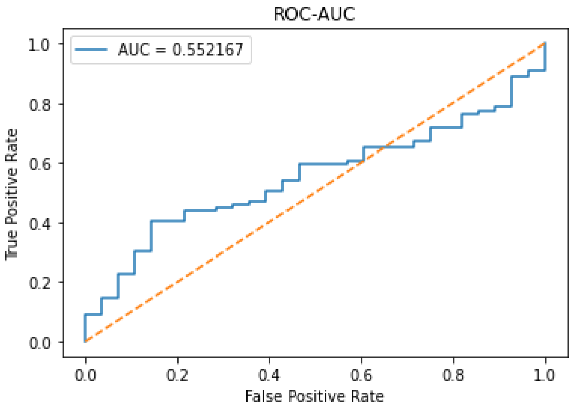
Bottle

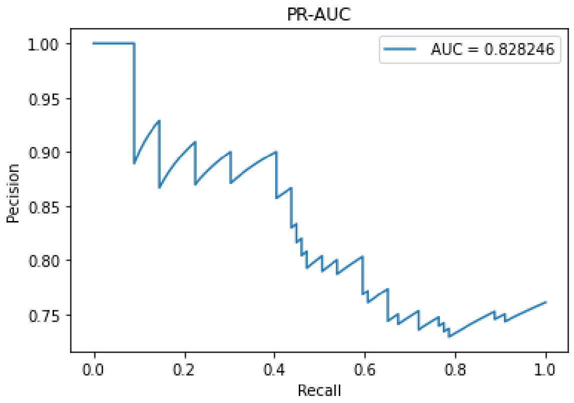


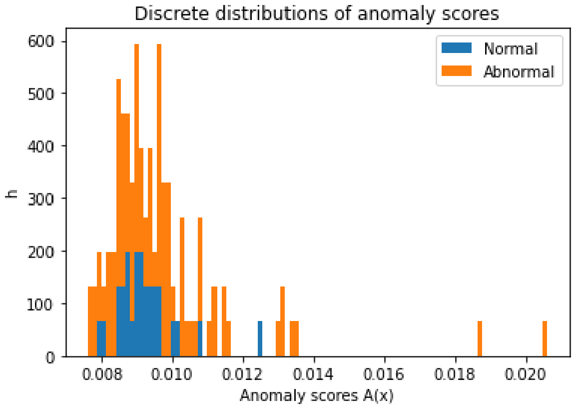




Carpet





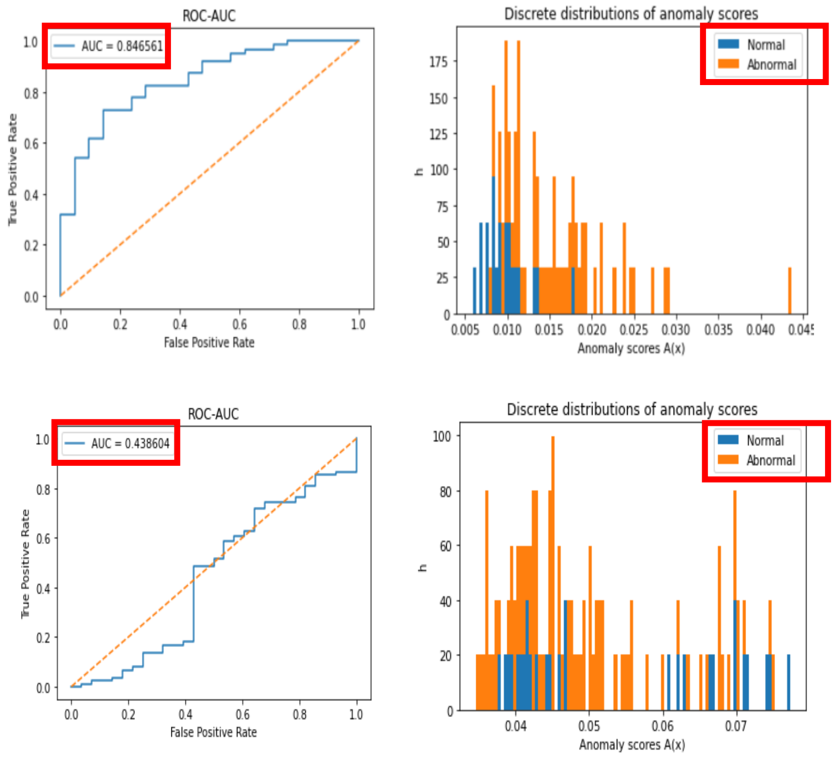


Comparison

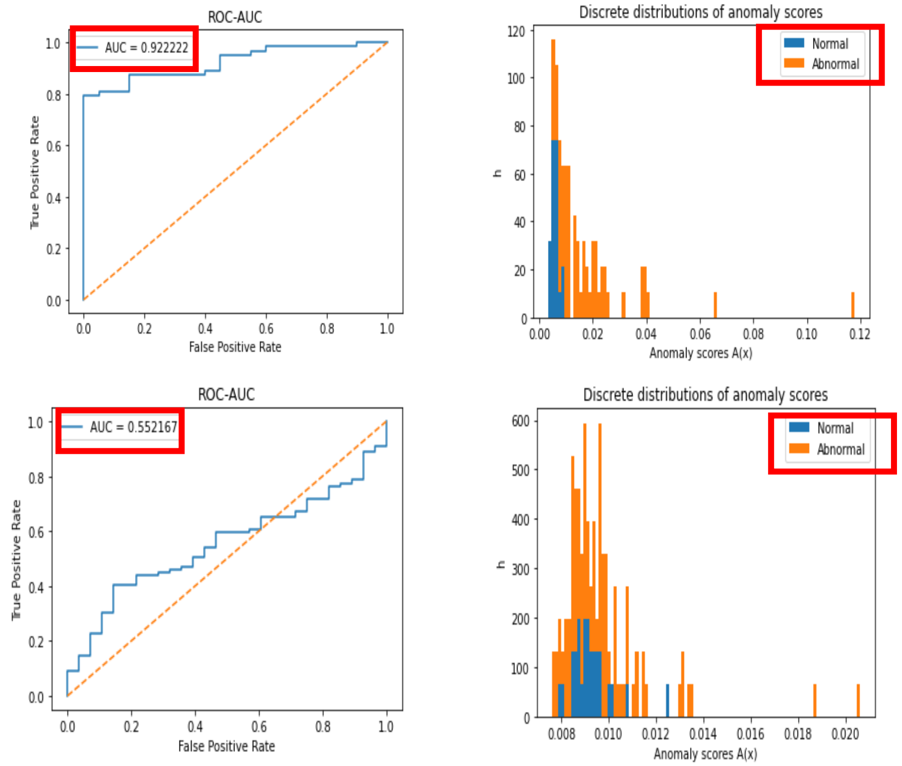
❖ AutoEncoder

Bottle

Carpet



❖ f-AnoGAN

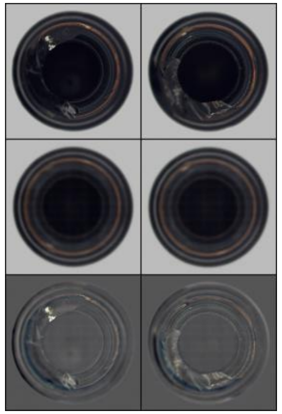


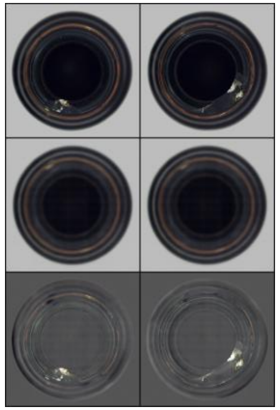
- Both models had better detection performance in the class “Bottle” compared to the class “Carpet”.
  - In the case of the bottle the abnormality detection performance was good, but why was the performance not good in the case of the carpet?

Comparison

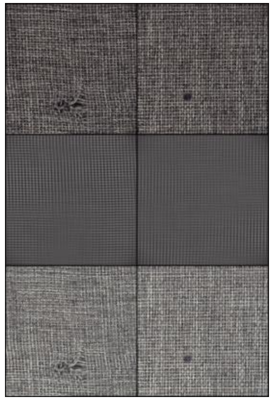
AutoEncoder

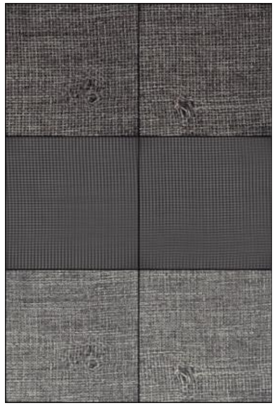
Bottle


broken large

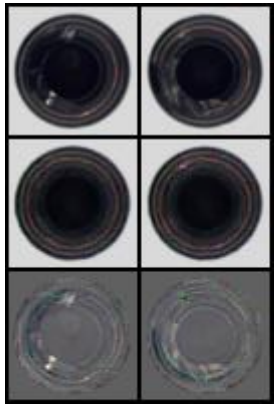

broken small

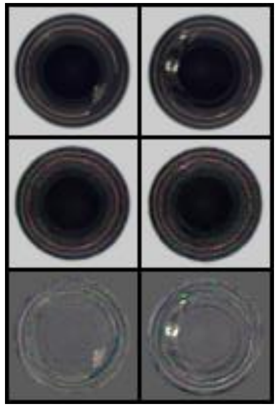
Carpet


cut


hole

f-AnoGAN


broken large


broken small


cut


hole

Limits & Contributions

❖ Limits

- For the class "Carpet", two models are unable to model all the subtle variations of the textural pattern, which results in a complete failure of the method.


hole


metal contamination

❖ Contributions

- If a threshold, which distinguishes abnormal things among products, are set up well, the models are useful to the manufacturing company

- In actual work sites, anomaly detection models of unsupervised learning can be very useful, because there are very few abnormal data compared to normal data.

Reference

- 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.
- Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., & Langs, G. (2017, June). Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In International conference on information processing in medical imaging (pp. 146-157).
- Bergmann, P., Fauser, M., Sattlegger, D., & Steger, C. (2019). MVTec AD--A comprehensive real-world dataset for unsupervised anomaly detection In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9592-9600).

Thank you

Q&A