GAN-based Anomaly Detection on FashionMNIST

Andrea Ferrari and Syrine Enneifer

Dataset : FashionMNIST

.Ten classes, each representing a different category of clothes (e.g. shoes, blouses, etc).

➔ **Normal input data**

Sampled from one class

➔ **Anomaly input data**

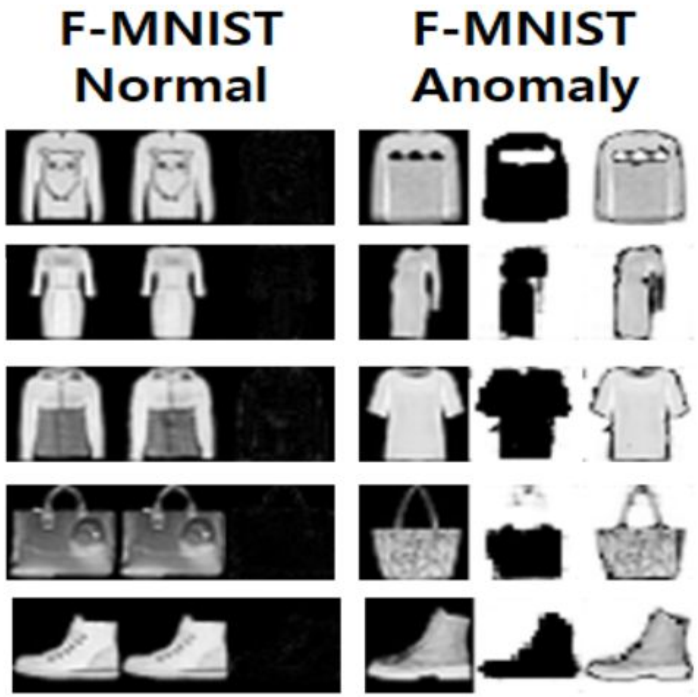
Sampled from the rest (i.e. 9 classes)

Our approach

Reconstruction based approach:

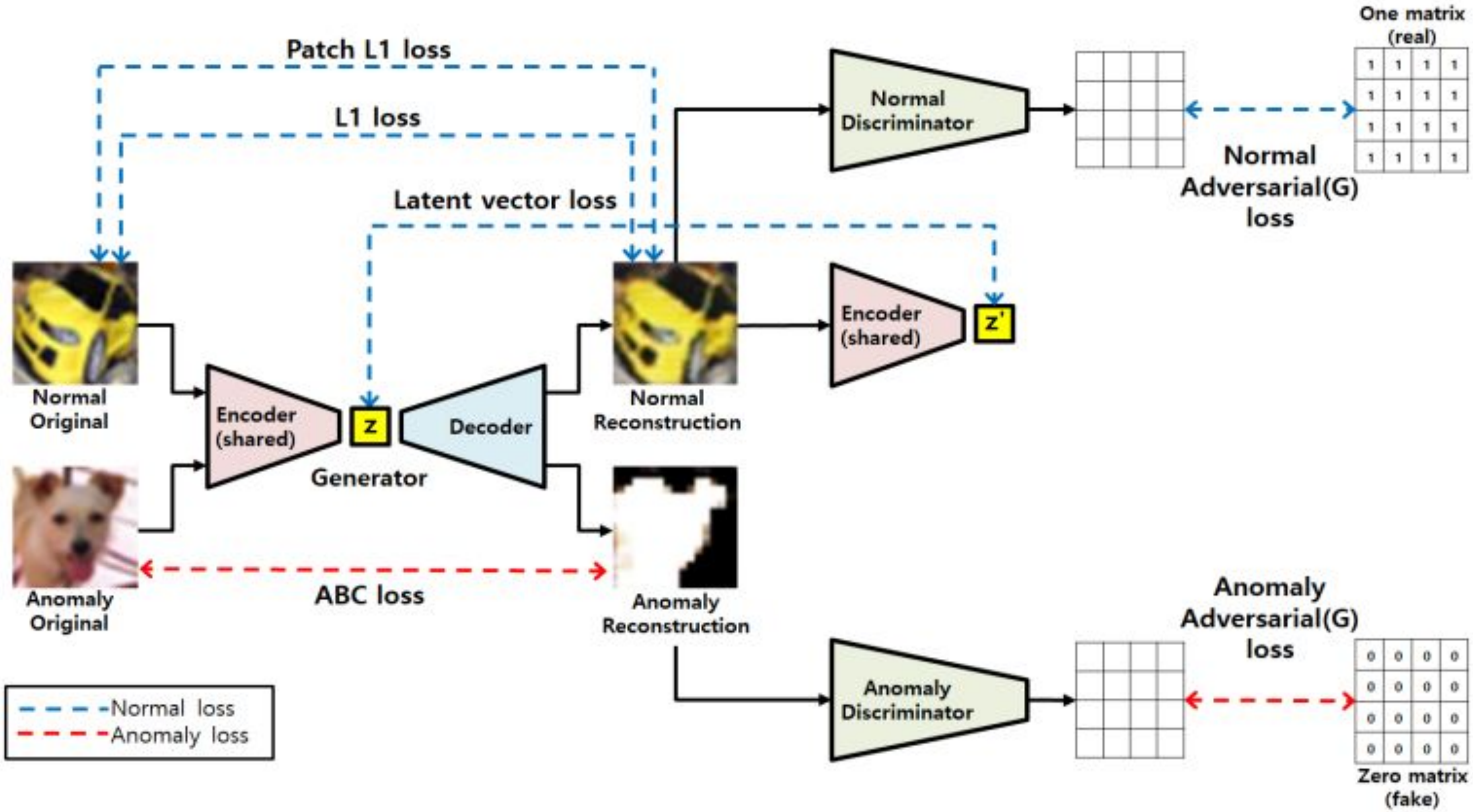
- Generator reconstructs normal input image with little to no noise
- Generator reconstructs anomaly input image with higher noise

→ To achieve this, adversarial training is introduced through 2 discriminators, who will predict whether the image is original or reconstructed.



GAN-based Anomaly Detection in Imbalance Problems, by Junbong Kim , Kwanghee Jeong , Hyomin Choi , and Kisung Seo

Model architecture



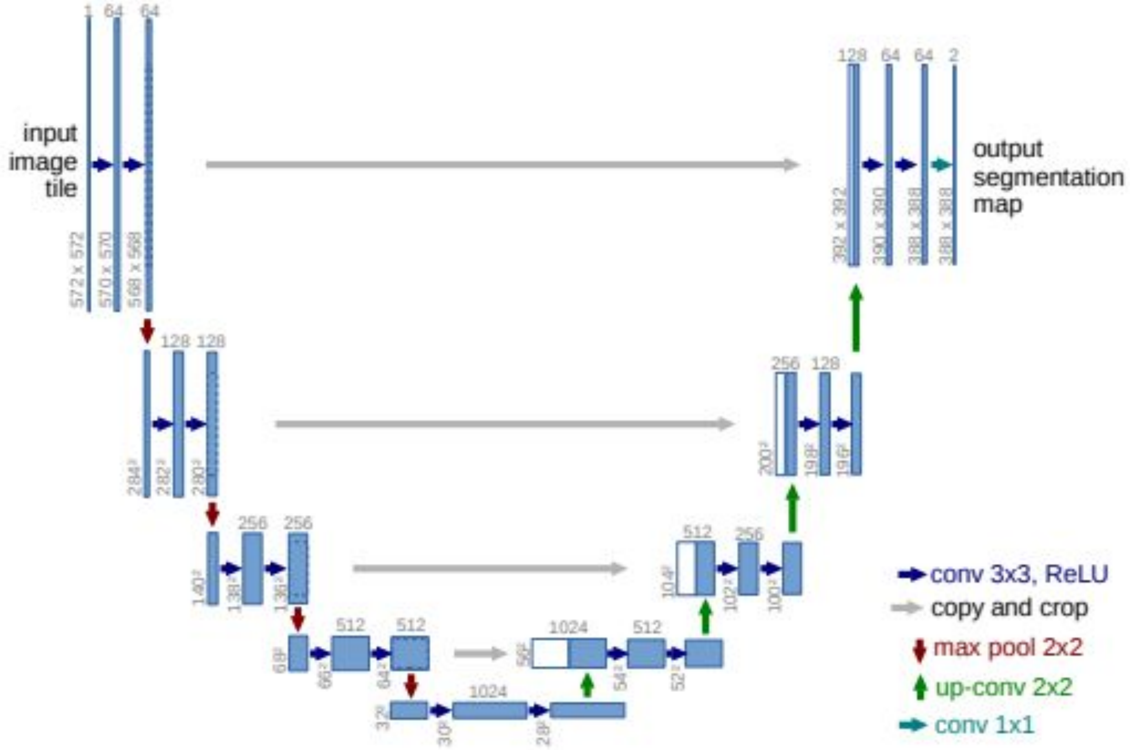
Loss functions

For this task, a total of different **8** losses are used to train our GAN :

- **6** for training the **generator**
- **1** for training the **normal discriminator**
- **1** for training the **anomaly discriminator**

→ Total loss function is a weighted combination of these loss functions.

Generator



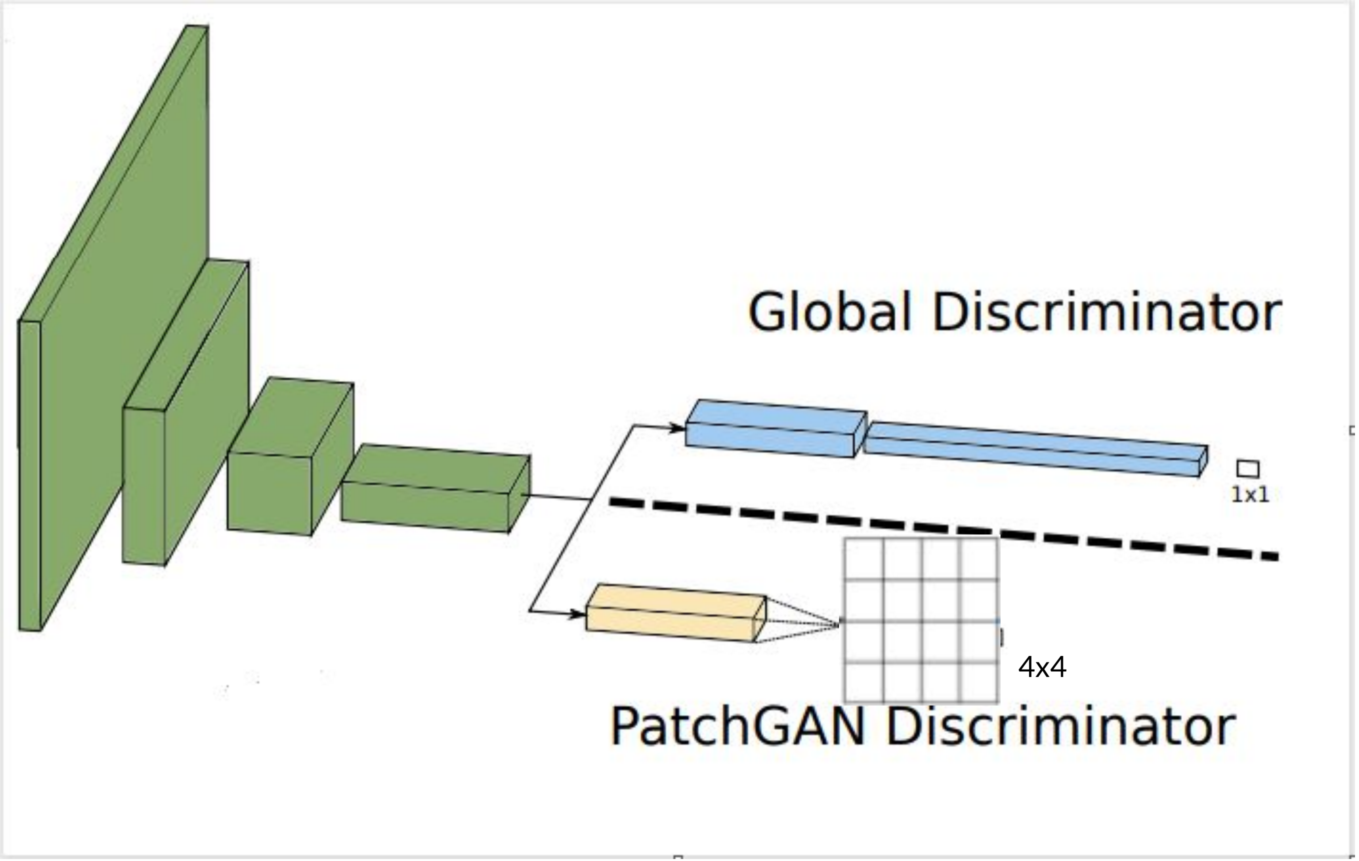
Variation of a U-net[1]

We do not use “max pooling” but strided convolution.

We have 4 convolutional layer, followed by 4 up-convolutions.

[1]: U-Net: Convolutional Networks for Biomedical Image Segmentation, by Olaf Ronneberger, Philipp Fischer, and Thomas Brox

Discriminator

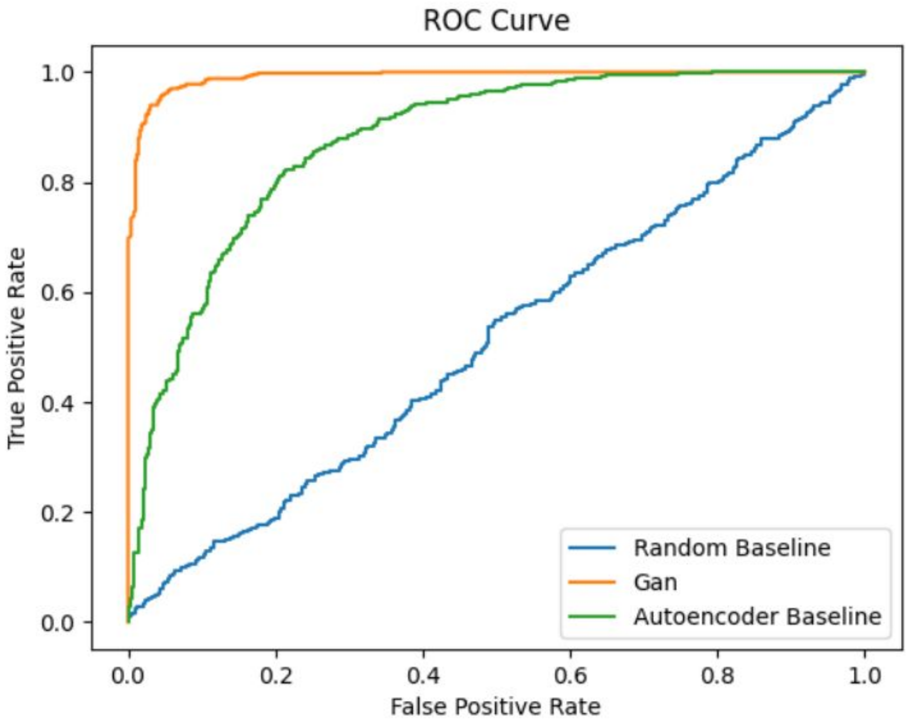


- First return is probability of the image being real or fake
- Second return is probability of the corresponding image patch being real or fake

Results

Main metric to evaluate performance in our case is the **AUROC** score

→ around 0.99155 in our case



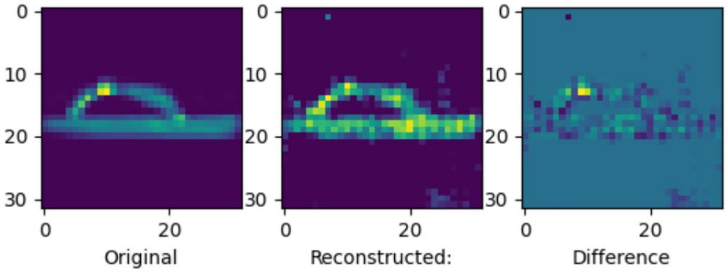
Baselines

→ Autoencoder trained on normal images (AUROC score = 0.8758)

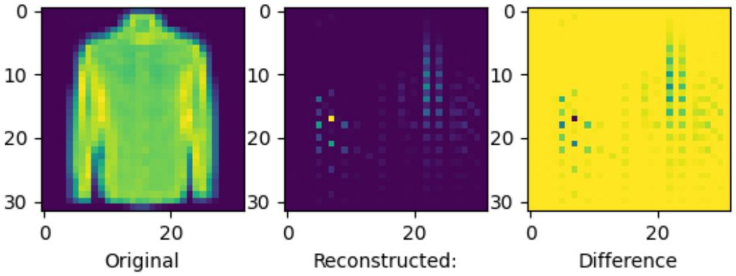
→ Random baseline (AUROC score = 0.49765)

Visualization of experimental results

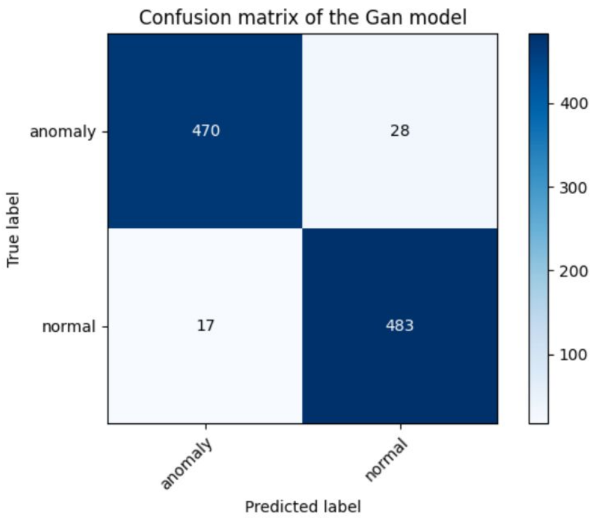
Normal :

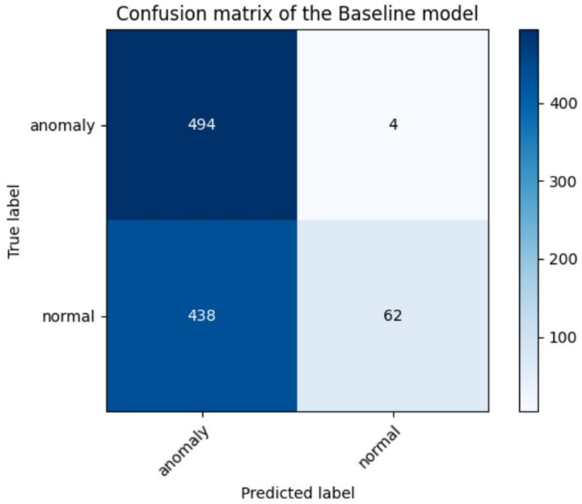


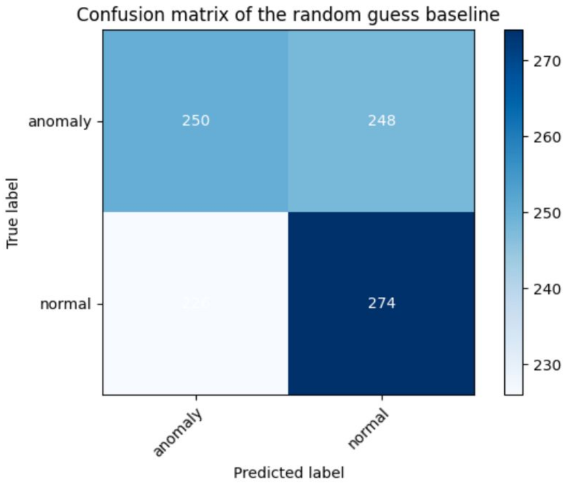
Anomaly :



Confusion Matrices







Conclusions

Our model reached optimal results, getting extremely close to our reference paper **'GAN-based Anomaly Detection in Imbalance Problems'** [3] that we tried to implement.

The model was very effective at distinguishing normal images from anomalous.

[3] : http://intlab.skuniv.ac.kr/paper/GAN-based_Anomaly_Detection_in_Imbalance_Problems.pdf