Machine Learning Homework 8
Anomaly Detection

ML TAs
mlta-2023-spring@googlegroups.com

Outline

- Review
- Task introduction
- Data
- Methodology
- Evaluation
- Baseline
- Report

Review

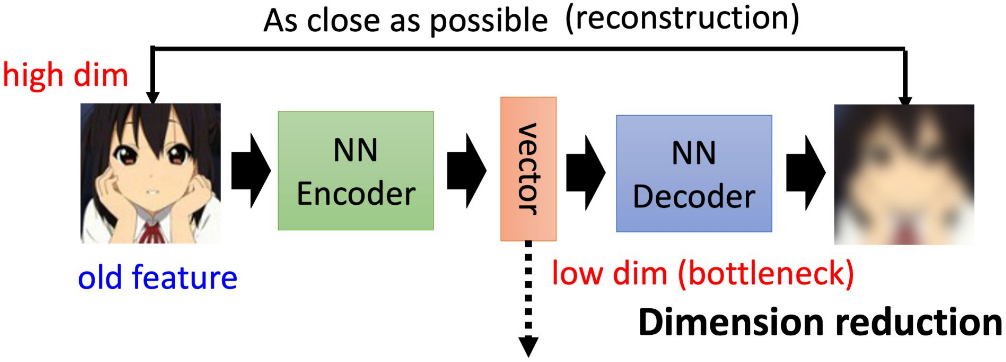
https://youtu.be/3oHlf8-J3Nc

https://youtu.be/JZvEzb5PV3U

Auto-encoder

Unlabeled Images

Sounds familiar? We have seen the same idea in Cycle GAN. ☺



Embedding, Representation, Code

New feature for downstream tasks

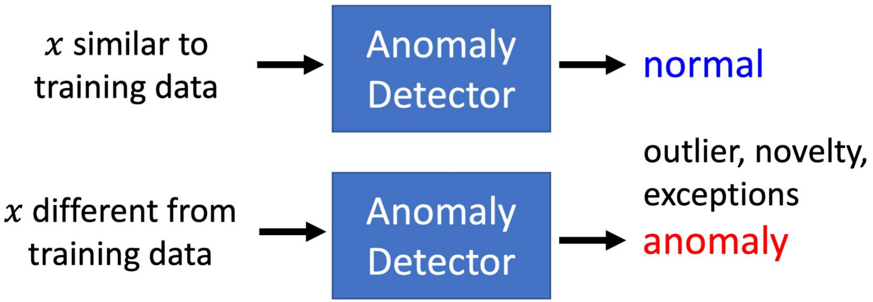
Review

https://youtu.be/3oHlf8-J3Nc

https://youtu.be/JZvEzb5PV3U

Anomaly Detection

- Given a set of training data $\{x^1, x^2, \cdots, x^N\}$
- Detecting input x is *similar* to training data or not.



Review

https://youtu.be/3oHlf8-J3Nc

https://youtu.be/JZvEzb5PV3U

Anomaly Detection

Binary Classification?

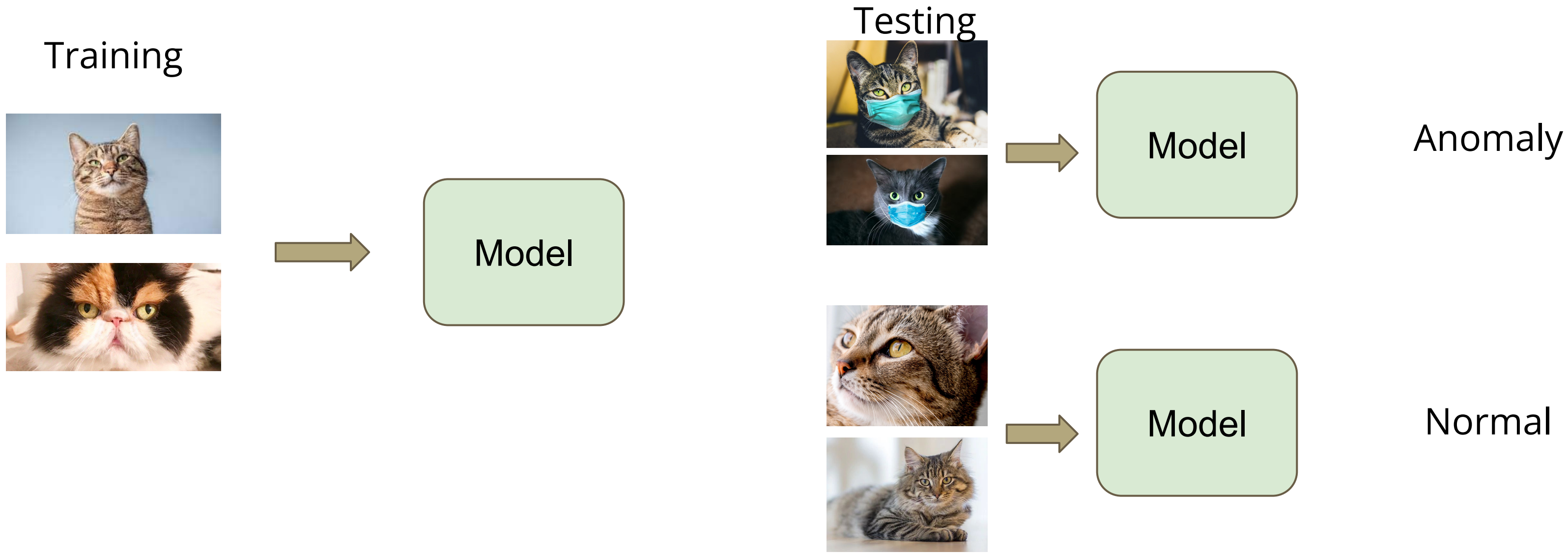
We only have one class.

Training auto-encoder

- Fraud Detection
  - Training data: credit card transactions, x: fraud or not
  - Ref: https://www.kaggle.com/ntnu-testimon/paysim1/home
  - Ref: https://www.kaggle.com/mlg-ulb/creditcardfraud/home
- Network Intrusion Detection
  - Training data: connection, x: attack or not
  - Ref: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
- Cancer Detection
  - Training data: normal cells, x: cancer or not?
  - Ref: https://www.kaggle.com/uciml/breast-cancer-wisconsin-data/home

Task Introduction

- Unsupervised anomaly detection
    - Training a model to determine whether the given image is similar with the training data.



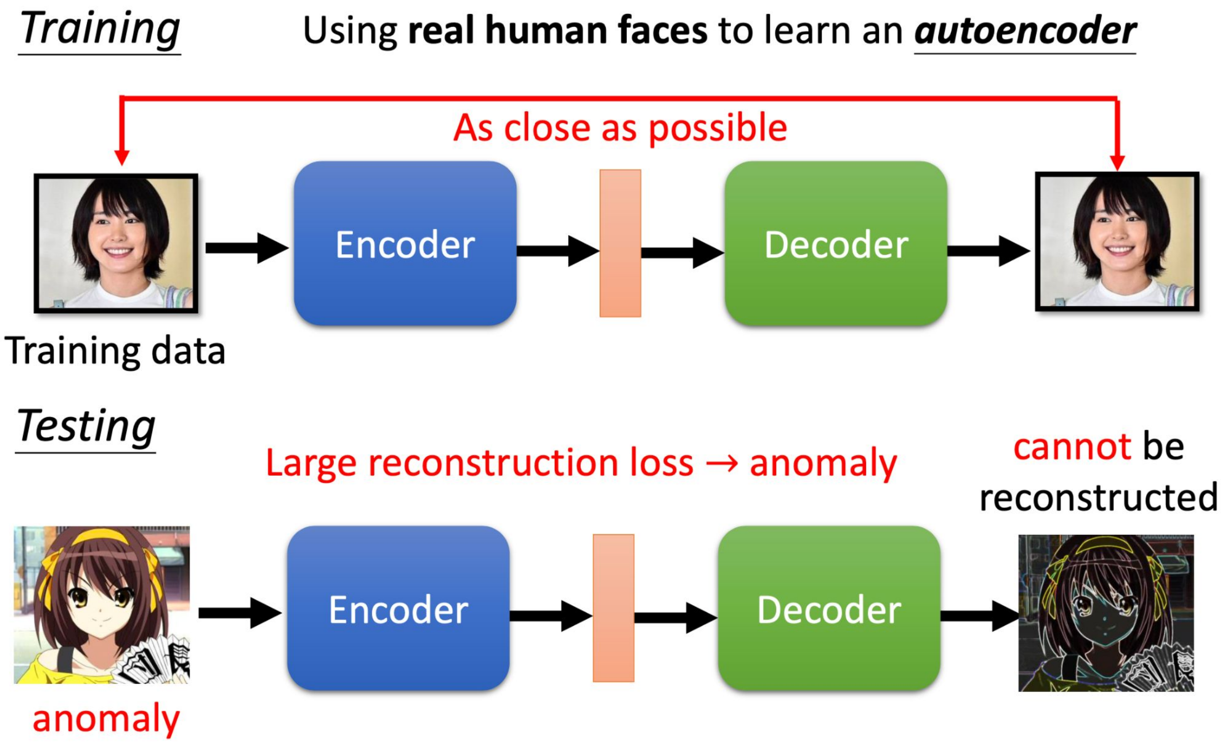
Data

- Training data
  - 100000 human faces
- Testing data
  - About 10000 from the same distribution with training data (label 0)
  - About 10000 from another distribution (anomalies, label 1)
- Format
  - data/

    |----- trainingset.npy

    |----- testingset.npy
  - Shape: (#images, 64, 64, 3) for each .npy file

Methodology



Methodology

- Train an autoencoder with small reconstruction error.
- During inference, we can use reconstruction error as anomaly score.
  - Anomaly score can be seen as the degree of abnormality of an image.
  - An image from unseen distribution should have higher reconstruction error.
- Anomaly scores are used as our predicted values.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

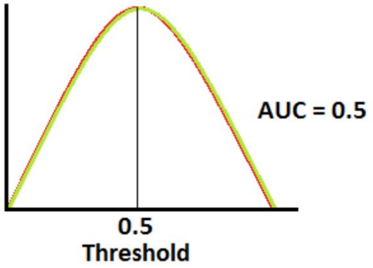
Evaluation - ROC AUC score

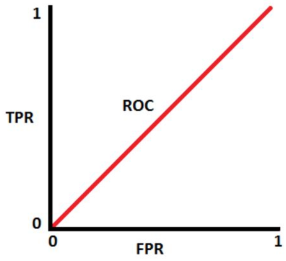
Why using ROC AUC score?

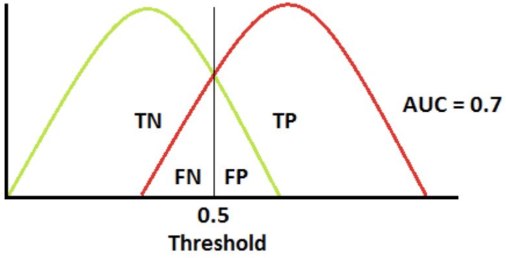
- If accuracy is applied, then a threshold is needed to determine the given image is an anomaly or not.
  - We only want a model that tells us how anomalous an image is.
  - e.g. MSE is a kind of anomaly score
- More about ROC curve
  - https://en.wikipedia.org/wiki/Receiver_operating_characteristic

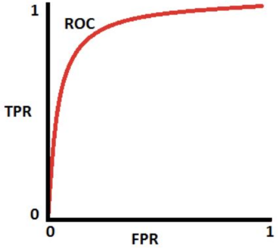
Evaluation - ROC AUC score

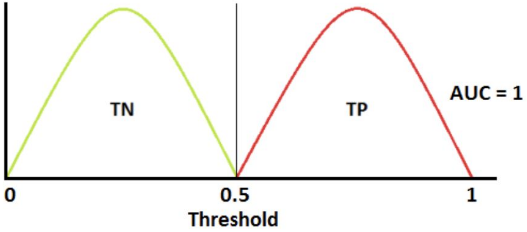
- TPR = TP / (TP + FN)
- FPR = FP / (FP + TN)
- AUC (Area Under Curve)

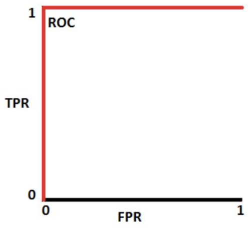












https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5

Evaluation - ROC AUC score Example

ID	Anomaly score	Label
0	11383	0
1	256676	1
2	862365	1
3	152435	0
4	848171	0

Sort by score

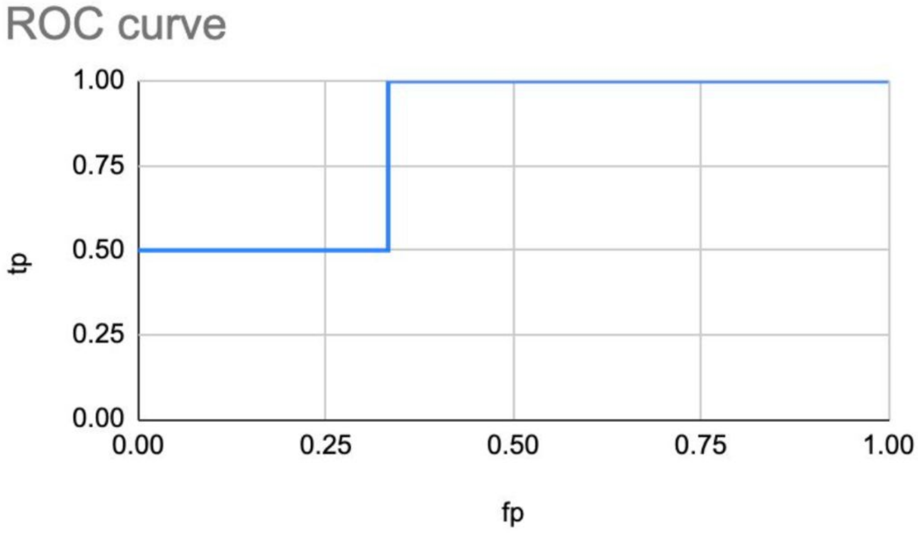
ID	Anomaly score	Label
2	862365	1
4	848171	0
1	256676	1
3	152435	0
0	11383	0

Evaluation - ROC AUC score Example

ID	Anomaly score	Label	fp before normalization	tp before normalization
2	862365	1	0	1
4	848171	0	1	1
1	256676	1	1	2
3	152435	0	2	2
0	11383	0	3	2

Evaluation - ROC AUC score Example

ID	Anomaly score	Label	fp	tp
0	11383	0	0	0.5
3	152435	0	0.333333	0.5
1	256676	1	0.333333	1
4	848171	0	0.666667	1
2	862365	1	1	1

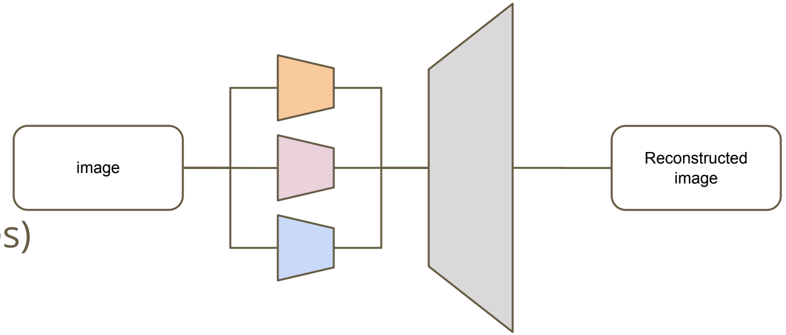


Area Under Curve: $0.5*\frac{1}{3} + \frac{2}{3} = 0.8333$

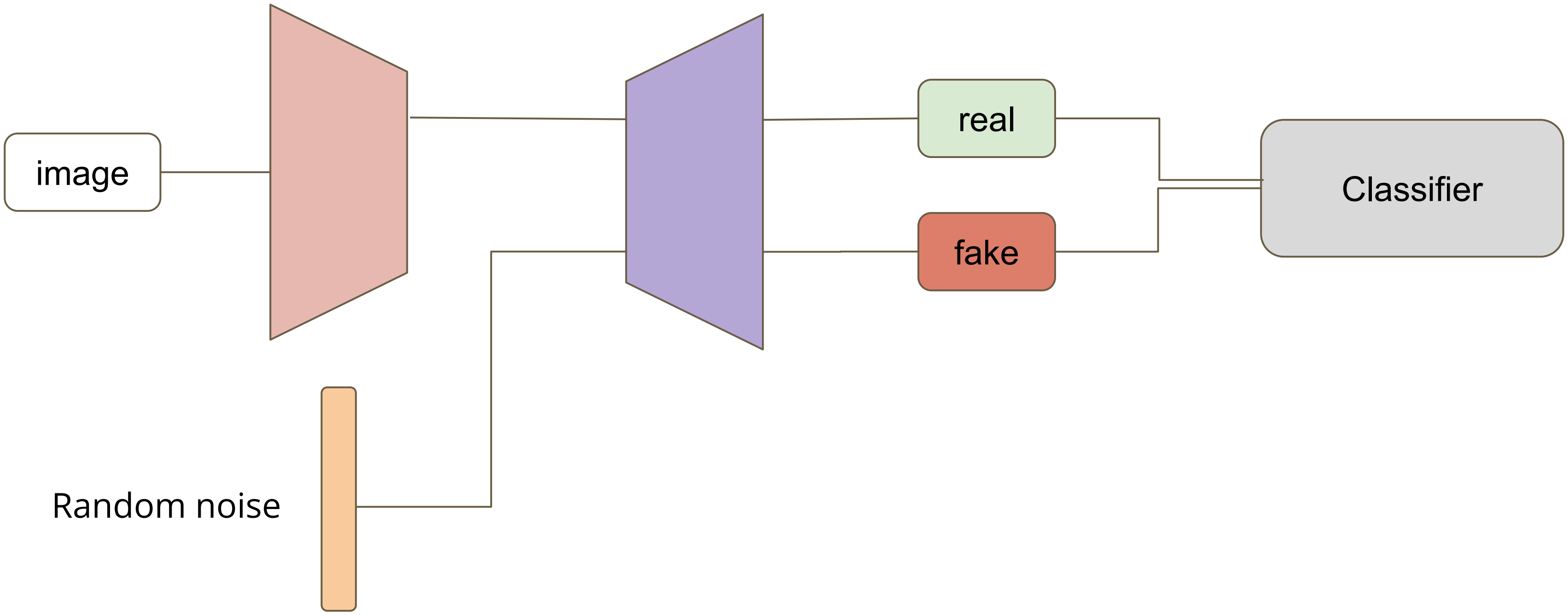
Baseline

- Simple
  - Sample code (10 minutes)
- Medium
  - Adjust model structure (15 ~ 20 minutes)
- Strong
  - Multi-encoder autoencoder (30 ~ 40 minutes)
  - Paper reference
- Boss
  - Add random noise and an extra classifier or use Resnet as encoder (30 ~ 40 minutes)
  - Paper reference
- Papers of anomaly detection

Multi-encoder autoencoder



Add random noise and extra classifier



Report

https://youtu.be/YNUek8ioAJk
https://youtu.be/8zomhgKrsmQ

1. Choose a variation of autoencoder. Show an image of the model architecture. Then, list an advantage and a disadvantage comparing with vanilla autoencoder. Also, put on the paper link as reference. Eg, denoising autoencoder, variational autoencoder, etc.
2. Train a fully connected autoencoder and adjust at least two different element of the latent representation. Show your model architecture, plot out the original image, the reconstructed images for each adjustment and describe the differences.

Report - Q2

For instance, let z be the output of encoder. Then we can adjust the first dimension of z as follows:

- z[0] = 2 * z[0]

Note: you should use the same autoencoder and only adjust the latent representation (output of encoder).

Grading

Simple Baseline (Public /Private)	+0.5 pts / +0.5 pts
Medium Baseline (Public /Private)	+0.5 pts / +0.5 pts
Strong Baseline (Public /Private)	+0.5 pts / +0.5 pts
Boss Baseline (Public /Private)	+0.5 pts / +0.5 pts
Code Submission	+2 pts
Report	+4 pts

Submission Format

- "ID,score" in the first row
- Followed by 19636 lines of "image ID,anomaly score"

```
ID,score
0,18.029802
1,29.577963
2,33.817013
3,36.073986
4,29.43562
```

Code Submission

- Submit your code to NTU COOL
  - We can only see your last submission
  - Do not submit the model or dataset
  - If your codes are not reasonable, your final grade will be x 0.9
  - You should compress your code into a single file
    - <student_id>_hw8.zip

Deadline

- Kaggle: 2023/05/19 23:59 (UTC+8)
- NTU COOL: 2023/05/19 23:59 (UTC+8)
- Gradescope: 2023/05/19 23:59 (UTC+8)

Link

- Kaggle: link
- Colab: link

Regulations

- You should NOT plagiarize, if you use any other resource, you should cite it in the reference.
- You should NOT modify your prediction files manually.
- Do NOT share codes or prediction files with any living creatures.
- Do NOT use any approaches to submit your results more than 5 times a day.
- Do NOT search or use additional data or pre-trained models.
- Your final grades x 0.9 if you violate of the above rules.
- Prof. Lee & TAs preserve the rights to change the rules & grades.

(*) Academic Ethics Guidelines for Researchers by the Ministry of Science and Technology (MOST)

If any questions, you can ask us via ...

- NTU COOL (Recommended)
  - https://cool.ntu.edu.tw/courses/24108/discussion_topics/196628
- Email
  - mlta-2023-spring@googlegroups.com
  - The title should begin with “[hw8]”
- TA hour
  - In-person: Each Friday during class
  - Online: Each Monday night on google meet
    - Link: Released on NTU Cool Discussion Board
    - 19:00 - 20:00 (Mandarin)
    - 20:00 - 21:00 (English)