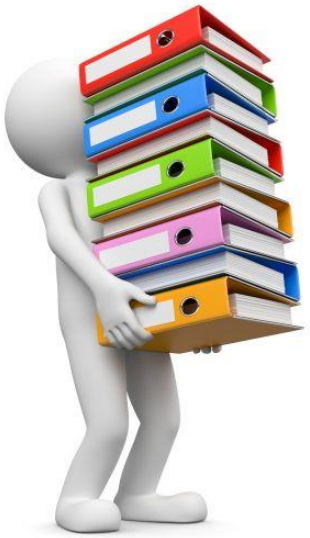
AUTO-ENCODER

Hung-yi Lee 李宏毅

Self-supervised Learning Framework



data

Self-supervised Learning

Pre-train

Auto-encoder

Self-supervised

Learned by tasks not requiring label data

Model for Task 1

Model for Task 2

Model for Task 3

Downstream Tasks

Outline

Basic Idea of Auto-encoder

Feature Disentanglement

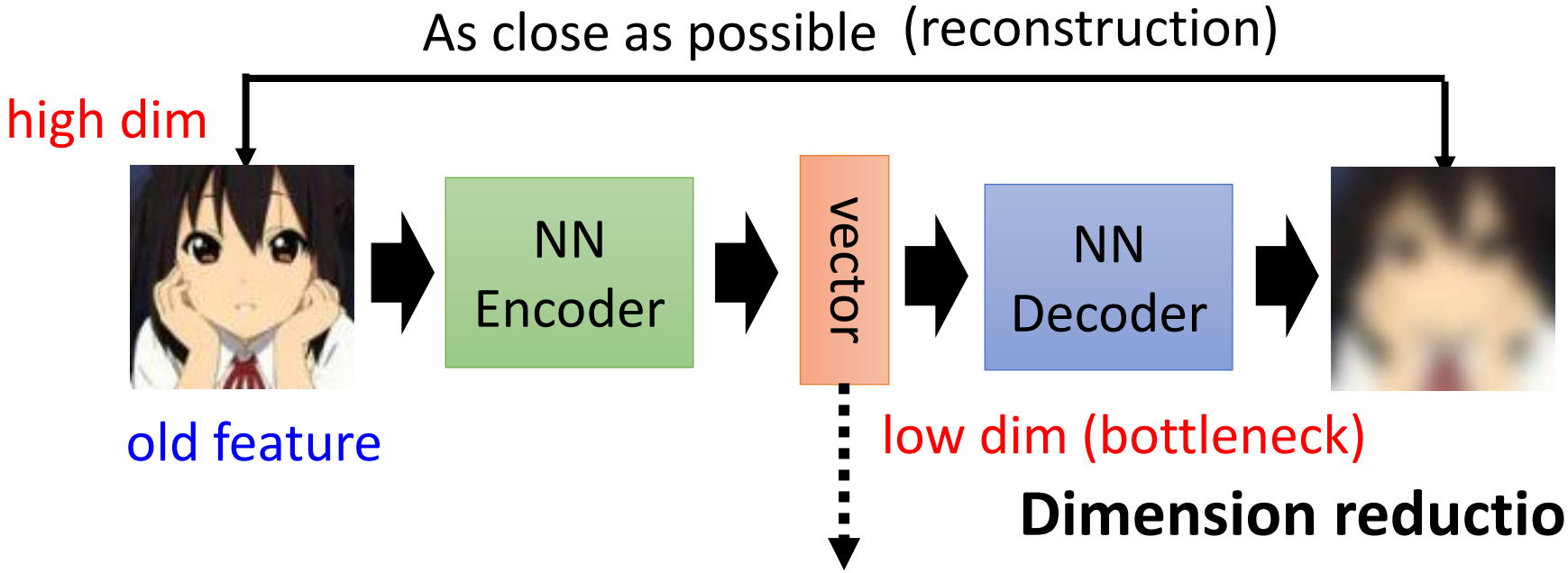
Discrete Latent Representation

More Applications

Auto-encoder

Unlabeled Images

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



More Dimension Reduction

(not based on deep learning)

https://youtu.be/iwh5o_M4BNU

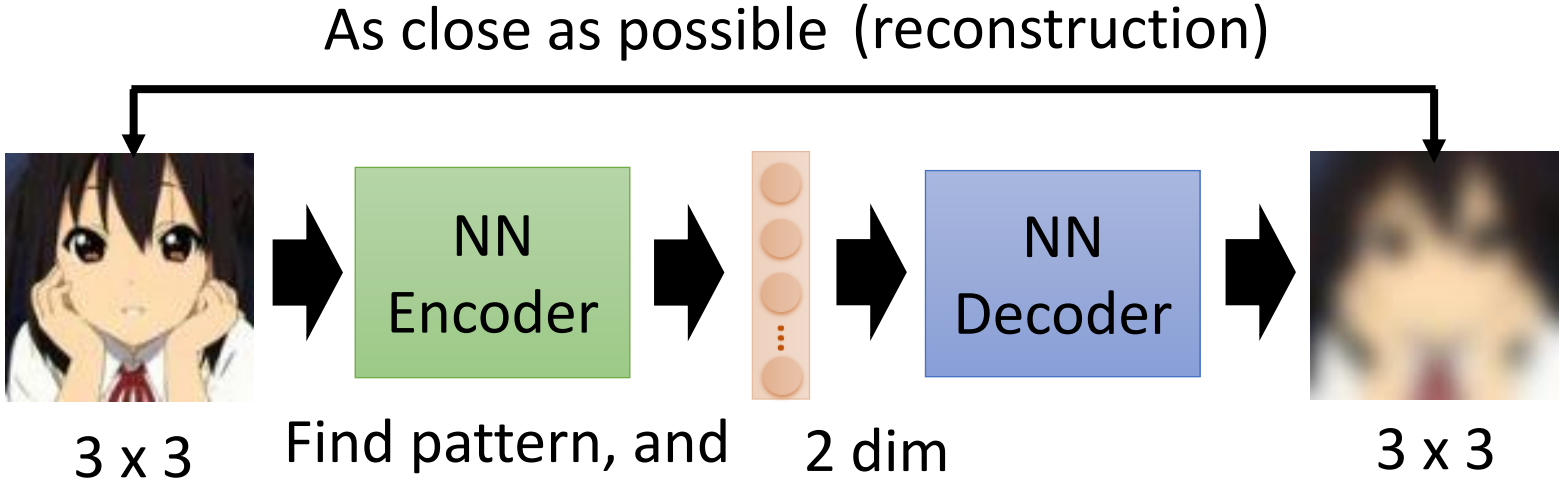
PCA

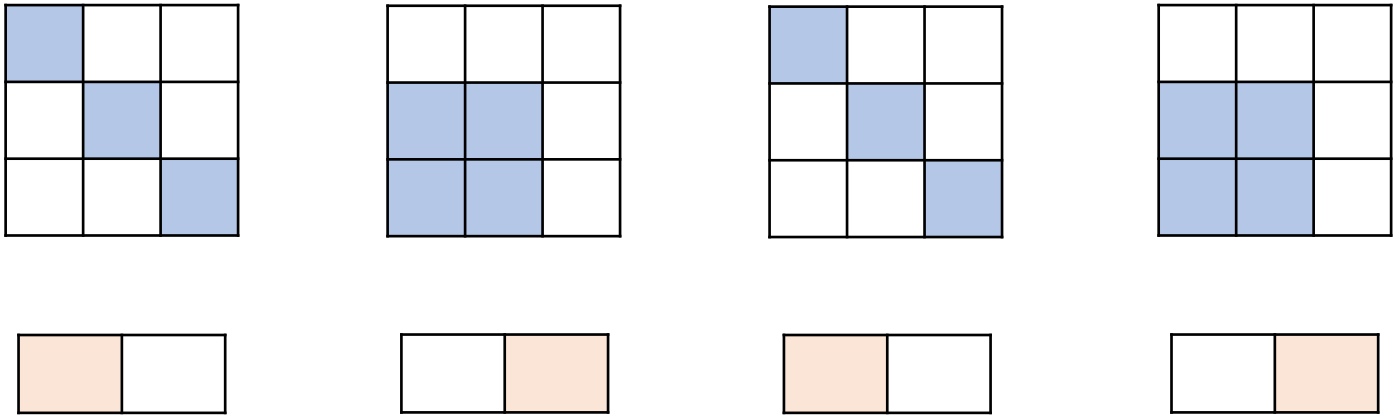
https://youtu.be/GBUEjkpoxXc

t-SNE

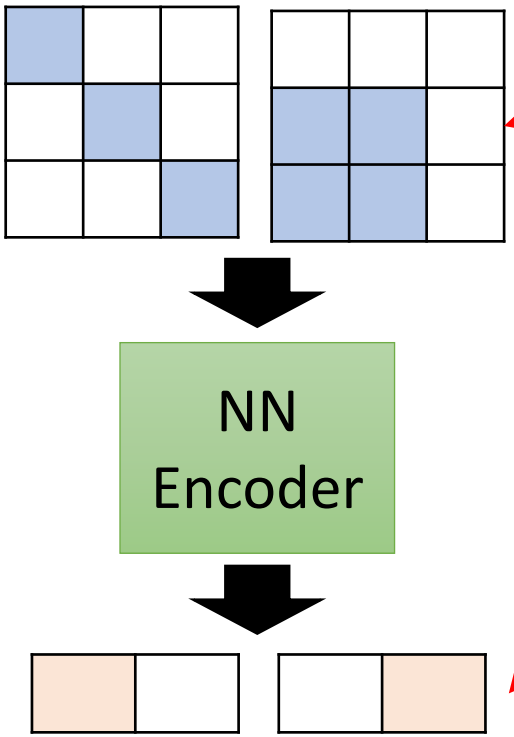
Why Auto-encoder?

Why Auto-encoder?



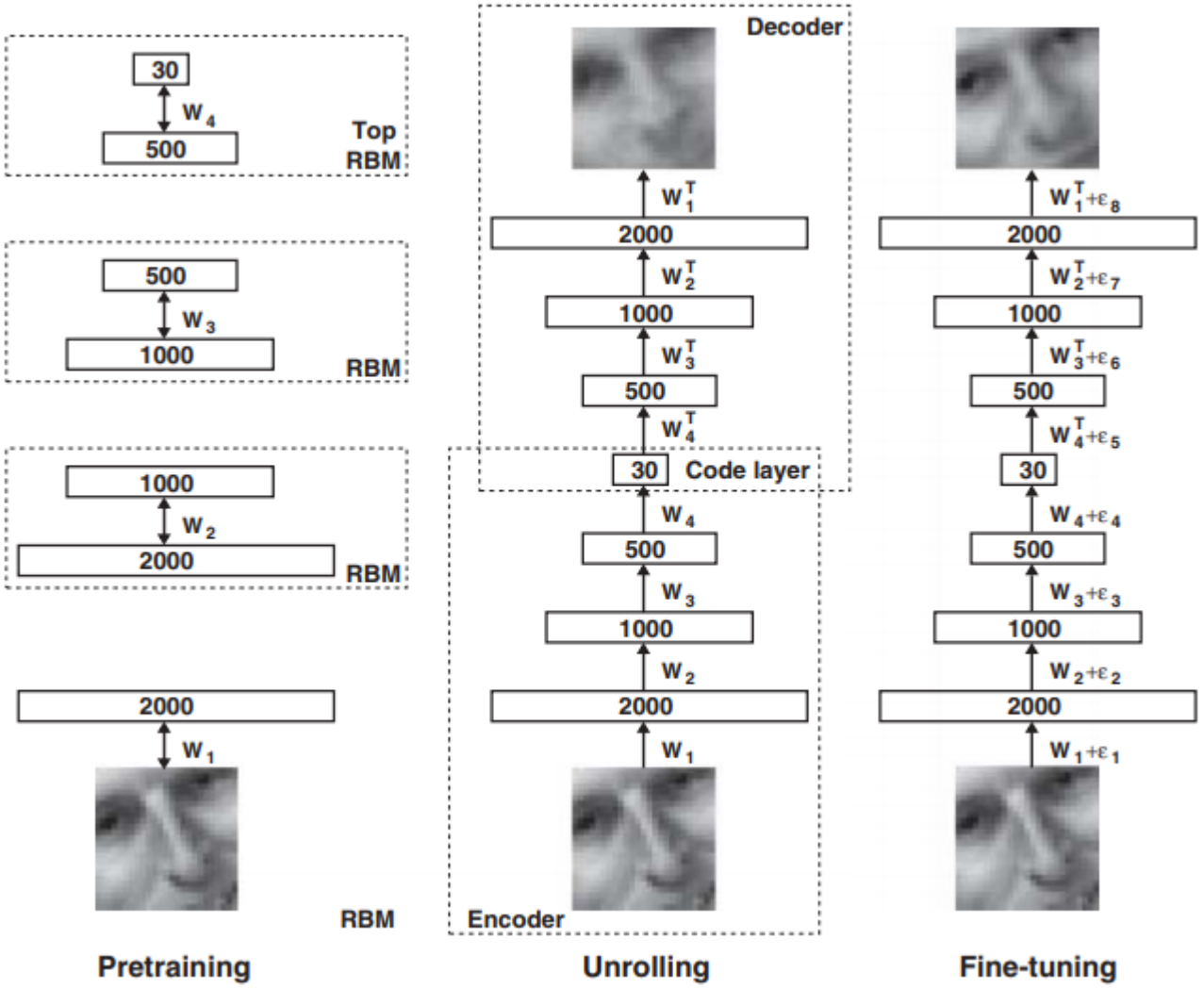


Why Auto-encoder?

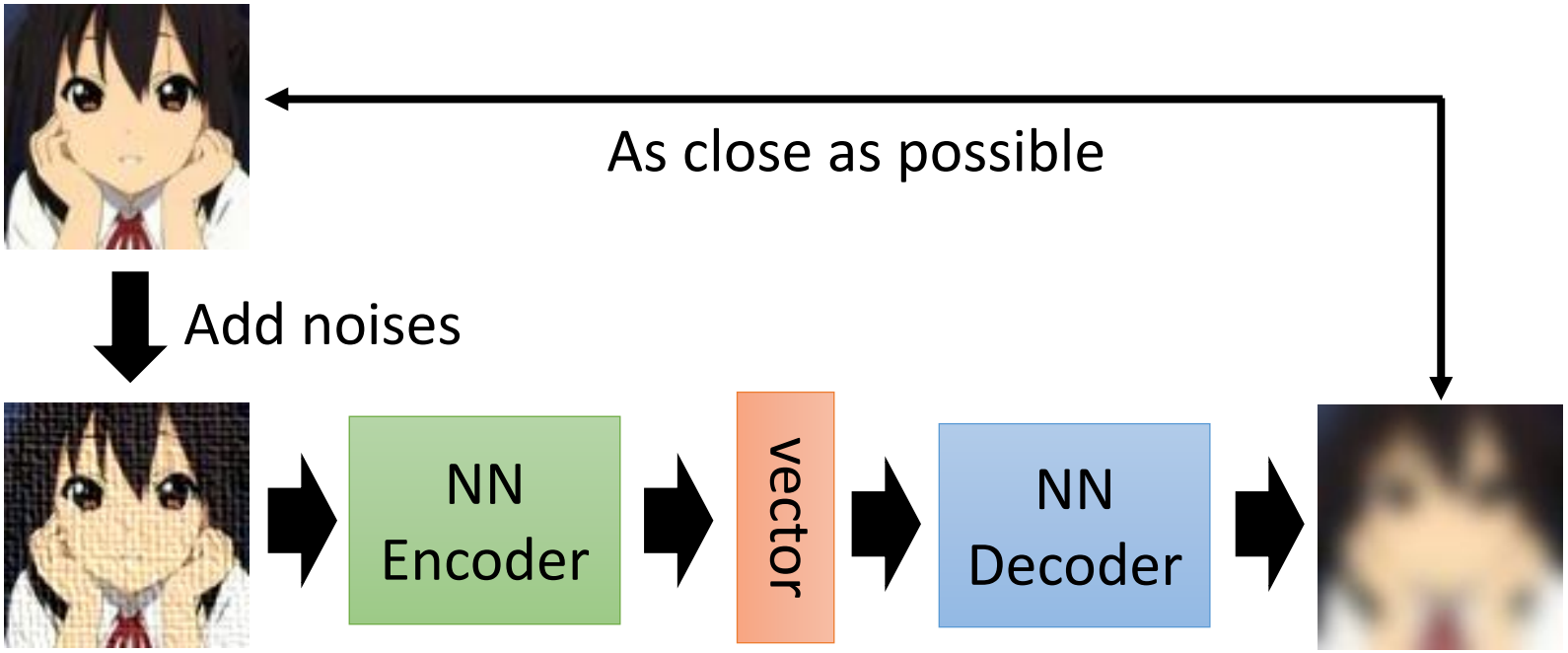


Auto-encoder is not a new idea

Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *Science* 313.5786 (2006): 504-507



De-noising Auto-encoder

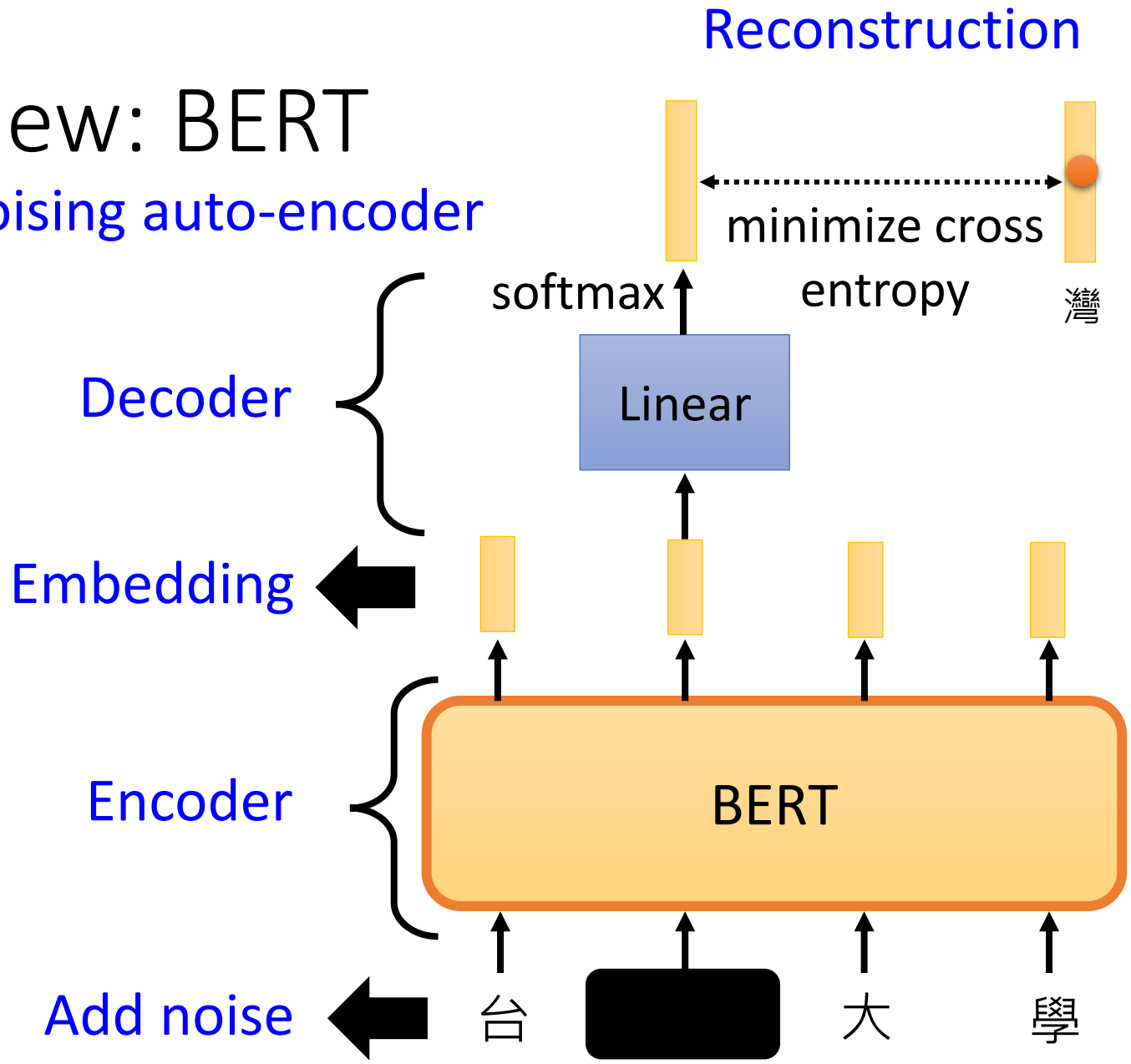


The idea sounds familiar? ☺

Vincent, Pascal, et al. "Extracting and composing robust features with denoising autoencoders." *ICML,* 2008.

Review: BERT

A de-noising auto-encoder



Outline

Basic Idea of Auto-encoder

Feature Disentanglement

Discrete Latent Representation

More Applications

Representation includes information of different aspects

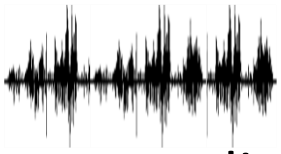
Encoder

Decoder

input image

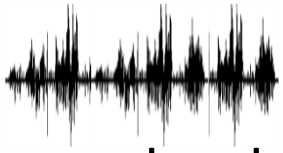
object, texture ….

reconstructed



Encoder

Decoder



input audio

content, speaker …

reconstructed

Encoder

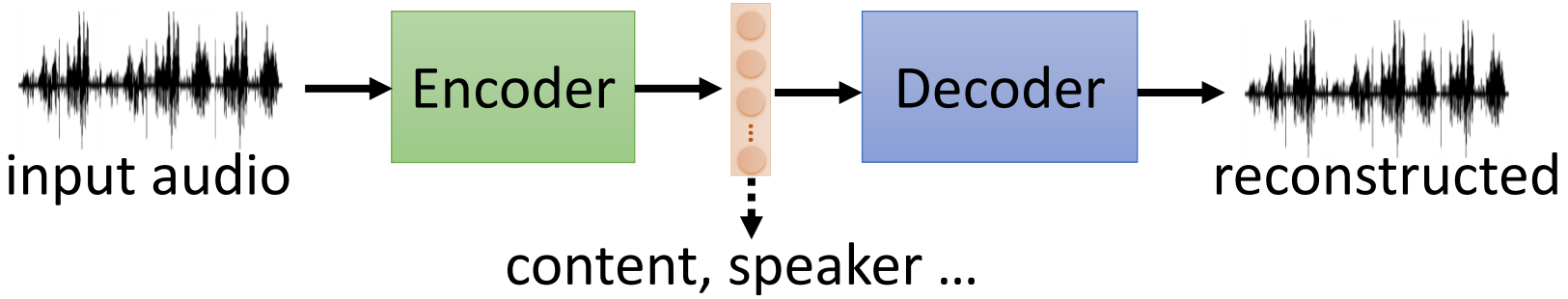
Decoder

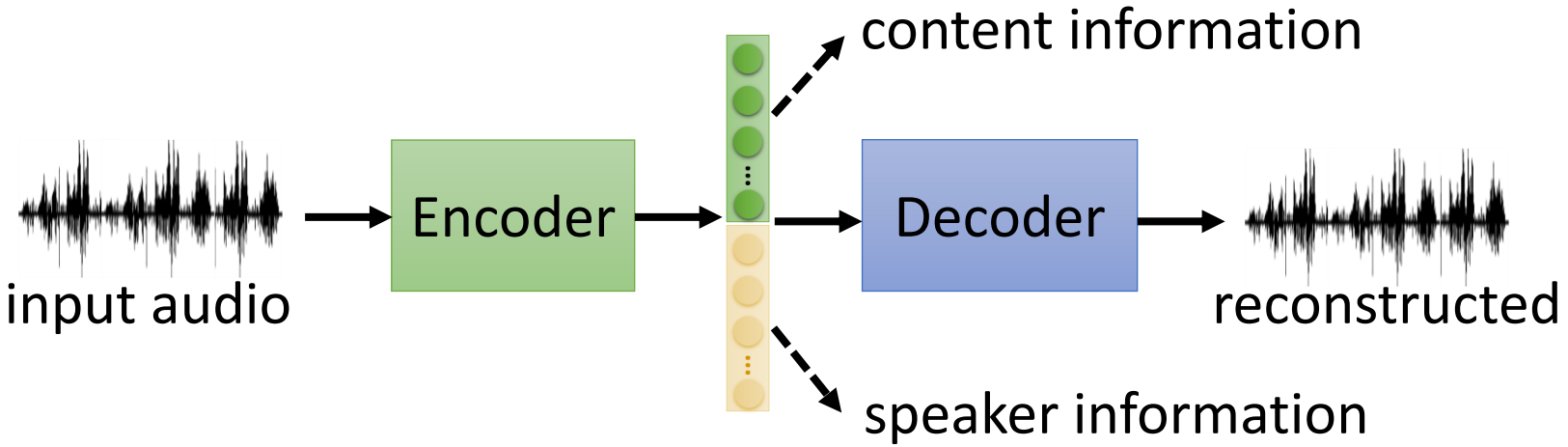
input sentence

syntax, semantic …

reconstructed

Feature Disentangle



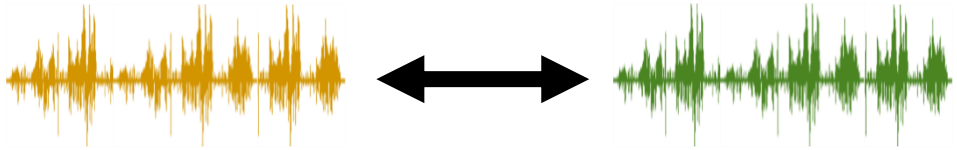


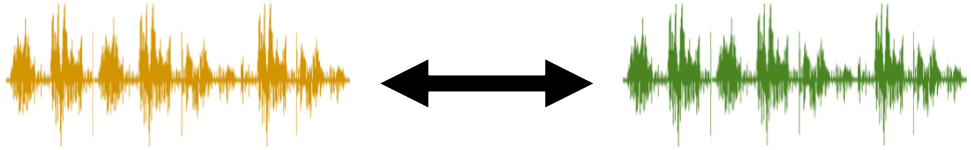
https://arxiv.org/abs/1904.05742
https://arxiv.org/abs/1804.02812
https://arxiv.org/abs/1905.05879

Application: Voice Conversion

In the past

Speaker A | Speaker B

How are you?  How are you?

Good morning  Good morning

Today

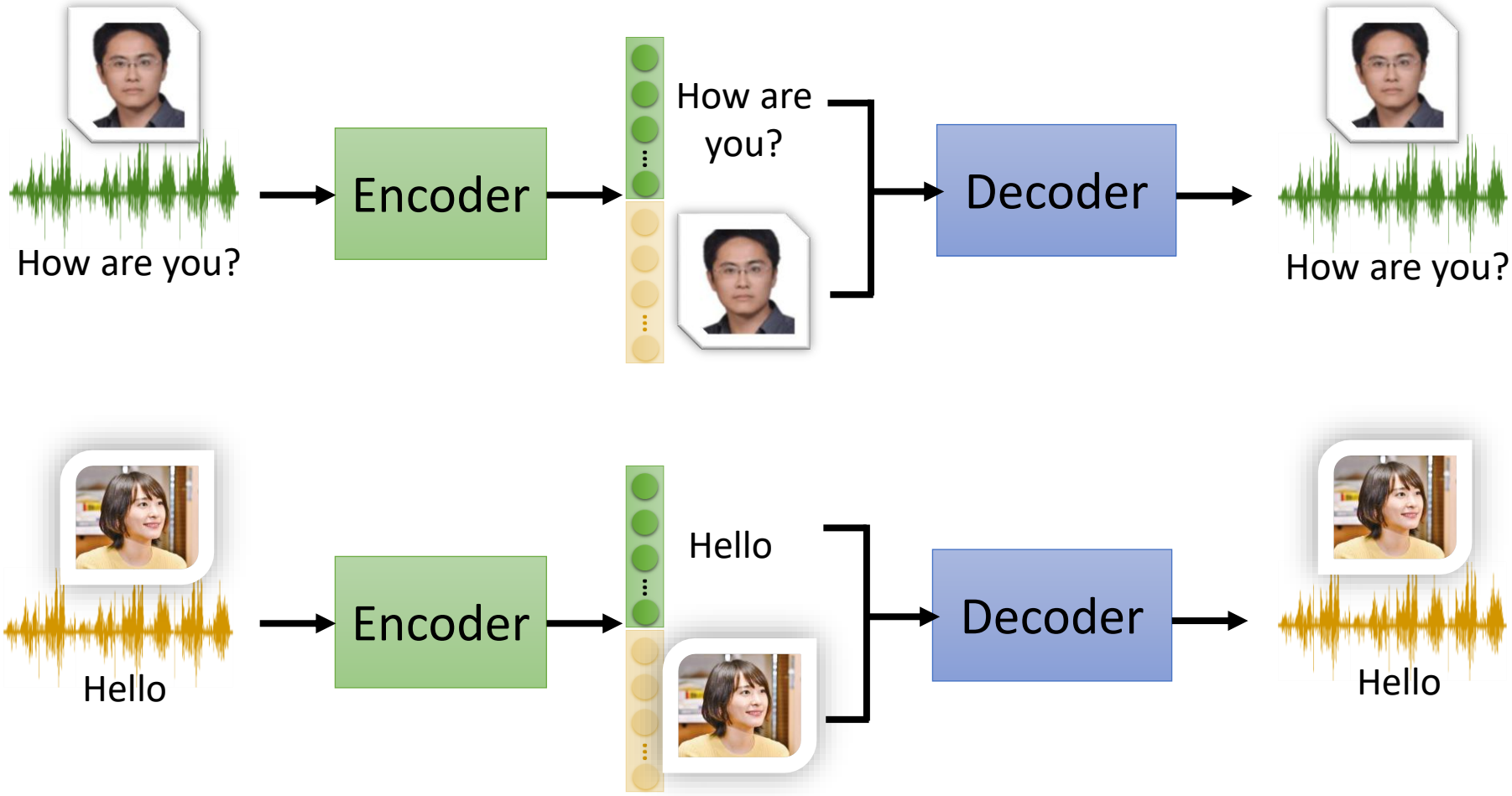
Speaker A | Speaker B

天氣真好 How are you?

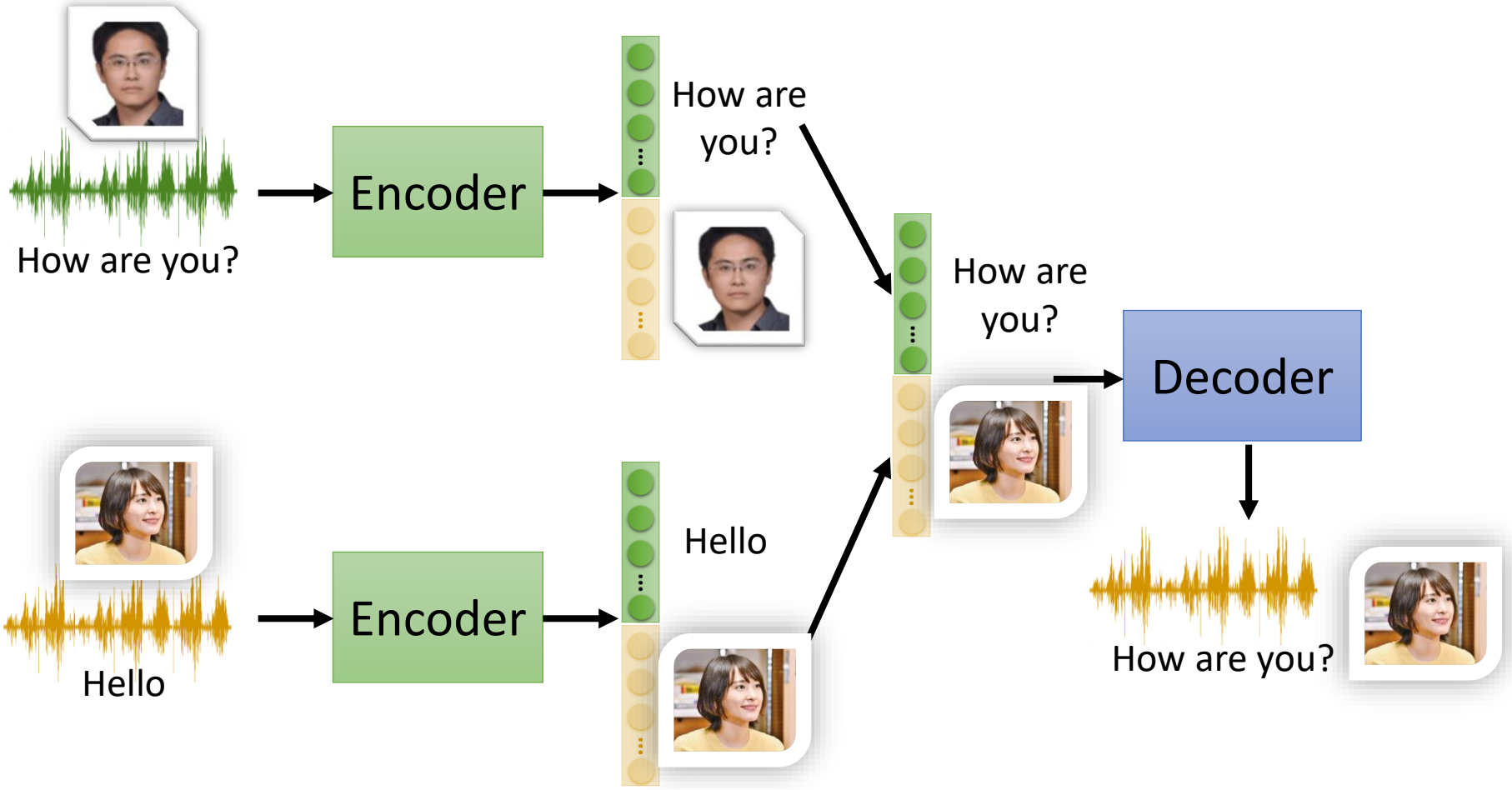
再見囉 Good morning

Speakers A and B are talking about completely different things.

Application: Voice Conversion

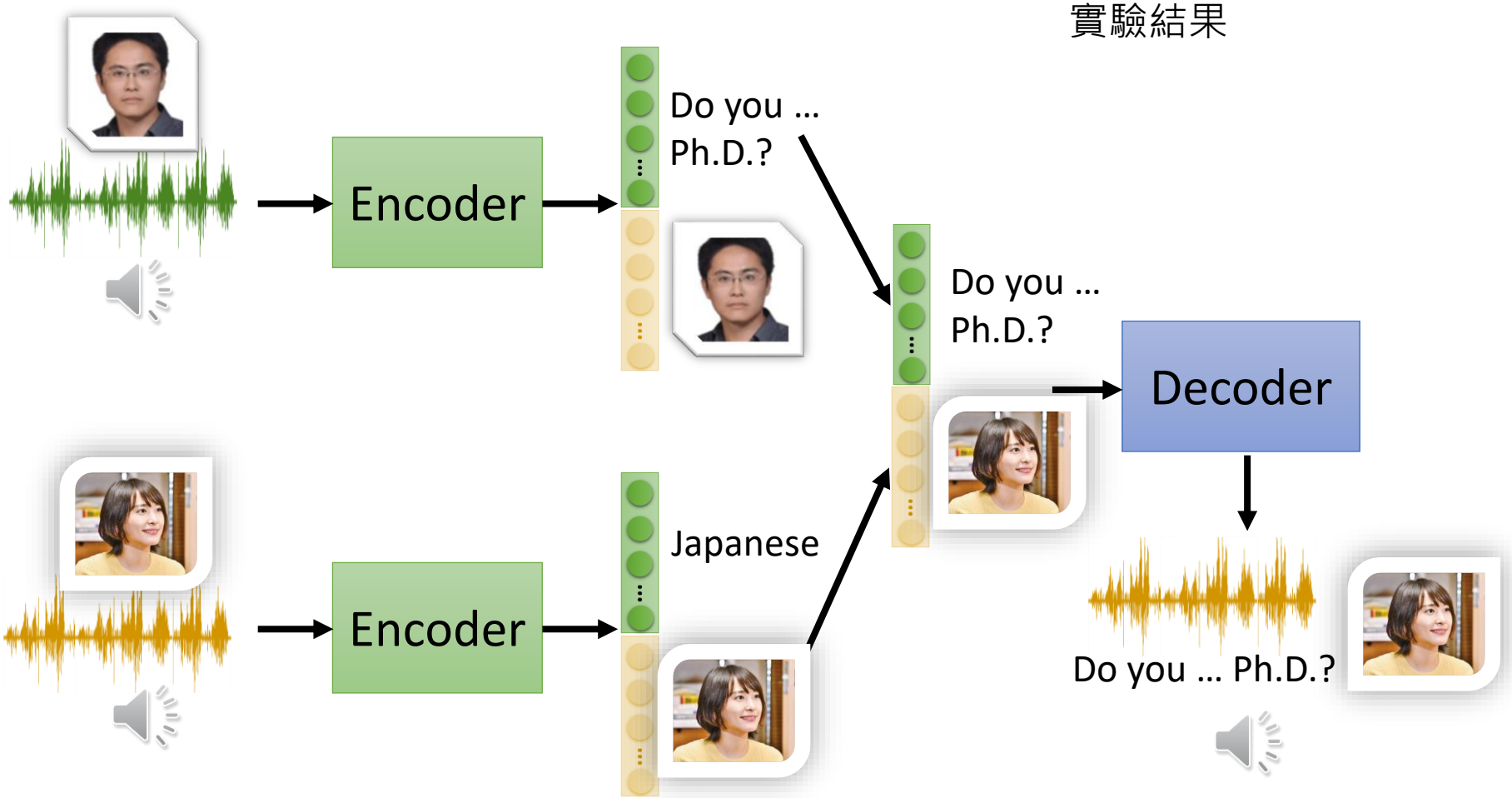


Application: Voice Conversion



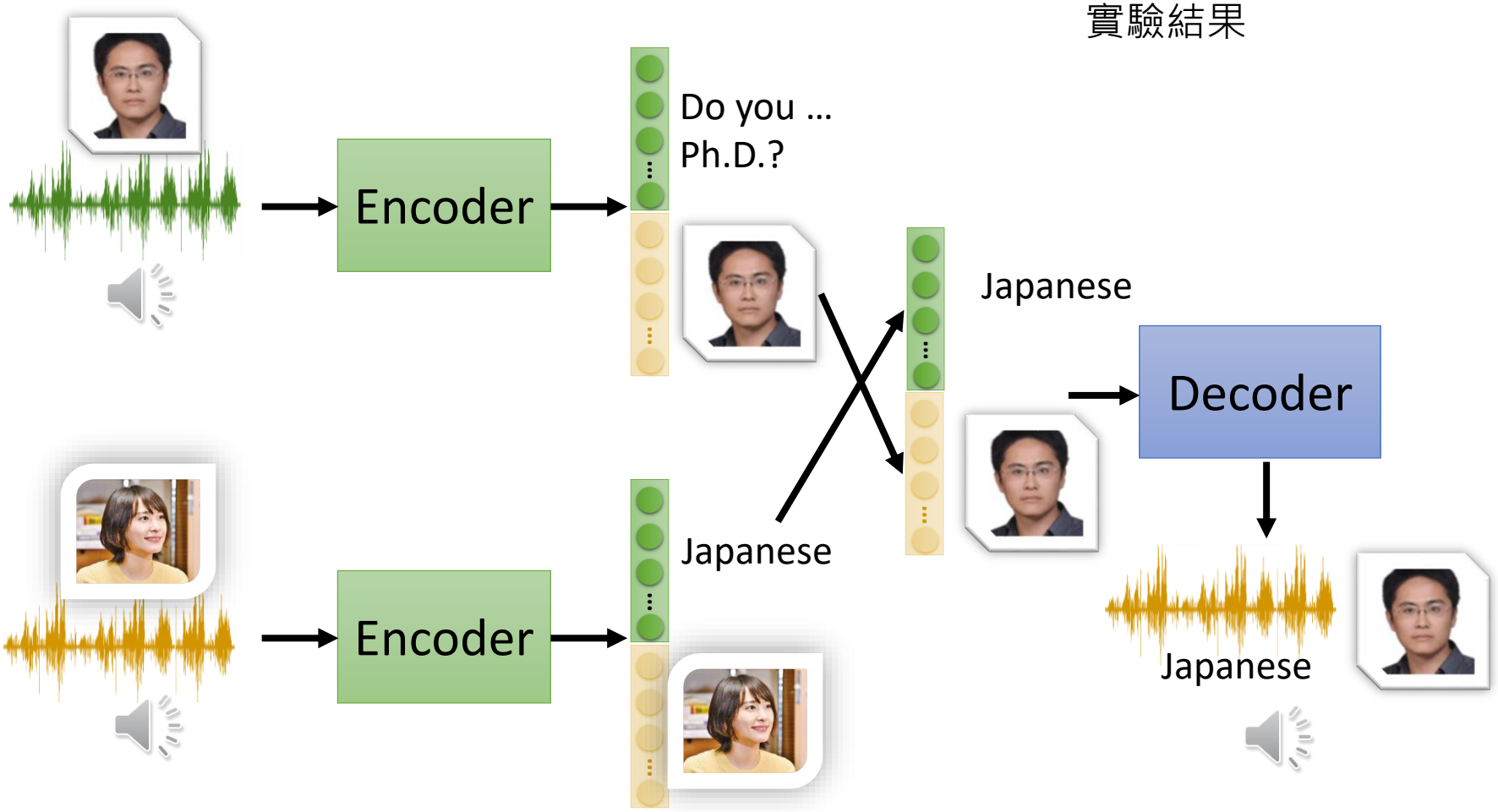
Application: Voice Conversion

感謝解正平同學提供
實驗結果



Application: Voice Conversion

感謝解正平同學提供
實驗結果



Outline

Basic Idea of Auto-encoder

Feature Disentanglement

Discrete Latent Representation

More Applications

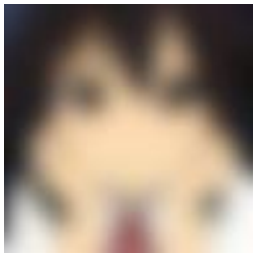
Discrete Representation

Real numbers

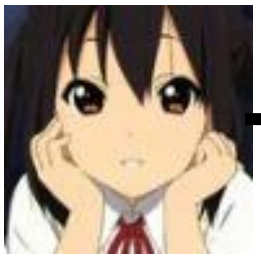
NN Encoder

0.9
0.1
0.3
0.7

NN Decoder



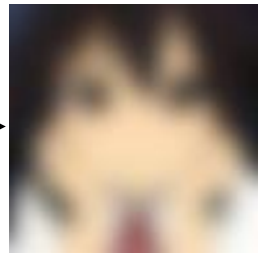
Binary



NN Encoder

1
0
0
1

NN Decoder



glasses

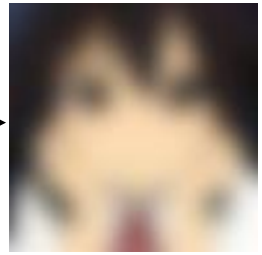
female

One-hot

NN Encoder

0
0
1
0

NN Decoder



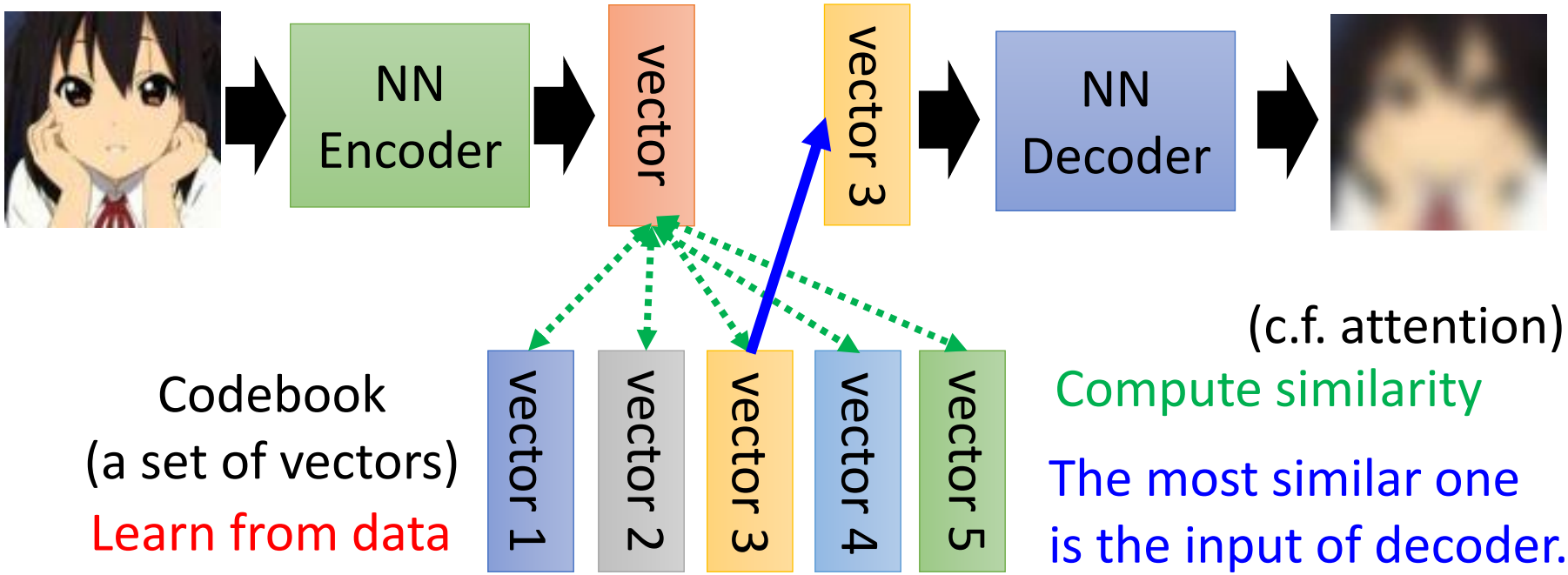
class

也許可以做到unsupervised classification

Discrete Representation

https://arxiv.org/abs/1711.00937

- Vector Quantized Variational Auto-encoder (VQVAE)



Codebook
(a set of vectors)
Learn from data

(c.f. attention)
Compute similarity

The most similar one
is the input of decoder.

For speech, the codebook represents phonetic information

這些codebooks可能就能represent最基本的

https://arxiv.org/pdf/1901.08810.pdf

https://arxiv.org/abs/1810.02851

Text as Representation

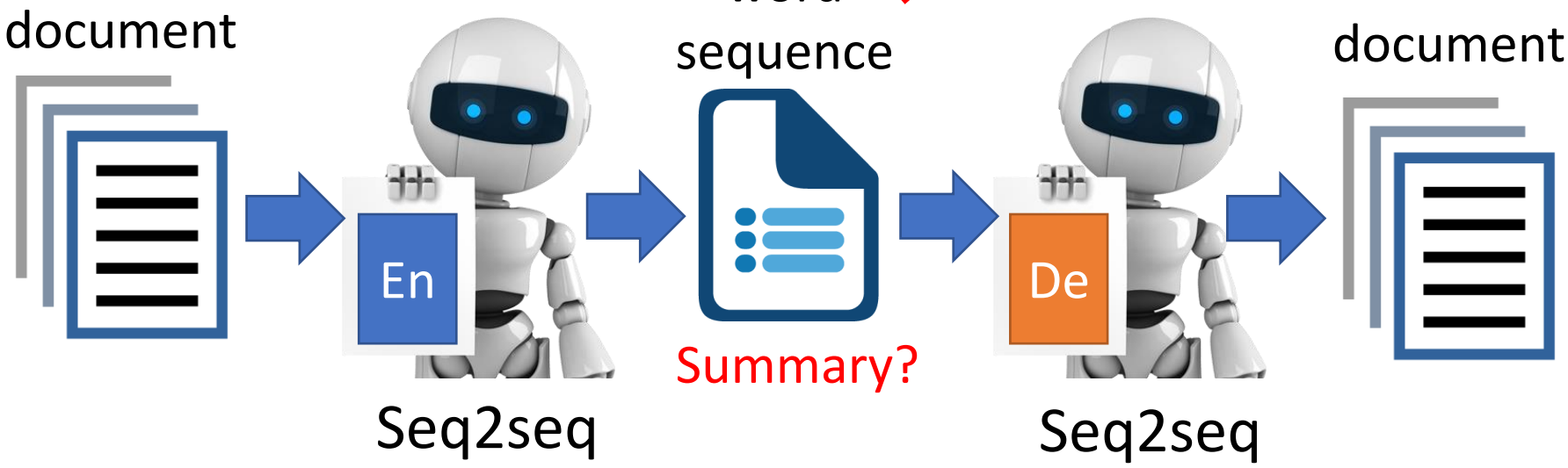
Only need a lot of documents to train the model

seq2seq2seq auto-encoder

Unsupervised Summarization

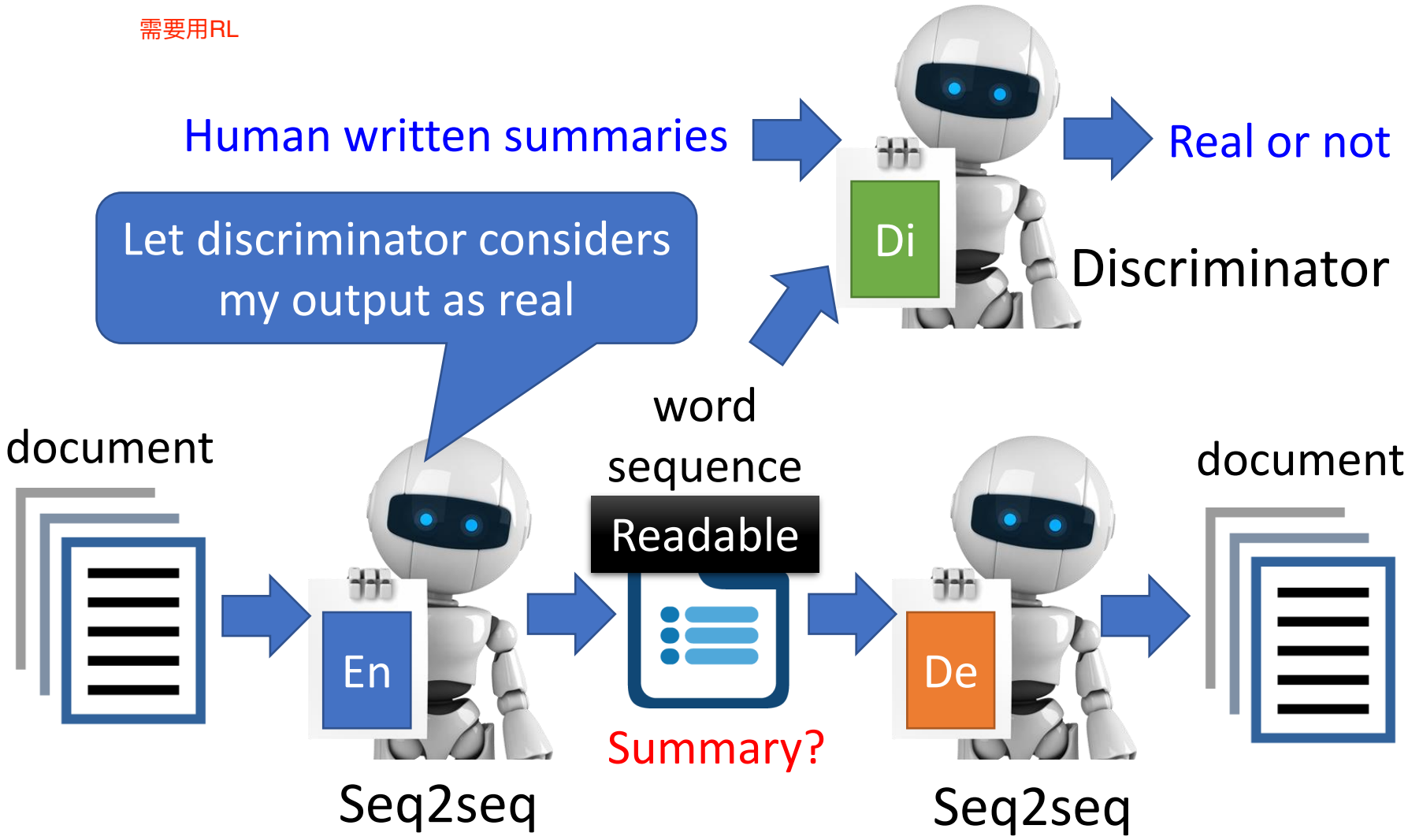
train的起來
但是中間的embedding，人類看不懂

not readable …



This is cycle GAN ☺

Text as Representation



感謝 王耀賢 同學提供實驗結果

Text as Representation

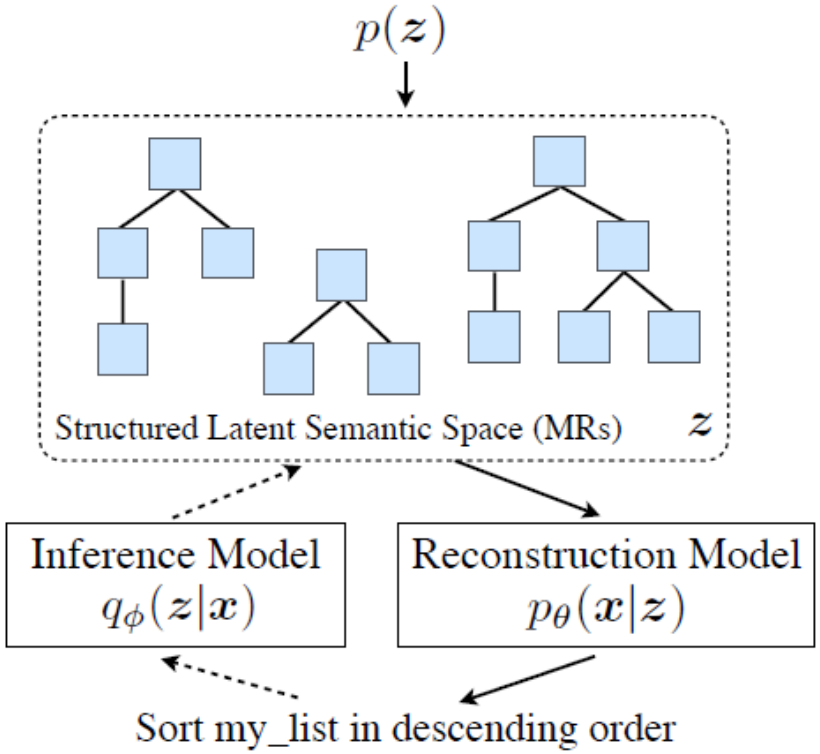
- ***Document***:澳大利亞今天與13個國家簽署了反興奮劑雙邊協議,旨在加強體育競賽之外的藥品檢查並共享研究成果 ……
- ***Summary***:
  - Human:澳大利亞與13國簽署反興奮劑協議
  - Unsupervised:澳大利亞加強體育競賽之外的藥品檢查
- ***Document***:中華民國奧林匹克委員會今天接到一九九二年冬季奧運會邀請函,由於主席張豐緒目前正在中南美洲進行友好訪問,因此尚未決定是否派隊赴賽 ……
- ***Summary***:
  - Human:一九九二年冬季奧運會函邀我參加
  - Unsupervised:奧委會接獲冬季奧運會邀請函

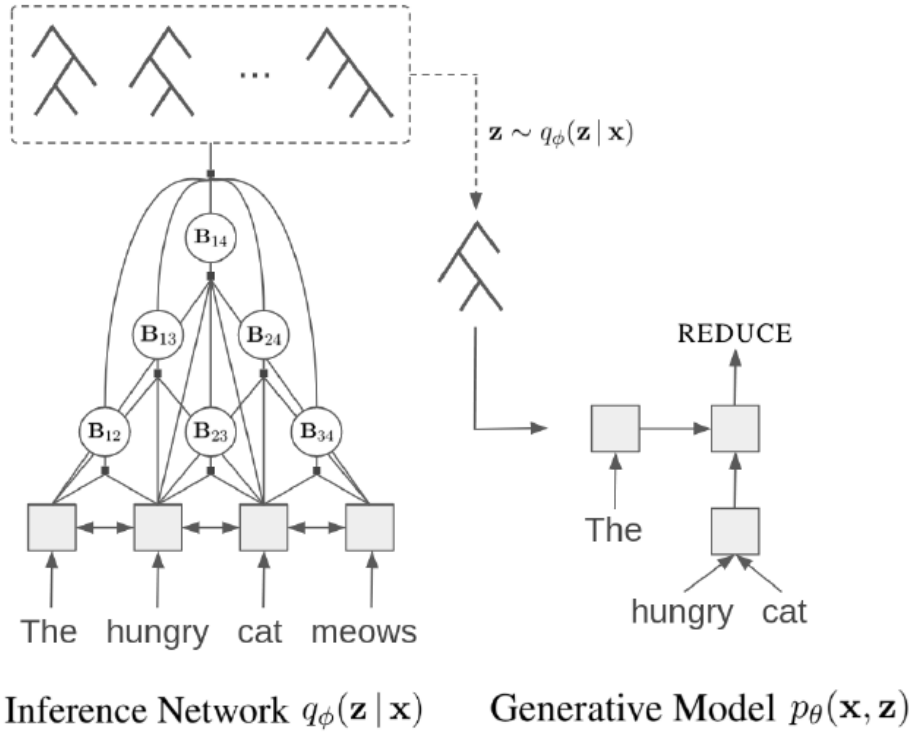
感謝 王耀賢 同學提供實驗結果

Text as Representation

- ***Document***:據此間媒體27日報道,印度尼西亞蘇門答臘島的兩個省近日來連降暴雨,洪水泛濫導致塌方,到26日為止至少已有60人喪生,100多人失蹤 ……
- ***Summary***:
  - Human:印尼水災造成60人死亡
  - Unsupervised:印尼門洪水泛濫導致塌雨
- ***Document***:安徽省合肥市最近為領導幹部下基層做了新規定:一律輕車簡從,不準搞迎來送往、不準搞層層陪同 ……
- ***Summary***:
  - Human:合肥規定領導幹部下基層活動從簡
  - Unsupervised:合肥領導幹部下基層做搞迎來送往規定:一律簡

Tree as Embedding





https://arxiv.org/abs/1806.07832

https://arxiv.org/abs/1904.03746

Outline

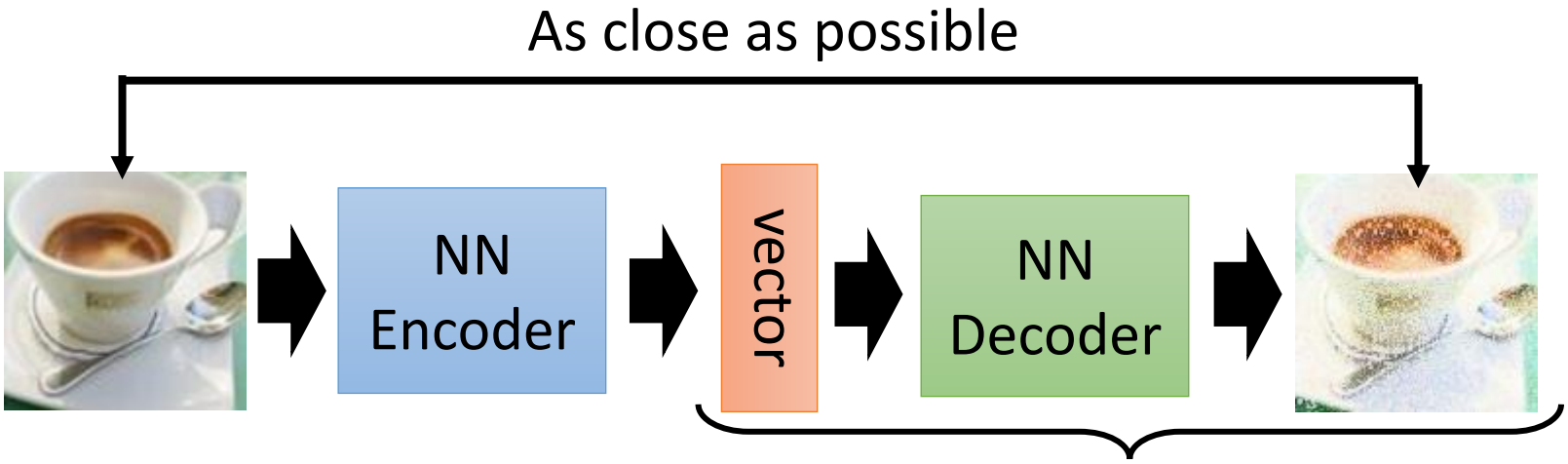
Basic Idea of Auto-encoder

Feature Disentanglement

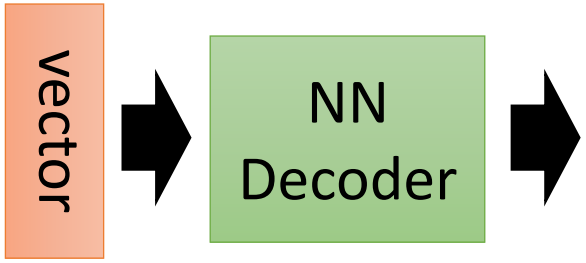
Discrete Latent Representation

More Applications

Generator

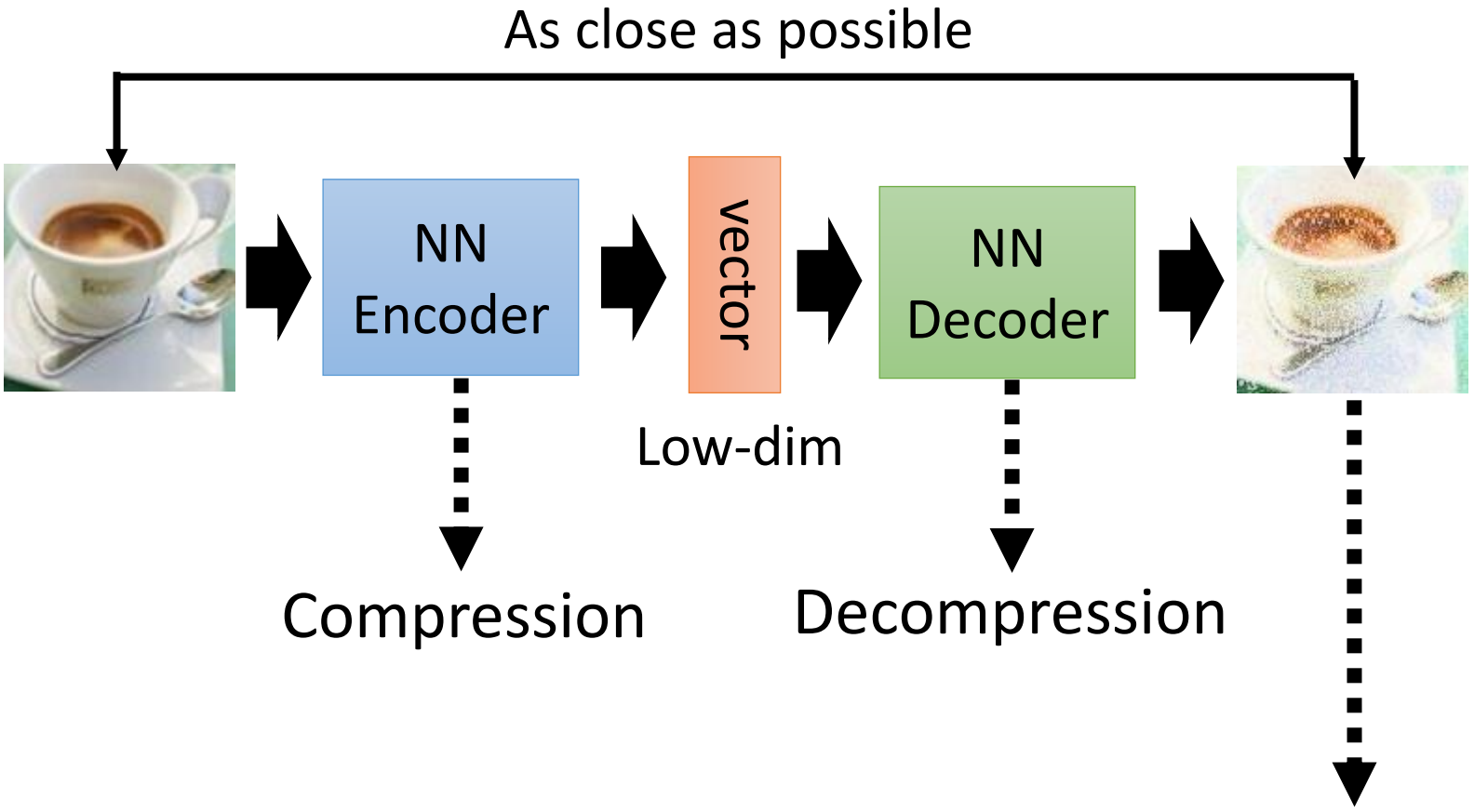


Randomly generate a vector from a distribution



With some modification, we have **variational auto-encoder (VAE).**

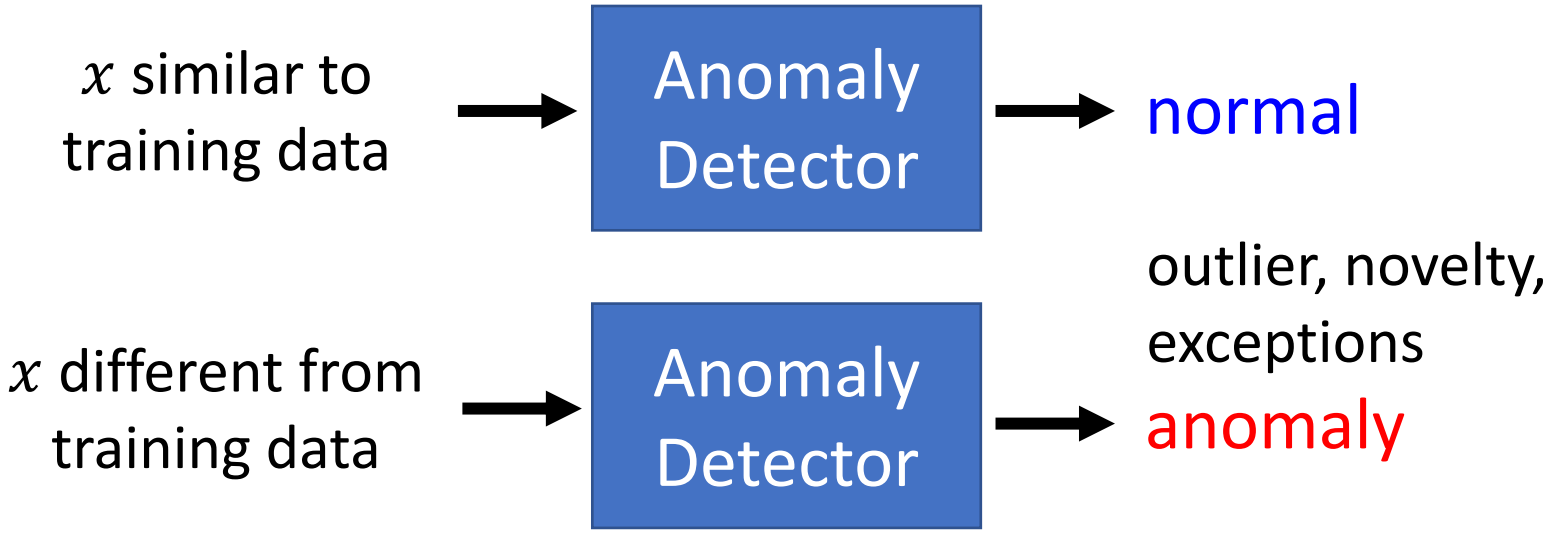
Compression



https://arxiv.org/abs/1708.00838
https://arxiv.org/abs/1703.00395

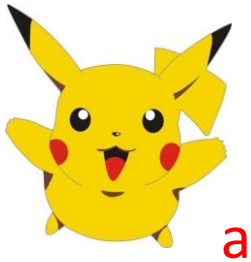
Anomaly Detection 作業

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



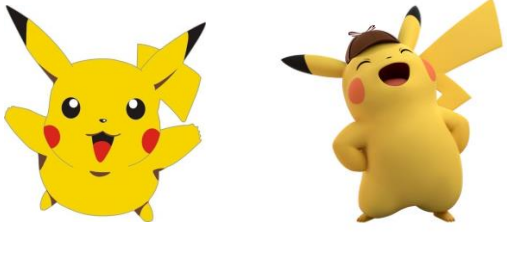
Anomaly Detection

Training Data:



anomaly

Training Data:



anomaly

Training Data:

anomaly

我們不容易蒐集到不正常的資料

Binary Classification?

We only have one class.

Training auto-encoder

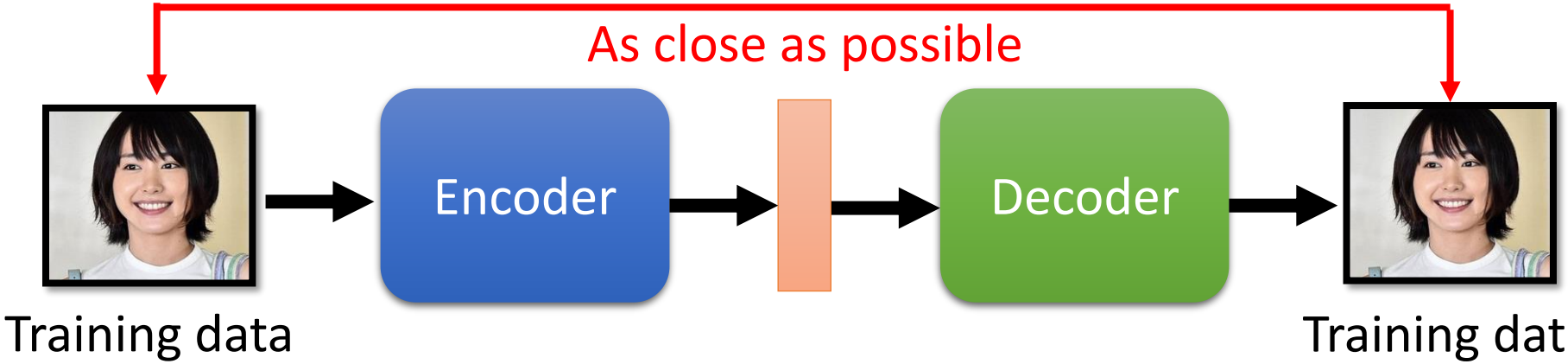
Anomaly Detection

- 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

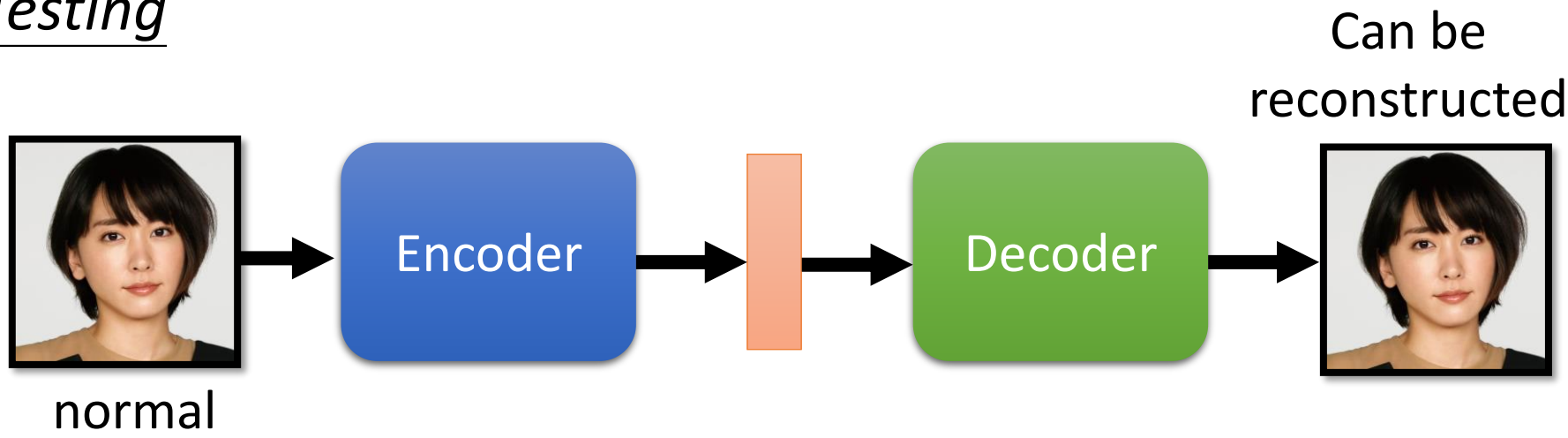
Approach: Auto-encoder

Training

Using **real human faces** to learn an ***autoencoder***



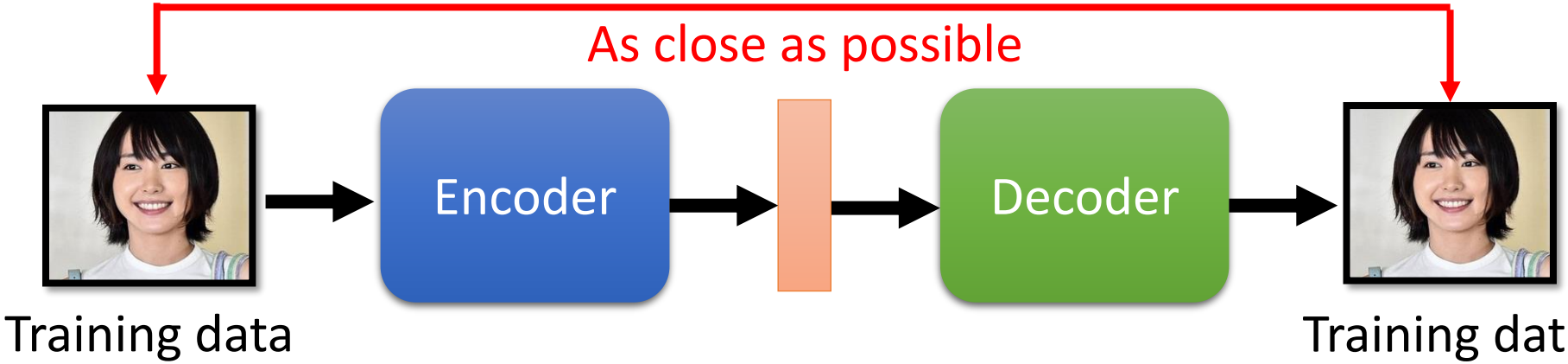
Testing



Approach: Auto-encoder

Training

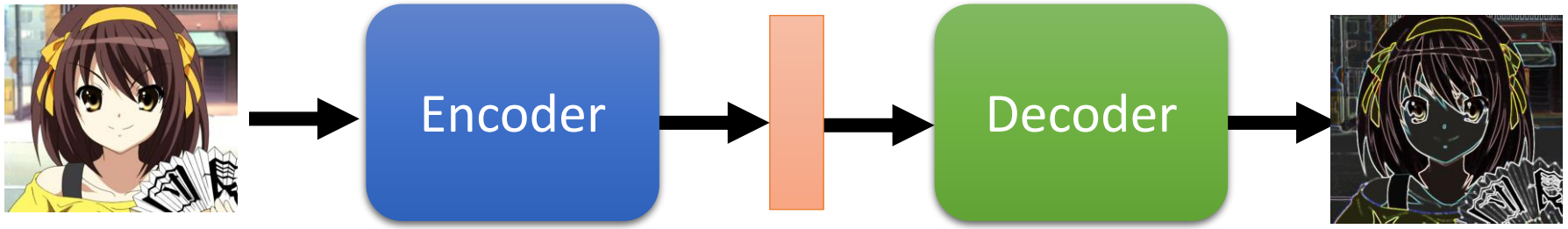
Using **real human faces** to learn an ***autoencoder***



Testing

Large reconstruction loss → anomaly

cannot be reconstructed



More about Anomaly Detection

- Part 1: https://youtu.be/gDp2LXGnVLQ
- Part 2: https://youtu.be/cYrNjLxkoXs
- Part 3: https://youtu.be/ueDlm2FkCnw
- Part 4: https://youtu.be/XwkHOUPbc0Q
- Part 5: https://youtu.be/Fh1xFBktRLQ
- Part 6: https://youtu.be/LmFWzmn2rFY
- Part 7: https://youtu.be/6W8FqUGYyDo

Concluding Remarks

- Basic Idea of Auto-encoder
- Feature Disentanglement
- Discrete Latent Representation
- More Applications