Big Data Research Project

USING VAE AND AE ENCODER AS FEATURE EXTRACTOR FOR FACE VERIFICATION

Supervised by

Alexandre Fournier Montgieux

Students:

ISAKOVA Dilbar

BOUTEBAL Sihem

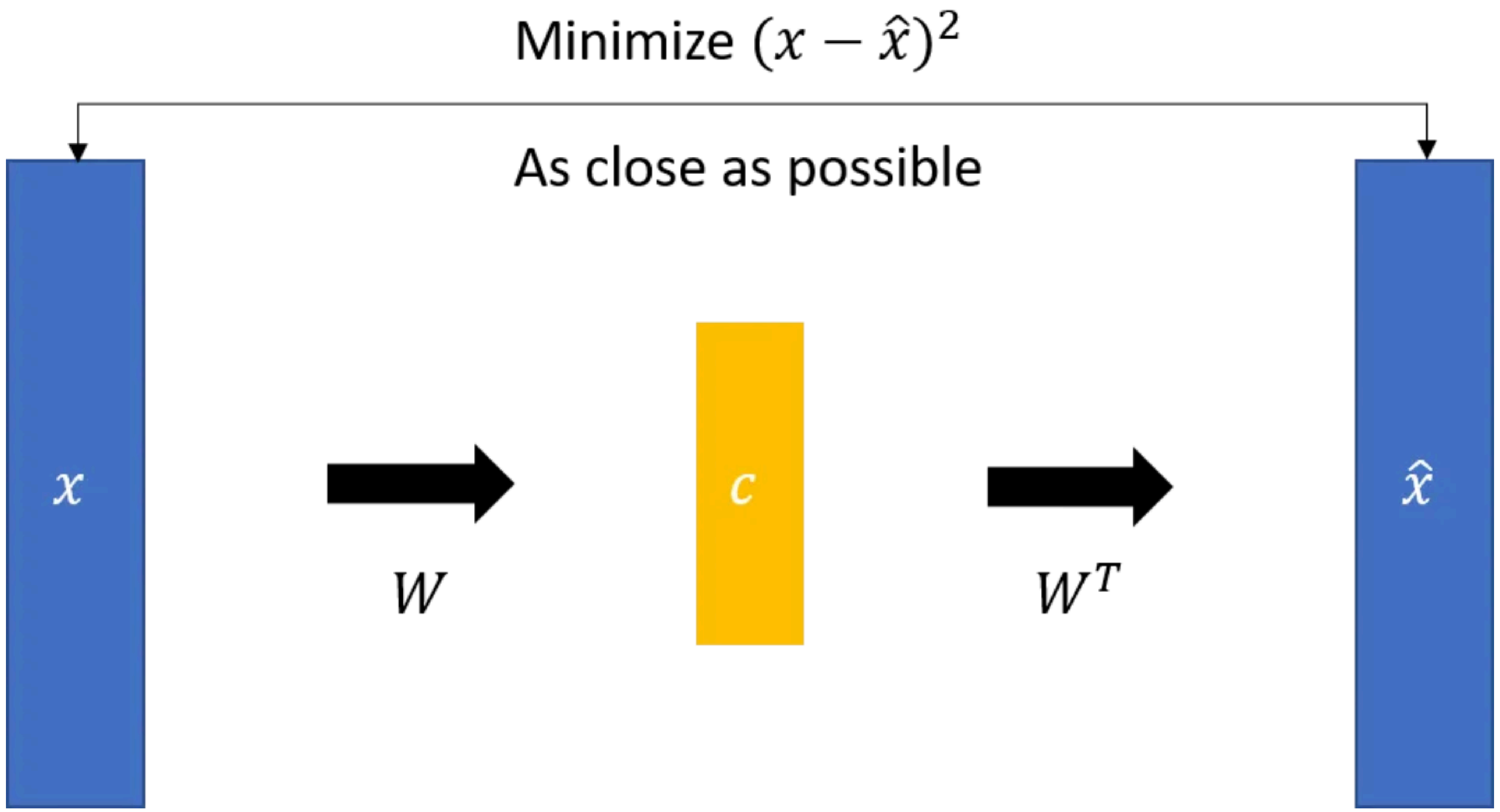
MEDJAHED Melissa

November 6th, 2024

CONTENT

Big Data Research Project

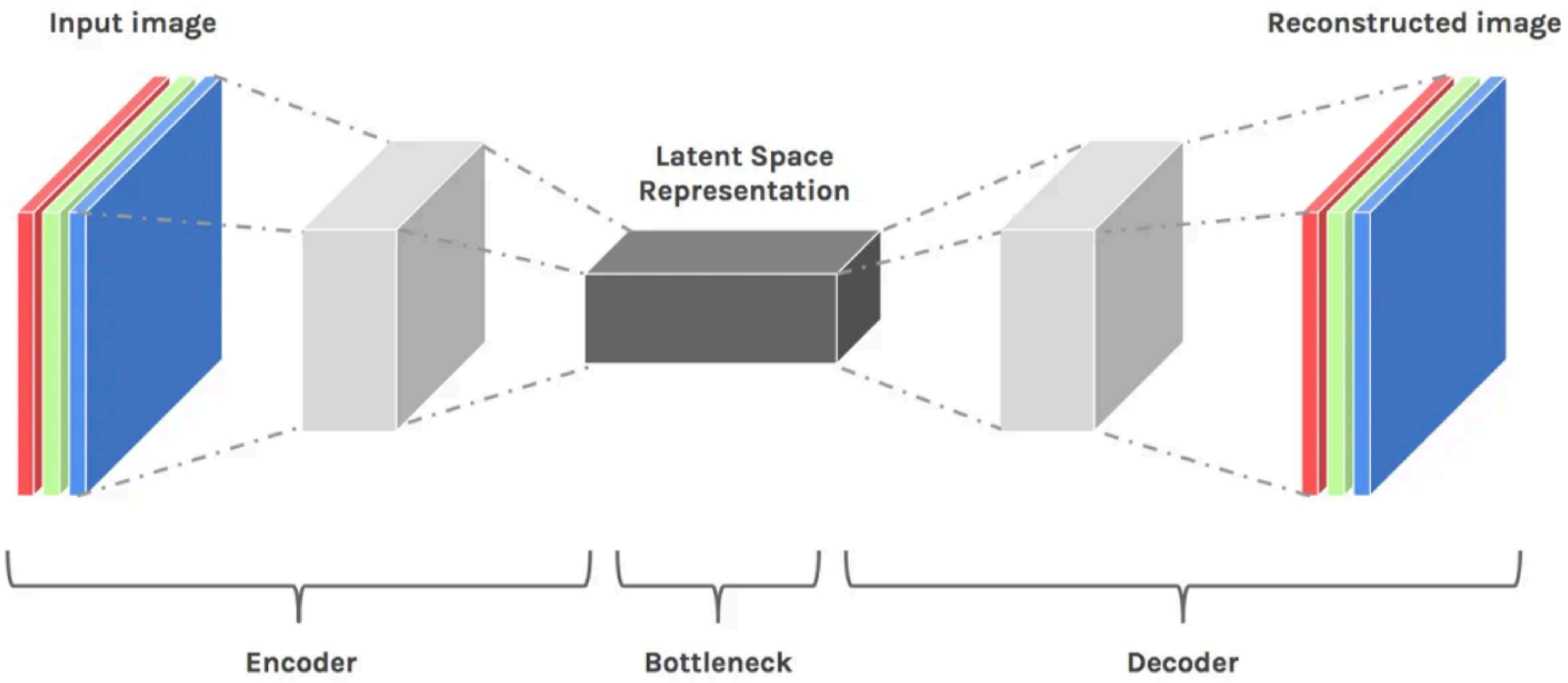
DIMENSIONALITY REDUCTION, PCA AND AE



PCA is to find a W through SVD (singular value decomposition) so that the matrices x and x hat be as consistent as possible.

HOWEVER, THE DIFFERENCE BETWEEN AE AND PCA IS THAT AE USES NEURAL NETWORKS INSTEAD OF SVD.

INTRODUCTION TO VAE



The basic scheme of a variational autoencoder

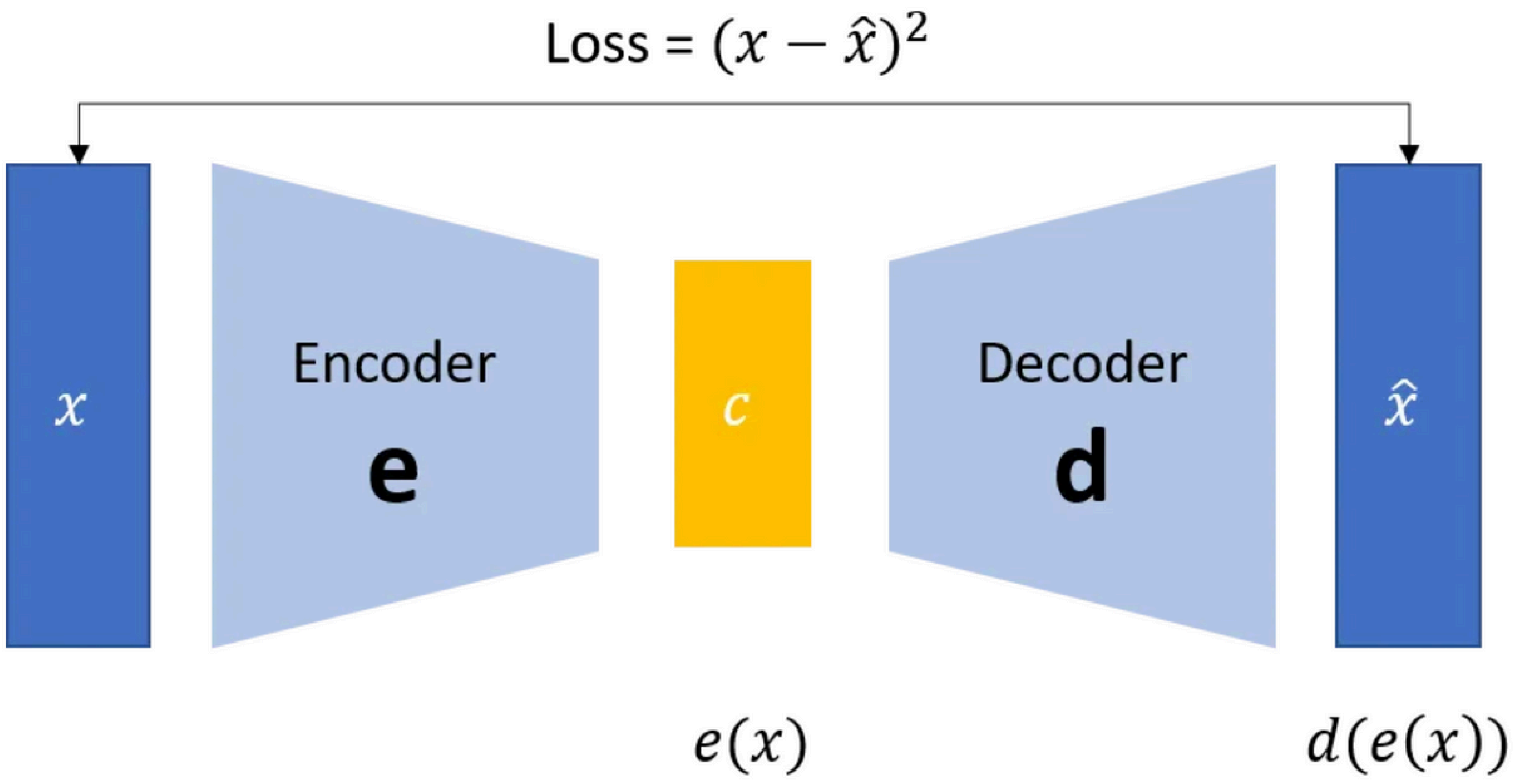
1) The model receives x as input
2) The encoder compresses it into the latent space
3)The decoder receives as input the information sampled from the latent space and produces x′ as similar as possible to x

CentraleSupélec

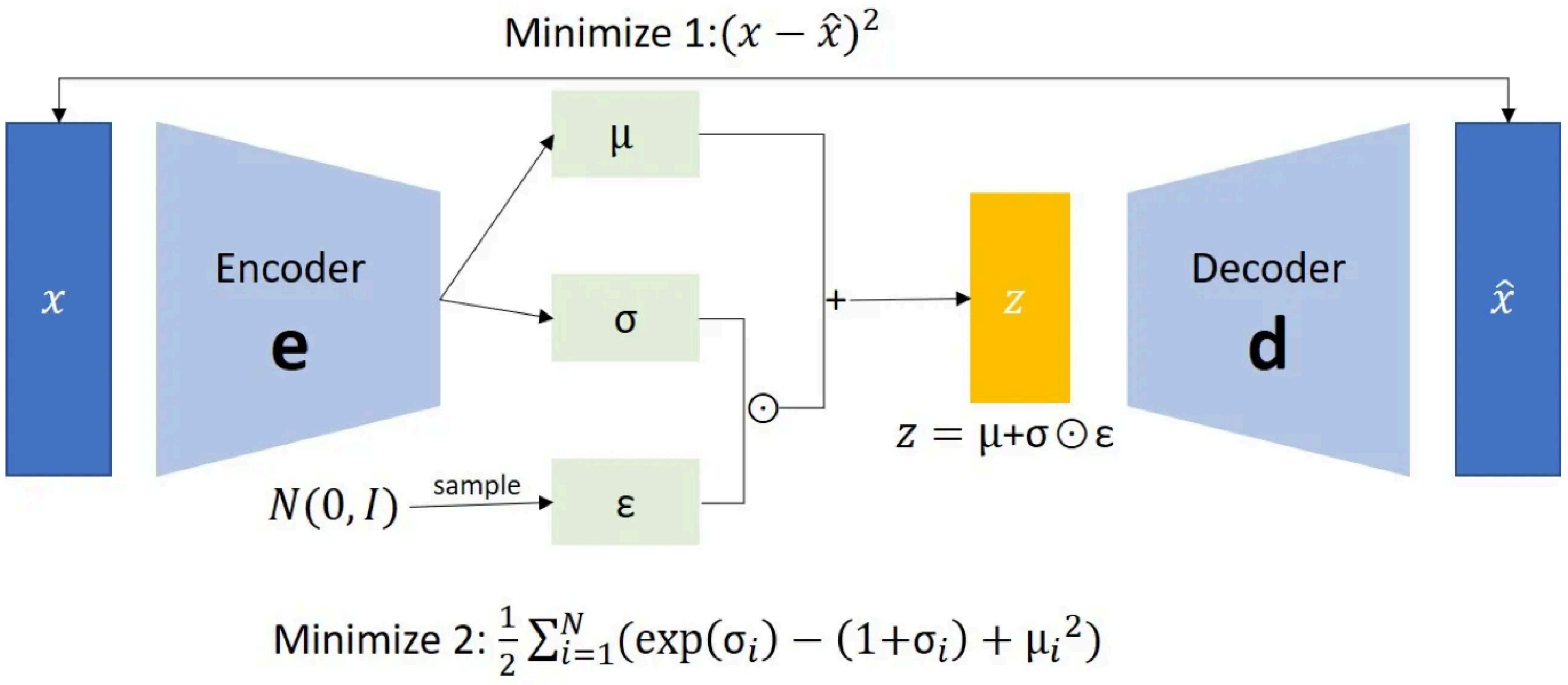
WHAT IS AUTOENCODER?

1. Compress **(encode)** input data to its essential features.

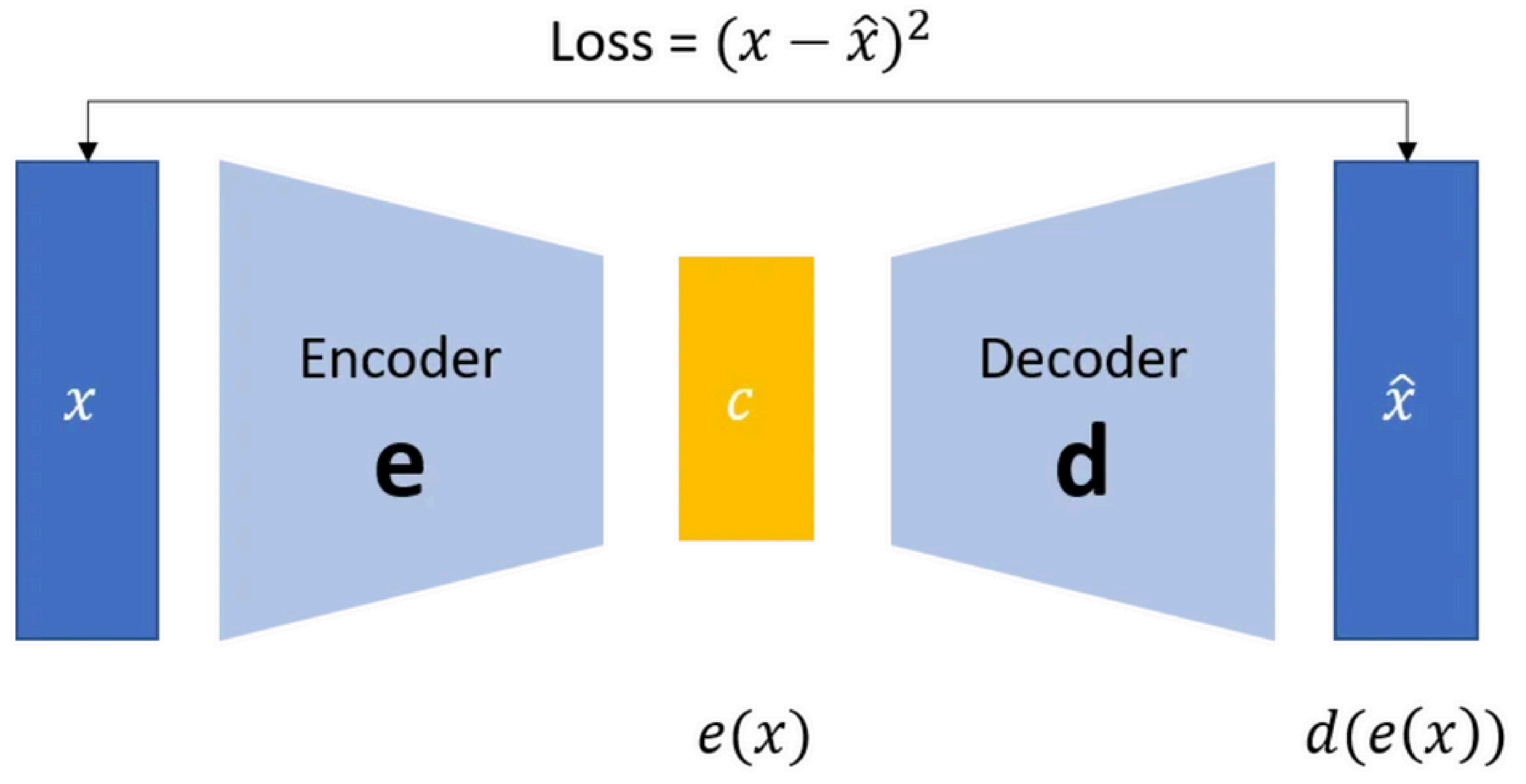
2. Reconstruct **(decode)** the original input from this compressed representation.



WHAT IS THE VARIATIONAL AUTOENCODER?

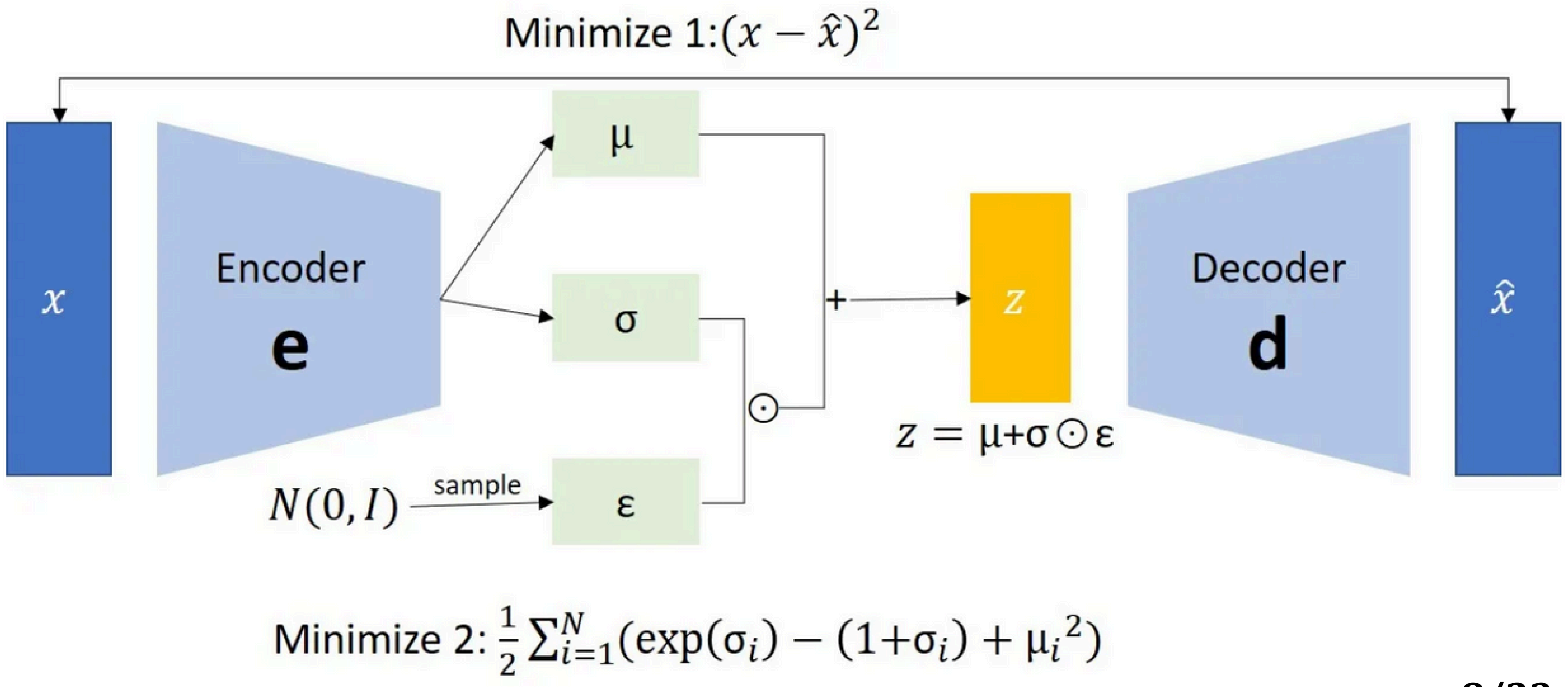


WHAT IS THE DIFFERENCE?



1. Most autoencoders learn discrete latent space models, VAEs learn continuous latent variable models.

2. VAEs model two different vectors: a vector of means, “μ,” and a vector of standard deviations, “σ.”



CentraleSupélec

LITERATURE REVIEW – TECHNIQUES FOR FR & FV

1. Holistic methods :

Eigenfaces (PCA)

- Reduce face image data to eigenfaces
- Focus on variations in facial structure

Fisherfaces (LDA)

- Focus on features that maximize class separability
- More robust to lighting variations

LITERATURE REVIEW – TECHNIQUES FOR FR & FV

2. Feature-based methods :

Hidden Markov Model

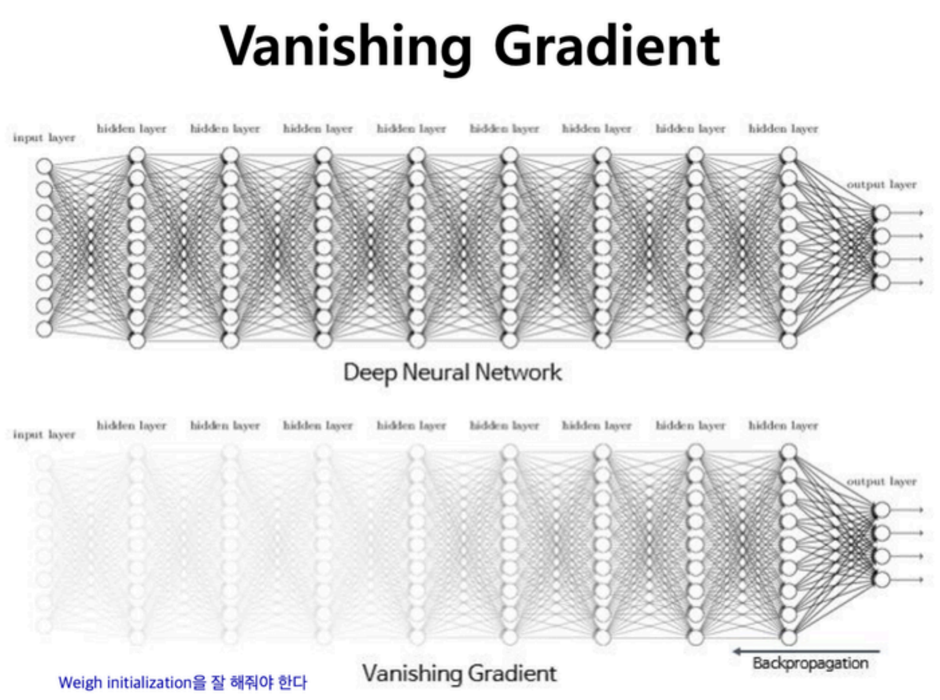
- Dividing the face into several regions
- Treating regions as a sequence of observations

CNN

- Deep learning models that learn hierarchical features from images
- Each layer detects increasingly complex patters

RESNETS

Problem with CNNs



Resnets

Results

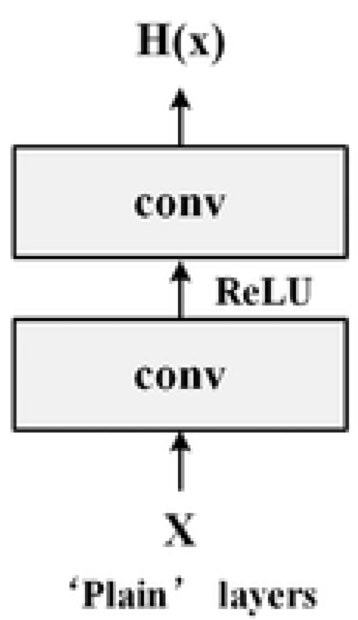
The 152-layer ResNet achieved a top-5 error rate of 3.57% on the ImageNet Dataset.

Contributed to multiple 1st-place wins in the COCO 2015 competition in tasks like object detection, segmentation.

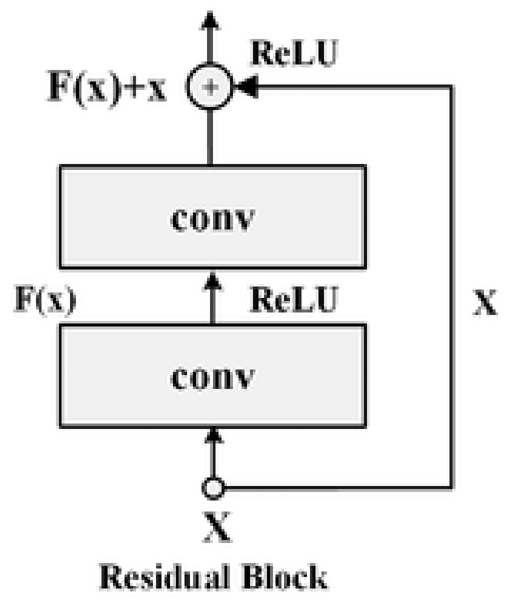
On the CIFAR-10 dataset, ResNets showed their ability to train networks with over 1,000 layers effectively.

RESNETS : OVERVIEW

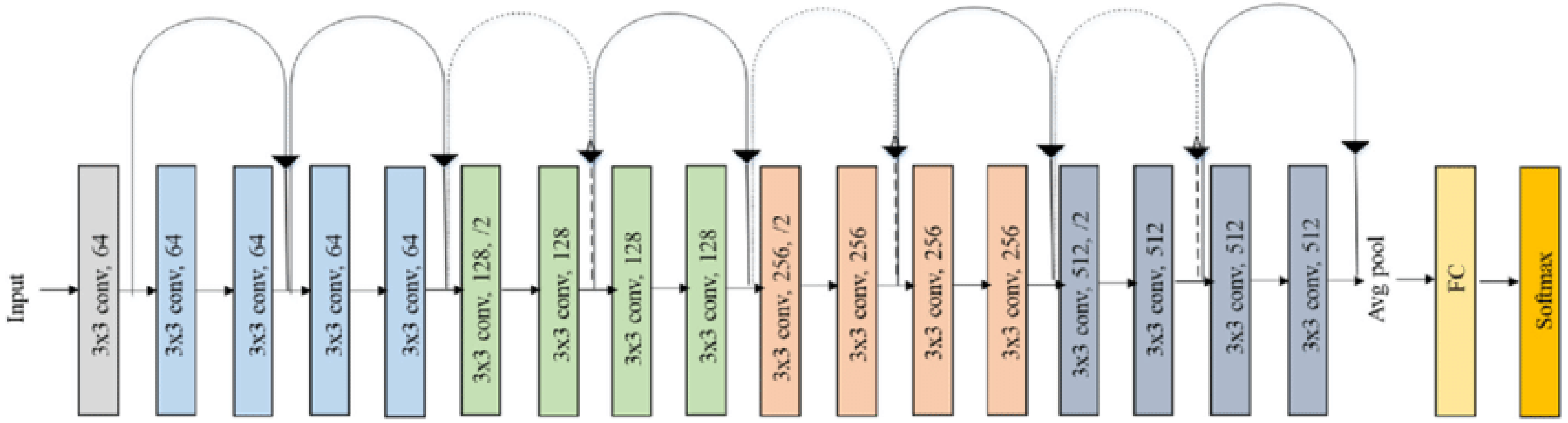
Plain Layer



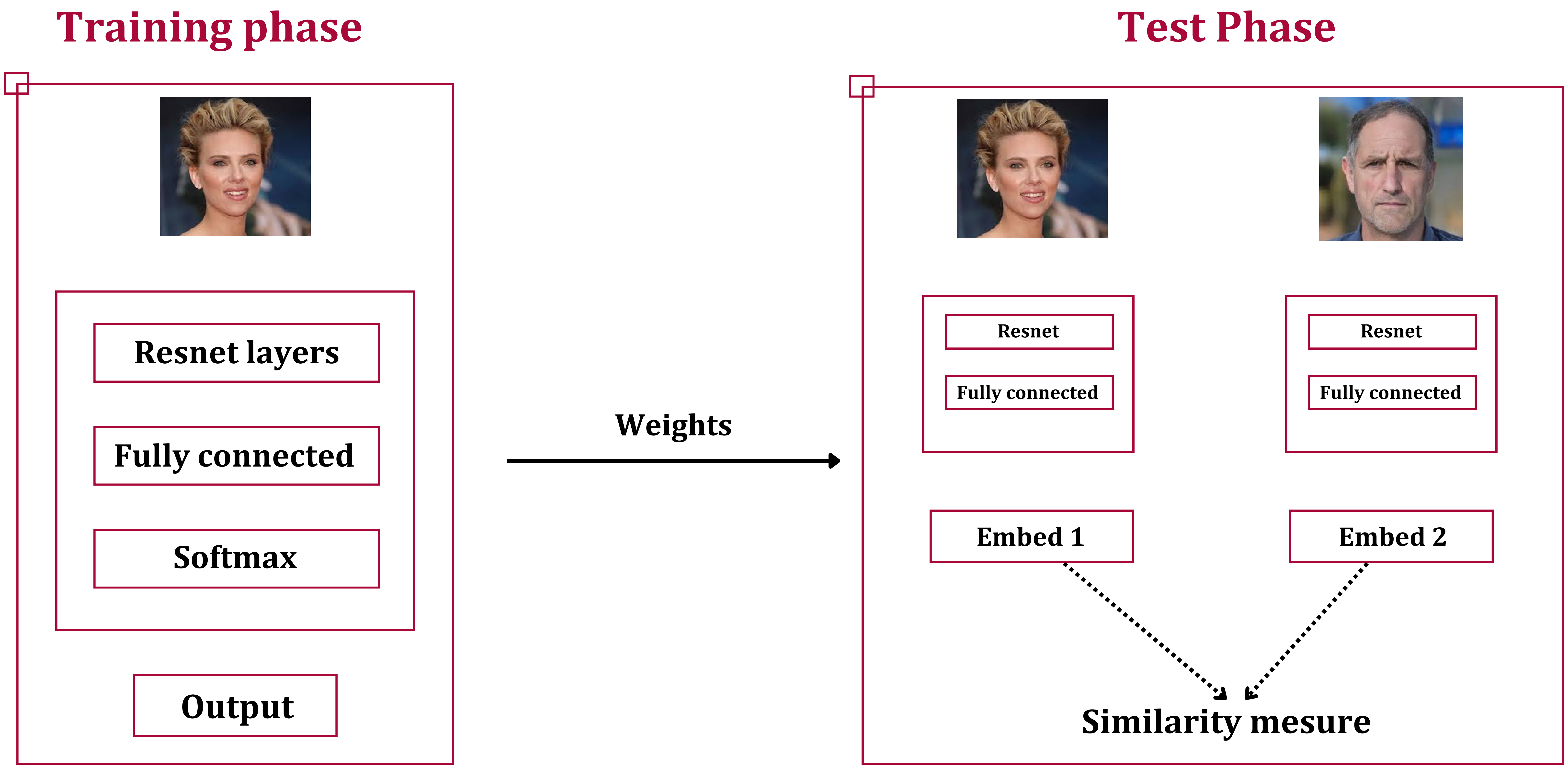
Resnet block



Exemple of Resnet18



RESNETS : FACE VERIFICATION TASK



PROJECT ANNOUNCEMENT

Objectives

Perform Face verification task using Variational Auto-encoders and compare the results with the baseline

Dataset

Casia Dataset : Introduced by Yi et al. in Learning Face Representation from Scratch. it contains 494,414 face images of 10,575 real identities collected from the web.

Loss

Arcface Loss

LOSS FUNCTION

ArcFace loss:

- Introduces an angular margin to the softmax loss
- Improves separability between classes

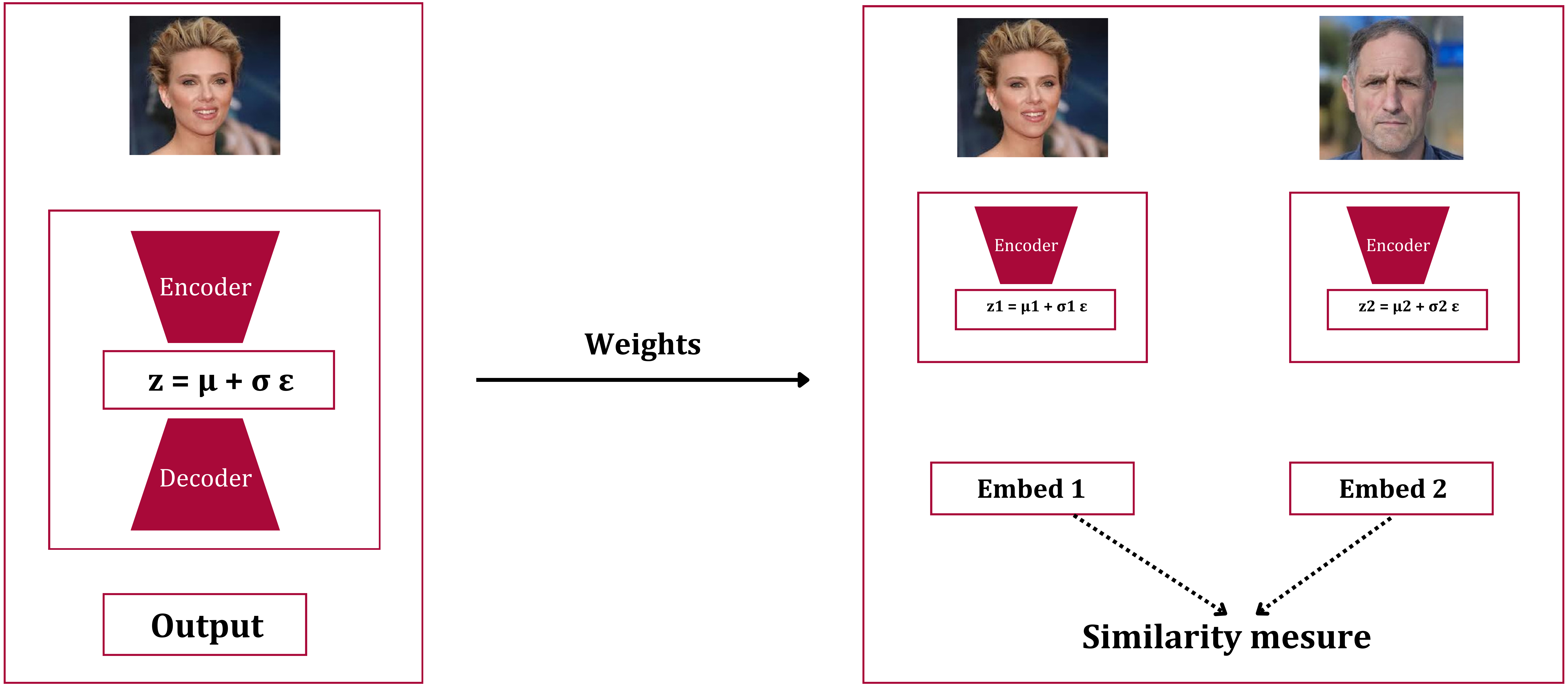
$$L = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos(\theta_{y_i}+m))}}{e^{s(\cos(\theta_{y_i}+m))}+\sum_{j=1,j\neq y_i}^{C}e^{s\cos\theta_j}}$$

- Replace the euclidean distance with an angular distance between classes
- Used to increase the interclass margin of embeddings

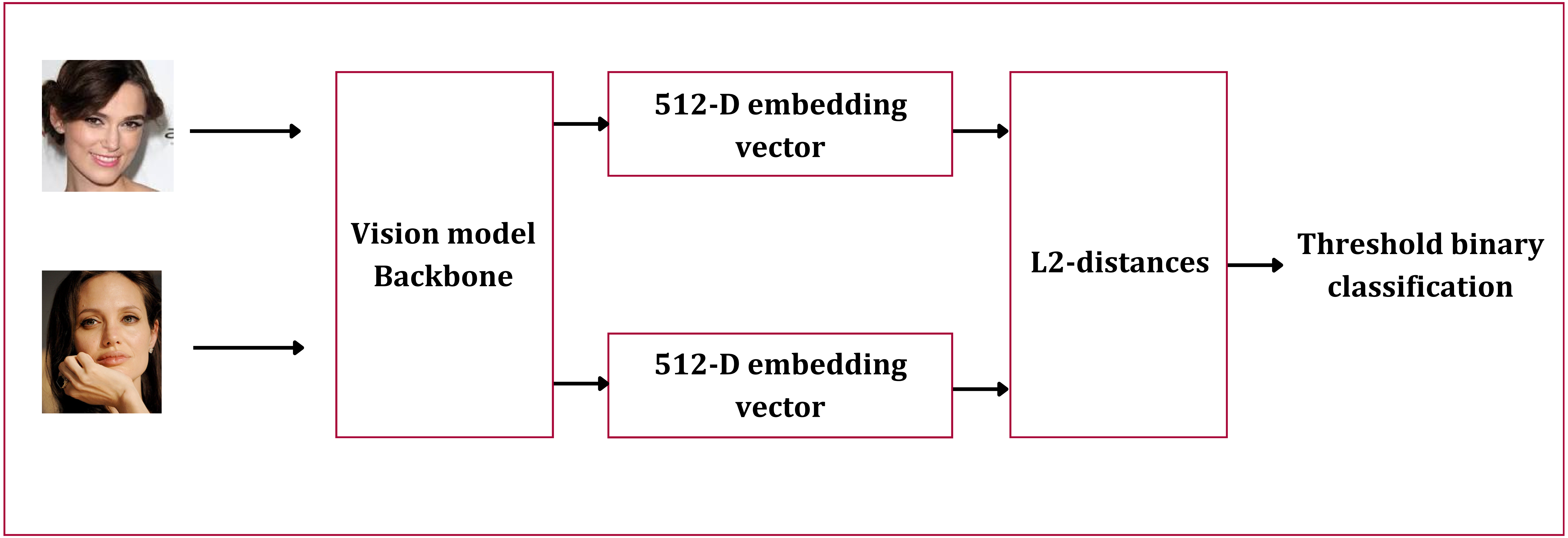
PROJECT ANNOUNCEMENT

Training phase

Test Phase

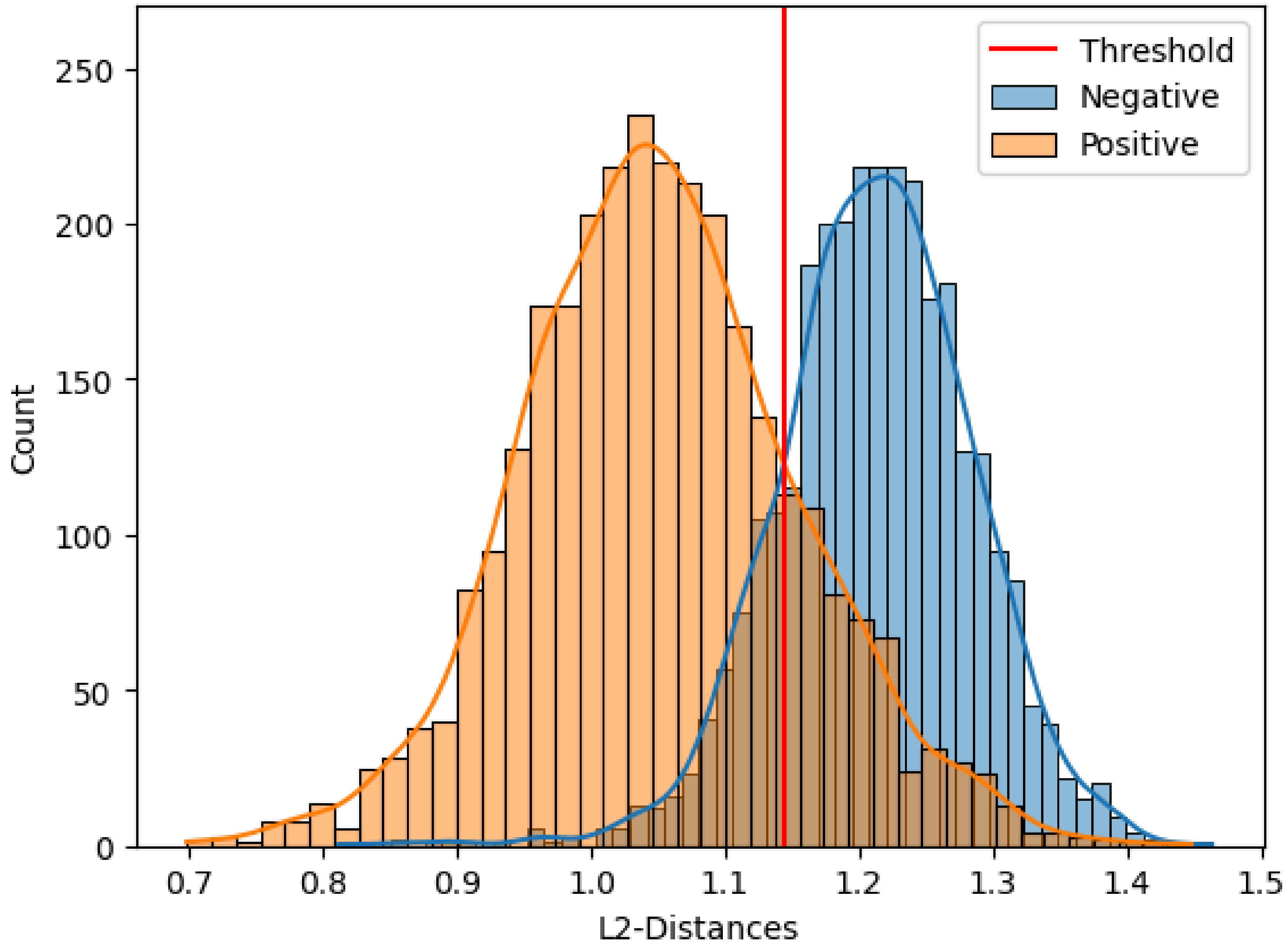


EVALUATION PIPELINE



EVALUATION PIPELINE

Using LFW-c dataset:

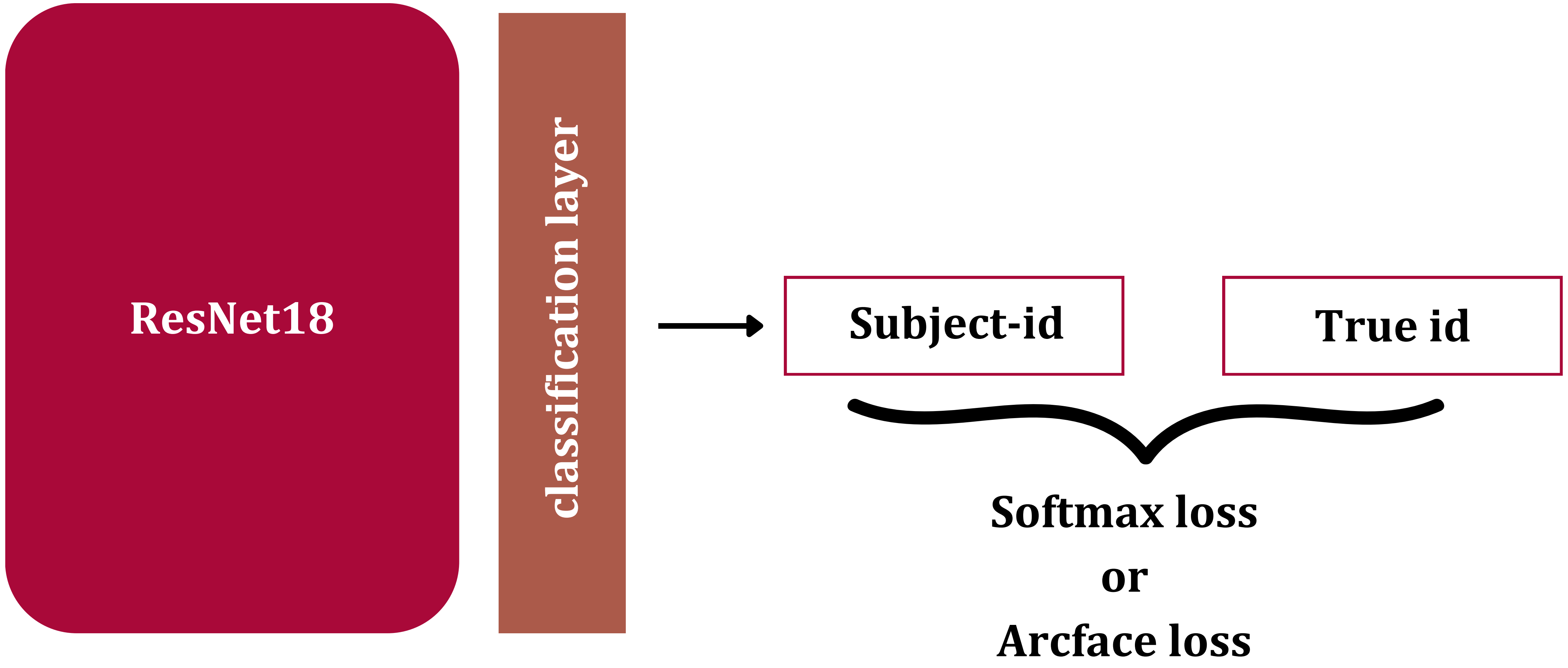


- 10 fold cross validation
- mean threshold : 1.145
- mean accuracy : 0.85

BASELINE METHODS

ResNet18 training (Arcface vs. Softmax loss):

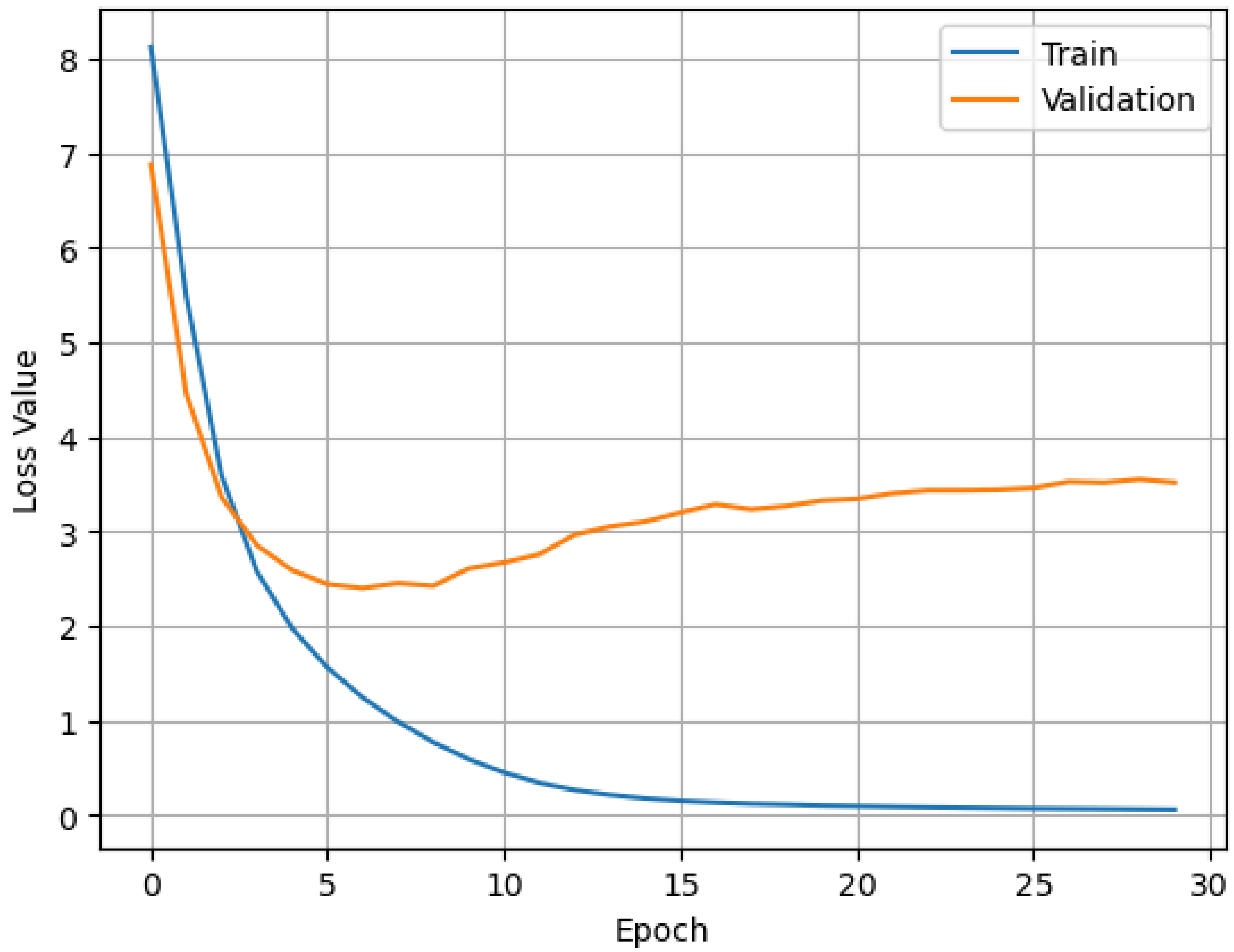
Casia Face Dataset

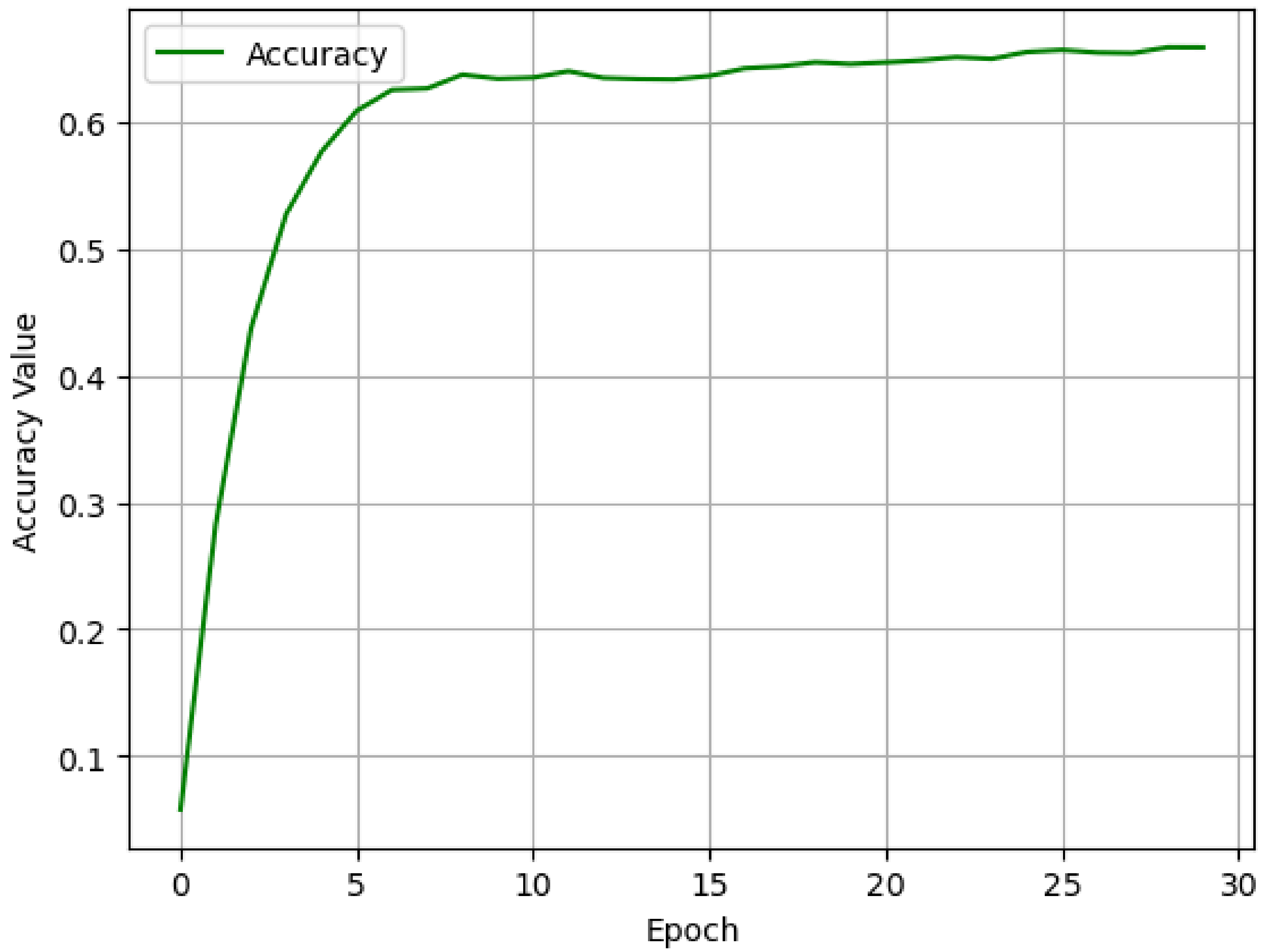


CentraleSupélec

BASELINE METHODS

Training results with softmax loss

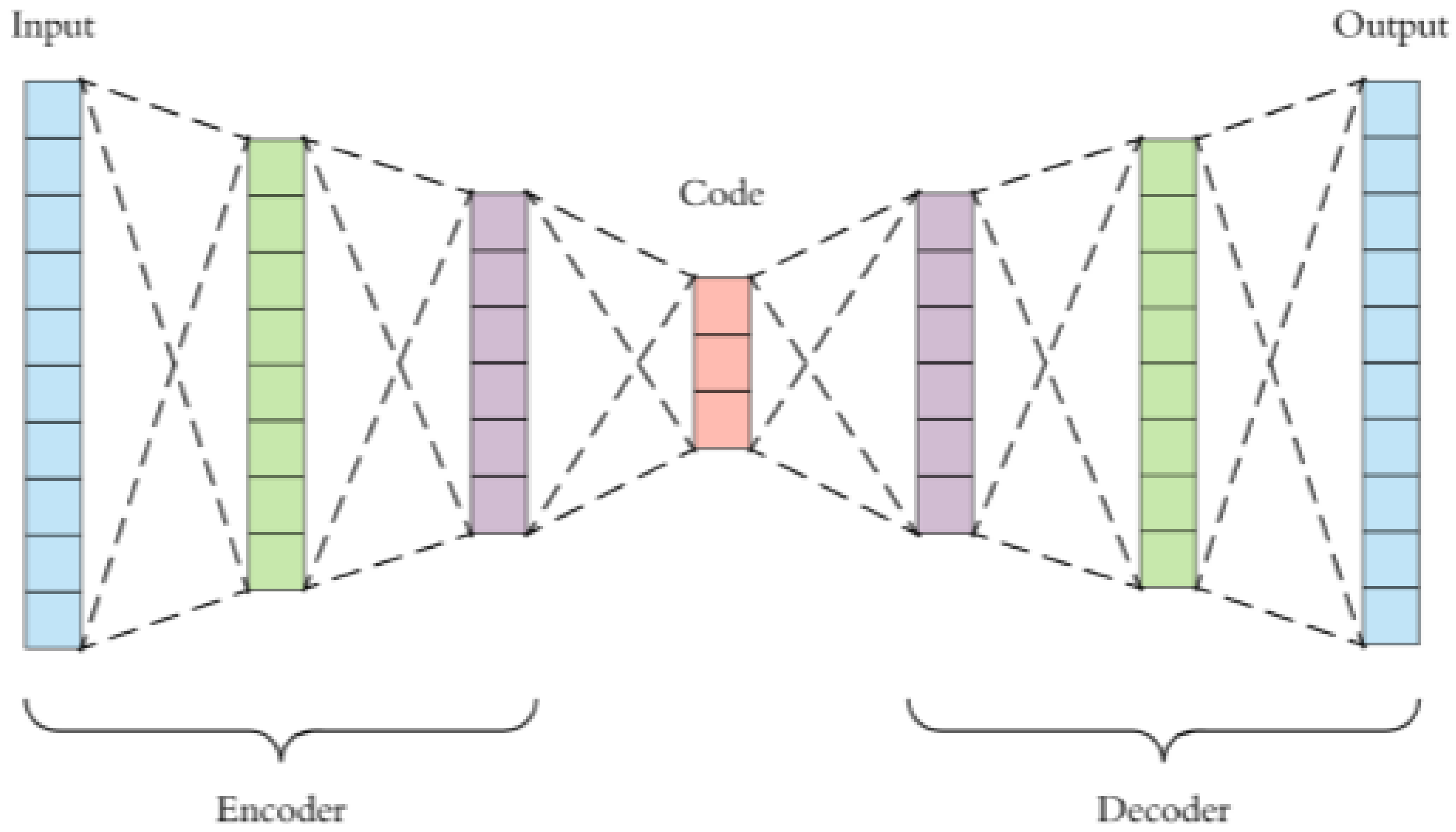




- Batch size : 64
- Optimizer : Adam
- Learning rate: 1e-4
- Final accuracy : 0.65

CentraleSupélec

BUILDING VAE FROM SCRATCH



- Fully convolutional architecture
- Minimize reconstruction loss
- Encoder: feature extractor
- Goal: learning compact representations of faces

Suggestion:

The decoder could be used for data augmentation during training.

COMPLETED STEPS:

- Litterature review ✓
- Evaluation pipeline ✓
- Training Resnet using softmax loss ✓
- Training Resnet using arcface loss

BIBLIOGRAPHIE

[1] *Enoch Solomon et al.* **Unsupervised Deep Learning Image Verification, 2023**
[2] *Irfan Dwiki Bhaswara.* **Exploration of autoencoder as feature extractor for face recognition system**
[3] *Alhadi B. Pinkie et al.* **Deep learning in image classification using residual network (resnet) variants for detection of colorectal cancer**
[4] *Raja G. et al.* **Autoencoders**
[5] *Vincent Dumoulin and Francesco Visin.* **A guide to convolutional autoencoders, 2018**
[6] *Eyael S. et al.* **Autoencoder based face verification system**
[7] *ARA V. et al.* **Face detetction and recognition using hidden markov models**
[8] *Jiankang Deng et al.* **Arcface: Additive angular margin loss for deep face recognition**
[9] *W. Zhao et al.* **Face recognition: A literature survey**
[10] *Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun.* **Deep residual learning for image recognition. Microsoft Research**
[11] *Sepp Hochreiter.* **The vanishing gradient problem during learning recurrent neural nets and problem solutions**
[12] *Dubravka Jevtić and Marijeta Slavković.* **Face recognition using eigenface approach**
[13] *Diederik P. Kingma and Max Welling.* **Auto-encoding variational bayes, 2014**
[14] *Erik Learned-Miller et al.* **Labeled Faces in the Wild: A Survey. Advances in Face Detection and Facial Image Analysis, 2016**
[15] *Chuan Li and Michael Wand.* **Precomputed real-time texture synthesis with markovian generative adversarial networks, 2016**
[16] *M. Liu et al.* **Deep learning for face recognition: A critical review, 2020**
[17] *Joao P. Hespanha Peter N. Belhumeur and David J. Kriegman.* **Eigenfaces vs. fisherfaces: Recognition using class specific linear projection, 2019**
[18] *Richard K. G. Kaori T. Rikiya Y., Mizuho N.* **Convolutional neural networks: an overview and application in radiology**
[19] *Kevin L. Sasha T., Diogo A.* **Resnet : Generalizing residual architectures**
[20] *Yi Sun, Xiaogang Wang, and Xiaoou Tang.* **Hybrid deep learning for face verification**