武漢大學 WUHAN UNIVERSITY

Non-Parallel Voice Conversion Using CycleGAN-VC3, Random CNN and Baidu API

group number：NO.7

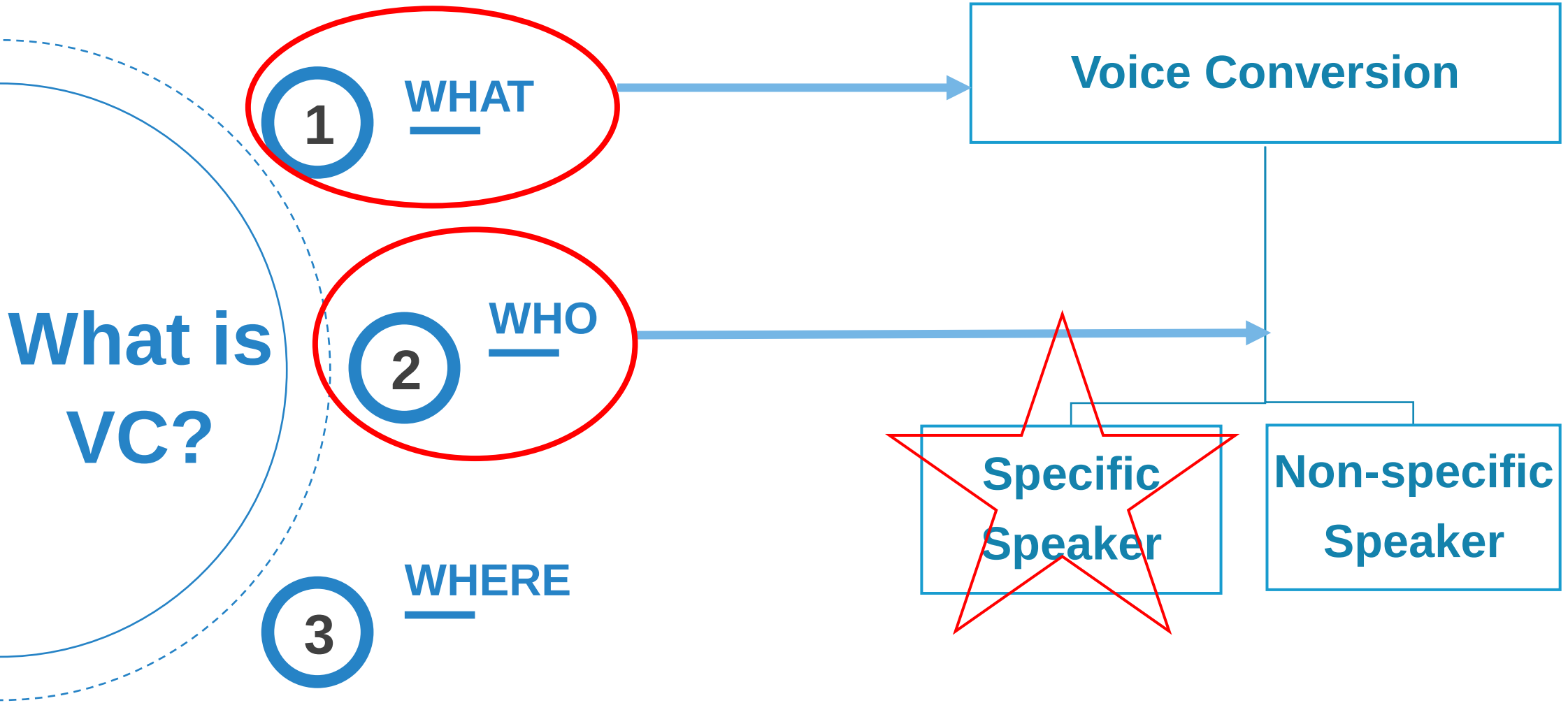
date：2021.06.04

自强、弘毅、求是、拓新

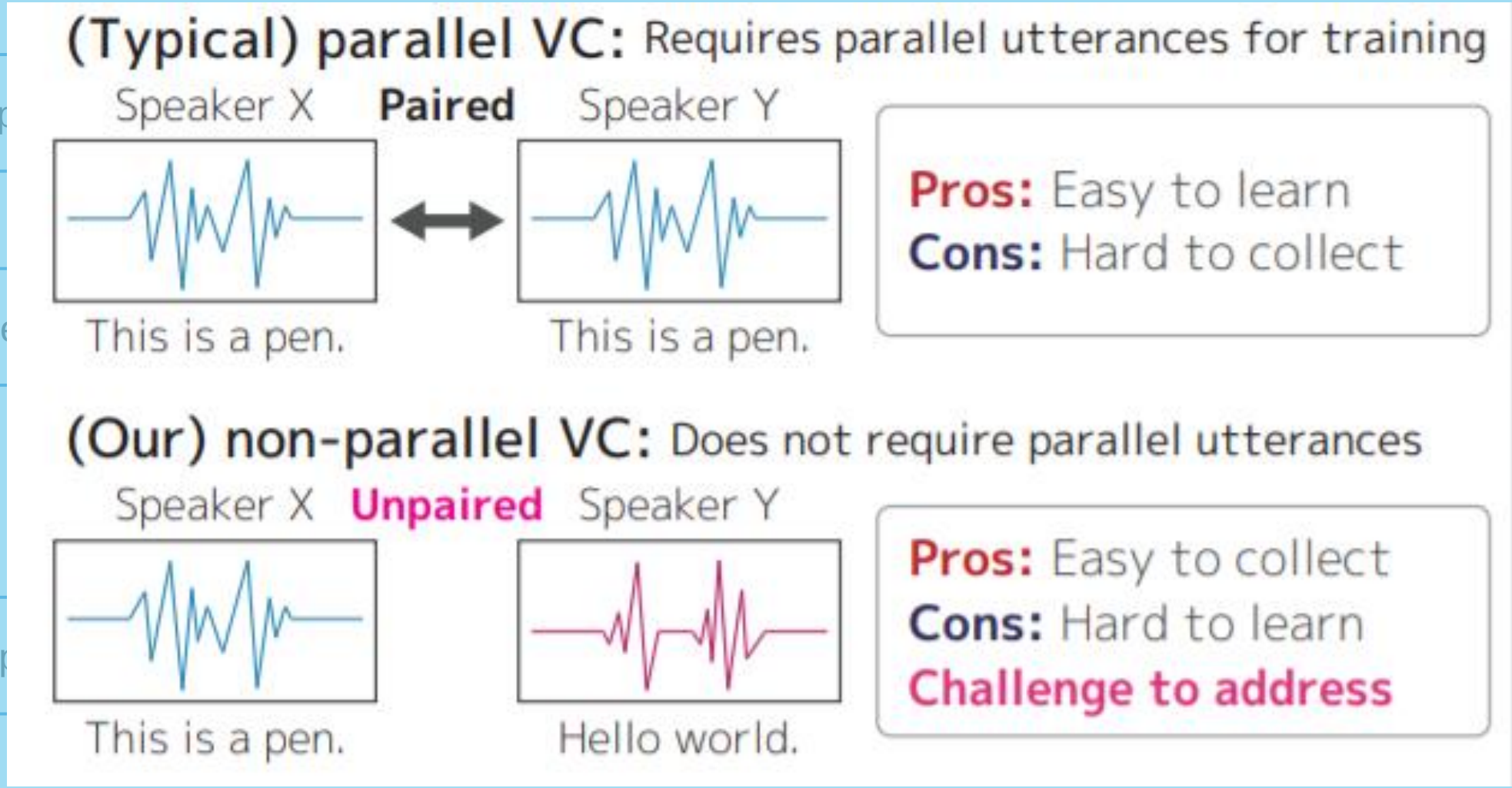
01

Introduction to Voice Conversion

1.1 What Is VC?


What is VC?
1
WHAT
2
WHO
3
WHERE
Voice Conversion
Specific Speaker
Non-specific Speaker

1.2 Basic Framework



1.3 Categories: Parallel VC

Statistical Methods

- DNN
- GMM

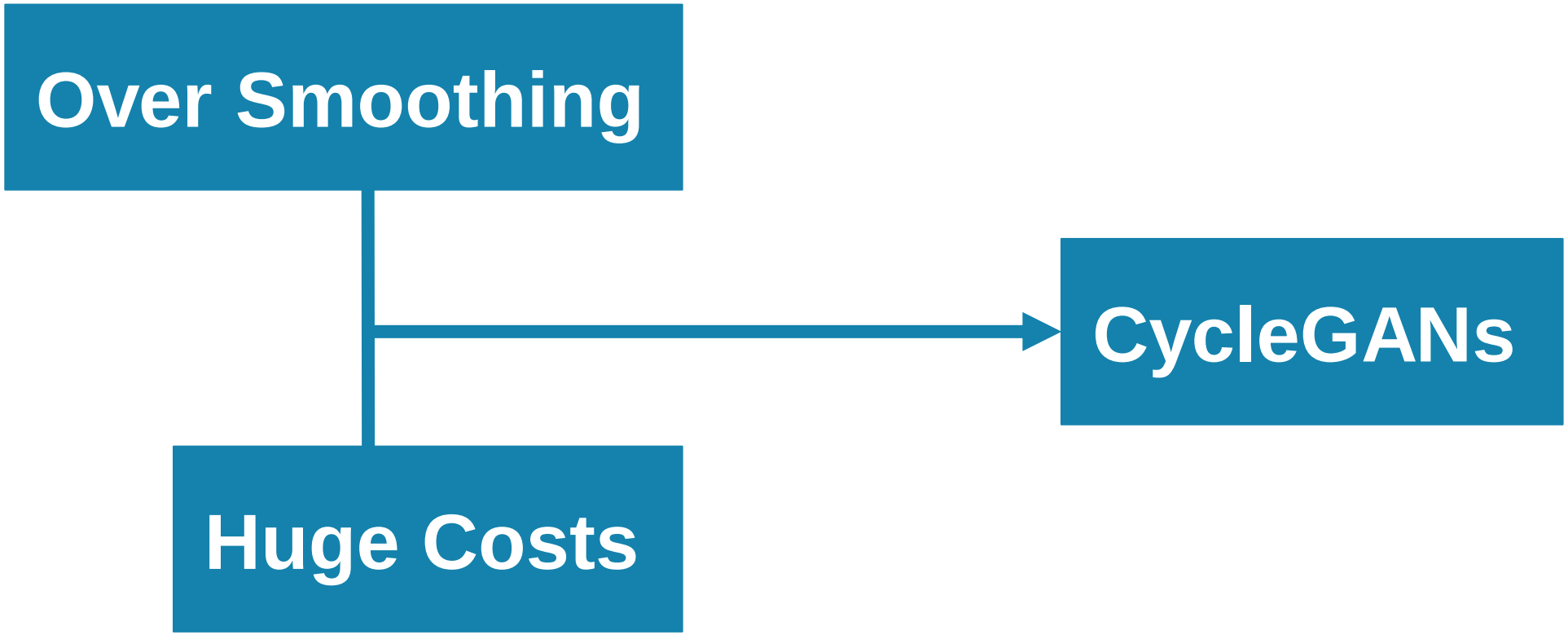
Neural Network based Method

- RBM
- CNN
- GANs

Exemplar-based Methods

- NMF

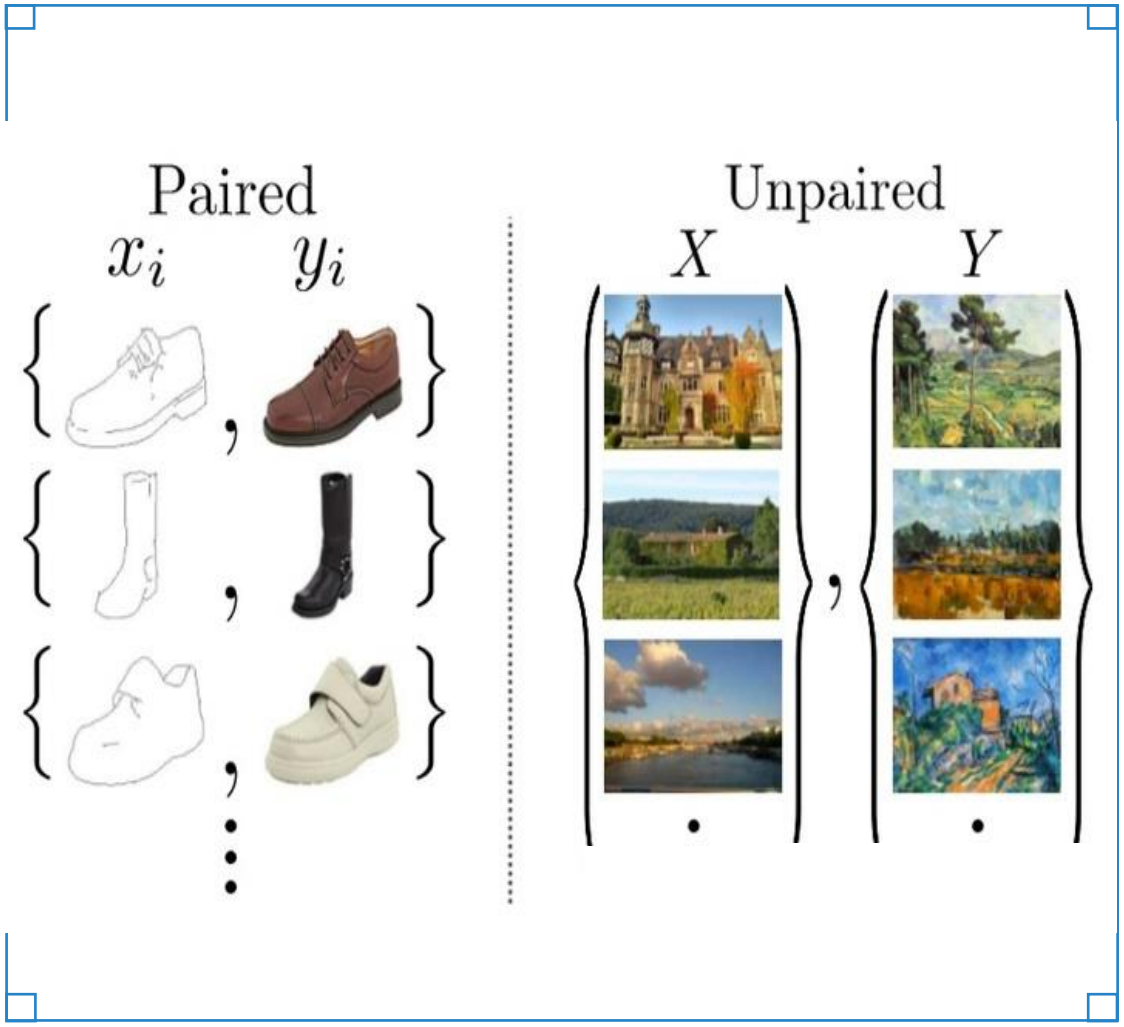
1.3 Categories: Non-parallel VC



02

Principles of CYCLEGAN-VCs

2.1 Introduction to CycleGANs



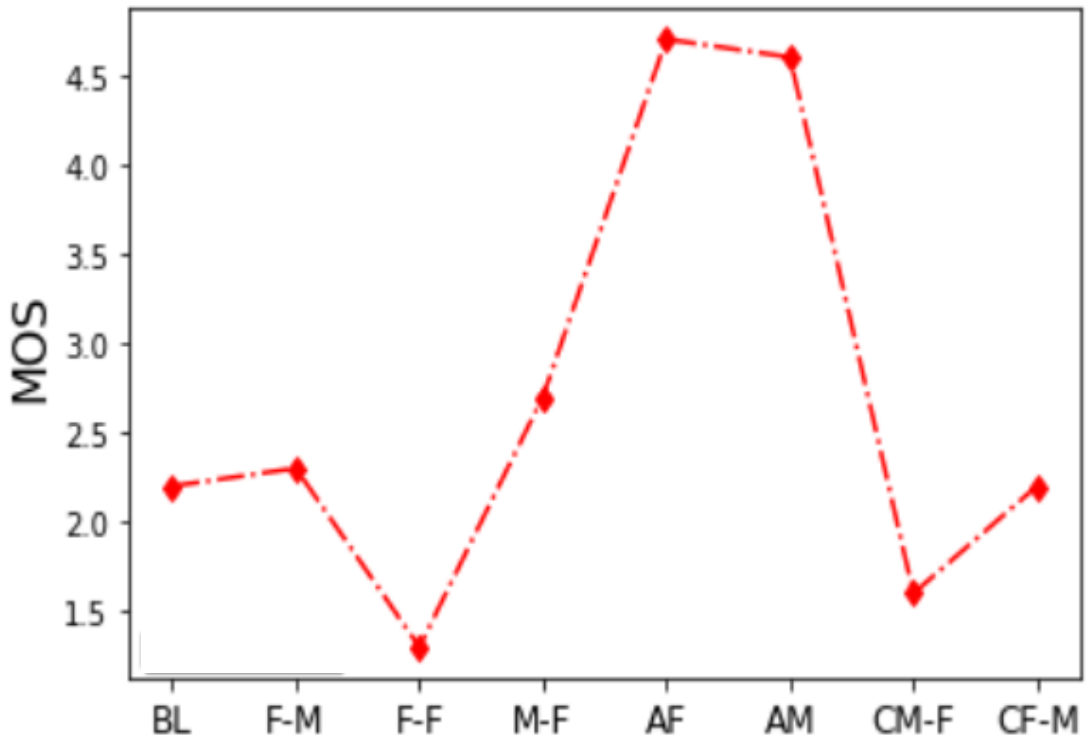
What is CycleGAN?

- cycle-consistent adversarial network
- for unpaired image to-image translation
- in cycle

03

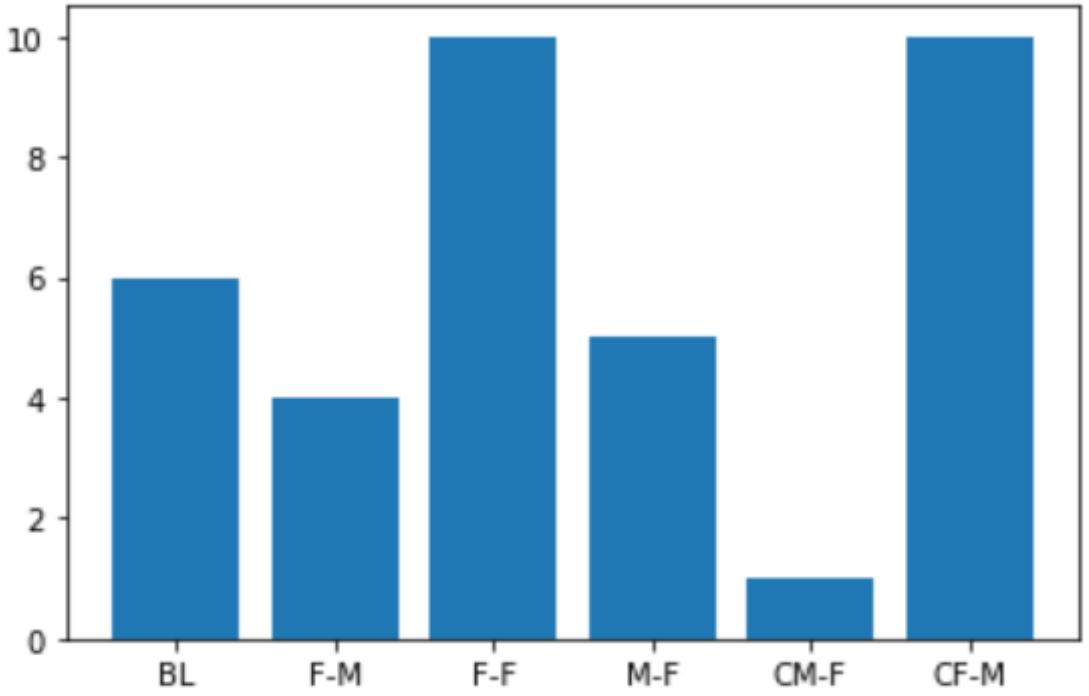
Results and Conclusions

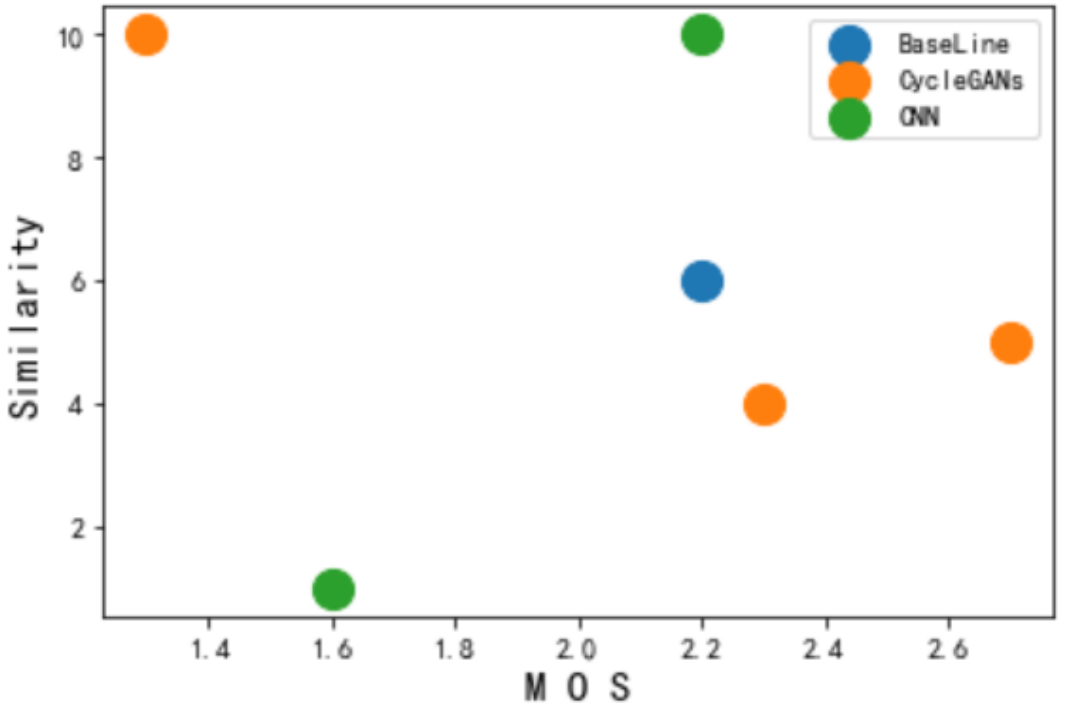
3.1 MOS



- Baidu API gets the highest score
- inter-gender VC & intra-gender VC
- Results of Random CNN?

3.2 ABX


10
8
6
4
2
0
BL
F-M
F-F
M-F
CM-F
CF-M


BaseLine
CycleGANs
CNN
Similarity
MOS
10
8
6
4
2
1.4
1.6
1.8
2.0
2.2
2.4
2.6

武漢大學
WUHAN UNIVERSITY

Thanks for Your Attention

自强、弘毅、求是、拓新