Personalized Singing Voice Beautifier

林泳鵬、汪宣甫、陳韋傑

Outline

- The Team
- Motivation
- Method
- Goals

The Team

- **Team Members**
  - R11944074 汪宣甫(Student)
  - R11922016 林泳鵬(Student)
  - R12922051 陳韋傑(Auditor)
- **GPU resources**
  - Tesla V100-SXM2-16GB/32GB **x(1 ~ 2)**
- **Music-related background** : None
- **Previous ML experiences** : Machine Learning / Deep Learning related courses

Motivation

- Our innovation aims to train a **personalized voice beautifier**.
- We referred to the relevant paper [1], but identified some limitations in its training data.
    - The training pair data : (professional, amateur) singing voice.
    - We observed that amateur singing surpasses that of ordinary individuals.
- So we want to investigate some methods that can generate data pairs which are more suitable to our scenarios.
- And use those optional datas to train a personalized voice beautifier.

[1] : Liu, Jinglin, et al. "Learning the beauty in songs: Neural singing voice beautifier." arXiv preprint arXiv:2202.13277 (2022).

Motivation (cont.)

- Example of their dataset.
  - Professional
  - Amateur
  - Mine

Method (cont.)

Learning the Beauty in Songs: Neural Singing Voice Beautifier

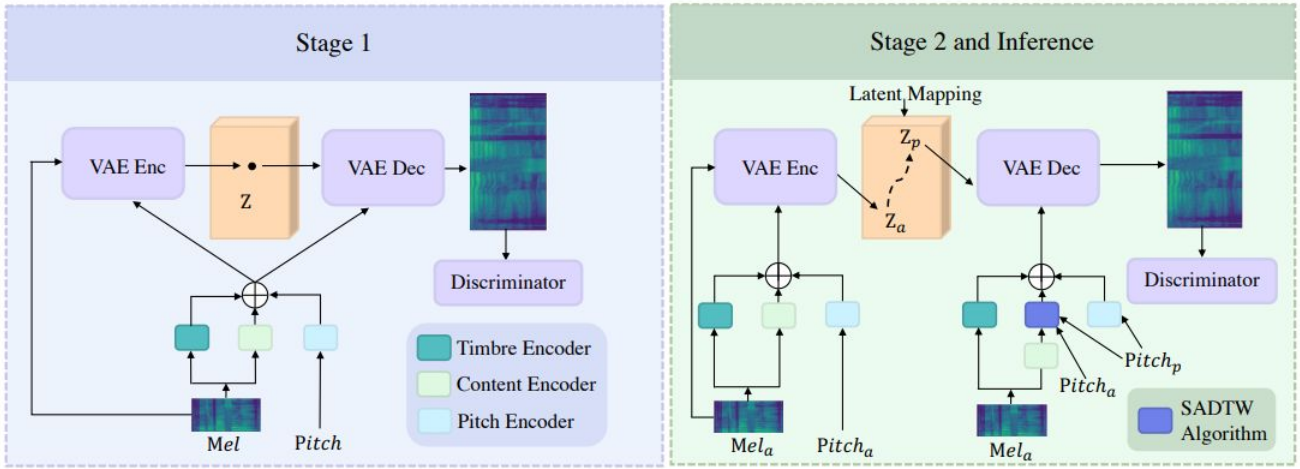
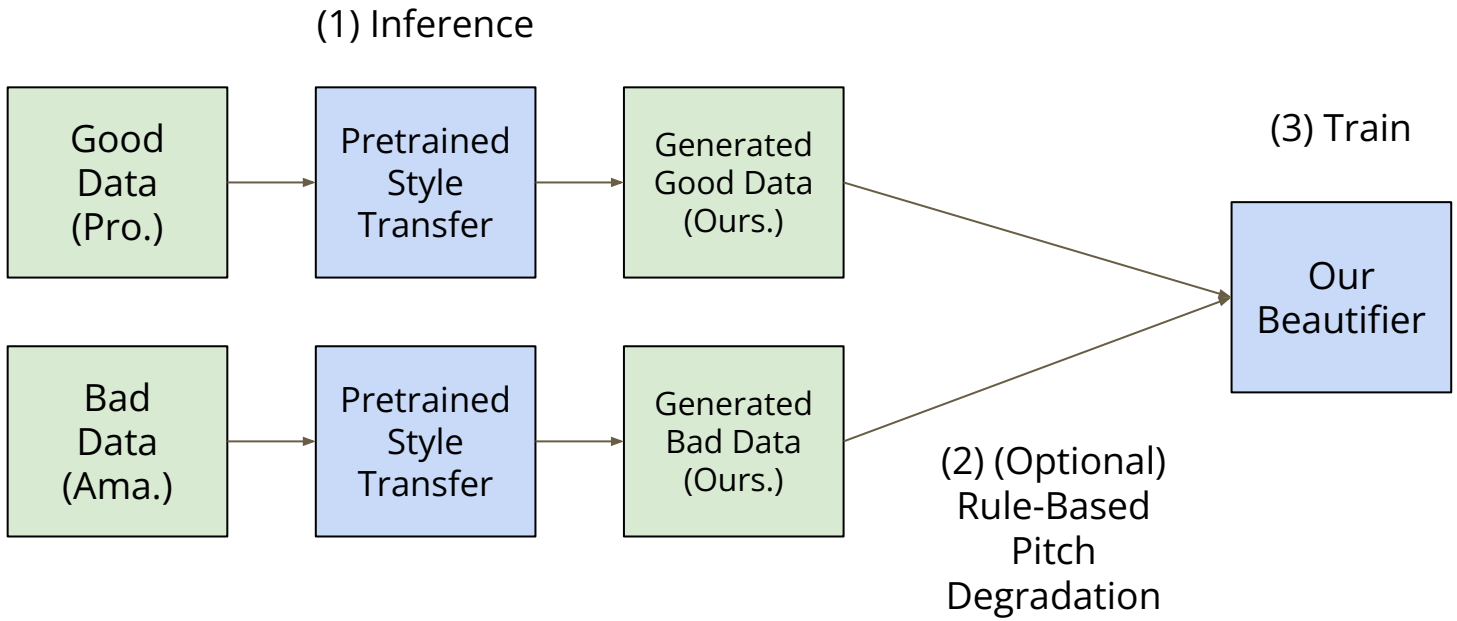
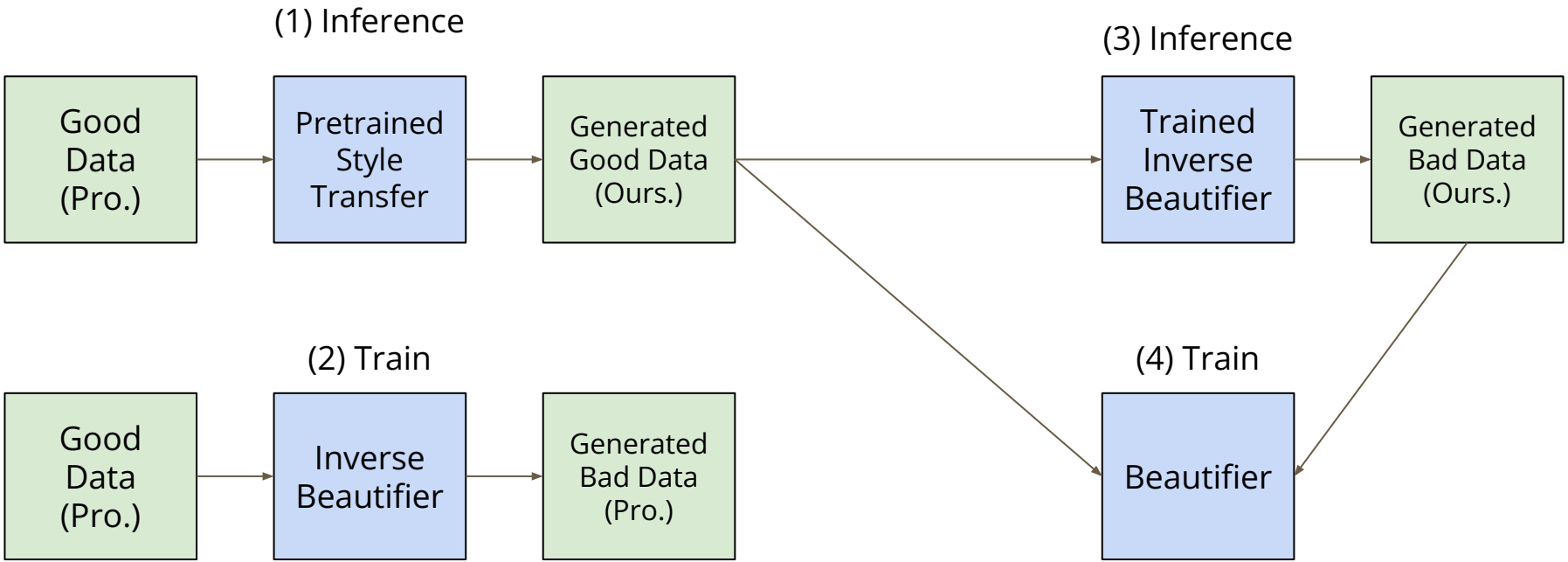


Figure 1: The overview of NVSB. The training process consists of 2 stages, and the second stage shares the same pipeline with the inference stage. "VAE Enc" means the encoder of CVAE; "VAE Dec" means the decoder of CVAE; "Mel" means the mel-spectrogram; "**z**" means the latent variable of the vocal tone; the "a"/"p" subscript means the amateur/professional version.

Method (cont.)


(1) Inference
Good Data (Pro.)
Pretrained Style Transfer
Generated Good Data (Ours.)
Bad Data (Ama.)
Pretrained Style Transfer
Generated Bad Data (Ours.)
(2) (Optional) Rule-Based Pitch Degradation
(3) Train
Our Beautifier

Method (cont.)


(1) Inference
Good Data (Pro.)
Pretrained Style Transfer
Generated Good Data (Ours.)
(3) Inference
Trained Inverse Beautifier
Generated Bad Data (Ours.)
(2) Train
Good Data (Pro.)
Inverse Beautifier
Generated Bad Data (Pro.)
(4) Train
Beautifier

Goals

- Goal 1 : Train our simple Beautifier
  - Use public dataset
  - Style transfer to match our timbre
- Goal 2 : Inverse Beautifier training to get more pair data
  - Data collection from public music platform
- Goal 3 : Train our Powerful Beautifier
  - Use more data from Inverse Beautifier training to fine-tune model