

HỆ THỐNG CHUYỂN ĐỔI BÀI NHẠC TIẾNG VIỆT SANG PHIÊN BẢN TIẾNG ANH

Nguyễn Thiên Bảo - 23520127

Tóm tắt

- Lớp: CS519.Q11.KHTN
- Link Github: [CS519.Q11.KHTN](#)
- Link YouTube video: https://youtu.be/6FoznYZu4_s

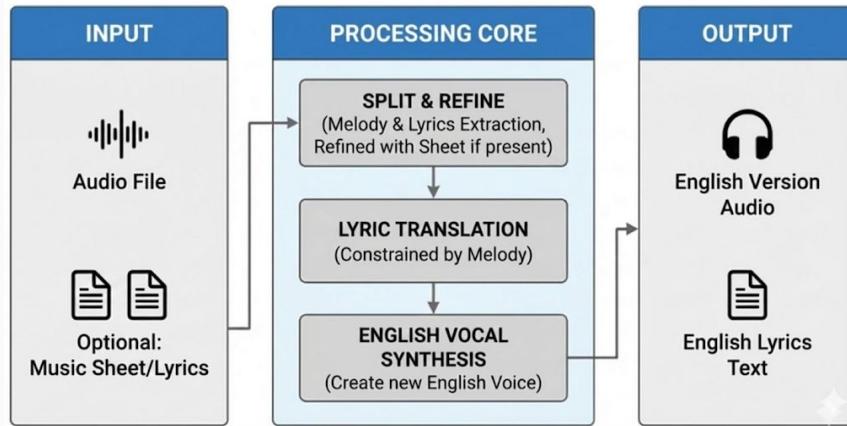


Nguyễn Thiên Bảo - 23520127

Giới thiệu

Nhiều người học tiếng Anh mong muốn nghe lại bài hát mình đã thuộc lòng giao điệu và ý nghĩa dưới một ngôn ngữ khác nhằm vừa giải trí vừa học ngôn ngữ/ nhà sáng tạo nội dung gấp rào cản khi muốn cover bài hát...

Hệ thống đề xuất: Pipeline tự động nhận file audio (tùy chọn: lyrics/music sheet) → tách nhạc nền - giọng hát → tách và dịch lời → sinh giọng hát tiếng Anh theo giao điệu gốc.



Thách thức:

- Sự khác biệt rõ rệt giữa ngôn ngữ đơn âm, có thanh điệu (Tiếng Việt) và ngôn ngữ đa âm, nhấn trọng âm (Tiếng Anh).
- Tiếng Việt chưa được chú trọng nghiên cứu nhiều như các ngôn ngữ lớn khác như trung, hàn, nhật,... đặc biệt là trong lĩnh vực dịch thuật âm nhạc.

Mục tiêu

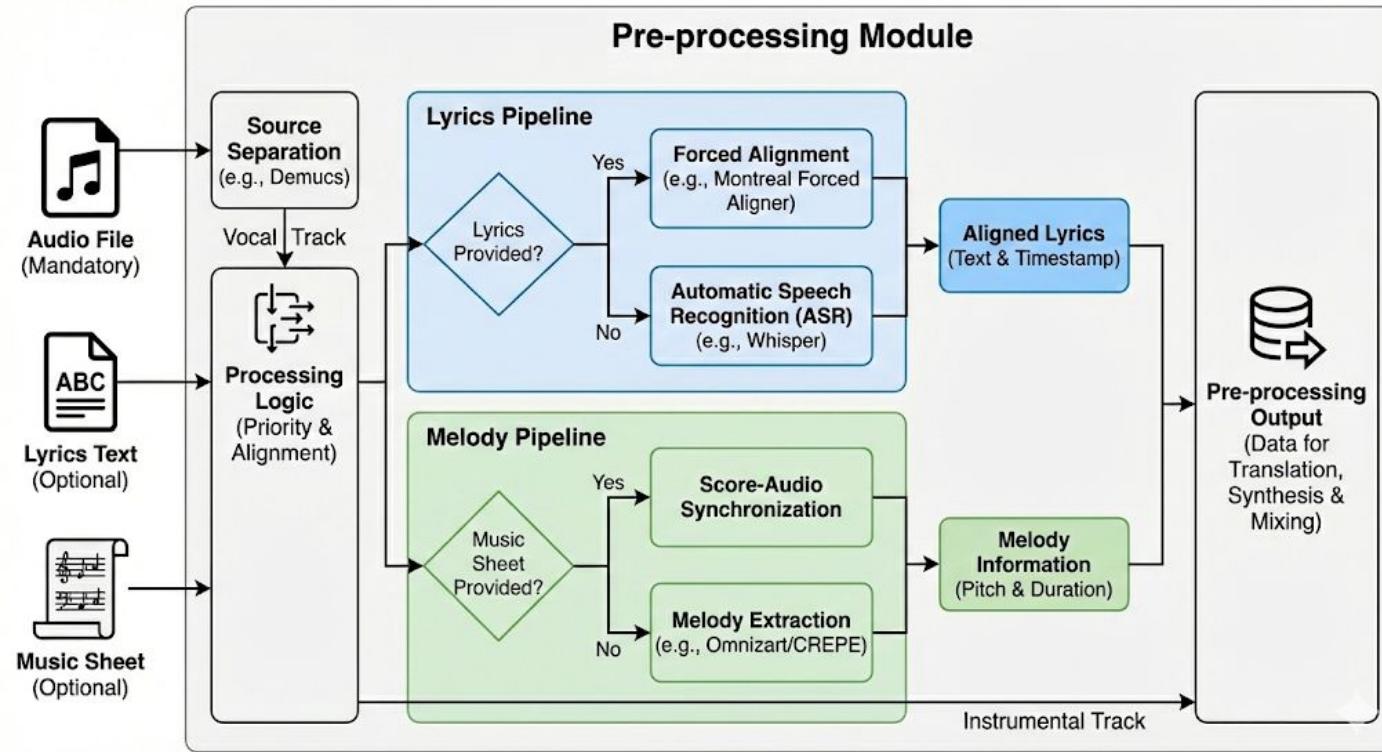
- Xây dựng được một hệ thống tự động chuyển bài hát tiếng Việt sang English Version.
- Dịch lyrics tiếng Việt → tiếng Anh **giữ cấu trúc hát** qua các đánh giá:
 - Thời lượng câu tương ứng
 - Nhịp và tiết tấu khớp melody
 - Văn & nhấn âm phù hợp
 - Độ chính xác ngữ nghĩa
- Tạo giọng hát tiếng Anh theo **melody gốc**. Xuất ra:
 - English lyrics
 - Audio English Version hoàn chỉnh.

Nội dung và Phương pháp

- 1. Xây dựng hệ thống tiền xử lý đầu vào tự động:**
 - tách vocal và nhạc nền từ audio,
 - trích xuất melody và lời hát từ phần vocal/music sheet & lyrics.
- 2. Đề xuất mô hình dịch lyrics có ràng buộc âm nhạc:**
 - giữ nội dung và số âm tiết tương đương với bản gốc,
 - nhấn nhá phù hợp, đảm bảo phối vần trong từng đoạn.
- 3. Xây dựng mô hình tổng hợp giọng hát (Singing Voice Synthesis)**
được điều kiện hóa theo melody gốc và lyrics tiếng Anh đã căn chỉnh.
- 4. Tích hợp thành hệ thống hoàn chỉnh** cho ra English Version.

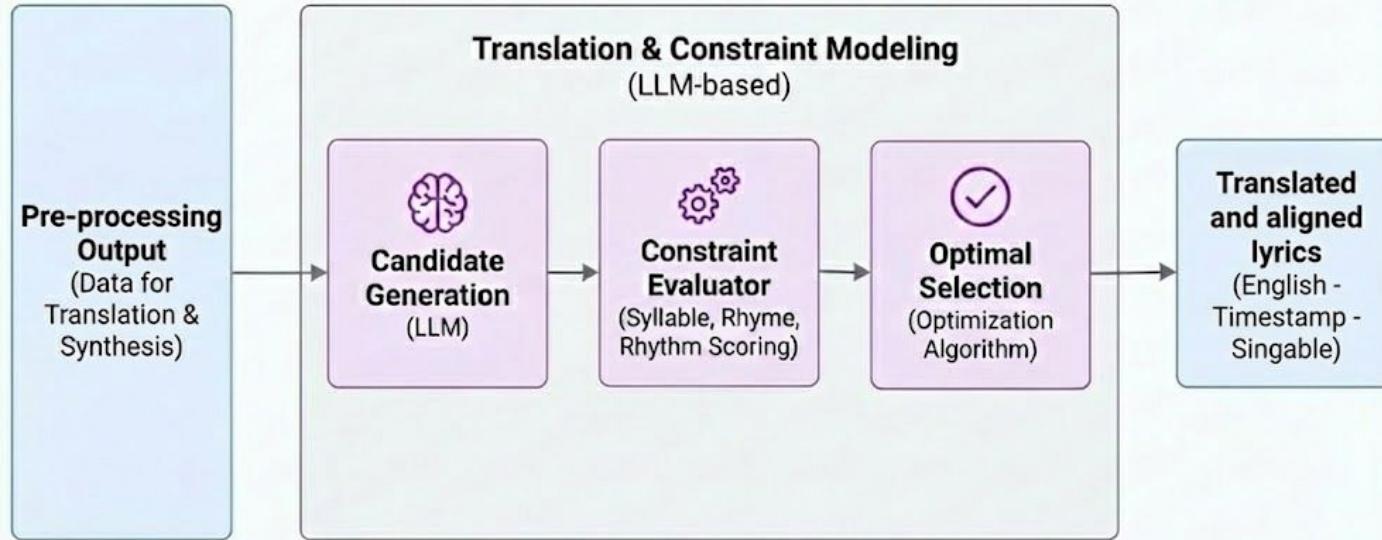
Nội dung và Phương pháp

Hệ thống tiền xử lý đầu vào tự động



Nội dung và Phương pháp

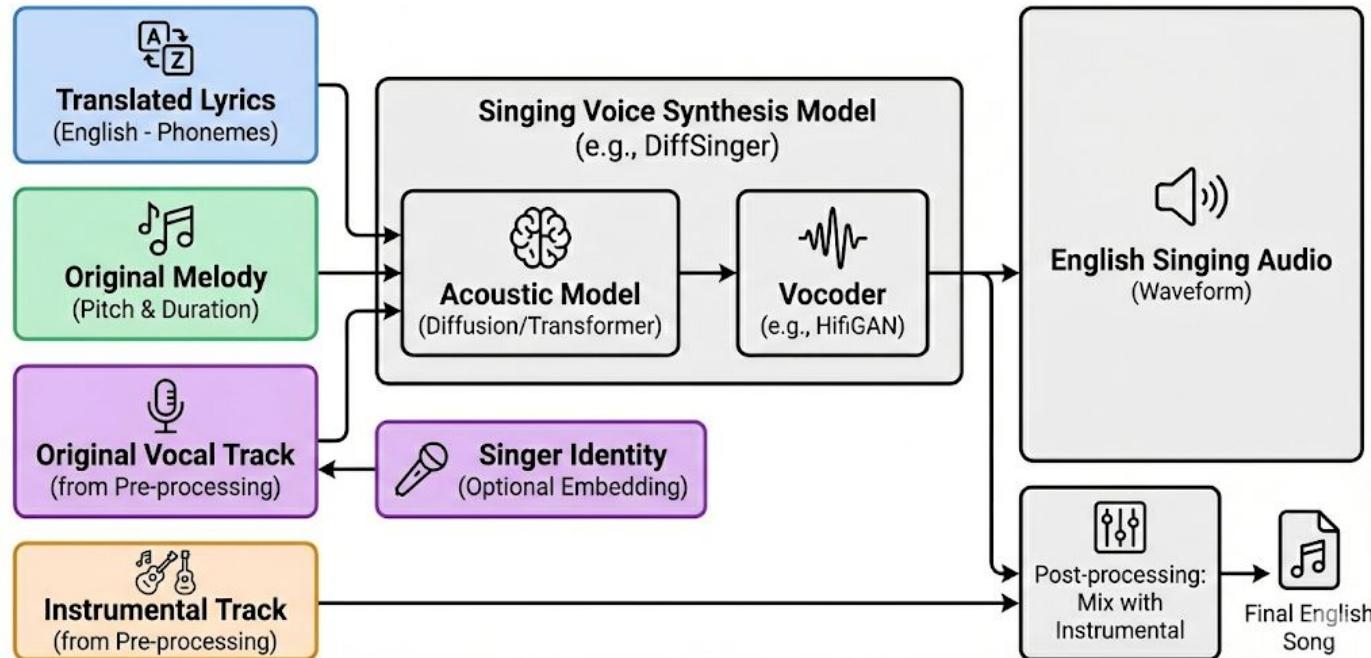
Mô hình dịch lyrics có ràng buộc âm nhạc Music-Constrained Lyric Translation Model



Nội dung và Phương pháp

Mô hình tổng hợp giọng hát

Singing Voice Synthesis (SVS) Module



Kết quả dự kiến

Hệ thống dễ sử dụng, các module phải được ghép nối mượt mà, sao cho từ file audio (music sheet/lyrics optional) có thể xuất ra:

- bộ lời tiếng Anh hoàn chỉnh, phù hợp để hát trên nền nhạc gốc.
- Audio English Version giữ giai điệu bản gốc.

Xây dựng web để thao tác với hệ thống.

Tài liệu tham khảo

- [1]. Linan Ou, Xiaojuan Ma, Min-Yen Kan, Ye Wang: Songs Across Borders: Singable and Controllable Neural Lyric Translation. ACL (1) 2023: 447-467
- [2]. Jinglin Liu, Chengxi Li, Yi Ren, Feilong Chen, Zhou Zhao: DiffSinger: Singing Voice Synthesis via Shallow Diffusion Mechanism. AAAI 2022: 11020-11028
- [3]. Simon Rouard, Francisco Massa, Alexandre Defossez: Hybrid Transformers for Music Source Separation. ICASSP 2023: 1-5
- [4]. Yu-Te Wu, Berlin Chen, Li Su: Multi-Instrument Automatic Music Transcription with Self-Attention-Based Instance Segmentation. IEEE/ACM Trans. Audio Speech Lang. Process. 29: 2798-2811 (2021)
- [5]. Chengxi Li, Kai Fan, Jiajun Bu, Boxing Chen, Zhongqiang Huang, Zhi Yu: Translate the Beauty in Songs: Jointly Learning to Align Melody and Translate Lyrics. Findings of the Association for Computational Linguistics: EMNLP 2023: 27-39