

# Study of transformer-based TTS and its embedded implementation

---

Sian-Yi Chen

Advisors : Tay-Jyi Lin and Chingwei Yeh

# Outline

## Action item

- 完成 Transformer-based TTS 的嵌入式系統(尚未完成)

## Status report

- 前情提要
  - VCC2020 baseline 版本使用腳本執行，因為使用腳本使用多種語言難以直接轉換成C，因此將過程中主要的程式獨立取出，並從Linux系統移植到Windows環境下，確保無誤。
- 實作進度
  - 要完成demo，系統須具備ASR、TTS、PWG三個部分，目前從TTS的部分開始實作，實作細節放下一頁
    - ASR
      - 輸出文字並轉換成拼音
      - 轉換成json格式
      - 套用x-vector
    - (進行中) ■ TTS
    - PWG
      - cmvn
      - normalize
      - decode

# TTS 實作細節

- 執行過程

- 逐行確認函數功能與輸出結果
- 判斷各程式碼是否為所需功能，如log檔就捨去
- python中可以藉由import使用其他檔案中所定義好的function，程式中import眾多function，因此需一層一層往裡面翻確認功能

- 後續規劃

- 用C讀取json格式的input檔，目前進行到一半
- 目前已確認第4步驟定義模型各層輸入輸出大小，各層是什麼模型，完成讀取json檔後，使用C定義一個Transformer模型

tts\_decode.py

decode.py

1. 設定隨機種子
  2. 讀取模型超參數檔
  3. 讀取 tts\_decode.py 程式參數
  4. 定義模型 (Transformer)
  5. 讀取模型已訓練的參數(bin檔)
  6. 讀取 INPUT (JSON 檔)
- ·  
·

TTS 程式執行步驟