

Hybrid VC

Student : Sian-Yi Chen

Advisor : Tay-Jyi Lin and Chingwei Yeh

● Online VC

□ 問題與解決方案

常見問題	解決方法
專有名詞種類範圍太廣	限制專業領域，預先加入提示，使用 Google API 提供的 Speech adaptation 解決
多語言混雜 (中英文交錯)	Google API 提供偵測語言功能，最多可以辨識四種語言
噪音	Google API 具備雜訊處理功能，無須額外處理
即興口語 (註1)	為 Google 演算法問題，無法改善
破音字	不轉換成中文字，改轉換成羅馬拼音，再由羅馬拼音轉換成 Google 語音

(註1) 即興口語：

沒有讀稿、背稿，平時說話就稱為即興口語，有著音韻變化、語助詞、副語言、資訊內容較為零散等特性，會導致辨識率下降

● Offline VC

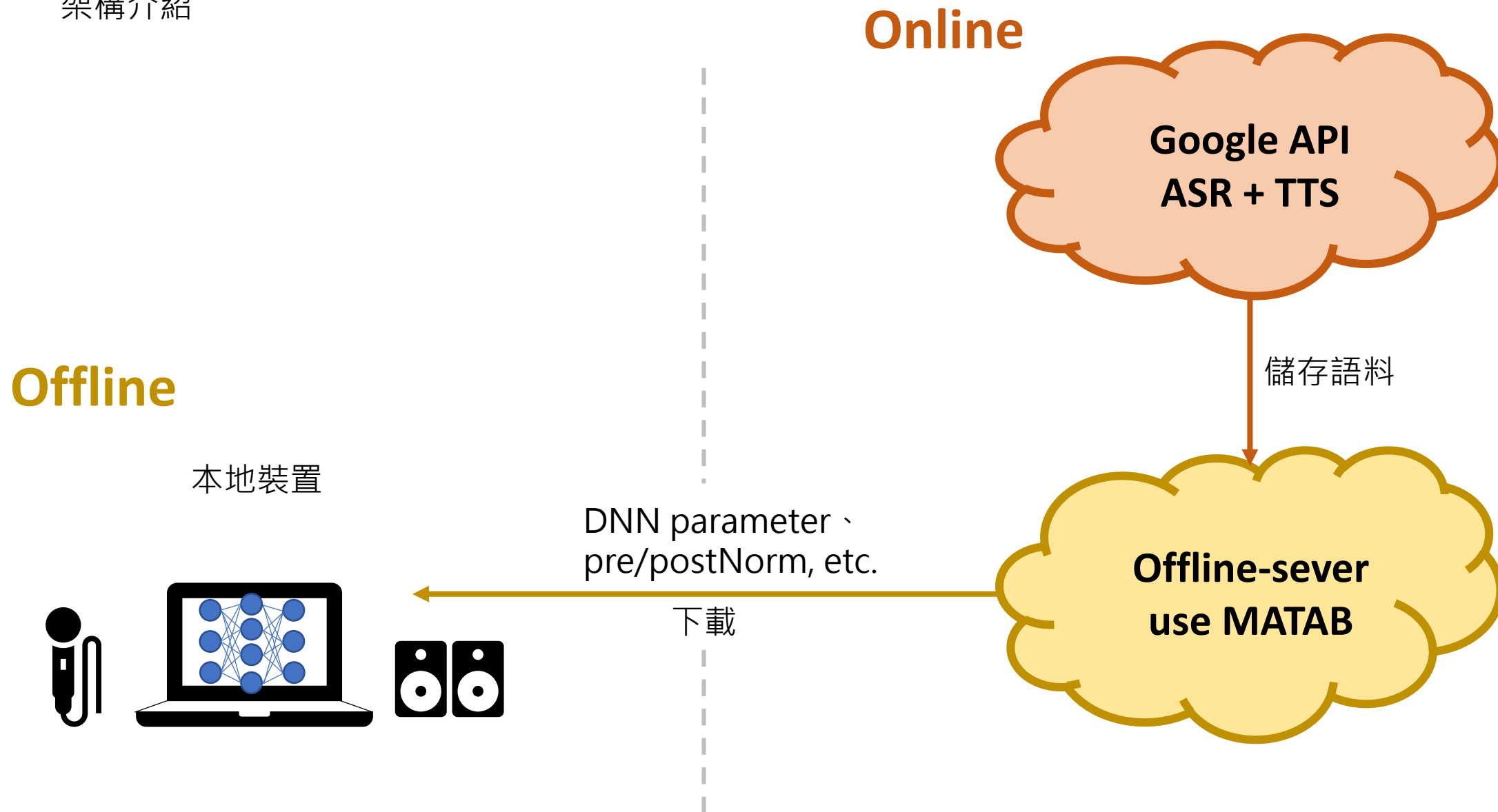
□ 架構介紹

□ 問題回覆

1. 如何得知新語料可改善舊有模型
2. 模型改善時機
3. 如何更換模型

Offline VC

架構介紹



Offline VC

問題回覆

- 如何得知新語料可改善舊有模型

1. 先前東昇與泰翔先將 320 句與其他數據畫出 21 區比例圖再利用 Frechet distance 方法去計算兩兩相似度，但此方法為相對概念，目前想到的量化方法為將計算結果，若結果表示為 1~5，將 1~5 皆產出一個模型，利用主觀聽覺與 PESQ 去定出一個指標作為好的標準
2. 將新增的語料畫出 21 區比例圖後與 320 句比例圖做相似度比較，若達第一點的標準就替換掉現有模型

- 模型改善時機

1. 此問題涉及到資料量多寡，在 Azure 官網中語音模型有提到訓練資料量需收集到半小時至一小時，才會有較好的轉換效果
2. 320 句若每句為 5 秒，時間長度近似為 27 分鐘，因此推估新增語料量應與 320 句語料量相似才足夠畫出比例圖與其做比對

- 如何更換模型

1. 從 offline-sever 下載下來的所有參數 (DNN model、Norm 參數) 均為文檔，可直接在裝置端將參數檔做更換