

Off line 語音轉換器規劃及進度

Student : Sian-Yi Chen

Advisor : Tay-Jyi Lin and Chingwei Yeh

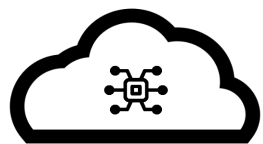
Outline

1. 語音轉換器進度
 1. 系統架構圖
 2. 所需資源
 1. API 過了試用期需付費
 2. 所需時間 (時程表)
 3. 程式架構圖
 1. 裝置端 python 所需技術
 2. 裝置端 java 所需技術
 3. 雲端技術尚未找齊
2. 實作架構
 - 已架好 python 環境
3. 下週規劃
 - 輸入音檔使用 SpeechRecognition 套件達成 Speech-to-Text，並輸出文字

■ 系統架構

有網路

轉換成 google 小姐



Google cloud platform

- ASR(Automatic Speech Recognition)
- TTS(Text To Speech)

1. input
2. upload
3. download
4. play

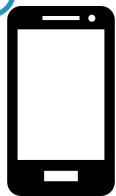
①



電腦：
使用python

優：較為快速可以看到成品
缺：並非一個可攜帶裝置

②



手機：
Android studio平台
使用 Java 語言

優：為一可攜帶的裝置
缺：語言不熟悉需時間學習

無網路



Google cloud platform

- Compute Engine 架 VM
(放目前實驗室的 YMDVC)

1. input
2. upload
3. store & convert (雲端操作)
4. store converted input (雲端操作)
5. download converted input & DNN model
6. play

■所需資源(1/2)

裝置：

1. 電腦，使用 Python 製作初版雛形
2. Android 手機，使用 Android Studio 平台 Java 語言實作
(沒接觸過 Java 語言，需要要多時間熟悉)

雲端：

使用 Google cloud platform API (90天免費試用+300元美金抵免額)

1. Speech-to-Text
 - 每個月可免費處理 60分鐘資料量，之後以 15 秒為單位計費
2. Text-to-Speech
3. Compute Engine 架設一台虛擬機器，將 DVC 1.0 語音資料與轉換程式放在雲端
 - 每月30G硬碟空間、1GB網路輸出量

■所需資源 - 時程表(2/2)

[illegible]

■ 程式架構(1/3) (使用 **Python** 製作雛形)

裝置：

```
麥克風輸入音檔          // speech_recognition.microphone()，函式可使麥克風輸入
是否連上網際網路        // 使用系統指令“os.system()”函數 ping 網域，來測
if ( 能連網 ) {          試是否有網路，回傳 0 or 1
    upload
    Speech-to-Text        // SpeechRecognition 套件
    Text-to-Speech        // &
    儲存音檔              // gTTS 套件
    下載 DNN 模型參數      // 使用 paramiko 套件使用 ssh 連線遠端 VM 複製檔案
    寫 python 版 DVC1.0
} else if (不能連網){
    使用儲存的 DNN 1.0 模型執行儲存的參數 // 寫 DVC 1.0 python 版本，需處理傳回來的參數
}
播放音檔                  // 使用 pygame 套件 mixer() 函式播放音檔
```

■ 程式架構(2/3) (使用 Java 實作手機 APP)

裝置：

```
麥克風輸入音檔          // Android 提供 MediaRecorder 達成手機收音、儲存、播音
是否連上網際網路        // 加入權限，再使用內建提供函數判斷是否連網
if ( 能連網 ) {
    Speech-to-Text        // google 有提供 libraries
    Text-to-Speech
    儲存音檔
    下載 DNN 模型        // 利用 ssh 遠端連線到 Google VM 複製檔案
} else if (不能連網){
    使用儲存的 DNN 1.0 模型執行儲存的參數    // 寫 DVC 1.0 java 版本，需處理傳回來的參數
}
播放音檔
```

■ 程式架構(3/3) (使用 Compute Engine 的 VM)

雲端：

開 google 虛擬機

將 YMDVC 放上雲端運行

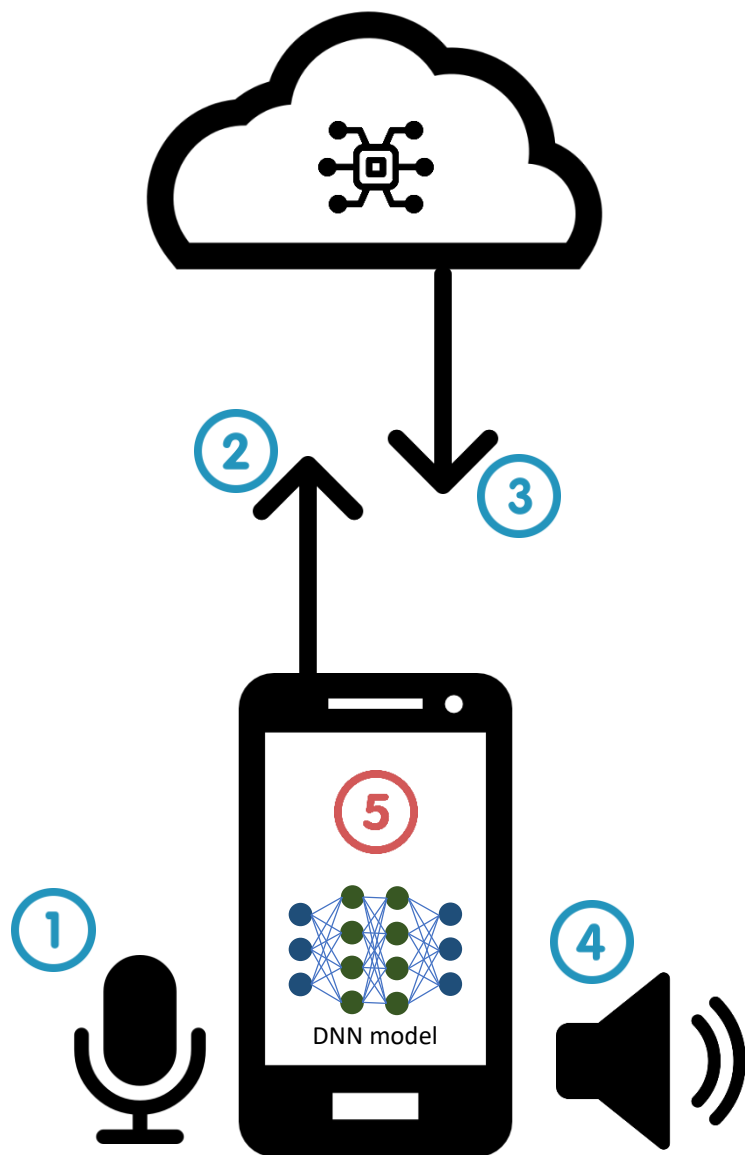
接收、轉換並儲存

將轉換後參數儲存並處理

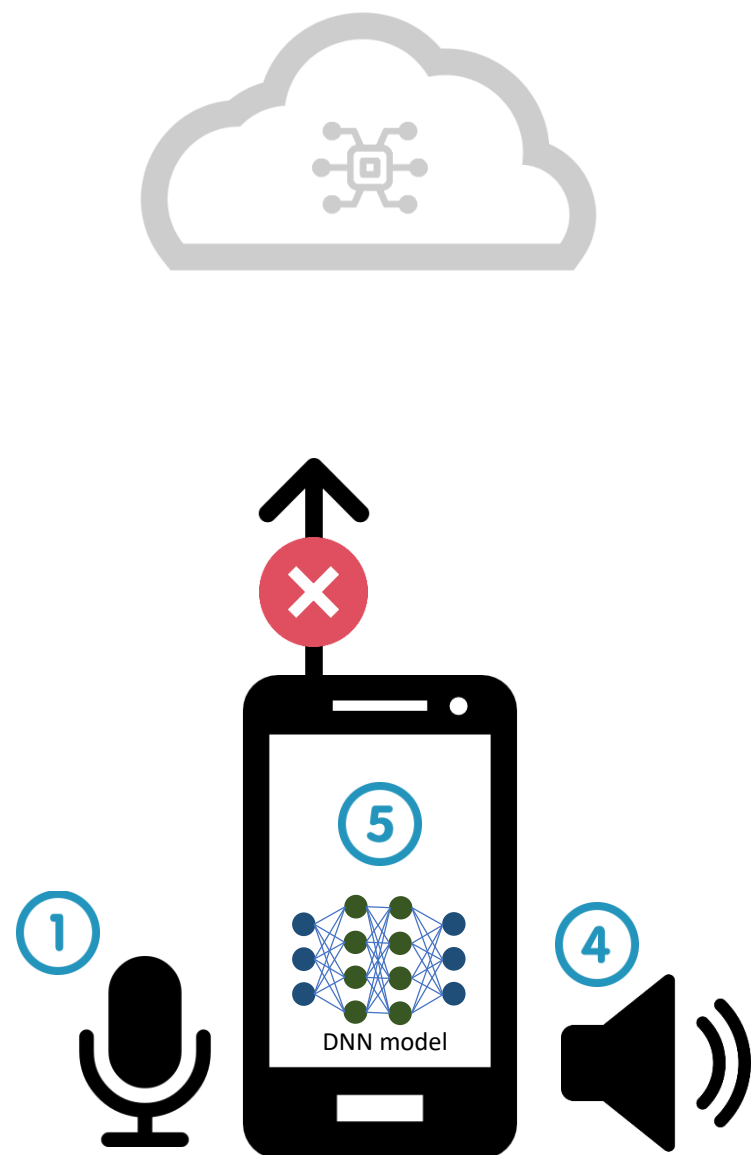
■ 新產品：

離線語音轉換器

- ① 輸入音訊
- ② 上傳至雲端運算
- ③ 下載轉換後聲音
- ④ 輸出音訊
- ⑤ 下載雲端模型並儲存
- ⑤ 利用DNN模型轉換



雲端運算並支援離線轉換



離線轉換