

Personalized offline DVC

Student : Sian-Yi Chen

Advisor : Tay-Jyi Lin and Chingwei Yeh

■ 會議記錄改前

陳憲億

Personalized offline DVC

報告摘要

- purpose
- implementation
- demo

討論內容

- Outline 的 wording 有點奇怪，改 purpose，implementation，demo
- 從 user 的角度去訂定 scenario，拿到這個 device 功能是什麼，該怎麼用
- 理想的產品是什麼，能展現什麼功能，價值在哪，需要怎樣的動作
- 想清楚產品的定位，講清楚難的部分，如何驗證他，論文要做的是這個

Action item

- 先規劃一版理想的 demo scenario，再討論實作細節

■ 會議記錄改後

陳憲億

報告事項

- Personalized offline DVC

報告摘要

- Creation purpose
 1. 實現一手持裝置可以不管在何時何地將 A(輸入) 聲音轉換成 B(目標) 聲音
- Implementation flow
 1. 在 PC 上完成功能流程
 2. 做一個手持裝置 on line
 3. 將第二部的輸入輸出傳上雲端並等待運算完下載 model
 4. Off line 時，第三步的 model 可以發揮作用
- Demo scenario
 1. 操作裝置
 2. 產品特點
 3. 展示流程

討論內容

- Outline 更改成 purpose、implementation、demo
- 從 user 的角度去訂定 scenario，拿到這個 device 功能是什麼，該怎麼用
- 理想的產品是什麼，能展現什麼功能，價值在哪，需要怎要的操作
- 想清楚產品的定位，講清楚難的部分，如何驗證他，論文要做的是這個

Action item

- 從使用者的角度做說明 (初次使用與穩定狀態時)
- 產品定位 (highlight 最重要的部分)

Outline

個人化離線語音裝置

● Product descriptions

□ Introduce

- 它是一個具有離線模式的即時錄放音手持裝置
- 功能為可以在何時何地將 A(輸入) 聲音轉換成 B(目標) 聲音

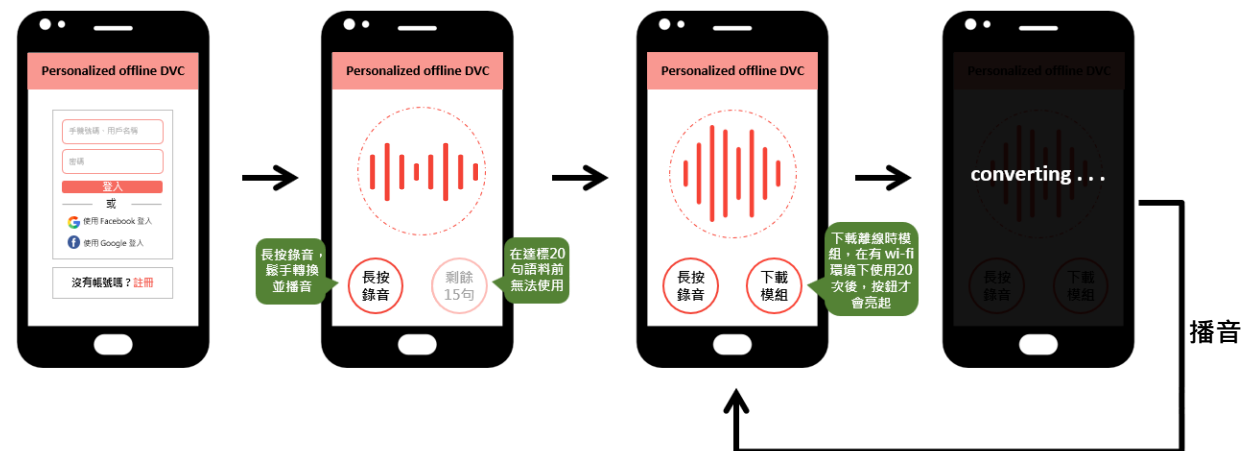
□ User scenario

- 初次使用
 1. 登入 / 註冊
 2. 導覽 (綠色框框)
 3. 在有網路的狀態下，請說 20 句後才能使用離線功能 (離線功能上顯示剩餘句數，ex：說了 5 句，顯示剩餘 15 句)
 4. 長按錄音，放開則進入等待畫面進行轉換並播音
- 達穩定狀態時
 1. 離線功能轉換效果會與使用者使用次數增加而上升
 2. 當達標 20 句時，按下下載模組即可在沒有網路等地進行錄放音轉換

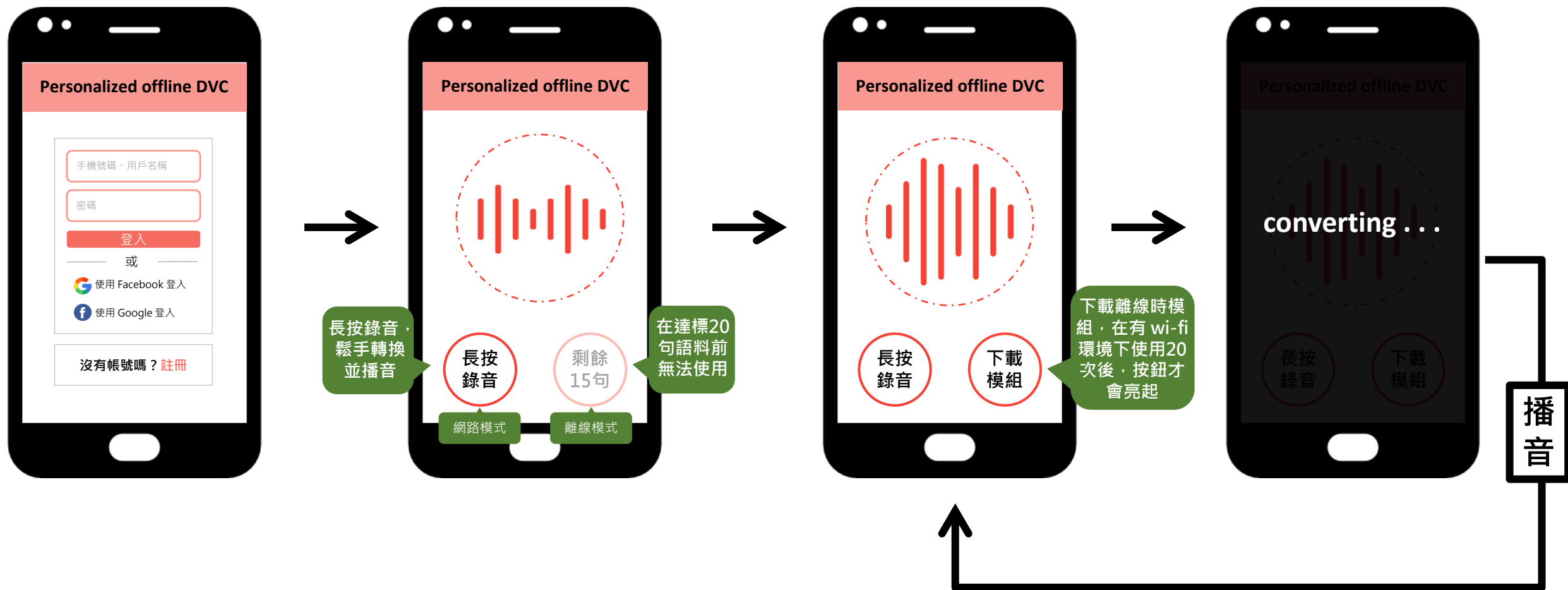
● key technology

- YMDVC 上雲端，因為需要讓它變成自動化，作業系統需要從 windows 變使用 Linux
 - 自動將資料對齊並訓練，最後再將訓練完參數做整理
- 本地端程式將輸入音訊推上雲端運算後下載參數與模型

● Demo scenario



產品預想圖



附錄

■ 程式架構 (裝置端)

1. 麥克風輸入音檔
2. 是否連上網際網路

if (能連網) {

3. 將輸入的音檔上傳至雲端虛擬機
4. Speech-to-Text
5. Text-to-Speech
6. 將轉換後的音檔上傳至雲端虛擬機
- if(觸發條件)
7. 下載 DNN 模型參數
8. 寫 Python 或是 Java 版 DVC

} else if (不能連網) {

9. 使用儲存的 DNN 模型執行儲存的參數

}

10. 播放音檔

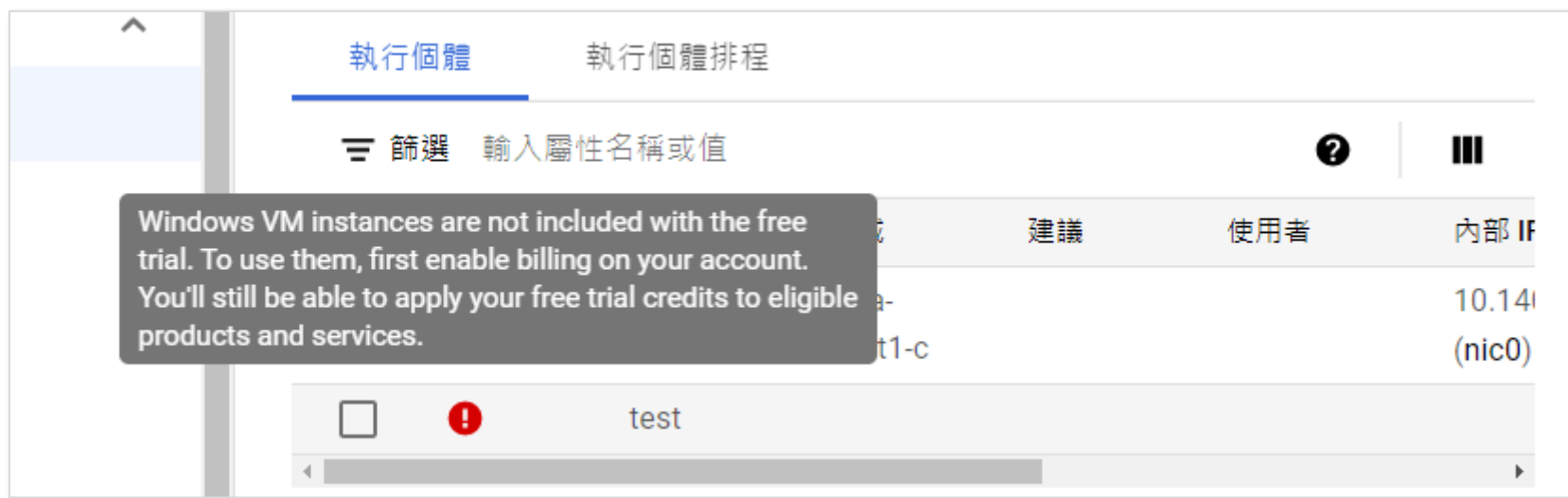
- 初次使用，請在有網路的狀況下至少說 20 句話再使用離線功能，原本想設定 320 句，但次數太多會降低使用意願
- 觸發條件 (按下 button 或是更新下載)
 - 模型參數手動下載or 自動下載
 - 轉換/下載期間跳出預估等待時間

■ 程式架構 (雲端)

1. 開 Google 虛擬機
2. 架設 YMDVC 執行環境
3. 利用裝置端錄製的聲音 & 轉換後的聲音 **自動**進行 DTW **對齊**
4. 將對齊後聲音進行 **自動訓練** (一個禮拜或是一個月一次)
5. 將取得的模型 **自動**做資料**整理** (模型轉換 pre/post norm, DNN model)
6. 將轉換完模型放進資料夾內方便裝置端存取

■ GCP VM 無法使用 Windows

Compute Engine (VM 執行個體)



執行個體 執行個體排程

篩選 輸入屬性名稱或值

Windows VM instances are not included with the free trial. To use them, first enable billing on your account. You'll still be able to apply your free trial credits to eligible products and services.

建議	使用者	內部 IP
		10.141.1.10
		(nic0)

test

Google Cloud Free計劃



計劃覆蓋

您的免費試用額度適用於所有Google Cloud資源，包括 [Google Maps Platform](#)的使用情況，但以下情況除外：

- 您無法將GPU添加到您的VM實例。
- 您不能請求增加配額。有關Compute Engine配額的概述，請參閱 [資源配額](#)。
- 您無法創建基於Windows Server 映像的VM實例。

要執行上面列表中的任何操作，您必須 [升級到付費Cloud Billing帳戶](#)。

GCP VM

```
fg6ts30@off-line-server:~ - Google Chrome
ssh.cloud.google.com/projects/dev-acolyte-307405/zones/asia-east1-c/instances/off-line-server?authuser=1&hl=zh_T...
Downloading grpcio-1.36.1-cp36-cp36m-manylinux2014_x86_64.whl (4.1 MB)
| 4.1 MB 66.7 MB/s
Collecting keras-applications>=1.0.8
  Downloading Keras Applications-1.0.8-py3-none-any.whl (50 kB)
  | 50 kB 8.9 MB/s
Collecting gast==0.2.2
  Downloading gast-0.2.2.tar.gz (10 kB)
Collecting h5py
  Downloading h5py-3.1.0-cp36-cp36m-manylinux1_x86_64.whl (4.0 MB)
  | 4.0 MB 52.3 MB/s
Collecting setuptools>=41.0.0
  Downloading setuptools-54.2.0-py3-none-any.whl (785 kB)
  | 785 kB 53.0 MB/s
Collecting markdown>=2.6.8
  Downloading Markdown-3.3.4-py3-none-any.whl (97 kB)
  | 97 kB 9.6 MB/s
Collecting werkzeug>=0.11.15
  Downloading Werkzeug-1.0.1-py2.py3-none-any.whl (298 kB)
  | 298 kB 66.6 MB/s
Collecting importlib-metadata
  Downloading importlib_metadata-3.10.0-py3-none-any.whl (14 kB)
Collecting cached-property
  Downloading cached_property-1.5.2-py2.py3-none-any.whl (7.6 kB)
Collecting zipp>=0.5
  Downloading zipp-3.4.1-py3-none-any.whl (5.2 kB)
Collecting typing-extensions>=3.6.4
  Downloading typing_extensions-3.7.4.3-py3-none-any.whl (22 kB)
Using legacy 'setup.py install' for gast, since package 'wheel' is not installed.
Using legacy 'setup.py install' for termcolor, since package 'wheel' is not installed.
Using legacy 'setup.py install' for wrapt, since package 'wheel' is not installed.
Installing collected packages: zipp, typing-extensions, six, numpy, importlib-metadata, cached-property, wheel, wer
kzeug, setuptools, protobuf, markdown, h5py, grpcio, absl-py, wrapt, termcolor, tensorflow-estimator, tensorboard,
opt-einsum, keras-preprocessing, keras-applications, google-pasta, gast, astor, tensorflow
  Running setup.py install for wrapt ... done
  Running setup.py install for termcolor ... done
  Running setup.py install for gast ... done
Successfully installed absl-py-0.12.0 astor-0.8.1 cached-property-1.5.2 gast-0.2.2 google-pasta-0.2.0 grpcio-1.36.1
h5py-3.1.0 importlib-metadata-3.10.0 keras-applications-1.0.8 keras-preprocessing-1.1.2 markdown-3.3.4 numpy-1.19.
5 opt-einsum-3.3.0 protobuf-3.15.7 setuptools-54.2.0 six-1.15.0 tensorboard-1.15.0 tensorflow-1.15.0 tensorflow-est
imator-1.15.1 termcolor-1.1.0 typing-extensions-3.7.4.3 werkzeug-1.0.1 wheel-0.36.2 wrapt-1.12.1 zipp-3.4.1
[fg6ts30@off-line-server ~]$
[fg6ts30@off-line-server ~]$
[fg6ts30@off-line-server ~]$
```

目前建置了 CentOS 7，並使其可以使用 pip install 套件

■ 參考文獻

Google Cloud 建立服務帳戶：

<https://support.google.com/cloudidentity/answer/7378726?hl=zh-Hant>

第一次開Google VM就上手 – Compute Engine操作簡介：

<https://blog.cloud-ace.tw/compute-engine/start-a-google-vm-compute-engine/>

使用其他 SSH 用戶端登入到 Google Compute Engine 的 Linux VM：

<https://dotblogs.com.tw/supershowwei/2017/08/23/120313>

在 CentOS 7 上安裝 Python3

<https://kirin.idv.tw/python-install-python3-in-centos7/>

■ 程式架構(1/3) (使用 **Python** 製作雛形)

麥克風輸入音檔

是否連上網際網路

if (能連網) {

 upload

 Speech-to-Text

 Text-to-Speech

 儲存音檔

 下載 DNN 模型參數

 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

}

播放音檔

// speech_recognition.microphone() , 函式可使麥克風輸入

// 使用系統指令 “os.system()” 函數 ping 網域 , 來測試是否有網路 , 回傳 0 or 1

// SpeechRecognition 套件

// &

// gTTS 套件

// 使用 paramiko 套件使用 ssh 連線遠端 VM 複製檔案

// 寫 DVC 1.0 python 版本 , 需處理傳回來的參數

// 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)

麥克風輸入音檔

是否連上網際網路

if (能連網) {

Speech-to-Text

Text-to-Speech

儲存音檔

下載 DNN 模型

} else if (不能連網){

使用儲存的 DNN 1.0 模型執行儲存的參數

}

播放音檔

// Android 提供 MediaRecorder 達成手機收音、儲存、播音

// 加入權限，再使用內建提供函數判斷是否連網

// google 有提供 libraries

// 利用 ssh 遠端連線到 Google VM 複製檔案

// 寫 DVC 1.0 java 版本，需處理傳回來的參數