

AEye

「眼」明「手」快

# SPCC AI Project Group

Justin Tang

Matthew See

Angus Wong

Anthony Chan

Dione Cheung



溝通的能力，是否理所當然？

如果有一天，你再不能  
理解別人在說什麼.....你會？

# 真實案例



一個美國的聾啞媽媽  
Shae 的網上分享

# 疫情之前



特殊學校上課



透過手語溝通

# 疫情之後



口罩出現



感到迷惘、困惑

AEye

「眼」明「手」快

# 目錄

01 研究背景

02 項目方案

03 技術難點

04 技術詳情

05 未來方向和目標

# 01 研究背景

# 手語：不止於「手」



開心



非常開心

面部表情是手語的一部分

# 手語：不止於「手」



「風」

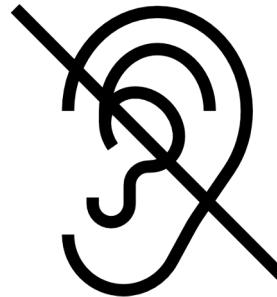
口型輔助溝通

# 背景



疫情：長期配戴口罩

# 背景



目標受眾：香港聾啞人士

# 背景



聾啞人士依賴手語溝通

# 受眾：聾啞人士

15.5萬

香港聽障人士人數

3900

香港手語使用人數

## 「新常態」下：



易被大眾忽略



口罩阻礙手語分辨



難以正常溝通

# 背景

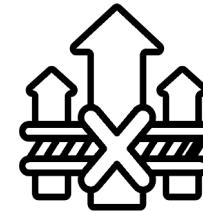


現存解決方法未能有效解決問題

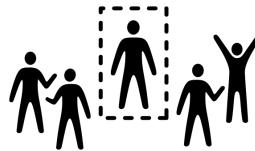
# 現存解決方法 —— ClearMask



價格昂貴：每個\$20 HKD



以塑膠作材質：影響呼吸



普及性：顯得**奇特**



保護性成疑

# 02 項目方案

# 項目總覽

目標



幫助香港聾啞人士  
在疫情下能順利以  
手語溝通

# 我們的觀察

我們做出不同口型  
的時候，

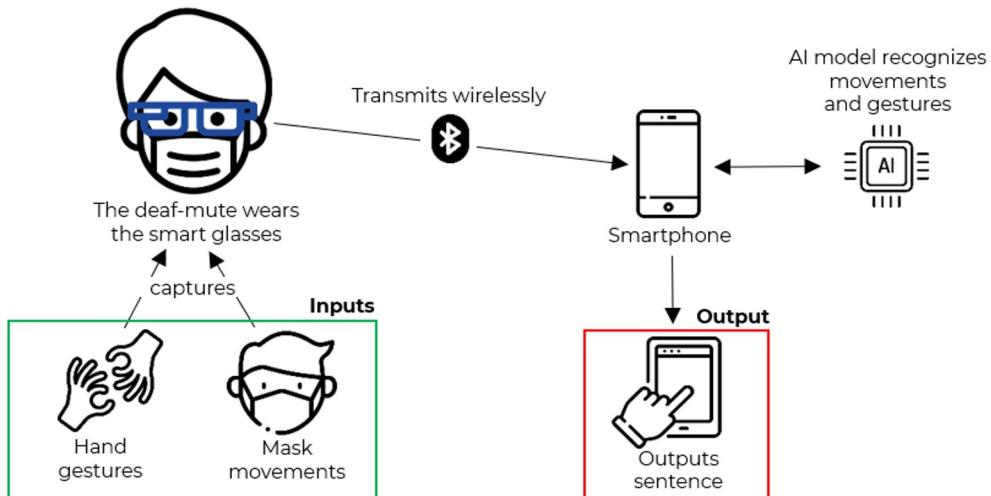
我們的口罩會產生  
不同程度的拉伸

—「戴」—「露」？

嘗試下：「你」與「我」



# 運作方式



透過鏡頭拍攝動作

將資料傳輸至手機

透過AEye程式辨別詞語

顯示翻譯結果

# 03 技術難點

# 技術難點

去年組別的設計

拍下手語動作

發到Jetson Nano

處理數據

發到手機

顯示結果

## 問題

識別速度慢

+

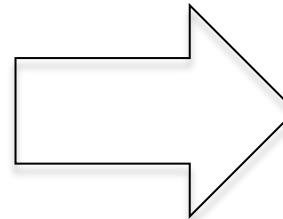
隨身攜帶一部  
Jetson Nano

## 目標

更簡單、順暢  
地辨認手語

## 難點

沒有香港手語的  
數據集



## 解決方案

自己製作數據集:

搜尋手勢 



錄製影片 

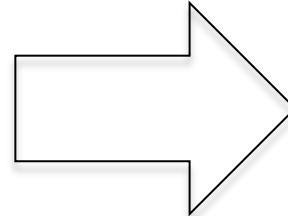
# 技術難點 ②

## 難點

現有翻譯器  
難以識別手勢：

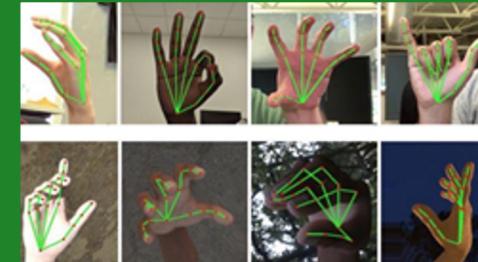
識別速度慢 

需要純色背景



## 解決方案

改變識別方式：  
把焦點放在全身動作



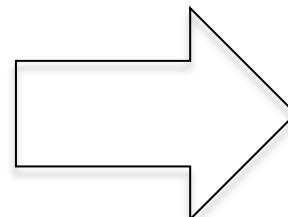
# 技術難點 ③

## 難點

沒有任何追蹤  
口罩的程式

## 解決方案

自己創造程序



# 04 技術詳情

# 數據集

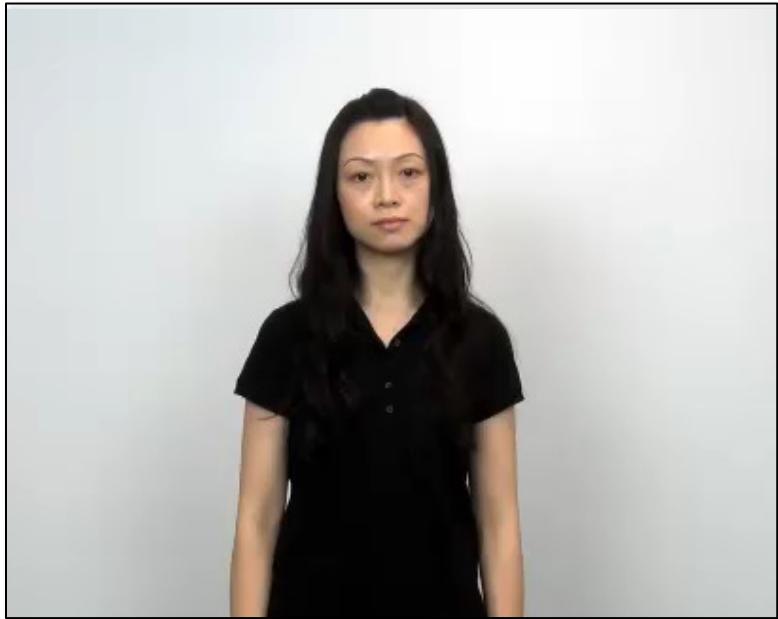
	詞語1	詞語2
A組	警員	警察
B組	用	拿
C組	善良	心跳
D組	計較	計算
E組	心急	緊張

5組相同手勢、需要依靠口型分辨

# A 組



警員

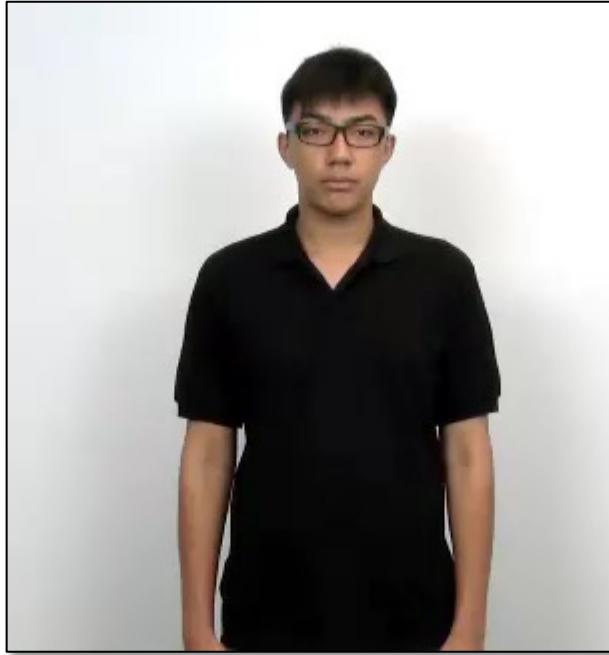


警察

# B 組

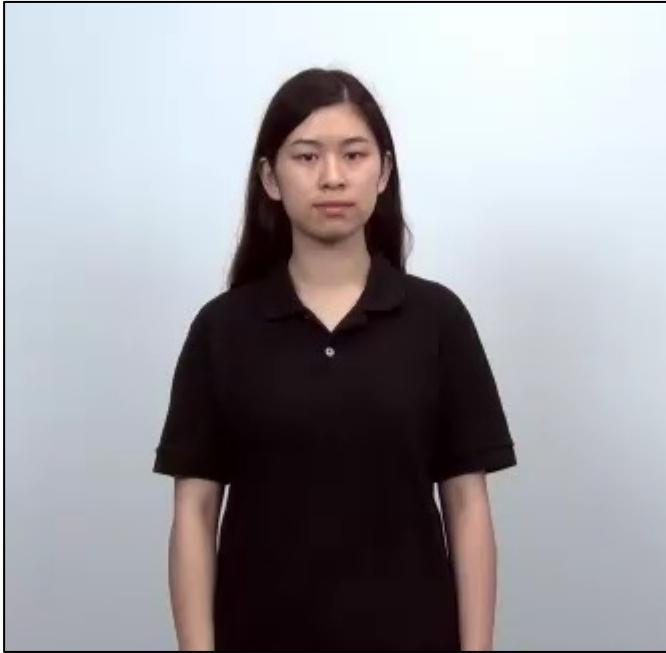


用



拿

# C 組

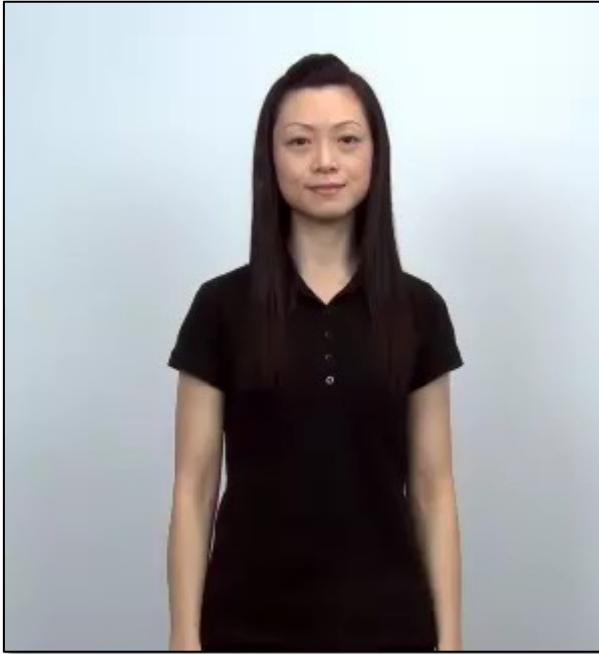


善良

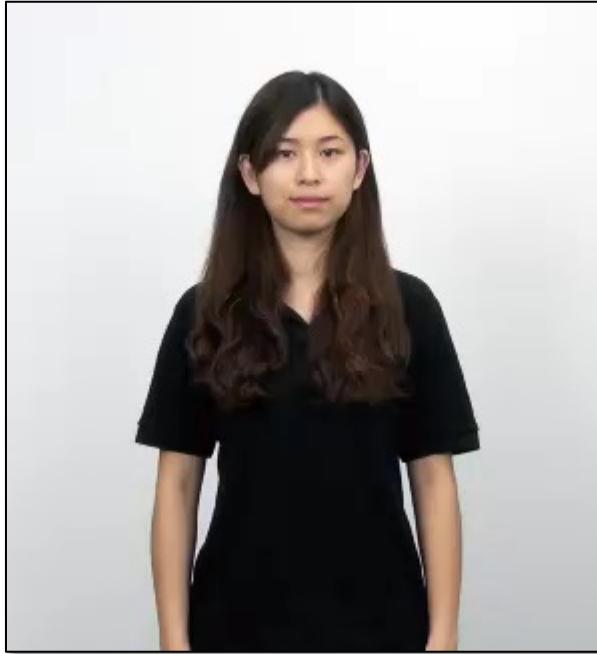


心跳

# D 組



計較

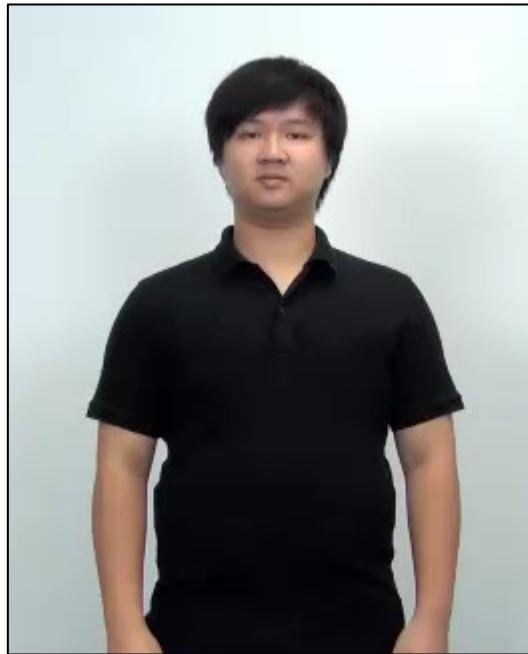


計算

# E 組



心急



緊張

# 運作

Google  
MediaPipe

+

LSTM  
模型

“However, most mobile phone operating systems **do not support 3D convolution** (which is a crucial part of the deep learning model), making **an app container** is needed for the deep learning model. Existing phones in the market are then either **not powerful enough or not fast enough.**” (Zhenxing and Yisiang et. al, 2020)





## Live ML anywhere

[MediaPipe](#) offers cross-platform, customizable ML solutions for live and streaming media.



**End-to-End acceleration:** Built-in fast ML inference and processing accelerated even on common hardware



**Build once, deploy anywhere:** Unified solution works across Android, iOS, desktop/cloud, web and IoT



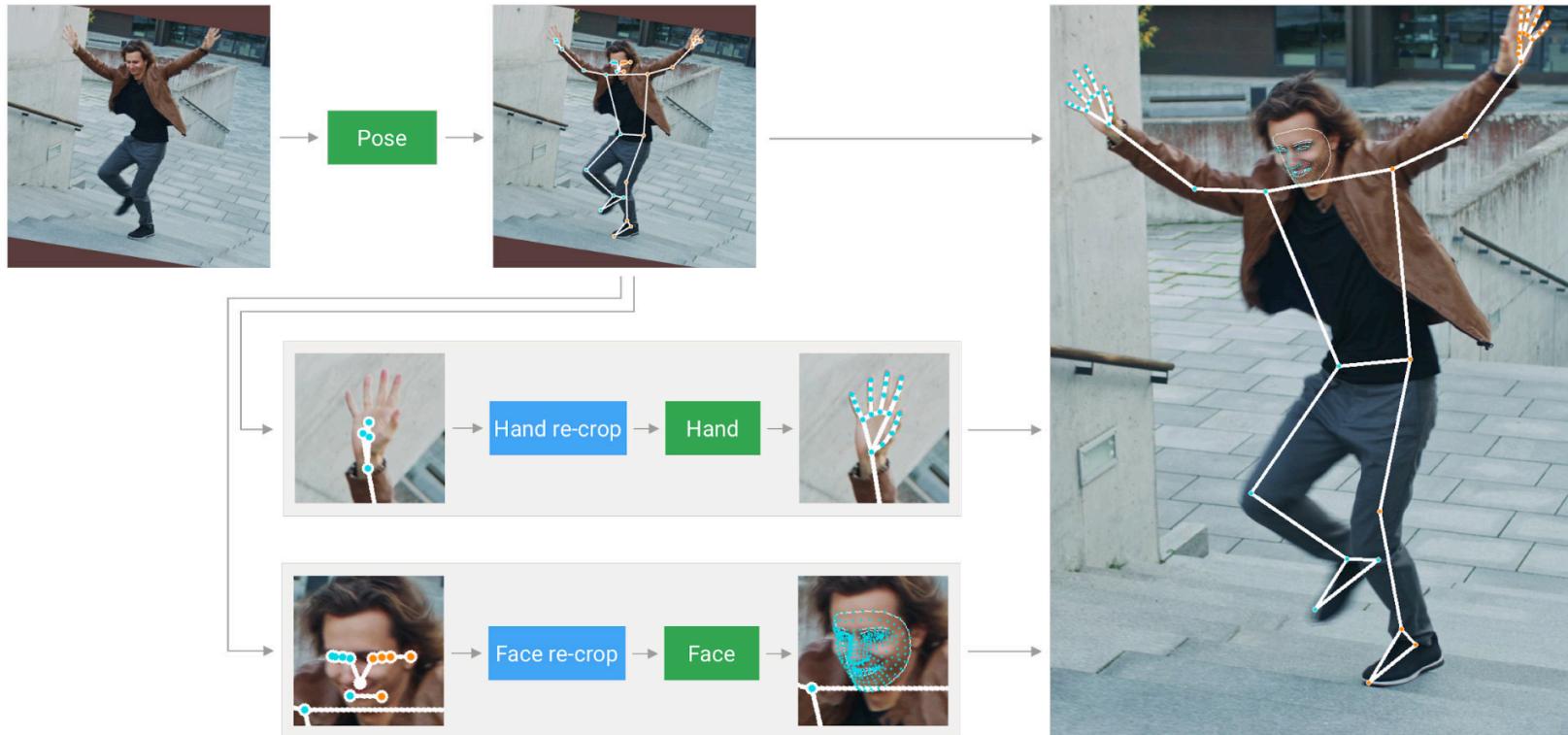
**Ready-to-use solutions:** Cutting-edge ML solutions demonstrating full power of the framework



**Free and open source:** Framework and solutions both under Apache 2.0, fully extensible and customizable



# MediaPipe Holistic



**468**

面部節點

**33**

姿勢節點

**21**

手部節點

(x,y,z) 坐標  
共1662 個



# 數據采集



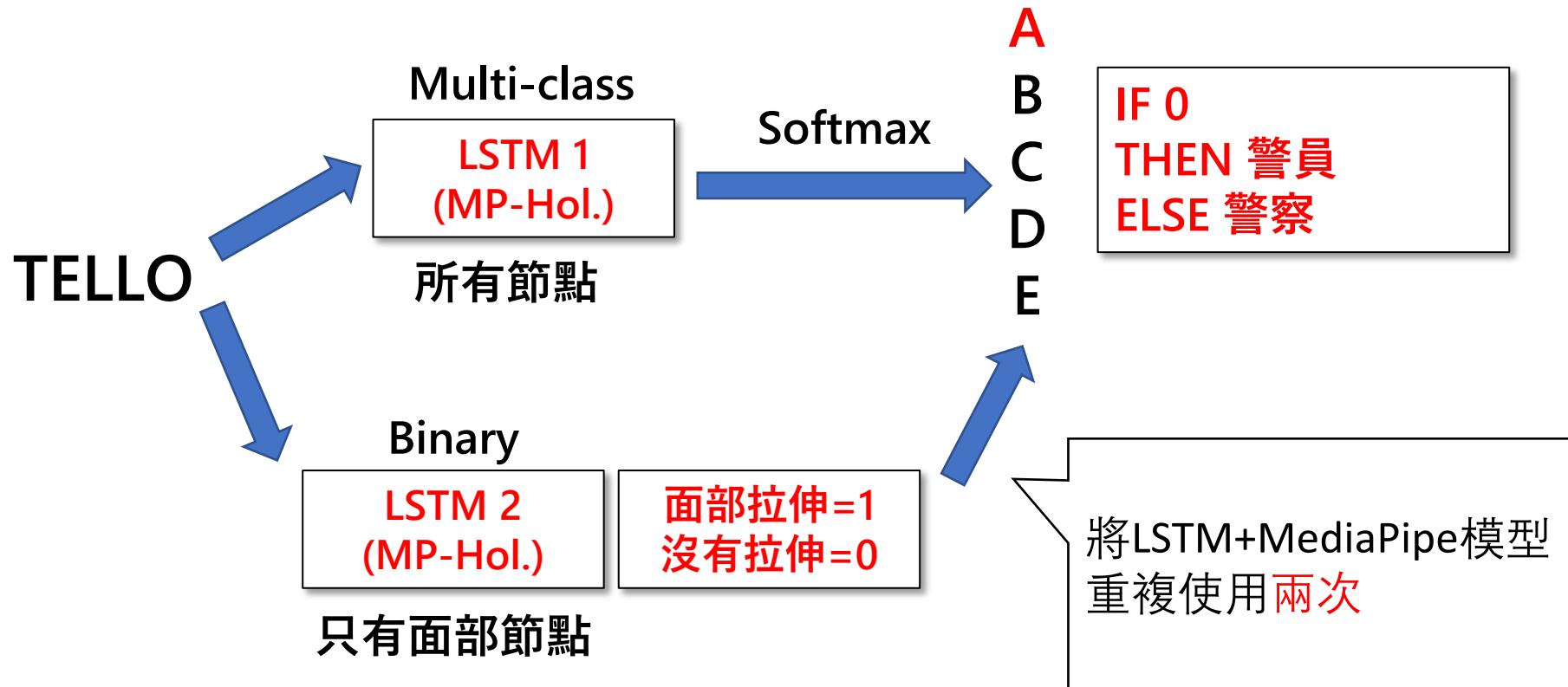


DJI TELLO  
圖傳系統

720p 鏡頭模組

太陽眼鏡

# 創 新 模 型



# DEMO

# LSTM 好處



需要較少數據



訓練速度快



即時識別

```
Epoch 1994/2000  
5/5 [=====] - 0s 25ms/step  
Epoch 1995/2000  
5/5 [=====] - 0s 25ms/step  
Epoch 1996/2000  
5/5 [=====] - 0s 24ms/step  
Epoch 1997/2000  
5/5 [=====] - 0s 24ms/step  
Epoch 1998/2000  
5/5 [=====] - 0s 25ms/step  
Epoch 1999/2000  
5/5 [=====] - 0s 25ms/step  
Epoch 2000/2000  
5/5 [=====] - 0s 24ms/step
```

Jupyter Notebook, Intel i7 10<sup>th</sup> Gen  
No GPU

莫拉

識別拉伸？

Model: "sequential\_23"

Layer (type)	Output Shape	Param #
lstm_69 (LSTM)	(None, 30, 64)	442112
lstm_70 (LSTM)	(None, 30, 128)	98816
lstm_71 (LSTM)	(None, 64)	49408
dense_68 (Dense)	(None, 10)	4160
dense_69 (Dense)	(None, 10)	2080
dense_70 (Dense)	(None, 1)	165

Total params: 596,741

Trainable params: 596,741

Non-trainable params: 0

Acc = 0.4-0.6 (隨機)

Reduce units from 64 to 32  
Detect less complex features

Model: "sequential\_23"

Layer (type)	Output Shape
--------------	--------------

lstm_69 (LSTM)	(None, 30, 64)
----------------	----------------

lstm_70 (LSTM)	(None, 30, 128)
----------------	-----------------

lstm_71 (LSTM)	(None, 64)
----------------	------------

dense_68 (Dense)	(None, 64)
------------------	------------

dense_69 (Dense)	(None, 32)
------------------	------------

dense_70 (Dense)	(None, 5)
------------------	-----------

Total params: 596,741

Trainable params: 596,741

Non-trainable params: 0



Model: "sequential\_4"

Layer (type)	Output Shape	Param #
--------------	--------------	---------

lstm_16 (LSTM)	(None, 30, 32)	183936
----------------	----------------	--------

lstm_17 (LSTM)	(None, 32)	8320
----------------	------------	------

dense_12 (Dense)	(None, 32)	1056
------------------	------------	------

dense_13 (Dense)	(None, 2)	66
------------------	-----------	----

Total params: 193,378

Trainable params: 193,378

Non-trainable params: 0

Acc = 0.97

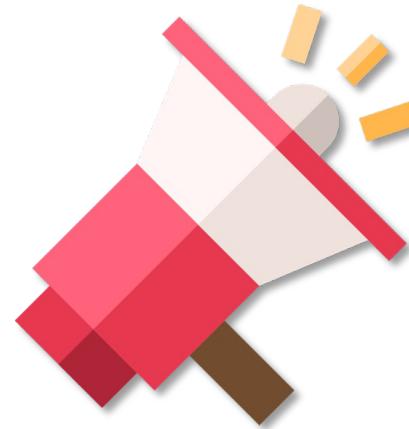
# 05 未來方向及目標

影響



讓聾啞人士如常溝通

影響



向大眾推廣手語

# 影響



提高對聾啞人士的關注

# 未來發展目標

本地推廣

增加詞庫

多類口罩分類

翻譯成句子

融入多國語言

Thank you!

AEye 「眼」明「手」快