

# 音樂曲風辨識

第七組

組員:

林耕澤 N26120870

羅文璟 N26120820

吳炯霖 N26120799

謝亞城 N26120309



# 目錄

1. 主要流程
2. 其他討論
3. 遇到的困難

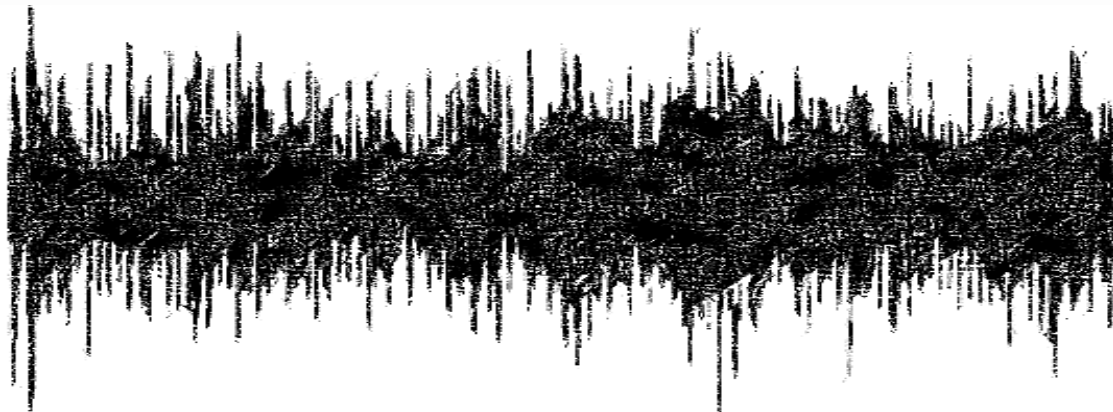
1

# 主要流程

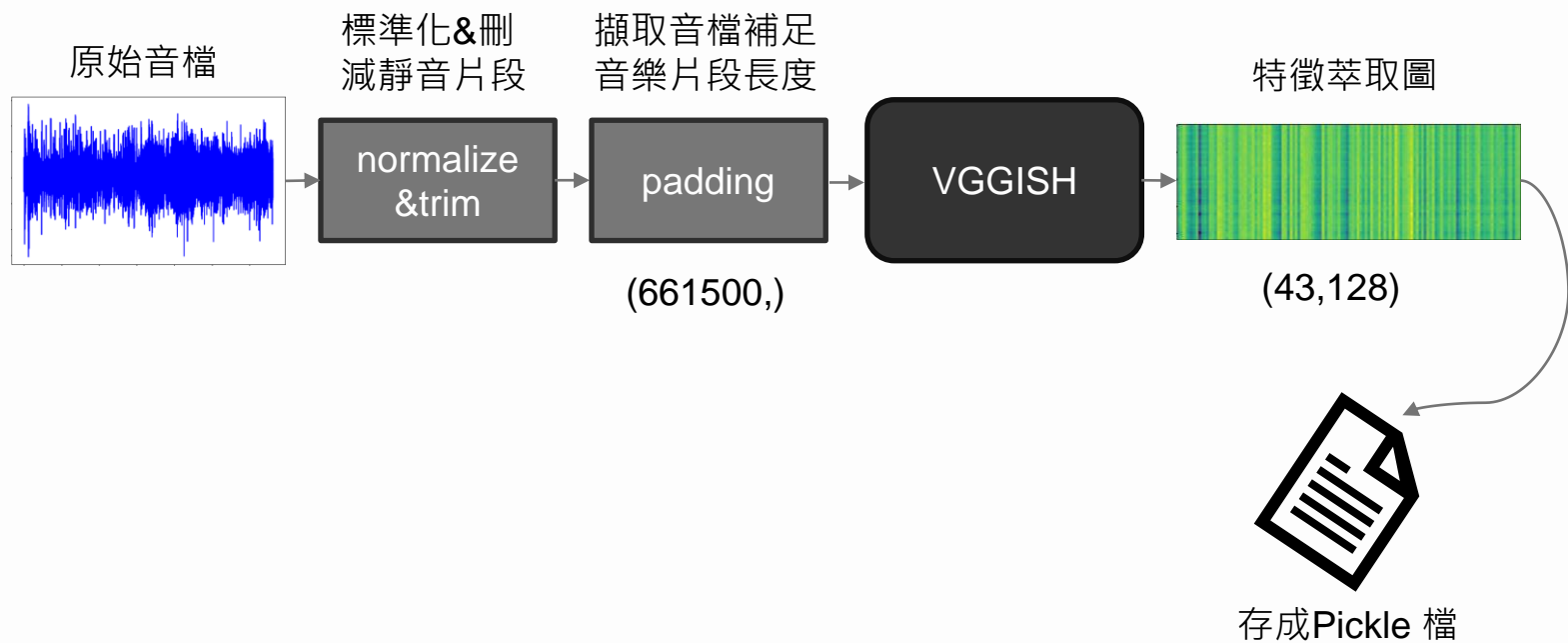
資料集、特徵萃取資料前處理、模型選擇、參數設定、模型表現集數據討論

# 資料集: GTZAN

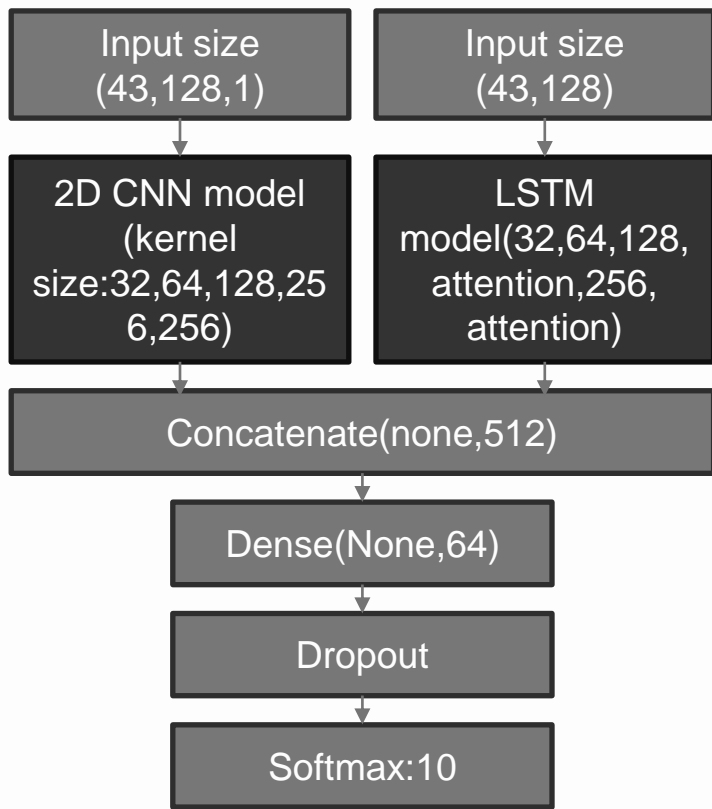
- 1000 個30秒片段音訊片段
- 10種曲風，每一種曲風100音訊片段
- 採樣頻率22050Hz
- WAV檔
- 種類包括:blue、classical、country、disco、hiphop、jazz、metal、pop、reggae、rock



# 特徵萃取以及資料前處理:



# Model: CNN\_LSTM\_A



## 其他嘗試

預訓練模型	非預訓練模型
<ul style="list-style-type: none"><li>• ResNet152V2</li><li>• Xception</li><li>• InceptionResNetV2</li></ul>	<ul style="list-style-type: none"><li>• Only CNN</li><li>• Only LSTM</li><li>• LSTM Autoencoder</li><li>• CNN Autoencoder</li></ul>

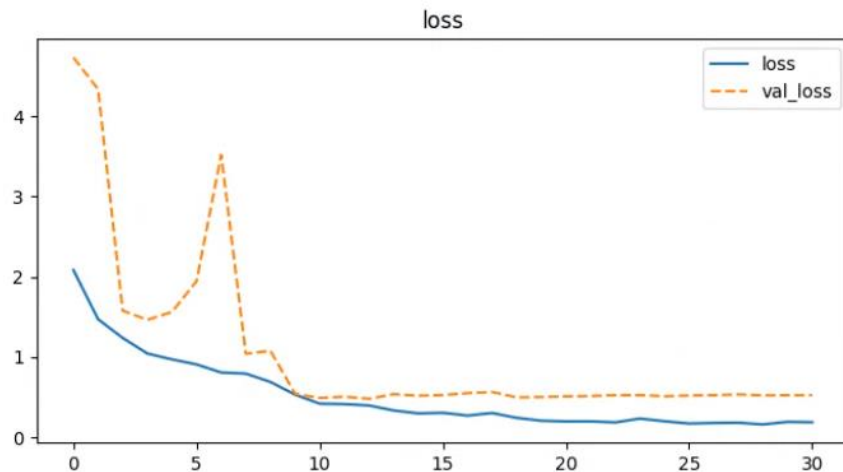
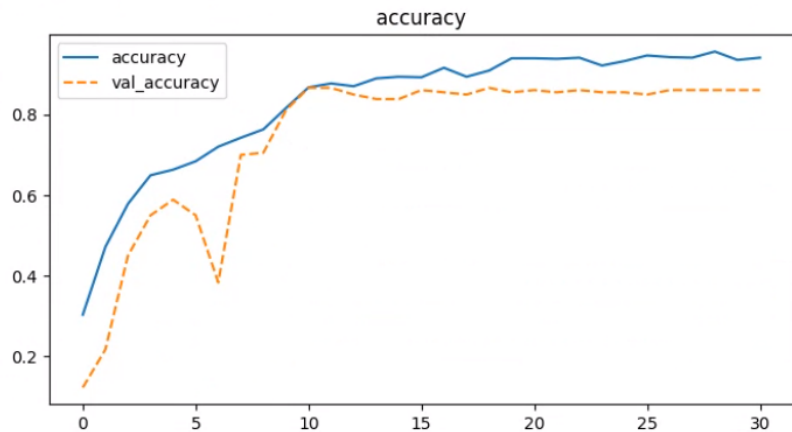
# 參數設定

- Batch Size: 8
- Initial lr: 0.01
- Epoch: 100
- Decay Epoch: 10
- Decay Factor: 0.1
- Loss Function: Categorical Cross-Entropy
- Optimizer: Adam

# 模型表現: 給我新的圖 要多印val的confusion

train accuracy: 96%

val accuracy: 86%

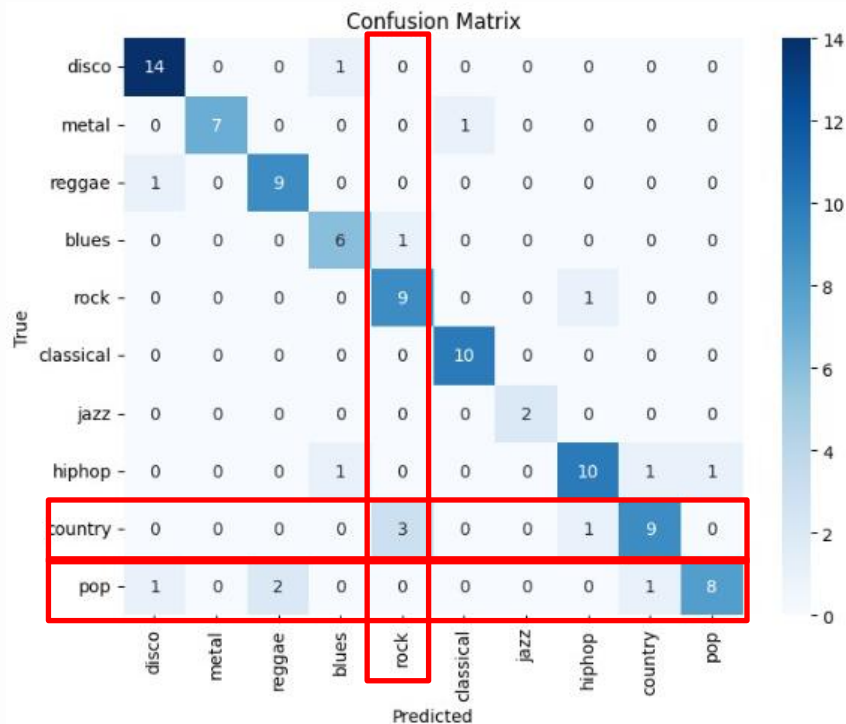




# Confusion Matrix

## Classification Report:

	precision	recall	f1-score	support
0	0.88	0.93	0.90	15
1	1.00	0.88	0.93	8
2	0.82	0.90	0.86	10
3	0.75	0.86	0.80	7
4	0.69	0.90	0.78	10
5	0.91	1.00	0.95	10
6	1.00	1.00	1.00	2
7	0.83	0.77	0.80	13
8	0.82	0.69	0.75	13
9	0.89	0.67	0.76	12
accuracy			0.84	100
macro avg	0.86	0.86	0.85	100
weighted avg	0.85	0.84	0.84	100



# 模型比較

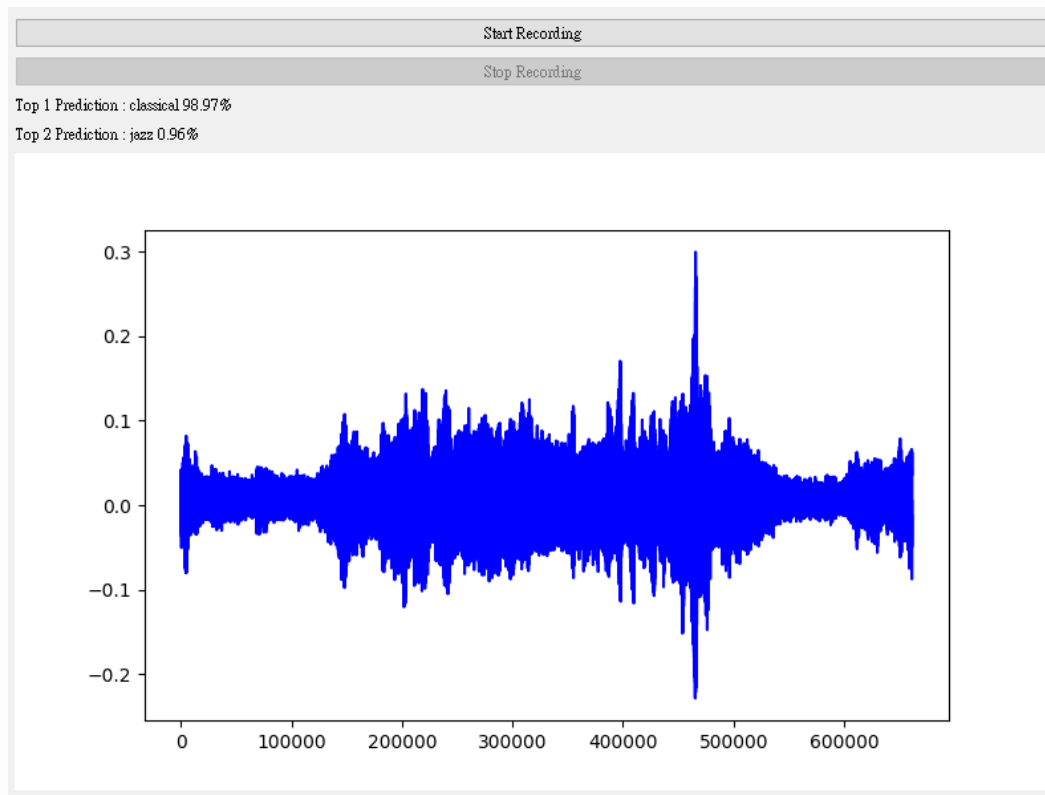
	CNN_LSTM_A	Only CNN	Only LSTM	Resnet152V2
val_accuracy	86%	84%	66%	36%
top2_accuracy	94%	92%	81%	53%
test_accuracy	84%	81%	61%	33%

推測原因:

(43,128,3)數據對大型pretrained 來說太小。

# Demo圖

48000Hz => 22050Hz



2

# 其他討論

其他在開發過程中的發現

# 針對採樣頻率探討

## CNN\_LSTM\_A為例

訓練資料採樣頻率	模型測試集準確率
22050Hz	85%
48000Hz	48%

### 推測原因:

1. 資料集不足(用Autoencoder、pretrained model測試效果稍好但一樣不佳)。
2. 因為GTZAN取樣頻率本身為22050Hz，48000Hz萃取會取出沒有意義的資料，使得模型無法訓練好。

# 針對資料集探討: Youtube Playlist

使用 youtube 爬取的資料集，在測試集準確度76%，但實際測試時效果與模型驗證時的76%相差甚遠。

## 推測原因: 資料集分布偏差

1. Youtube 歌單中的歌曲有些混雜(Blues歌單中不一定是Blues)。
2. 無從得知原始音訊採樣頻率(可能會因年代不同而有所不同)。
3. 一首歌曲可能有很多種風格融合。

## 為何不自己蒐資料集?

我們很多歌曲也分不出來是什麼風格。

# 遇到的困難

好可惡的arm64

# 遇到的問題: Docker部署

Docker打包不成功，因為Tensorflow-intel在pip 上沒有支援Linux版本

```
115.0 Collecting tensorflow-hub==0.16.1
115.0   Downloading tensorflow_hub-0.16.1-py2.py3-none-any.whl (30 kB)
115.3 ERROR: Ignored the following versions that require a different python version: 0.36.0 Requires-Python >=3.6,<3.10; 0.37.0 Requires-Python >=3.7,<3.10; 0.52.0 Requires-Python >=3.6,<3.9; 0.52.0rc3 Requires-Python >=3.6,<3.9; 0.53.0 Requires-Python >=3.6,<3.10; 0.53.0rc1 Requires-Python >=3.6,<3.10; 0.53.0rc2 Requires-Python >=3.6,<3.10; 0.53.0rc3 Requires-Python >=3.6,<3.10; 0.53.1 Requires-Python >=3.6,<3.10; 0.54.0 Requires-Python >=3.7,<3.10; 0.54.0rc2 Requires-Python >=3.7,<3.10; 0.54.0rc3 Requires-Python >=3.7,<3.10; 0.54.1 Requires-Python >=3.7,<3.10; 1.6.2 Requires-Python >=3.7,<3.10; 1.6.3 Requires-Python >=3.7,<3.10; 1.7.0 Requires-Python >=3.7,<3.10; 1.7.1 Requires-Python >=3.7,<3.10
115.3 ERROR: Could not find a version that satisfies the requirement tensorflow-intel==2.16.1 (from versions: 0.0.1)
115.3 ERROR: No matching distribution found for tensorflow-intel==2.16.1
115.3
```

將Docker包成amd64的image與Xavier的arm64環境不合，無法使用

```
cils0004@cils0004-desktop:/dev$ sudo docker run --rm -v /home/cils0004/HW1:/home/ubuntu/content -v /tmp/.X11-unix:/tmp/.X11-unix -e DISPLAY=$DISPLAY -e GDK_SCALE -e GDK_DPI_SCALE -v /dev/snd:/dev/snd --device /dev/snd/by-id --privileged=true -it lowen0909/tensorflow2_10
WARNING: The requested image's platform (linux/amd64) does not match the detected host platform (linux/arm64/v8) and no specific platform was requested
exec /usr/local/bin/python: exec format error
cils0004@cils0004-desktop:/dev$
```



# 遇到的問題: Docker部署

## 解決方法

使用助教過去的image為基底，往上加我們需要的套件。

```
# 下載 Docker Image  
sudo docker pull q36111265/music_emotion
```

# 遇到的問題: PyAudio ALSA 問題

Pyaudio是用來錄音的python模組，在Windows上能夠正常運作，不過到了Xavier上時卻會出現問題。

**無法辨識出羅技麥克風:**

如下錯誤顯示: (ALSA: **Advanced Linux Sound Architecture**)這些錯誤表明ALSA無法找到或評估特定 PCM (脈衝編碼調制) 設備的某些定義或配置。

```
ALSA lib pcm_dsnoop.c:641:(snd_pcm_dsnoop_open) unable to open slave
ALSA lib pcm_dmix.c:1089:(snd_pcm_dmix_open) unable to open slave
ALSA lib pcm.c:2642:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.rear
ALSA lib pcm.c:2642:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.center_lfe
ALSA lib pcm.c:2642:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.side
ALSA lib pcm.c:2642:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.hdmi
ALSA lib pcm.c:2642:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.hdmi
ALSA lib pcm.c:2642:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.modem
ALSA lib pcm.c:2642:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.modem
ALSA lib pcm.c:2642:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.phoneline
ALSA lib pcm.c:2642:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.phoneline
ALSA lib pcm_dsnoop.c:641:(snd_pcm_dsnoop_open) unable to open slave
ALSA lib pcm_dmix.c:1089:(snd_pcm_dmix_open) unable to open slave
ALSA lib pcm_dmix.c:1089:(snd_pcm_dmix_open) unable to open slave
```

# 遇到的問題: PyAudio ALSA 問題

## 嘗試解決辦法

1. 使用PyAudio列出所有音訊設備裝置，沒有列出webcam c270。
2. 確認在container中的dev/snd中可以存取到webcam usb裝置，但PyAudio stream 錄音的預設裝置使種沒有認出。
3. 查看alsa.config檔。

# 工作分配

訓練模型: 林耕澤、羅文璟

介面: 謝亞城、吳炯霖

Debug 針對部屬問題: 全部人

部署: 謝亞城、吳炯霖

PPT: 林耕澤

Thank You For Listening