

工研院-AI資料科學 期末專題報告

組別:第七組

團隊產品:Youtuber人臉辨識

團隊名稱:You是誰

團隊組員

01林采萱、04李佩瑾、16洪笙觀、26范紀予、37羅茗韋

報告日期

2022/08/02

Contents

01

開發動機

九天玄女的煩惱

02

資料前處理

資料蒐集與清理

03

模型部署

Yolov5s

04

Demo

Video

組員名稱



01林采萱

模型建立
爬蟲
簡報



37羅茗韋

網站前端
爬蟲



04李佩瑾

模型建立爬蟲



16洪笙觀

模型建立
爬蟲



26范紀予

模型建立
簡報
網站後端

團隊產品摘要

開發動機

Youtuber人臉辨識



開發動機-使用者遇到的問題

01 背景、個性

- 忙於通靈占卜，以此養家的媽媽
- 老花眼，不擅使用電子產品
- 但他的孩子卻沈迷於Youtube 頻道。

算命師

單親媽媽



九天玄女
唯一指定

一位7歲童
有代溝

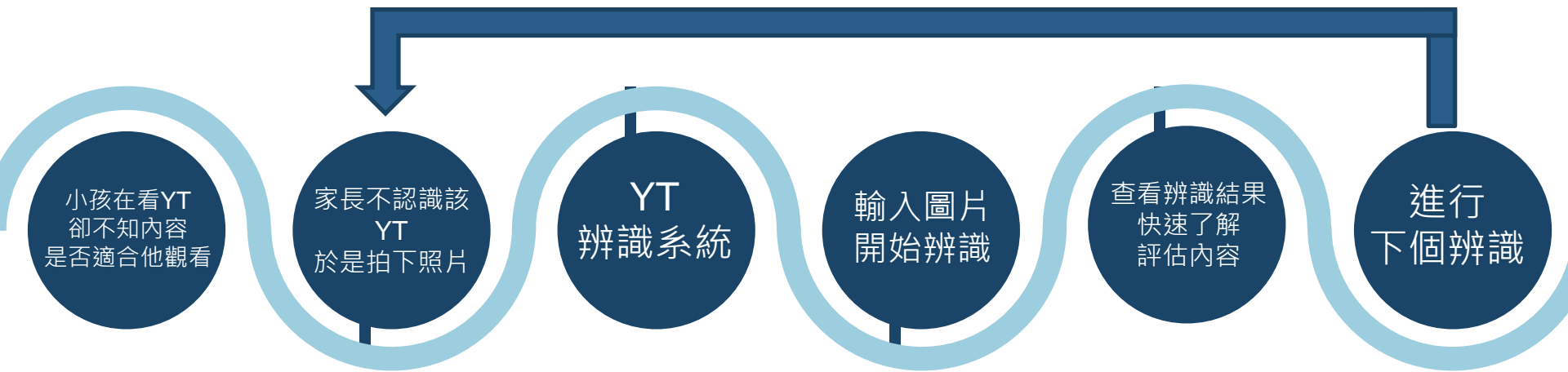
02 問題

- Youtuber越來越多，
- 孩子也是Youtube重度使用患者，
- 親子關係漸行漸遠，想了解他們在看甚麼，
- 卻不知這Youtuber是誰也無從查起。

03 END GOAL

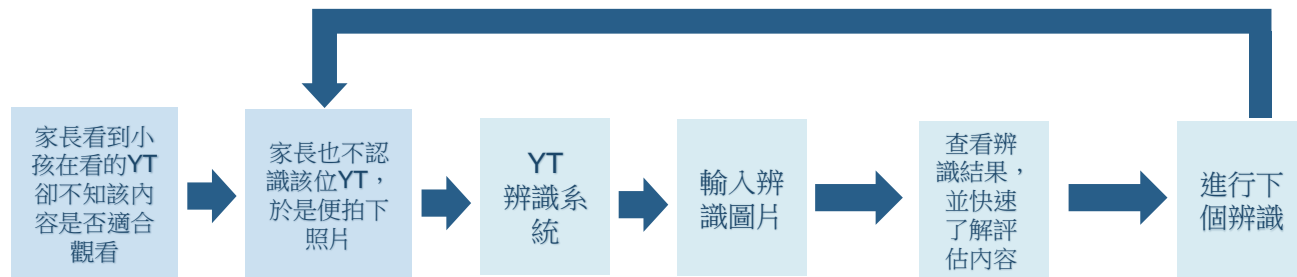
- 簡單快速的了解兒子平常在看的youtuber，藉此拉近親子距離
- 節省搜尋時間，在工作與親子間取得生活平衡。

使用流程



服務命名：

使用者服務旅程



前台服務或系統

➤ 介面操作

使用網站上面的簡易介面來進行上傳圖片等操作

後台服務或系統

➤ 偵測系統

有利用AI模型來進行圖片人臉辨識

➤ 搜尋系統

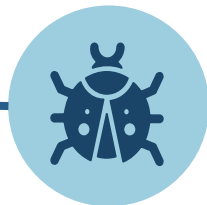
用戶搜尋之後，透過搜尋系統把AI模型辨識的結果，顯示給使用者



資料蒐集與處理

資料蒐集與處理

爬蟲



Web Crawl

因爬蟲後圖片品質不一，
最後選擇人工搜圖

人工找圖



IG+FB

搜圖同時也過濾掉不適合進行訓練，
背景也比較乾淨的圖片

資料標記



Labeling

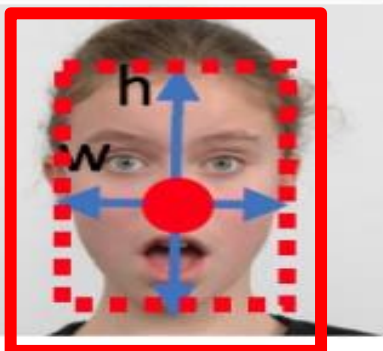
採人工標記/ code(自動標記)



Deployment



2D會比較好 for Yolo



Box:



Pose:



5 Landmarks:



Landmarks:



Mask:



3D mesh(Ours):



3D mesh:

More Informative



模型比較問題

模型使用



Nano
YOLOv5n

4 MB_{FP16}
6.3 ms_{V100}
28.4 mAP_{COCO}



Small
YOLOv5s

14 MB_{FP16}
6.4 ms_{V100}
37.2 mAP_{COCO}



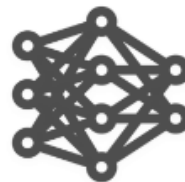
Medium
YOLOv5m

41 MB_{FP16}
8.2 ms_{V100}
45.2 mAP_{COCO}



Large
YOLOv5l

89 MB_{FP16}
10.1 ms_{V100}
48.8 mAP_{COCO}



XLarge
YOLOv5x

166 MB_{FP16}
12.1 ms_{V100}
50.7 mAP_{COCO}

Key Numbers



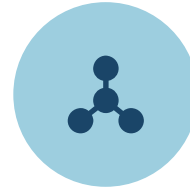
1500

photos



50

Youtubers



83.2

mAP



Key Numbers



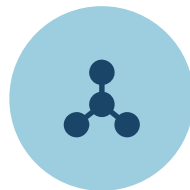
80%

Training Set



13%

Validation Set



7%

Testing Set



RESULT

custom_YOLOv5s summary: 232 layers, 7384065 parameters, 0 gradients, 17.2 GFLOPs

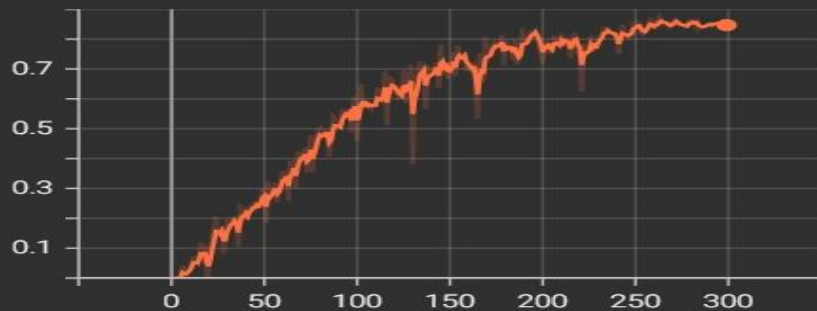
Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95: 100%
all	200	205	0.718	0.767	0.832	0.686



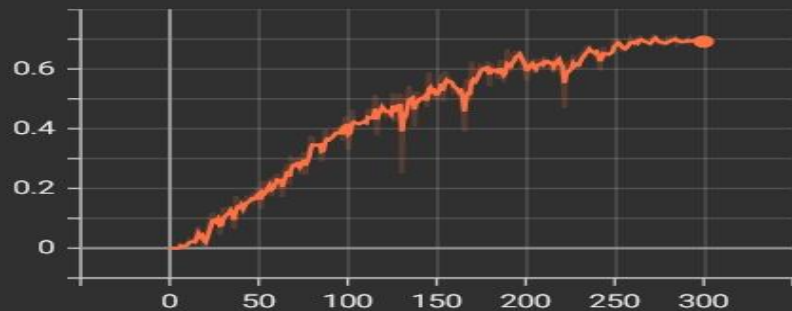
RESULT

metrics

metrics/mAP_0.5
tag: metrics/mAP_0.5



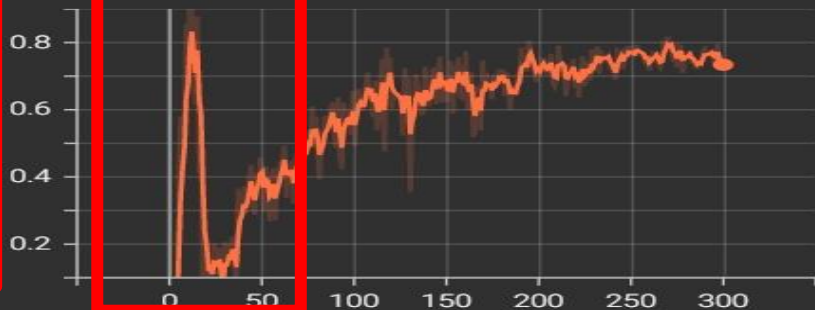
metrics/mAP_0.5:0.95
tag: metrics/mAP_0.5:0.95



metrics/precision
tag: metrics/precision



metrics/recall
tag: metrics/recall



實驗

1. 我們只有時間了解準確率問題
2. 但是Recall rate 很高的地方也是我們可以深入研究的
3. 交叉質(Intersection over union)

交疊比例並無顯著差異，所以這邊不探討

真正例: $TP = TruePositive$

真反例: $TN = TrueNegative$

假正例: $FP = FalsePositive$

假反例: $FN = FalseNegative$

則，查准率和查全率计算公式：

查准率: $Precision = \frac{TP}{TP+FP}$

查全率: $Recall = \frac{TP}{TP+FN}$

模型使用



Nano

YOLOv5n

4 MB_{FP16}
6.3 ms_{V100}
28.4 mAP_{COCO}



Small

YOLOv5s

14 MB_{FP16}
6.4 ms_{V100}
37.2 mAP_{COCO}



Medium

YOLOv5m

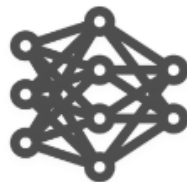
41 MB_{FP16}
8.2 ms_{V100}
45.2 mAP_{COCO}



Large

YOLOv5l

89 MB_{FP16}
10.1 ms_{V100}
48.8 mAP_{COCO}



XLarge

YOLOv5x

166 MB_{FP16}
12.1 ms_{V100}
50.7 mAP_{COCO}

Key Numbers



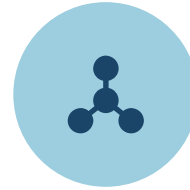
1500

photos



25

Youtubers



87.4

mAP



Key Numbers



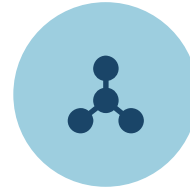
80%

Training Set



13%

Validation Set



7%

Testing Set




```
CPU times: user 27.8 s, sys: 3.27 s, total: 31 s
```

3	0.874	0.713
3	0.898	0.812
6	0.543	0.468
1	0.83	0.626
5	0.624	0.474
6	0.721	0.572
1	0.995	0.896
4	0.591	0.503
1	0.995	0.821
1	0.995	0.877
1	0.793	0.687
8	0.788	0.687
9	0.912	0.771
1	0.995	0.709
1	0.995	0.895
1	0.995	0.796
1	0.995	0.703
5	0.541	0.443
3	0.972	0.892
1	0.995	0.611
1	0.995	0.697
2	0.83	0.747
5	0.995	0.64
1	0.995	0.995
5	0.825	0.664
49	0.995	0.758
1	0.913	0.78

最後選擇modelS的原因



1500

photos



可辨識廣

Youtubers多才比較符合
我們所需要的情況



無顯著差異

mAP只有差距**4%**

部署

部署講解



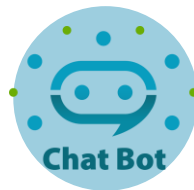
Heroku

Web-Server



網站的方式

現有版本



ChatBot

未來版本





Demo

結論

結論

1. 2D is better for yolo :

因為人臉是屬於3D 所以在辨識細節會比較困難(Transfer 會有難度)

2. 部署不容易:

因為Heroku 會有大小上限制所以在部屬上還需要調整，未來部屬到其他平台不會有這個問題。

3. 資料收集困難:

希望在未來可以解決這個問題



未來展往

- 1. State of the art :

所以我們嘗試使用yolo7模型，但發現模型過於巨大，GPU 容易不足

- 2. More:

優質的圖篇增加辨識準確率

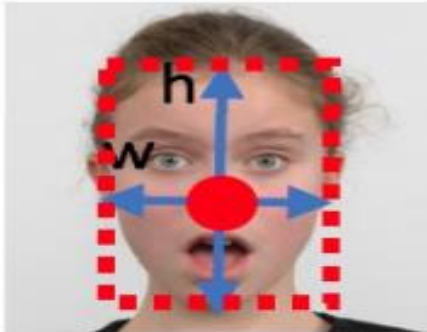
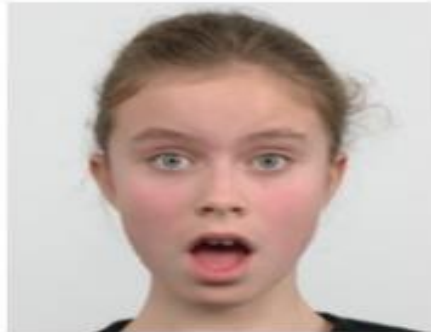
增加可參考的東西(增列新聞)

- 3. 自動化:

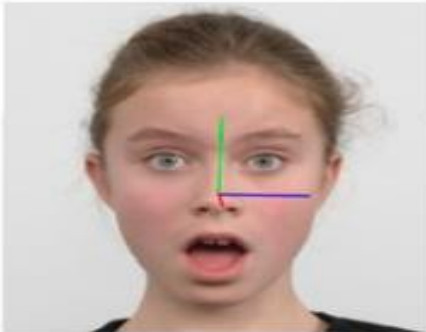
自動化標記(RetinaFace)



RetinaFace



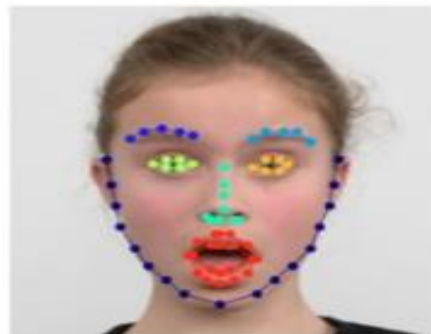
Box:
4 scalars



Pose:
7 scalars



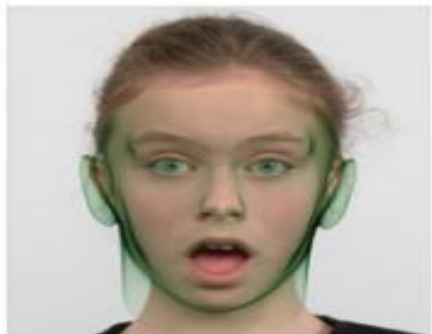
5 Landmarks:
10 scalars



Mask:
H x W matrix



3D mesh(Ours):
3 x 1k vertices

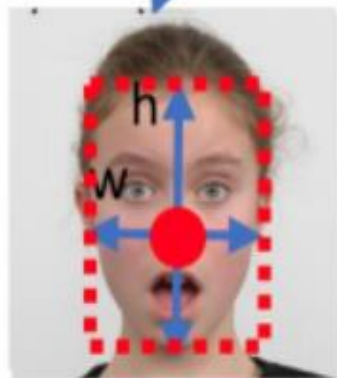


3D mesh:
3 x 53k vertices

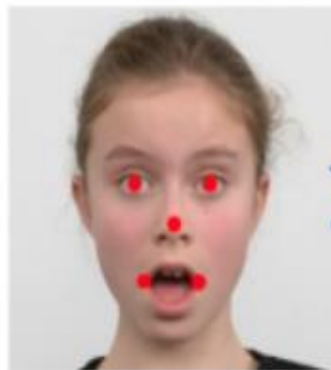
68 Landmarks:
136 scalars

More Informative

(1) More semantic points, more accurate box prediction



Face Detection
(one center point)



2D Face Alignment
(five points)

1k 3D points enhance
pose-invariant 5 points

Cheap 5 points enhance
robust 1k points



3D Face Reconstruction
(1k points)

(2) More challenging training scenario, more robust point prediction



+



=



68 vertices ([x, y, z]*68)

+

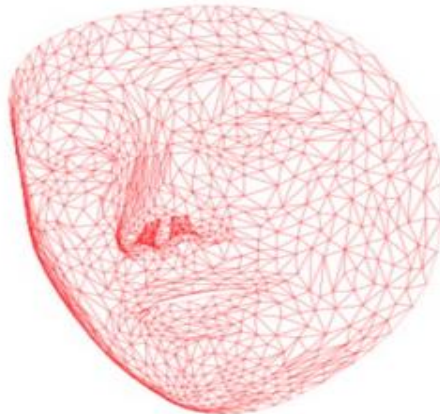
111 triangles (template)

=

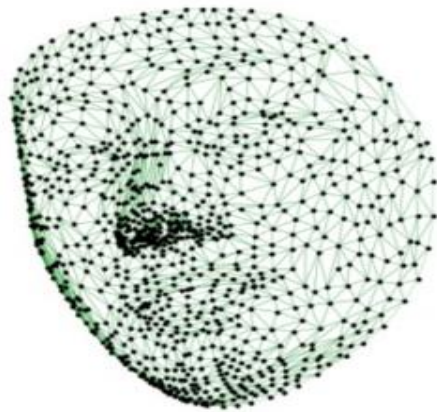
Mesh68



+



=



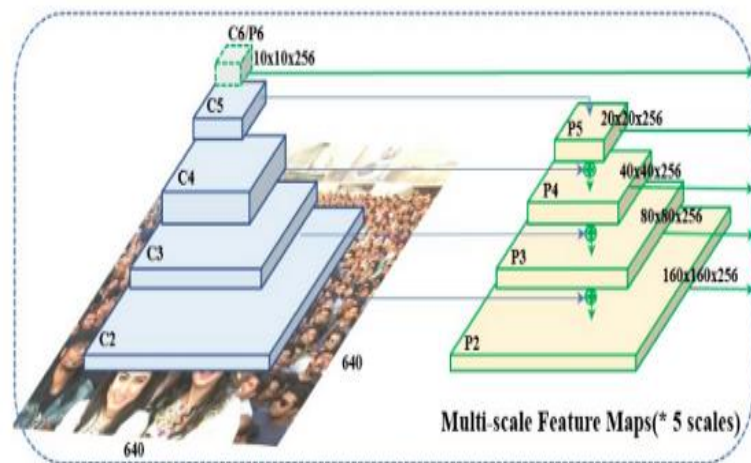
1035 vertices ([x, y, z]*1035)

+

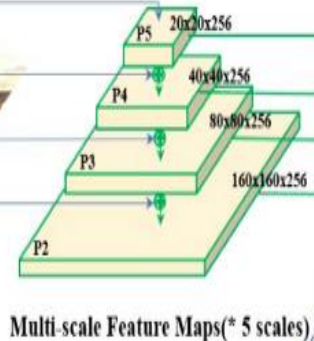
1999 triangles (template)

=

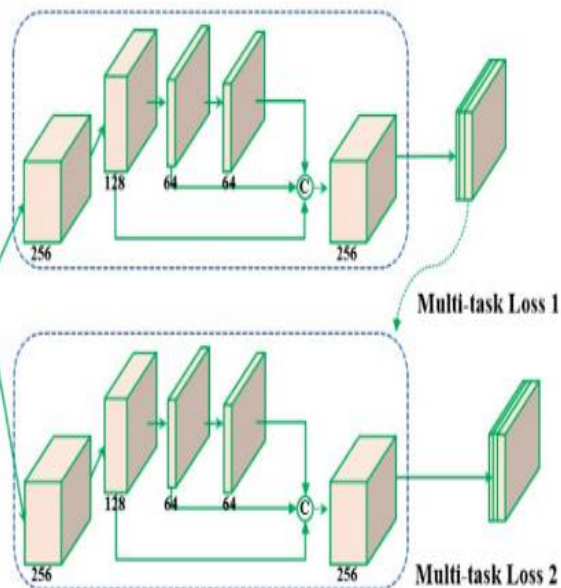
Mesh1k



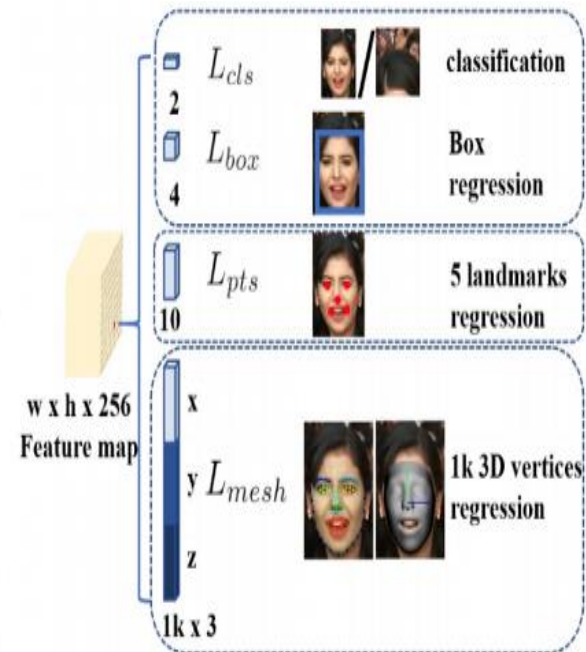
Feature Pyramid Network



(a) Network Structure



Cascade Multi-scale Context Head Module (* 5 scales * 2 cascade)



(b) Multi-task Loss

甘特圖

項目	階段	階段一、發想期							階段二、設計期		階段三、實踐期	
	月	七							七		八	
	日期	8		15				22			25	2
進度1、產品計畫												
1.1	資料蒐集											
1.2	資料處理											
進度2、產品設計研發												
2.1	模型訓練											
2.2	網頁部屬											
進度3、產品檢驗測試												
3.1	整體測試完成											
進度4、產品雛型製作												
4.1	產品修正											
4.2	產品原型											
進度百分比		0		20		40		60		80		100

Reference

<https://www.instagram.com/87acup/>
<https://www.instagram.com/alanchannel/>
<https://www.instagram.com/goodalicia/>
<https://www.instagram.com/brianptseng/>
https://www.instagram.com/ggu_kim/
https://www.instagram.com/thechef_fred/
<https://www.instagram.com/helloiamhook/>
<https://www.instagram.com/huang45490/>
<https://www.instagram.com/huangbrotherss/>
https://www.instagram.com/chef_jamestw/
<https://www.instagram.com/chillseph/>
https://www.instagram.com/kart_wang/
<https://www.instagram.com/liketai/ai/>
<https://www.instagram.com/malianjie0902/>
https://www.instagram.com/sanyuan_japan/
<https://www.instagram.com/pierreliupeil/>
<https://www.instagram.com/jesstang11/>
<https://www.instagram.com/youshengwang/>
https://www.instagram.com/big_star_ken/?hl=zh-tw
https://www.instagram.com/aries_8248/?hl=zh-tw
<https://www.instagram.com/raopeisheng/?hl=zh-tw>
https://www.instagram.com/jam_steak/
<https://www.instagram.com/coffee89921/>
<https://www.instagram.com/kevin0204660/?hl=zh-tw>
https://www.instagram.com/crown_du/
https://www.instagram.com/zamy_ding/?hl=zh-tw
https://www.instagram.com/dinter_1126/?hl=zh-tw
<https://www.instagram.com/bluepigeon0810/>
<https://www.instagram.com/dontkjoanne/>
<https://www.instagram.com/annie72127/>
https://www.instagram.com/gu_yuze/
<https://www.instagram.com/linda831212/?hl=zh-tw>
<https://www.instagram.com/froggychiu/>
<https://www.instagram.com/ginachiki/>
<https://www.instagram.com/goldfishbrain17/>
<https://www.instagram.com/hajimesyachodesu/?hl=ja>
<https://www.instagram.com/howhowhasfriends/>
<https://www.instagram.com/imseriou/?hl=zh-tw>
https://www.instagram.com/jam_pepper_lin/
<https://www.instagram.com/joemanweng/?hl=zh-tw>
<https://www.instagram.com/la.112814/?hl=zh-tw>
<https://www.instagram.com/angela.lai.908/>
https://www.instagram.com/yuzi_lu/
<https://www.instagram.com/jengsu/>
https://www.instagram.com/chang_ricky_life/
<https://www.instagram.com/soon6669/>
<https://www.instagram.com/toyzlol/?hl=zh-tw>
<https://www.instagram.com/tsaibrotherderboofan/>
<https://www.instagram.com/wawawaku/>
<https://www.instagram.com/uccu0323/?hl=zh-tw>
<https://www.youtube.com/watch?v=MdF6x6ZmLAY&t=186s>
<https://www.bilibili.com/video/BV1eu411r7Qu/>



Thank you

Q&A

The background features abstract, organic shapes in shades of dark blue, light blue, and light gray. These shapes are scattered across the white background, with some appearing as large, flowing forms and others as small, isolated dots. The overall aesthetic is modern and minimalist.