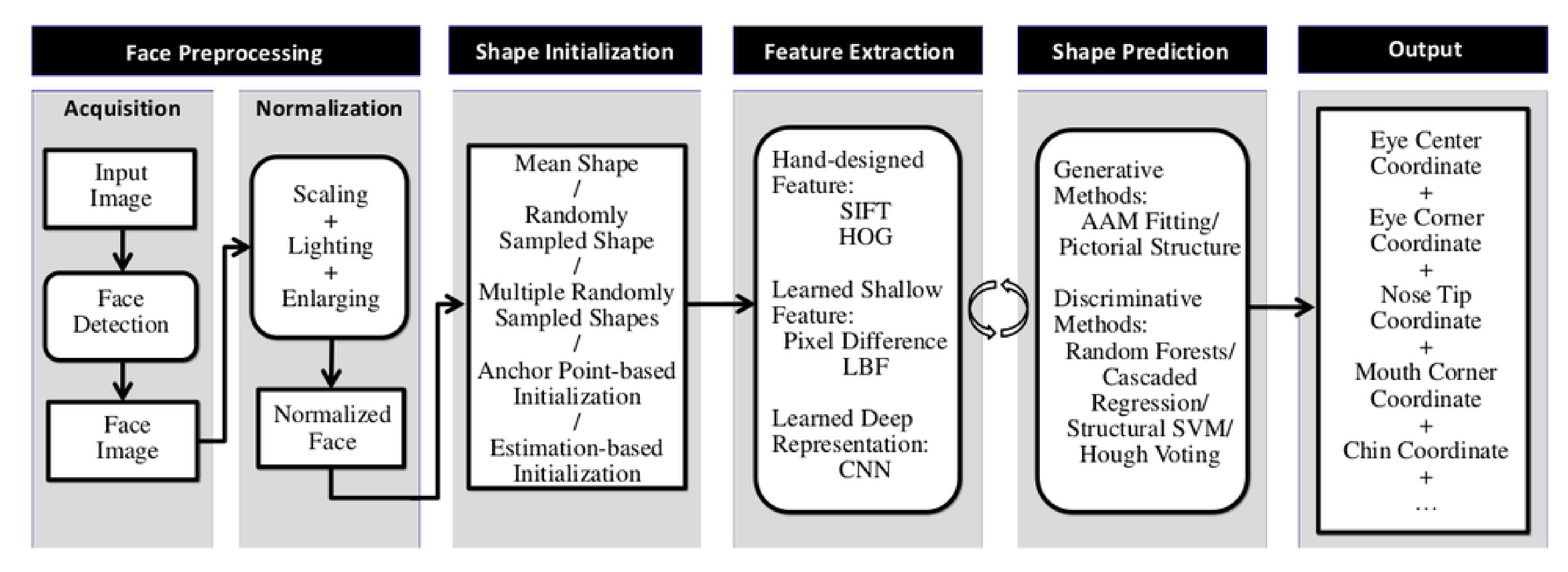
Look at Boundary

: A Boundary-Aware Face Alignment Algorithm

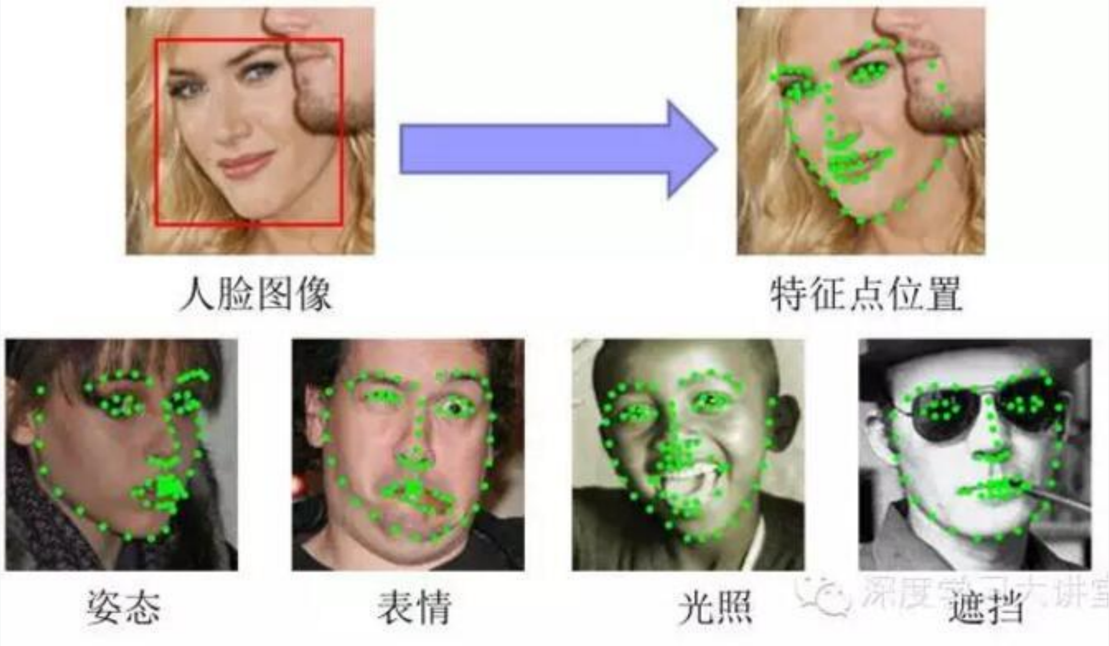
Before we start

Facial landmark localisation

人臉關鍵點定位是在人臉偵測的基礎上，依據輸入的影像，偵測面部的關鍵點。



Face alignment



人臉對齊可以看作是一張人臉影像搜尋人臉預先定義的點(landmark)，然後通過迭代的方式來細節化估計的形狀。在搜尋過程中，有兩種數據會被引用，一種就是人臉的外觀，另一種就是形狀。人臉對齊主要將人臉的特徵點檢測出來後標記。

Face landmark

Face alignment

CNN Convolution Neural Network

卷積神經網路

CNN是DL中最重要的一部分，在影像辨識中甚至能超越人類辨識

Convolution(卷積)

f, g的卷積公式如下：

連續形式:

$$(f * g)(n) = \int_{-\infty}^{\infty} f(\tau)g(n-\tau)d\tau$$

離散形式:

$$(f * g)(n) = \sum_{\tau=-\infty}^{\infty} f(\tau)g(n-\tau)$$

“卷”：對g函數進行翻轉

“積”：把g函數平移到n，這個位置對兩個函數的對應點相乘後相加

兩個函數的卷基本質上就是先將一個函數翻轉，然後進行滑動疊加。

(疊加在連續是指對兩個函數乘積求積分，離散是指加權求和)

Roll the dices

有兩個骰子，試問兩個骰子丟出去後兩個骰子點數和為四的機率是多少?

f	1	2	3	4	5	6

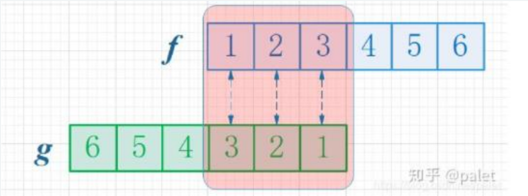
g	1	2	3	4	5	6

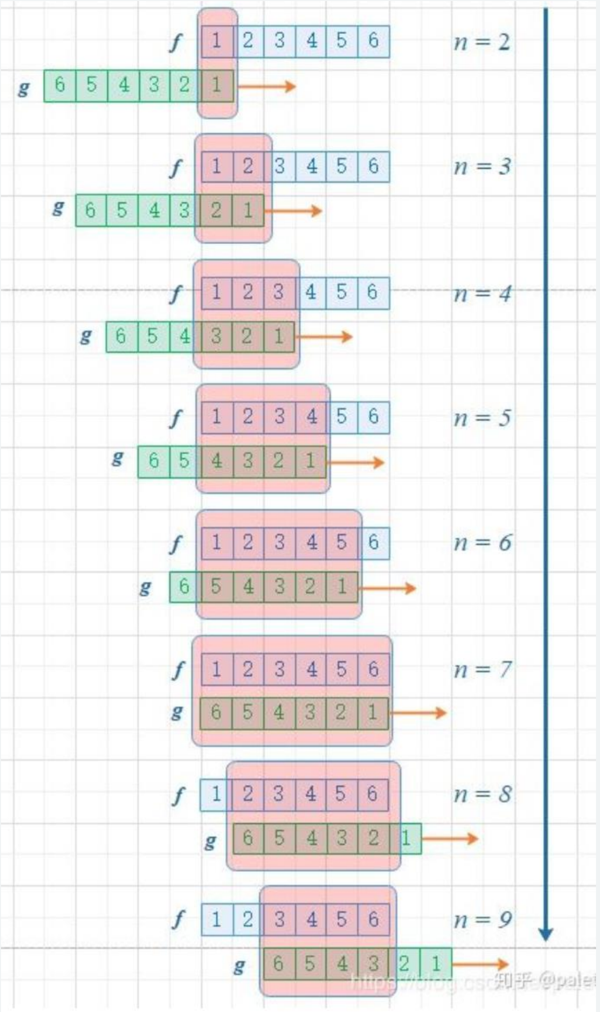
兩函數f,g，f(1)表示第一個骰子骰出1的機率...以此類推

$$f(1)g(3) + f(2)g(2) + f(3)g(1)$$

$$\Rightarrow \quad (f * g)(4) = \sum_{m=1}^{3} f(4-m)g(m)$$

先翻轉一下



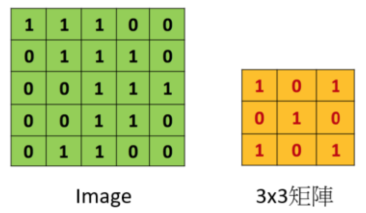


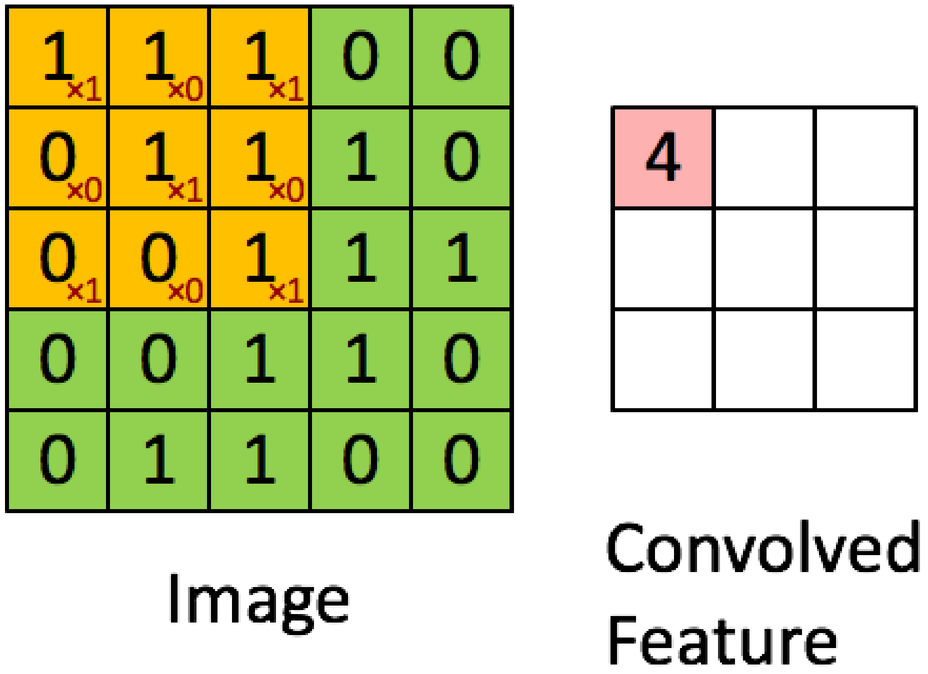
Convolution Layer

銳化特效，可以看出皮卡丘的邊緣跟毛都變得更細更清楚了，有一點點怪怪的XD

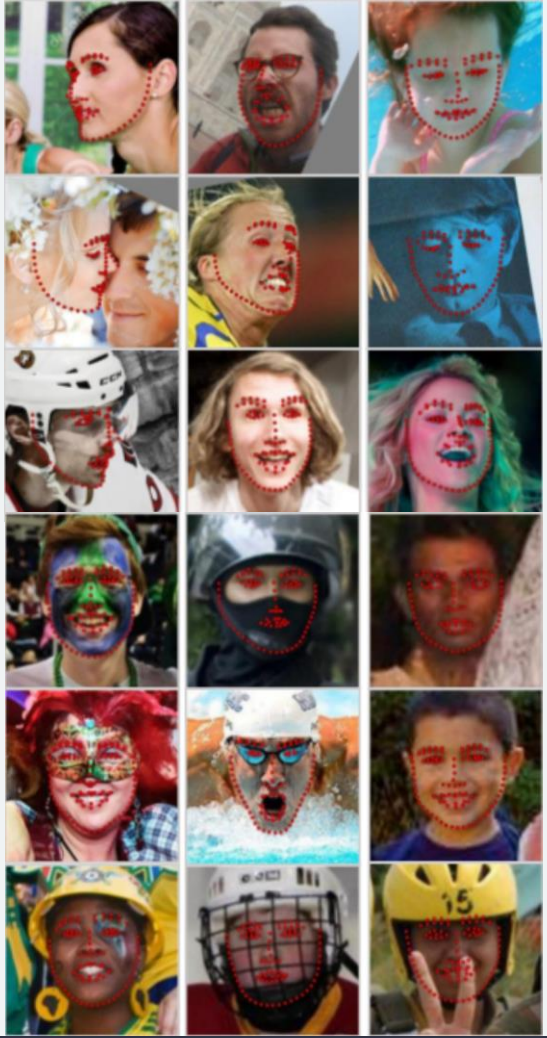
浮雕特效，看得出來除了皮卡丘的邊緣以外都變成灰色的了，變成浮雕的感覺了

每張圖片都可以表示成由pixel組成的陣列，美圖軟體的銳化、浮雕之類的美術特效就是用卷積達成的。**卷積就是在對圖片去做擷取特徵的動作，找出最好的特徵最後再進行分類**


1 1 1 0 0
0 1 1 1 0
0 0 1 1 1
0 0 1 1 0
0 1 1 0 0
Image
1 0 1
0 1 0
1 0 1
3x3矩陣


1 ×1 1 ×0 1 ×1 0 0
0 ×0 1 ×1 1 ×0 1 0
0 ×1 0 ×0 1 ×1 1 1
0 0 1 1 0
0 1 1 0 0
Image
4
Convolved Feature

Original Image	3x3 Kernel	After Image	
	[0 0 0] [0 1 0] [0 0 0]		原圖不變
	[1 1 1] [1 -7 1] [1 1 1]		銳化
	[-1 -1 -1] [-1 8 -1] [-1 -1 -1]		邊緣強化
	[-1 -1 0] [-1 0 1] [0 1 1]		浮雕



Intro

Wayne Wu , Chen Qian , Shuo Yang, Quan Wang , Yici Cai, Qiang Zhou
Tsinghua National Laboratory for Information Science and Technology (TNList),
Department of Computer Science and Technology, Tsinghua University
SenseTime Research
Amazon Rekognition

wwy15@mails.tsinghua.edu.cn caiyc@mail.tsinghua.edu.cn zhouqiang@tsinghua.edu.cn {qianchen, wangquan}@sensetime.com shuoy@amazon.com

CVPR 2018

Related work

在人臉對齊的文獻中，除了經典的演算法像是: ASMs, AAMs, CLMs 以及級聯迴歸方式以外，DCNNs(深度卷積神經網路)也取得了很好的效果，基於深度卷積神經網路的方法主要分為兩類: 座標迴歸模型和熱力圖迴歸模型。

座標迴歸模型是直接輸入影像和landmark之間的關係。

熱力圖迴歸模型是為每個landmark生成可能的熱力圖。

About

- 2D face alignment
- Unlike the conventional heatmap based and regression based method
- Boundary lines as geometric structure
- Remove the ambiguities in the landmark definition

Contribution

- Achieves 3.49% mean error on 300-W Fullset
- Achieves 3.92% mean error with 0.39% failure rate on COFW dataset
- Achieves 1.25% mean error on AFLW-Full dataset
- Propose a new dataset which includes poses, expressions, illuminations, makeups, occlusions, and blurriness.

The objective of the paper

Devise an effective face alignment algorithm to handle face with unconstrained pose variation and occlusion across multiple datasets and annotation protocols.

By utilising boundary lines as the geometric structure of a human face to help facial landmark localisation.

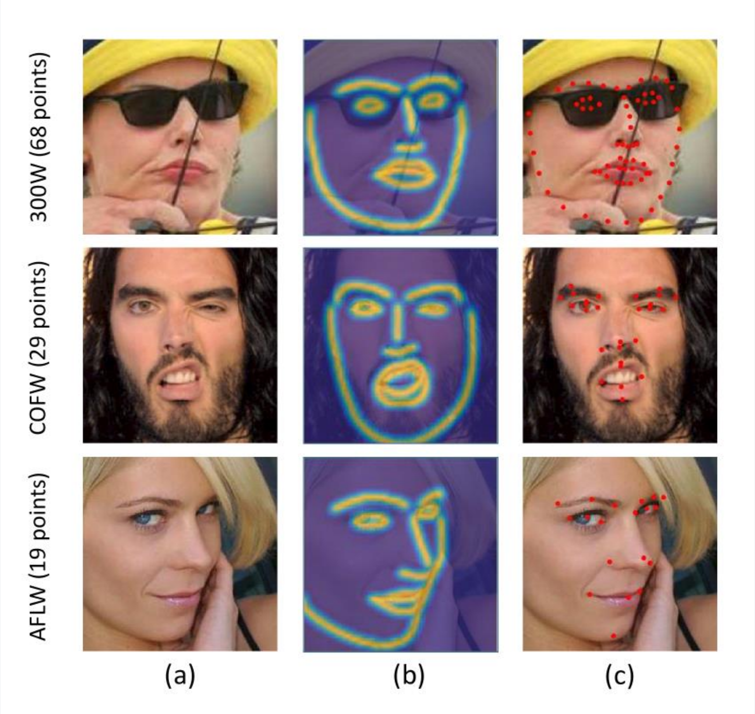
Abstract

1. Why using boundary ?
2. How to use boundary ?
3. What is the relationship between boundary estimation and landmarks localization ?

Q1.Why using boundary?

1. It is easier to identify facial boundaries comparing to facial landmarks under large pose and occlusion.
2. Facial landmarks other than corners can hardly remain the same semantical locations with large pose variation and occlusion.
3. Different annotation schemes of existing datasets lead to different number of landmarks and annotation schemes of future face alignment datasets can hardly be determined .

Boundary-aware algorithm



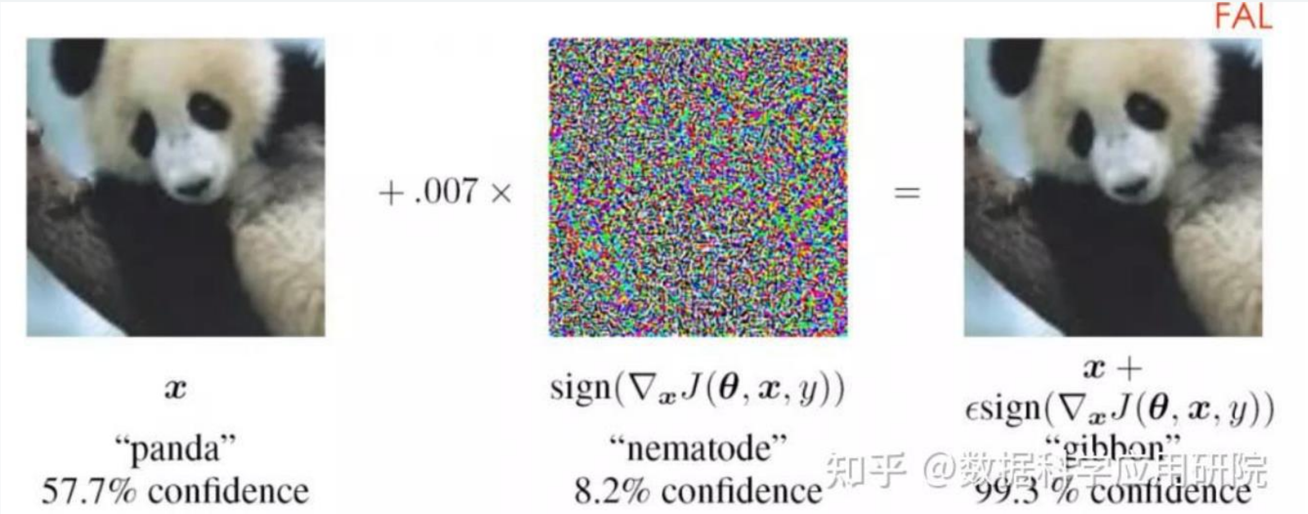
First estimate facial boundary heatmaps and then regress landmarks with the help of boundary heatmaps.

Q3.Relationship ?

To explore the relationship between facial boundaries and landmarks, we introduce adversarial learning ideas by using a landmark-based boundaries effectiveness discriminator. **Experiments have shown that the better quality estimated boundaries have, the more accurate landmarks will be.**

Adversarial data

對抗樣本指的是將真實樣本加入一些擾動後新生成的樣本，這些樣本可能會使機器判別錯誤而導致安全性問題。



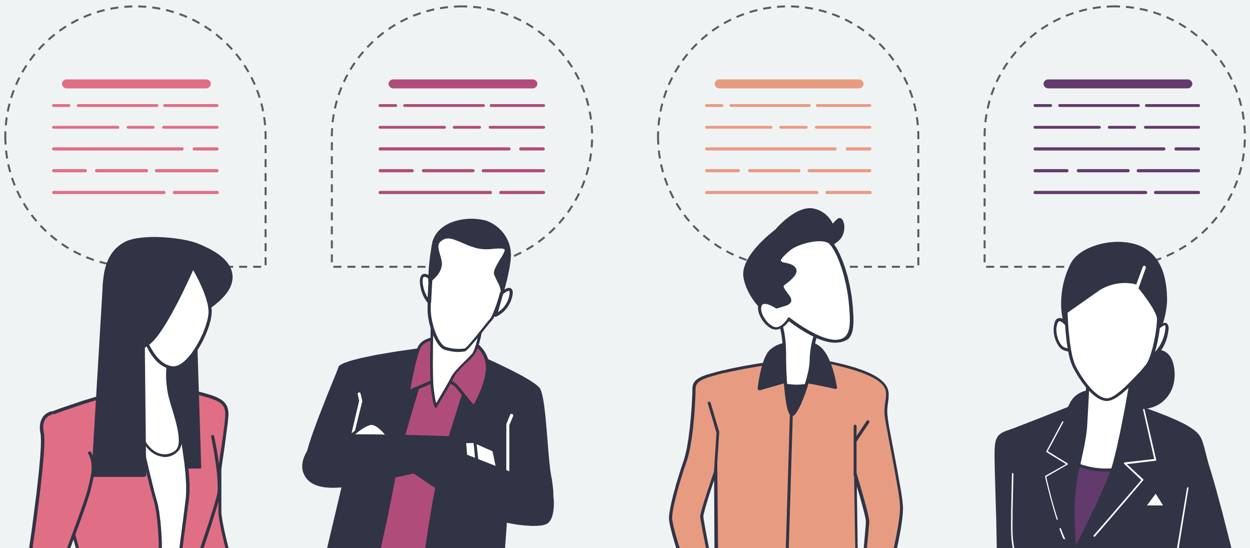
Adversarial learning

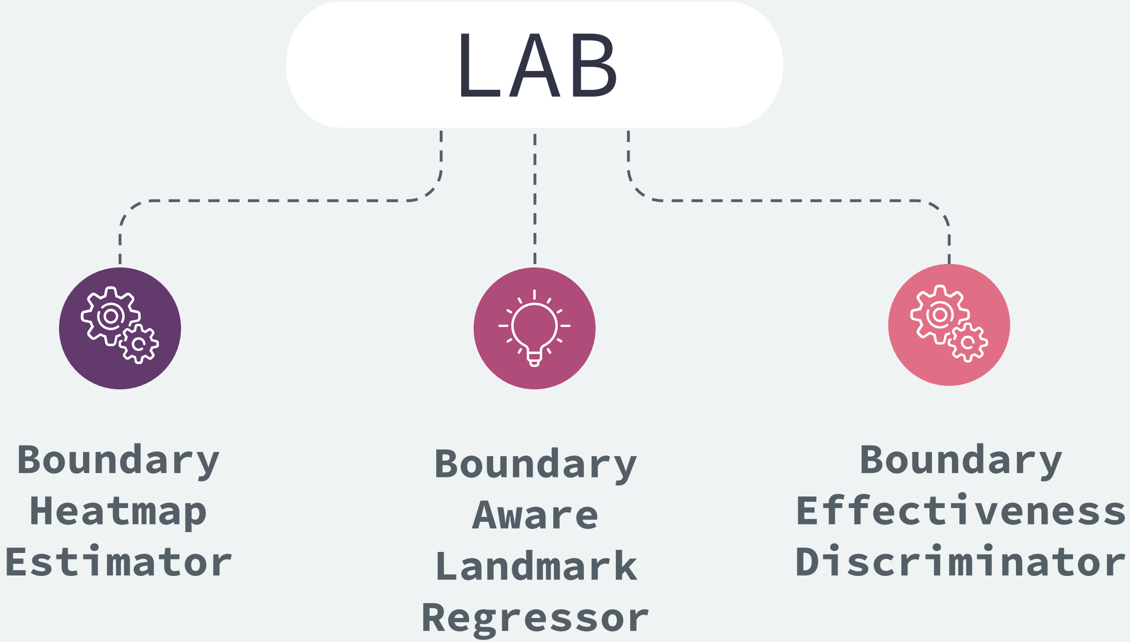
機器學習中的對抗樣本引起了極大的關注，於是相對應的問題也衍生而出，像是如何更有效的產生對抗樣本(對抗攻擊)、針對對抗樣本提供有效的防禦(對抗防禦)，這類型的問題以及方法都被稱為對抗學習。

對抗樣本的危機很大，尤其是對於無人駕駛、醫療診斷、金融分析這類方面的安全性至關重要。

LAB

A Boundary-Aware Face Alignment Algorithm

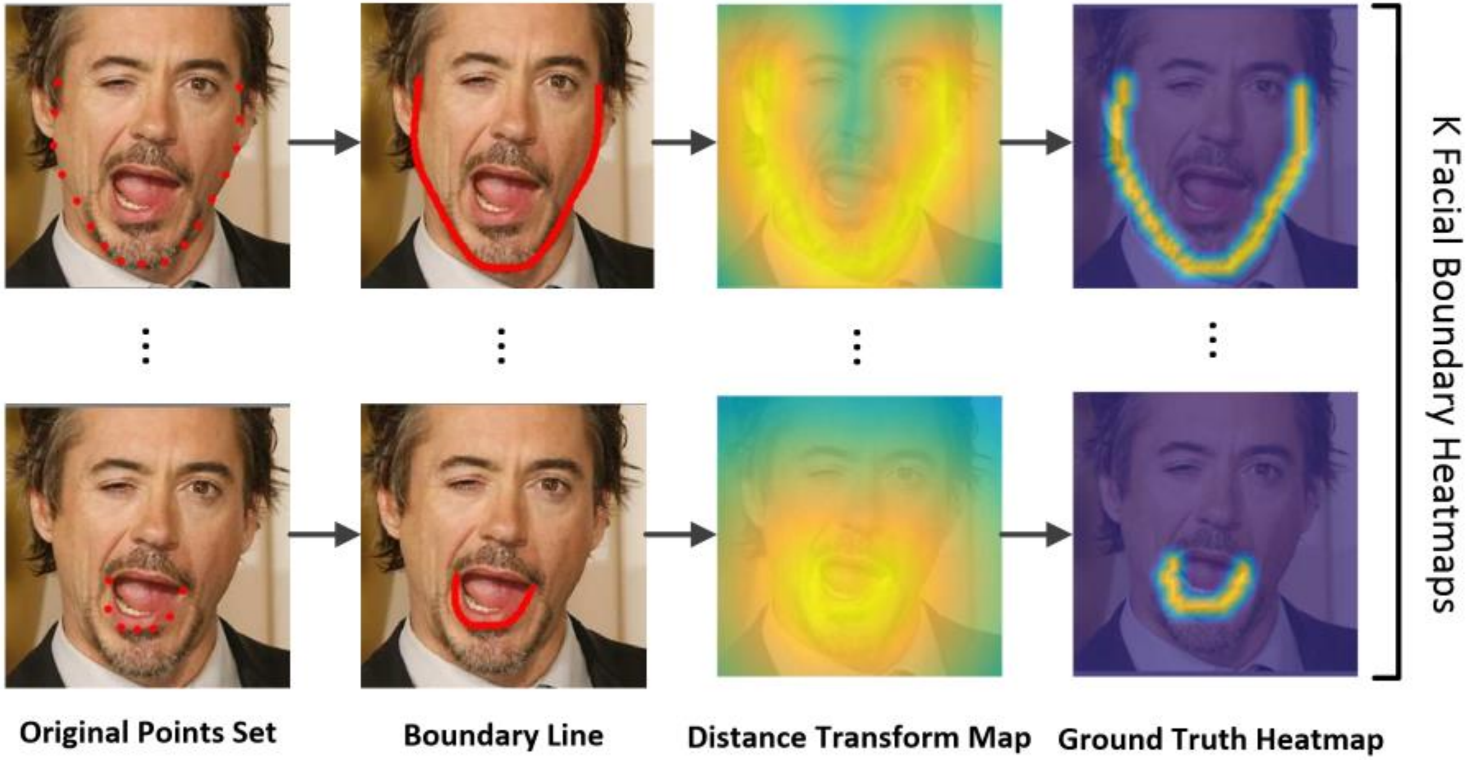



LAB
Boundary Heatmap Estimator
Boundary Aware Landmark Regressor
Boundary Effectiveness Discriminator

Boundary-aware landmark regressor

The details of boundary heatmap are defined as follow :

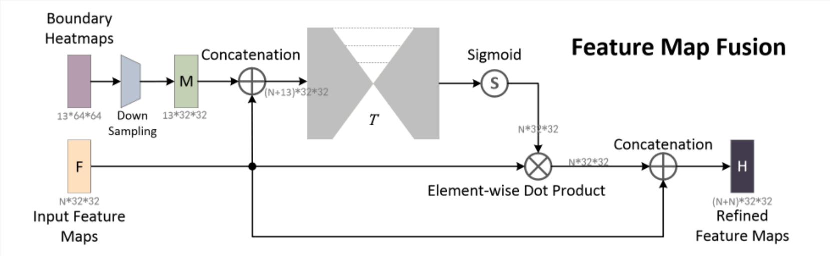
1. Given an image I, denote its ground truth annotation by L landmarks as $S = \{s_l\}_{l=1}^{L}$
2. K subsets of S (S_i) are defined to represent landmarks belongs to K boundaries respectively
3. For each boundaries, S_i is interpolated to get a dense boundary line.
4. Then a binary boundary map B_i, the same size as I, is formed by setting only points on the boundary line to be 1, others 0.
5. A distance transform is performed based on each B_i to get distance map D_i. We use a gaussian expression with standard deviation σ to transform the distance map to ground-truth boundary heatmap M_i.



K Facial Boundary Heatmaps
Original Points Set
Boundary Line
Distance Transform Map
Ground Truth Heatmap

Multi-stage boundary heatmap fusion

邊界熱力圖 M 與特徵圖 F 的融合



為了更好的利用邊界熱力圖中包含的大量資訊，作者提出了多層邊界熱力圖融合方案。本演算法以一個 4 階段 18 層的網路結構為基礎網路，在輸入層和每個階段上執行邊界熱力圖融合。實驗結果表明在基礎網路上執行的如何次數越多，得到的效果越好。

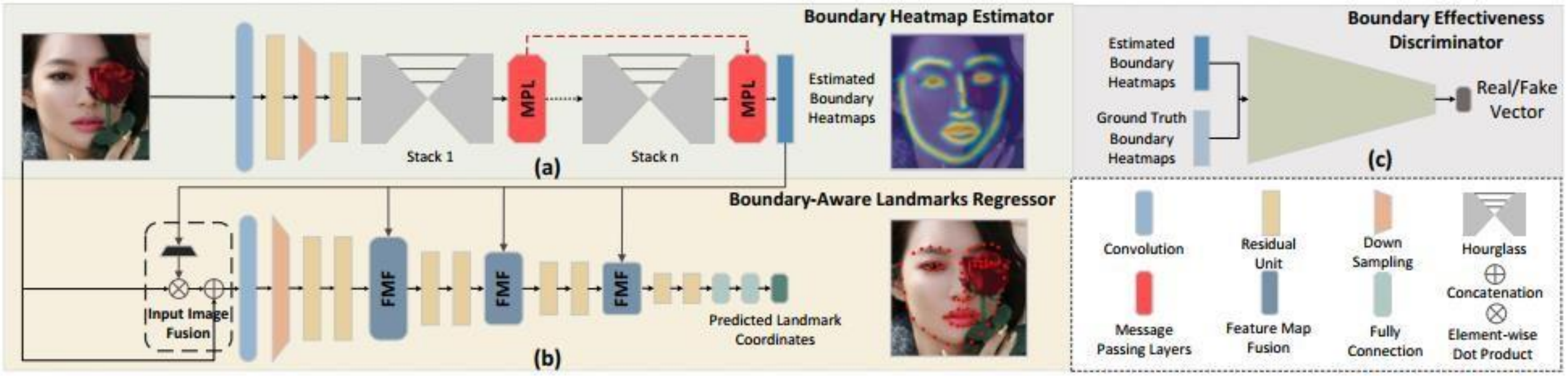
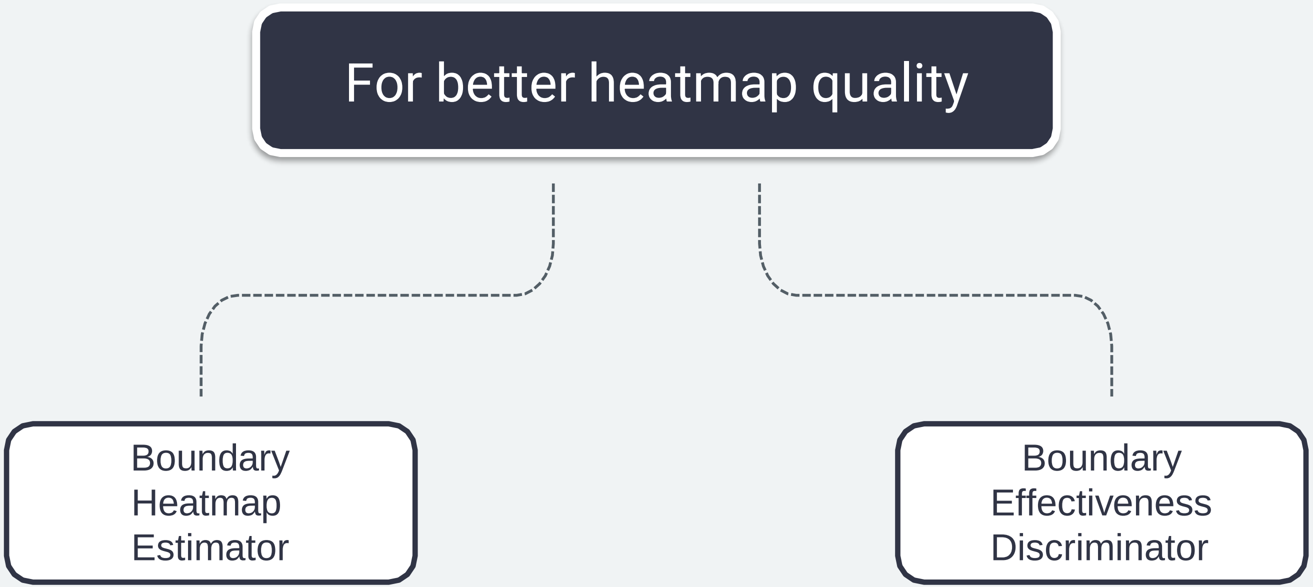
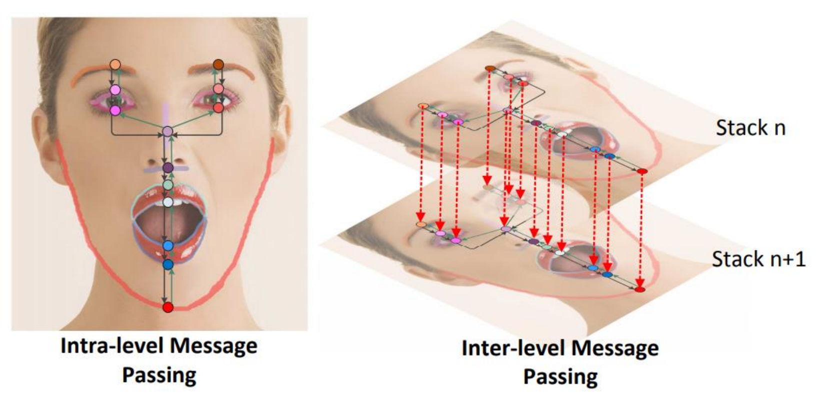


Figure 2: Overview of our Boundary-Aware Face Alignment framework. (a) Boundary heatmap estimator, which based on hourglass network is used to estimate boundary heatmaps. Message passing layers are introduced to handle occlusion. (b) Boundary-aware landmarks regressor is used to generate the final prediction of landmarks. Boundary heatmap fusion scheme is introduced to incorporate boundary information into the feature learning of regressor. (c) Boundary effectiveness discriminator, which distinguishes "real" boundary heatmaps from "fake", is used to further improve the quality of the estimated boundary heatmaps.

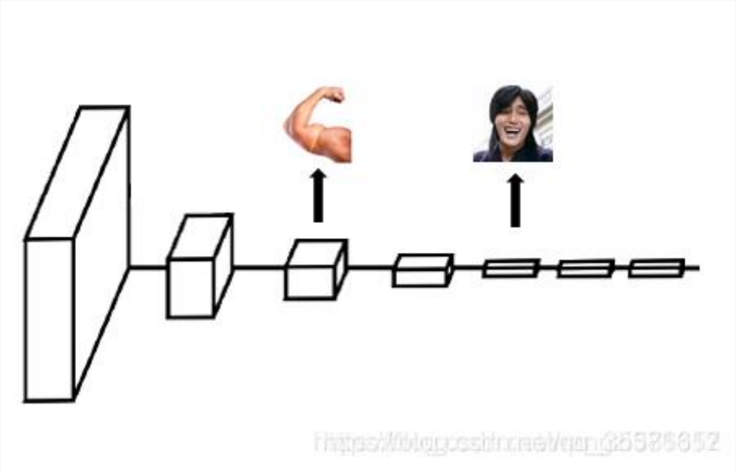

For better heatmap quality
Boundary Heatmap Estimator
Boundary Effectiveness Discriminator

Boundary heatmap estimator



We implemented message passing following. In this implementation, the feature map at the end of each stack needs be divided into K branches, where K is the number of boundaries, each represents a type of boundary feature map.

Stacked hourglass structure



Stacked HG網路是作用是通過估計人體關鍵點的熱度圖來實現關鍵點的定位。這種網路結構的特點是就是能夠充分利用多尺度的特徵映射，並以且輸入和輸入有相同的輸出。傳統CNN估計姿態的網路大多只使用最後一層的卷積特徵，這樣會造成訊息的丟失。事實上對於姿態估計，全身不同的關鍵點並不是在相同的feature map上具有最好的識別度，所以可能需要設計一個可以同時使用多個feature map的網路。

Message passing

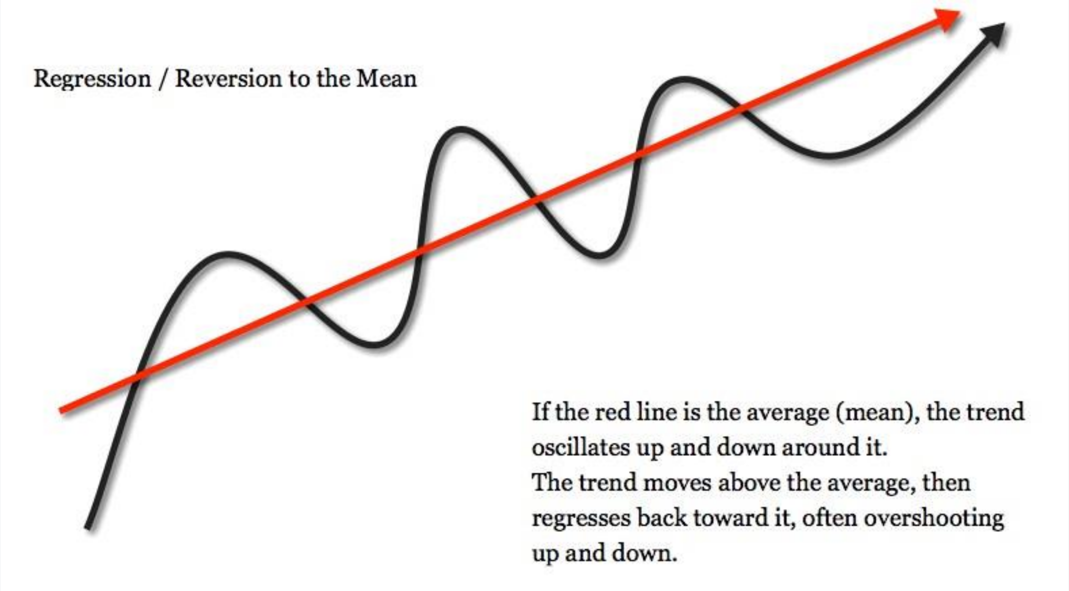
This process is visualised During occlusion, visible boundaries can provide help to occluded ones according to face structure.

-**Intra-level message passing** is used at the end of each stack to pass information between different boundary heatmaps.

-**Inter-level message passing** is adopted to pass message from lower stacks to the higher stacks to keep the quality of boundary heatmaps when stacking more hourglass subnets

Regression-to-the-mean problem

在統計學中，迴歸均值指的是如果變量在其一次測量時是極端值，則在下一次測量時會有趨近於平均的現象。

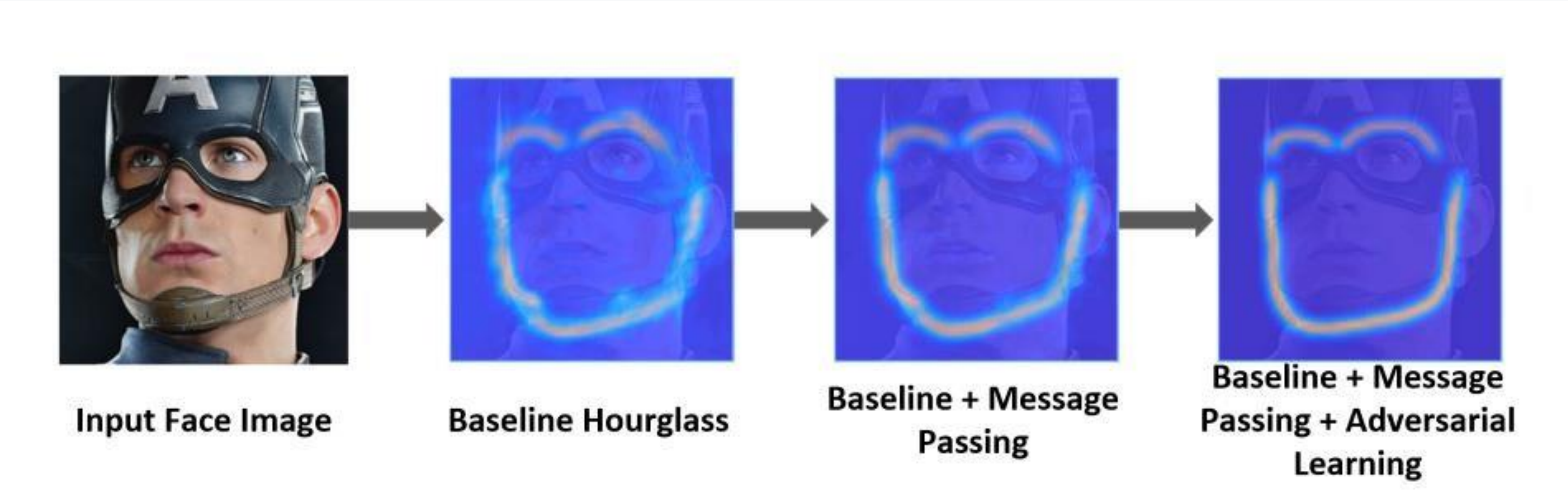


Boundary effectiveness discriminator

$$d_{\text{fake}}(\hat{M}, \hat{S}) = \begin{cases} 0, & \Pr_{s \in \hat{S}}(Dist(s) < \theta) < \delta \\ 1, & \text{otherwise} \end{cases}$$

The hard-to-define term "quality" of heatmaps has a very clear evaluation metric. If helping to produce accurate landmark coordinates, the boundary heatmap has a good quality. According to this, we propose a landmark based boundary effectiveness discriminator to decide the effectiveness of the generated boundary heatmaps.

Effectiveness of heatmap



Cross dataset face alignment

Because of the gap between annotation schemes, these datasets can hardly be jointly used.

Models trained on one specific dataset perform poorly on recent in-the-wild test sets.

From a new perspective, we take facial boundaries as an all-purpose middle-level face geometry representation. Facial boundaries naturally unify different landmark definitions with enough landmarks. And it can also be applied to help training landmarks regressor with any specific landmarks definition. The cross-dataset capacity is an important by-product of our methods. Its effectiveness is evaluated.

Resources

- Look at Boundary : A Boundary-Aware Face Alignment Algorithm paper
- Look at Boundary: A Boundary-Aware Face Alignment Algorithm code (github)
- Look at Boundary : A Boundary-Aware Face Alignment Algorithm project
- Stacked Hourglass Networks for Human Pose Estimation
- Structured Feature Learning for Pose Estimation

Reference

- [機器學習 ML NOTE]Convolution Neural Network 卷積神經網路
- 如何通俗易懂的解釋卷積?
- 機器學習中火爆的對抗學習是什麼，有什麼應用?
- 【人臉辨識】使用5 Facial Landmarks進行臉孔校正
- 《Look at Boundary: A Boundary-Aware Face Alignment Algorithm 》閱讀筆記
- 人臉關鍵對齊
- 人脸对齐（一）--定义及作用
- Stacked Hourglass Networks 理解
- [人脸关键点检测] Look at Boundary: A Boundary-Aware Face Alignment Algorithm
- [论文笔记] 人脸关键点检测方向系列论

Thanks!

Do you have any questions?