


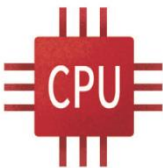


圖像識別與理解

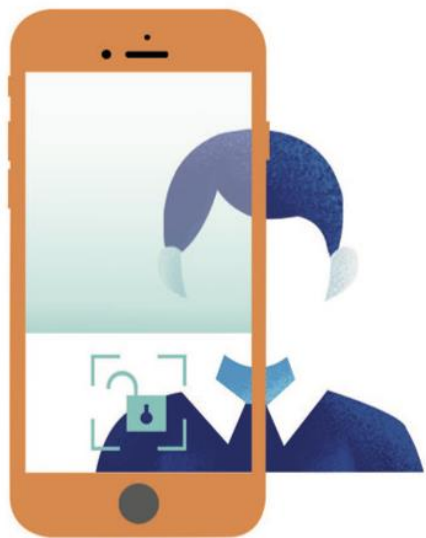
電腦能否**理解**圖像？ 或者說，電腦能否**識別**圖像中的內容？

機器視覺的基本構想：

以相機為眼，以電腦為腦，讓機器類比出人類的視覺認知能力。

	接受图像	理解认知
人类的视觉系统		
机器的视觉系统		

圖：人類視覺系統與機器視覺系統的類比



人脸解锁手机



识别图像中的文字



检测人群中的异常行为

電腦的人臉識別

給定一張照片，讓電腦判斷一下，這張照片上究竟有沒有人臉。



有人脸



有/无人脸



没有人脸

圖：人臉判別問題

判斷**有沒有人臉**是一個**二分類問題**

對於**人類**而言是非常簡單的而直觀的。

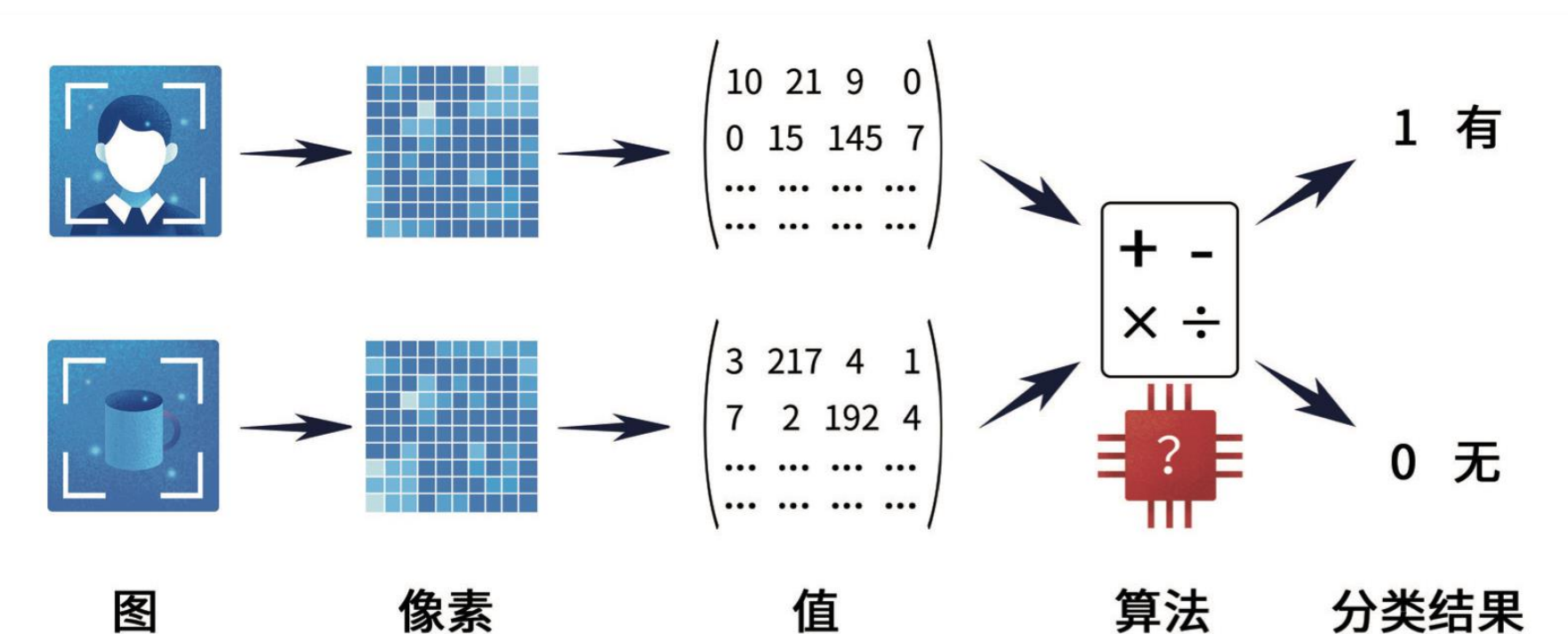
在**電腦**的眼中，
一幅圖像就是按照空間順序排列的一系列像素值。



電腦要根據這些像素值進行分類。

電腦如何根據這些像素值進行分類？

- 之前通過對輸入資料進行線性計算，一個線性分類器就能夠區分兩種類型的數據。



SenseStudy課程平臺 “人工智能入門（上）”

實驗5 – 8 圖像分類

```
train_x, train_y, test_x, test_y = load_CelebA_sample_dataset()

for i in range(5):

    sample = train_x[i*100]

    img_sample = list2numpy(sample)

    fig() + image(img_sample)

    print(train_y[i*100])
```



SenseStudy課程平臺 “人工智能入門（上）” 實驗5 – 8 圖像分類

```
train_x_flat = flatten(train_x)
test_x_flat = flatten(test_x)

model = linear_classifier()
model.train(train_x_flat, train_y)
pred = model.predict(test_x_flat)
acc = accuracy(pred, test_y)
print('The accuracy is: ',acc)

for i in range(len(test_x)):
    if pred[i] != test_y[i]:
        sample = test_x[i]
        img_sample = list2numpy(sample)
        fig() + image(img_sample)
        print('Prediction is: ',pred[i], ' The truth is: ',test_y[i])
```



SenseStudy課程平臺 “人工智能入門（上）” 實驗5 – 8 圖像分類

```
train_feat, test_feat = load_CelebA_features()  
model = linear_classifier()  
model.train(train_feat, train_y)
```

```
pred_y = model.predict(test_feat)  
test_acc = accuracy(pred_y, test_y)  
print('The accuracy is: ', test_acc)
```

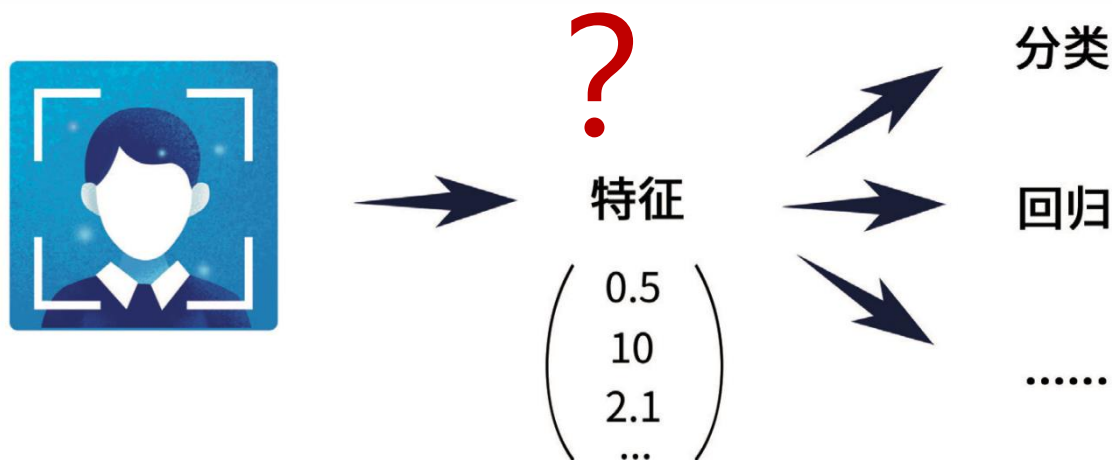
```
for i in range(len(test_feat)):  
    if pred_y[i] != test_y[i]:  
        sample = test_x[i]  
        img_sample = list2numpy(sample)  
        fig() + image(img_sample)  
        print('Prediction is: ', pred_y[i], ' The truth is: ', test_y[i])
```



基於特徵的演算法

在深度學習技術出現以前，**基於特徵的演算法**一直佔據著圖像識別演算法的主要地位。這類演算法通常包含兩個主要步驟：

1. 利用人工設計的演算法，從圖像中**提取**某種**特徵**。
2. 在這些特徵資料上**訓練**分類、回歸或其他機器學習**模型**，實現目標。



圖：特徵與機器學習模型

特徵是什麼？

鑽石的兩個“特徵” { 重量
成色 } 預測價格

一顆鑽石，除了重量與成色的特徵，還有產地、硬度、雜質的化學成分等資訊作為特徵。

Q：一個鑽石有多種物理屬性，應該提取哪些作為特徵呢？

提取與目標相聯繫的特徵。



數字圖像的特徵

對於數字圖像，電腦唯一能使用的資料就是它們的**像素值**。



設計**演算法**，對圖像中的像素值進行**運算**

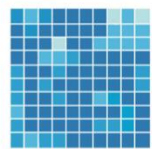


得到描述圖像的**特徵**（對解決問題有意義的特徵）



图

- 特征1: 全部像素的平均值=26
- 特征2: B区域与A区域像素的差值
- 特征3: A中红色通道与B中绿色通道的2倍值的和=99
-



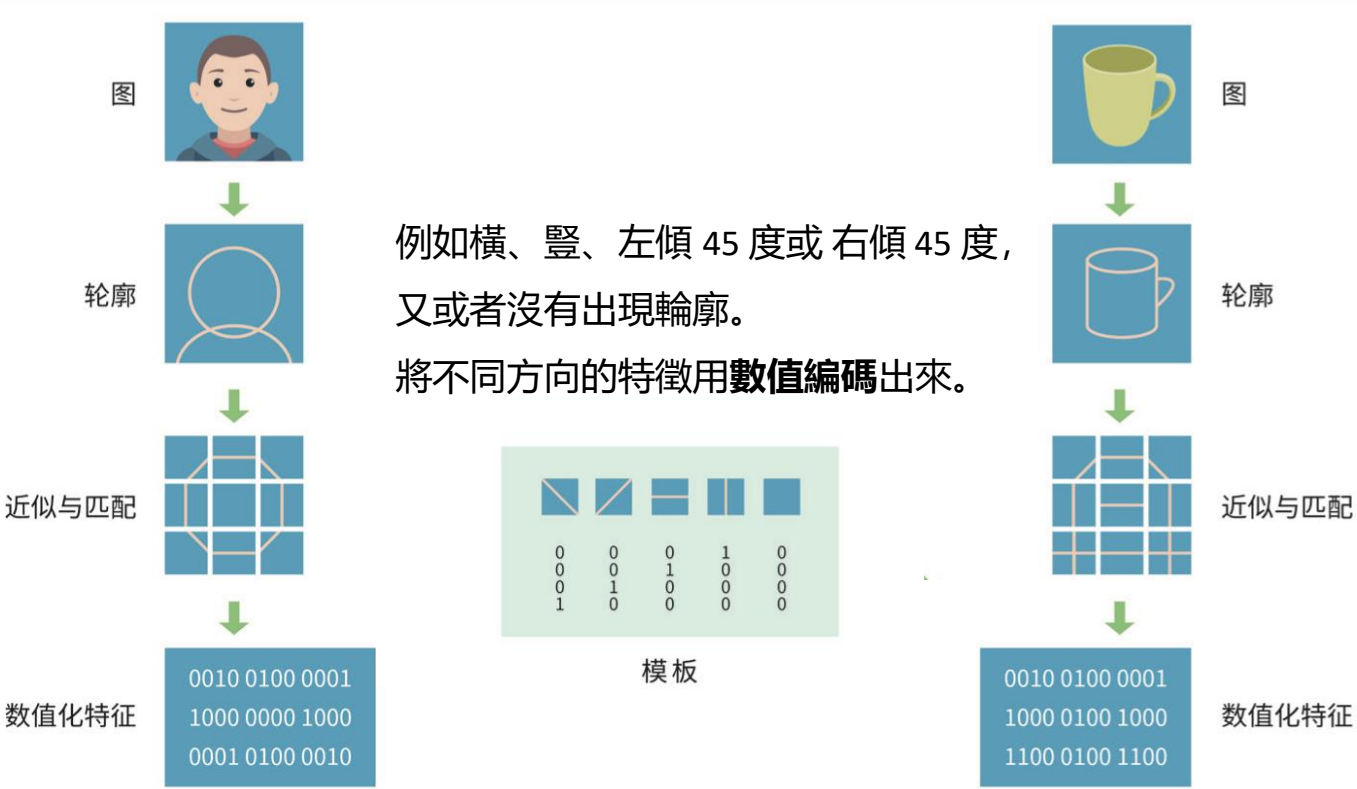
全部像素值

$$\begin{pmatrix} 26 \\ -10 \\ 99 \\ \dots \end{pmatrix}$$

特征值组成的特征向量

圖：圖像特徵提取的示意圖

人臉判別問題 → 使用**面部輪廓的特徵**就可以在在一定程度上判斷照片中是否有人臉



圖像中物體的輪廓通常是由**曲線**組成的。

為了描述一條輪廓曲線，將其**切分**成若干片段，並將每一段近似為一條直線段，並估計直線段的方向。

將所有段落對應的**特徵拼接**起來，就得到了完整的輪廓特徵。



利用上述演算法提取的特徵**保留了原圖中物體輪廓**的信息，過濾掉了諸如顏色、局部細節等信息。

原因：對於人臉判別問題而言，物體的輪廓更為重要。

目前在圖像識別中被廣泛使用的**方向梯度長條圖(HOG)**就是基於上述特徵設計想法發展出來的。

示例：在這個實驗中，利用函數 `extract_hog` 提取圖像的輪廓特徵，並訓練分類器判斷一張圖片是否為人臉。

```
img_list = []           # 圖片清單
label_list = []         # 類別列表
feat_list = []          # 空列表
for img in img_list:    # 對圖片清單中的每一張圖片
    feat = extract_hog(img) # 提取梯度長條圖特徵
    feat_list.append(feat)  # 將提取的特徵加入到特徵清單
model = LinearClassifier()
model.train(feat_list, label) # 利用提取的特徵訓練分類器
```


SenseStudy課程平臺 “人工智能入門（下）”

實驗7 – 4 線性分類器技術-圖像分類

```
train_data, train_lable, test_data = load_data()  
show_data(train_data)  
print(train_lable)
```

```
train_features = load_hog('train_features')  
test_features = load_hog('test_features')
```

```
model = LinearClassifier()  
model.train(train_features, train_lable)  
acc, Y_predict = model.pred(test_features)  
show_data(test_data)  
print("Prediction: ", Y_predict, "accuracy is: ", acc)
```



有效的圖像特徵 + 合適的機器學習模型 ➡ 解決了大量的視覺問題

基於特徵的方法的性能仍與人們的期望相差較遠

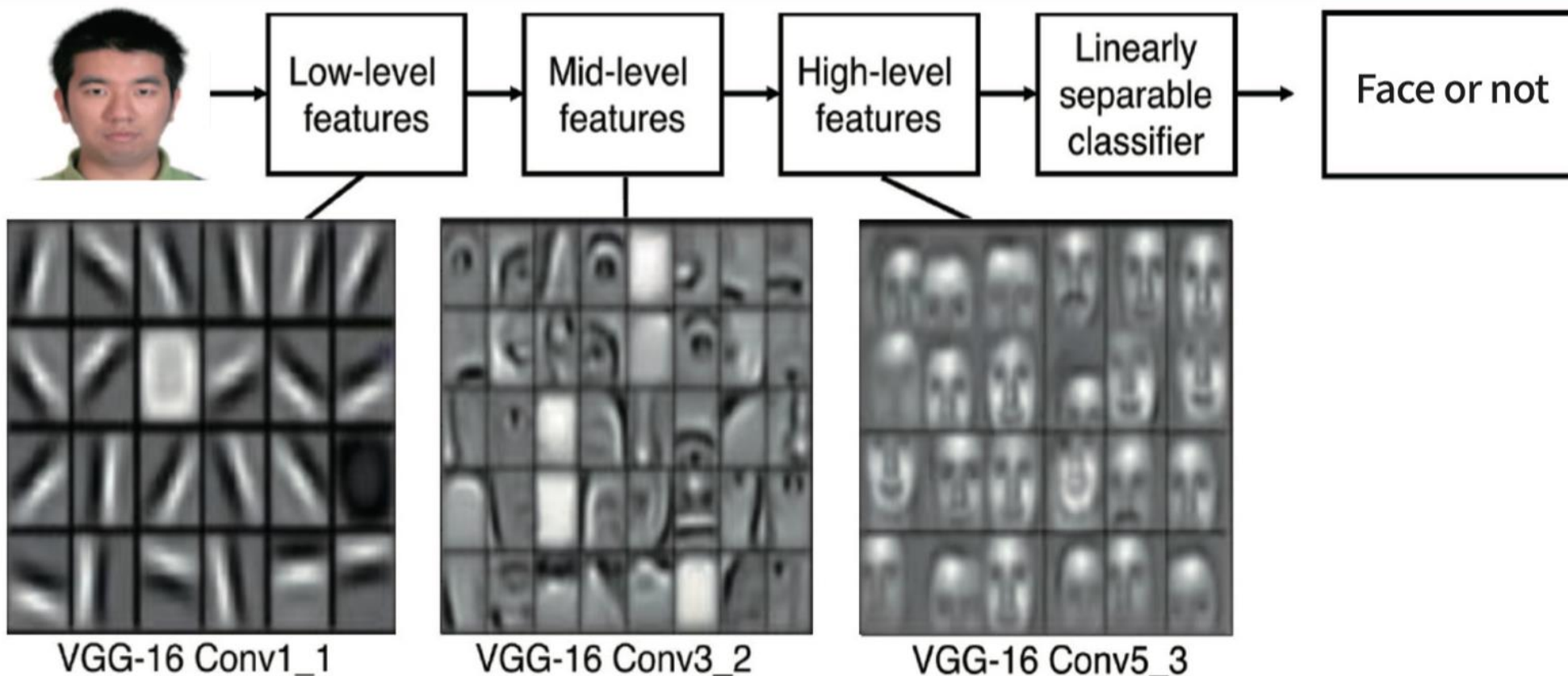
如上一章的人臉判別系統，由於輪廓特徵只描述了輪廓，忽略了顏色與細節，在判別一些輪廓和人臉相似的照片時很容易受到“欺騙”，做出錯誤的判斷。

原因

人工設計的特徵過於簡單

深度學習方法可以**自動從資料學習**有效的特徵提取演算法

卷積神經網路中的低、中、高級特徵



- 第一層輸出的特徵圖只保留了像素間的局部關係的資訊，例如邊緣、顏色等。稱這種特徵為**低層特徵**。
- 特徵可以體現人類認知的概念。這種接近人類理解的特徵被稱為**高層特徵**。

SenseStudy課程平臺 “人工智能入門（上）”

實驗4 – 7 認識卷積核

```
fig() + image(img_furn)
img_gray = rgb2gray(img_furn)

v_kernel = [
    [-1,0,1],
    [-1,0,1],
    [-1,0,1]
]
img_v = apply_kernel(img_gray, v_kernel)
img_v = abs(img_v)
fig() + image(img_v, cmap='gray')
```

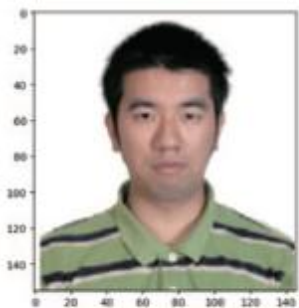




輸入圖像

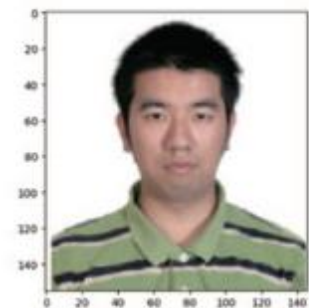
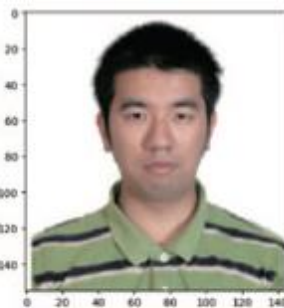
計算參數

對應的特徵輸出



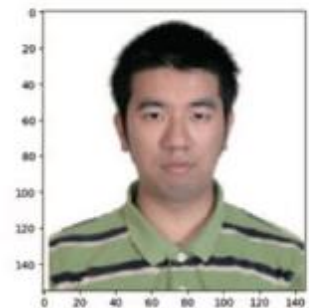
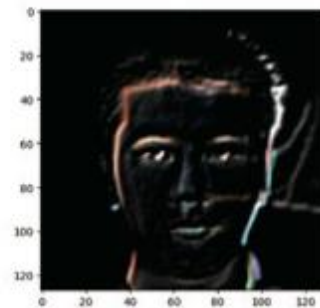
$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

參數1



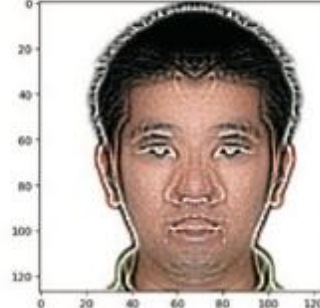
$$\begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

參數2



$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -7 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

參數3

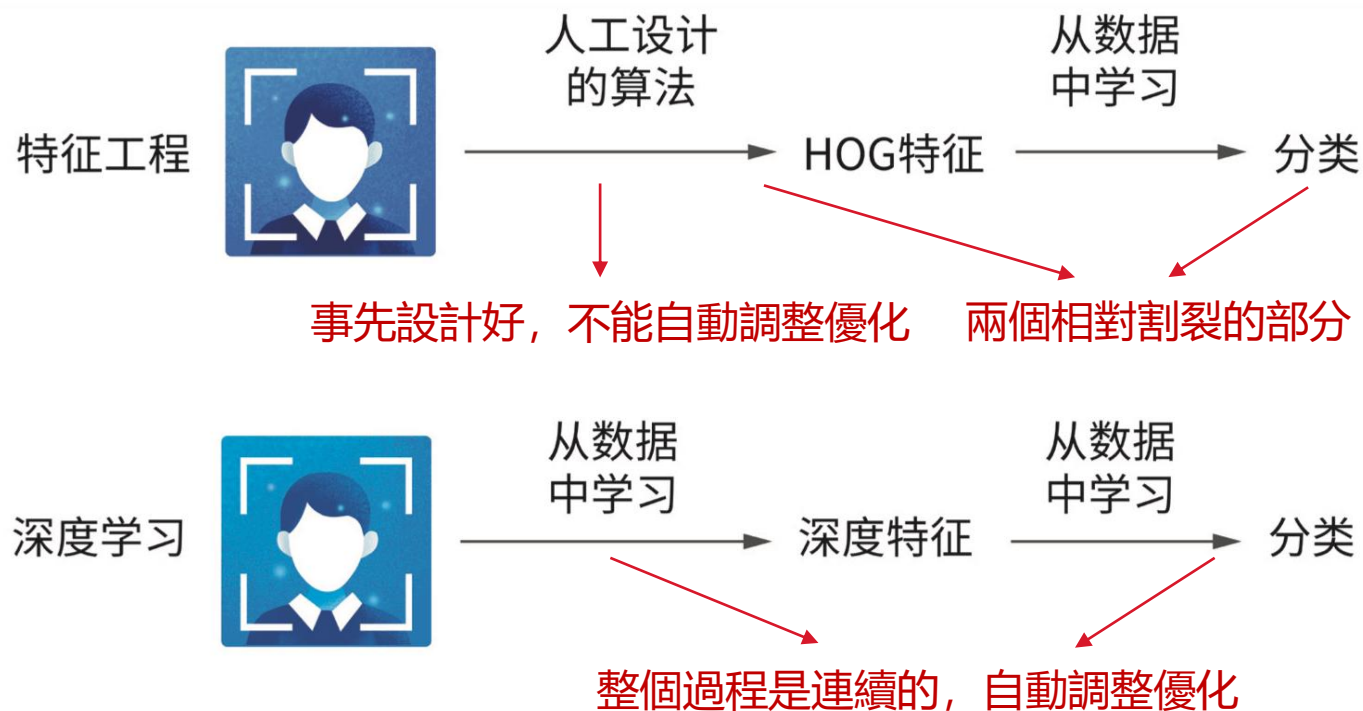


使用不同的參數，
提取出的特徵有所不同。

Q: 使用什麼樣的參數，才能提取出合適的特徵，進而完成人臉判別任務呢？

卷積神經網路可以從數據中學習出合適的參數。

只要提供足夠的資料對卷積神經網路進行訓練，卷積神經網路就可以完成設定的目標。



- 只要我們在**輸入端**提供訓練圖片，在**輸出端**給出對應的判定結果，神經網路就可以自動學習其中的參數。
這種從原始輸入到最終輸出構成一個可以**學習調優的整體**，
各模組在一個統一的學習過程中協同調節的過程稱為**端到端學習**。

訓練神經網路完成人臉判別問題的步驟

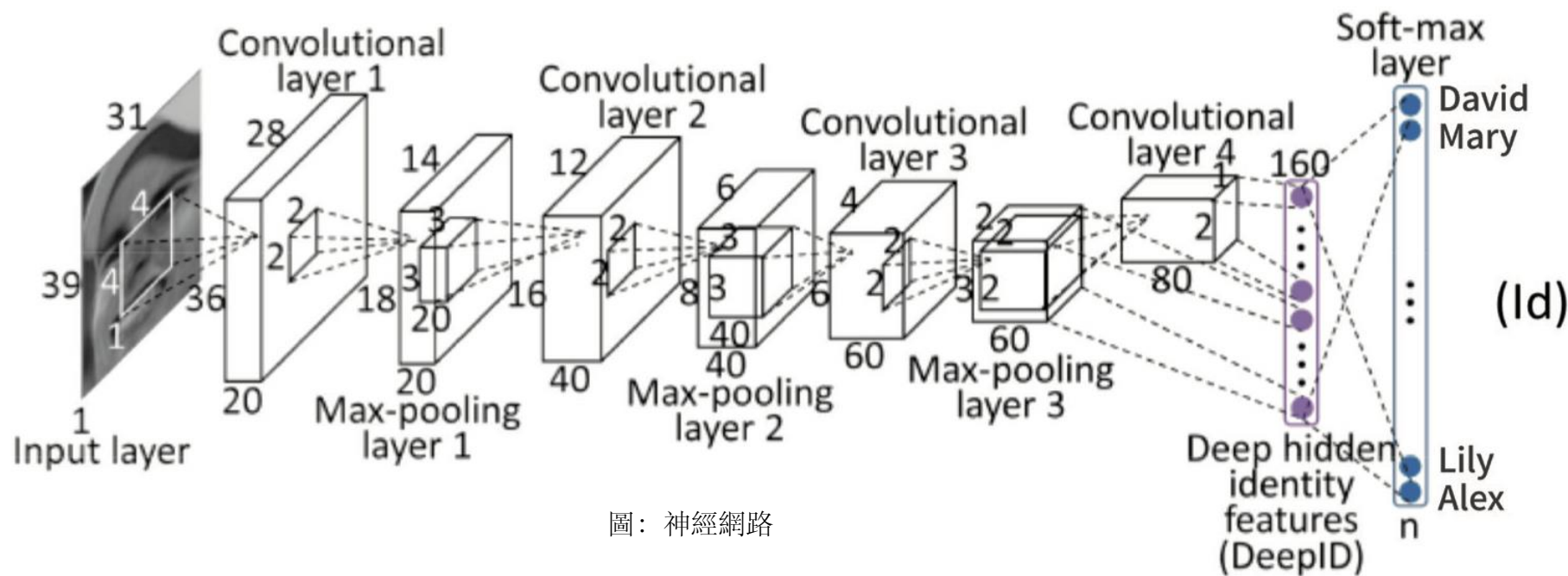
1. **載入**人臉圖像與非人臉圖像的**數據集**；
2. **初始化**一個卷積神經網路；
3. 使用資料集**訓練**神經網路；
4. 選取一張資料集以外的圖像，利用訓練好的神經網路判讀圖片中是否有人臉，最後**測試**我們訓練的神經網路判定的結果是否正確。

人臉屬性—深度學習特徵

神經網路具有很強大的特徵提取的能力。

簡單的人臉判別實驗只展示了深度神經網路的一個方面。

目前，神經網路已經能夠支援**上百萬人規模的人臉識別**。



圖：神經網路

上圖所示的神經網路先將人臉圖片轉換為一個**128 維度的特徵**，然後再對特徵進行分類。

人臉屬性—深度學習特徵

為了識別這麼多人，神經網路需要保證

不同人的圖片的特徵有所**差異**。

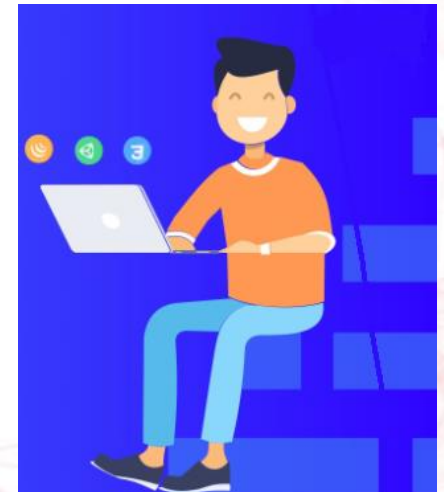
同一人的不同圖片的特徵基本**相同**。

經過大規模人臉資料訓練之後的那一層特徵會有很強的表達能力，每一個數字都可能是一個很強的特徵。

這些特徵可能對應著人臉中，
是男人還是女人，
是瓜子臉還是其他臉型，
皮膚是不是白的，
大多數的特徵，都是**神經網路自動學習**出來的。

SenseStudy課程平臺 “人工智能基礎（體驗版）” 實驗3 – 2 利用深度神經網路完成圖像分類

- 1、`cifar_train = data.get('cifar10-small', subset='train')`
`fig() + plot(cifar_train)`
- 2、`label_names = cifar_train.meta['label_names']`
`print(label_names)`
- 3、`net = CNNClassifier(in_shape=(32,32,3),backbone=ResNet(10),num_classes=10)`
- 4、`net.demo_train(cifar_train)`
- 5、`cifar_test = data.get('cifar10-small',subset='test')`
`img, label = cifar_test[300]`
`fig() + plot(img)`
`label_name = label_names[label]`
`pred = net.predict(img)`
`pred_name = label_names[pred]`
`print("Ground truth is {}, Prediction is {}".format(label_name,pred_name))`



SenseStudy課程平臺 “人工智能入門（下）”

實驗7 – 7 根據人臉進行性別分類

對於一個有60張人臉照片的訓練集：

```
train_dataset, test_dataset = load_gender_data()
name=train_dataset.name
for i in range(len(train_dataset.label)):
    img=read_gender_img(name[i])
    fig()+image(img)
    print('Label is:', train_dataset.label[i])
```

訓練一個分類器，並對 test_dataset 數據集中不同照片的性別進行分類。



ERROR! Predicted label is: 0 , but Ground Truth is: 1



SenseStudy課程平臺 “人工智能入門（下）”

實驗7 – 7 根據人臉進行性別分類



- 1、對於一個有10張戴眼鏡人臉照片和10張不戴眼鏡人臉照片的訓練集：
- 2、訓練一個分類器，並上傳任意照片對其戴眼鏡的屬性進行分類。