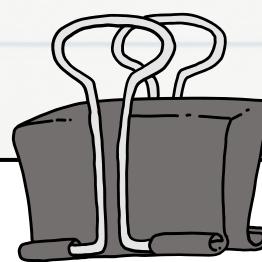
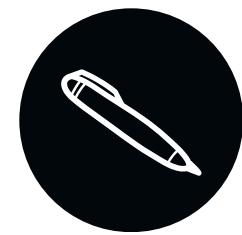


電腦視覺的進階技巧



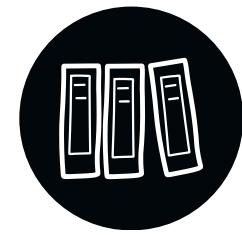


9-1 電腦視覺的基本任務



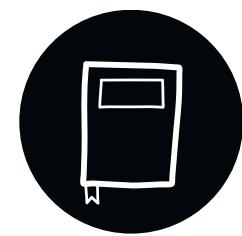
影像分類

為一張影像指定單一標籤或多標籤



影像分割

將一張圖片分割或劃分成不同區塊，
每個區塊都代表一個類別



物體偵測

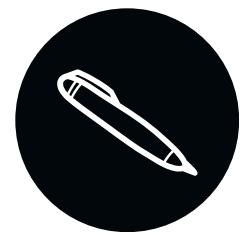
在影像中將我們要的物件周圍框出一個矩形（邊界框），
並為每個矩形標出一個類別





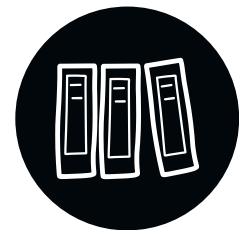


電腦視覺的小眾任務



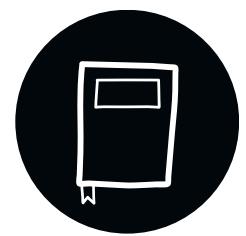
影像相似性評分

評估兩張影像視覺上的相似程度



關鍵點偵測

準確定位我們在影像中要找的屬性



姿勢預估

識別圖像或影片中物體或人體的位置、方向和姿勢



9-2影像分割

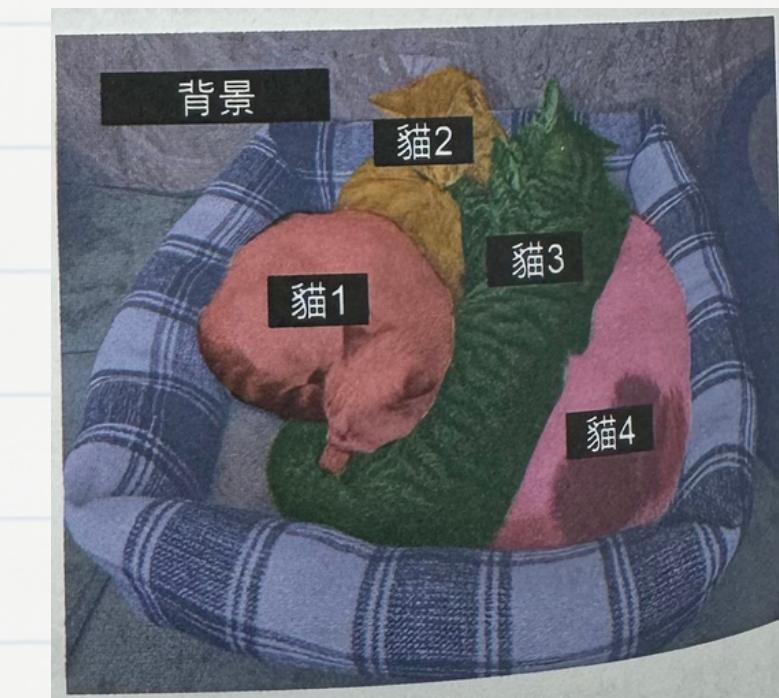
用模型為圖中的每個像素指定一個類別，從而將影像分割成不同區塊

- **語義分割**

將每個像素獨立歸入一個語意類別

- **實體分割**

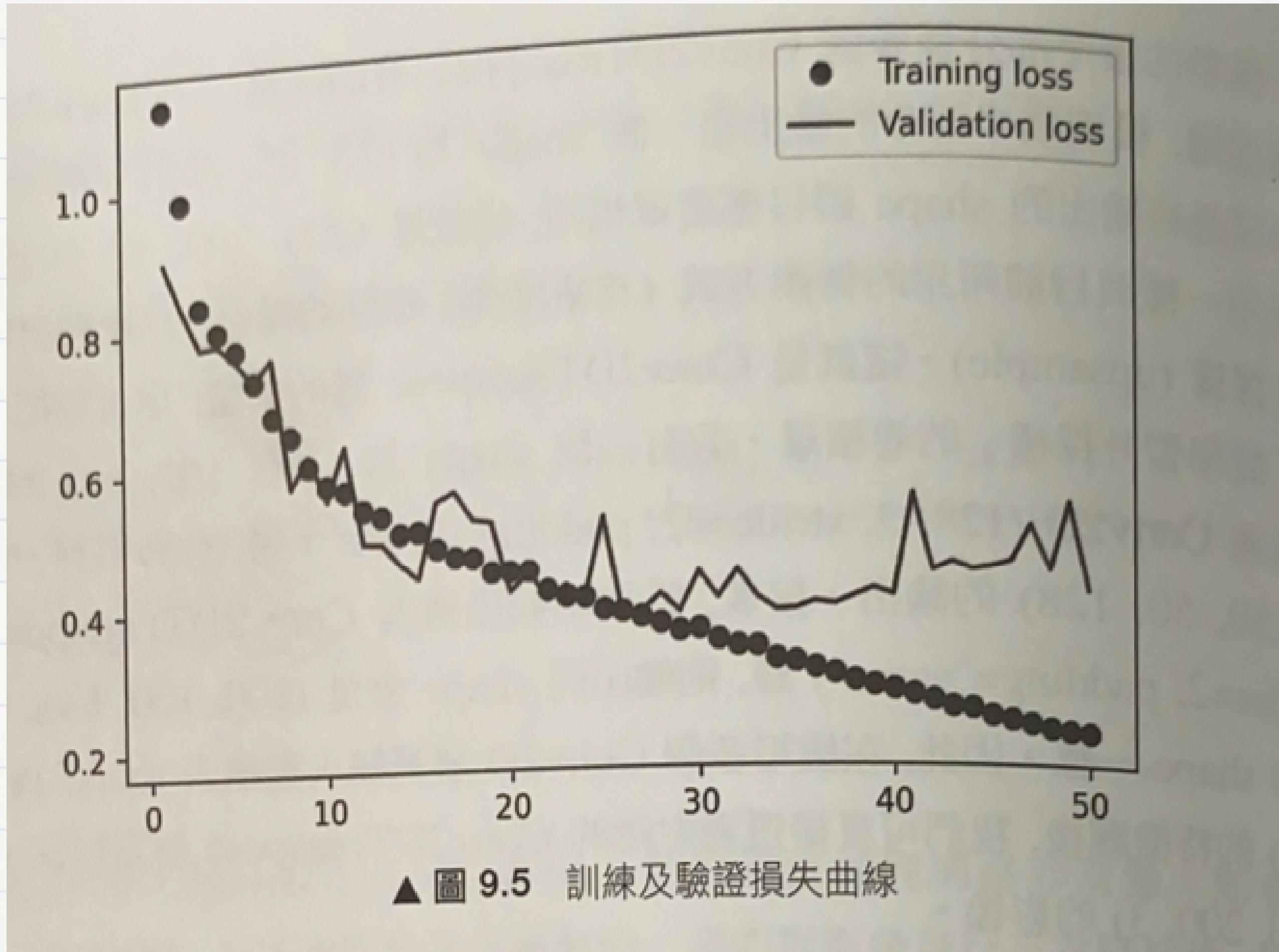
除了依類別將像素進行分類外，還要分析出獨立的物件實體

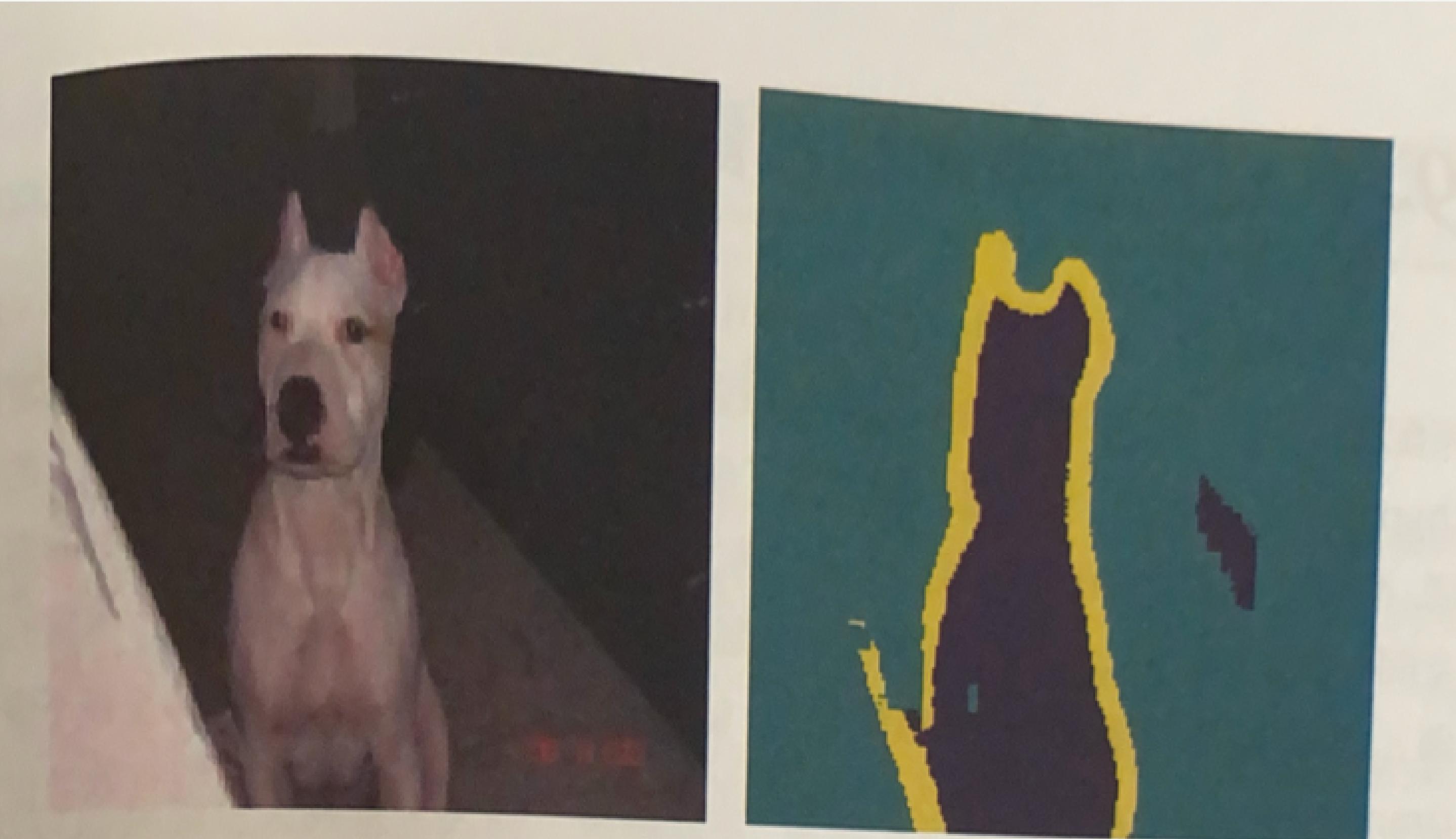




分割遮罩

相當於標籤，是一張與輸入影像大小相同的圖，有著單一的顏色channel，其中的每個整數值都對應於輸入影像中相應像素的類別





▲ 圖 9.6 測試影像及模型所預測的分割遮罩



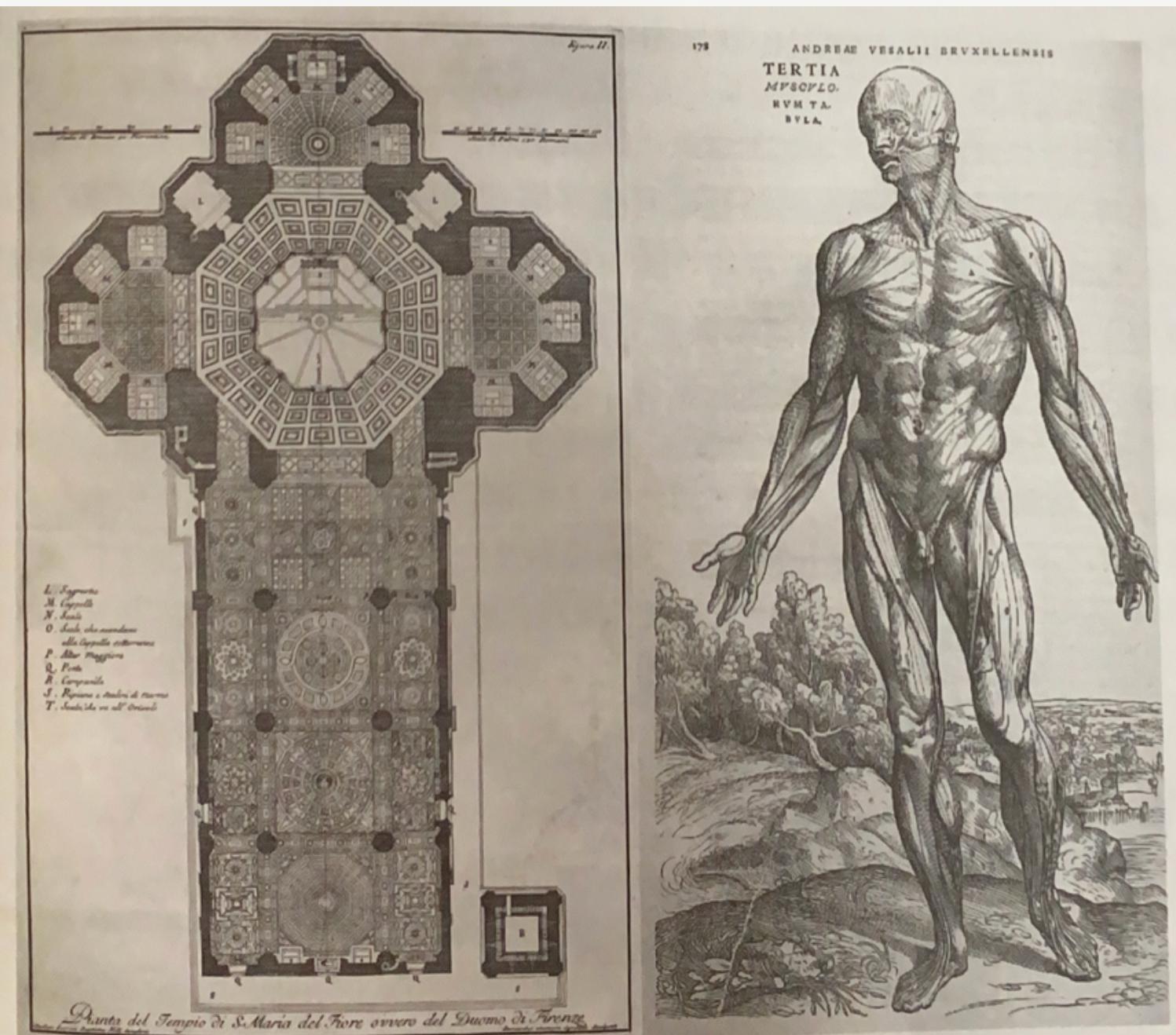
9-3現代卷積神經網路的架構模式

模型架構是創建過程中，所有選擇的結果。這些選擇定義了模型的假設空間。是一個包含了所有可能函數的空間，梯度下降，會在其中尋找出最合適的函數。和特徵工程一樣，一個好的假設空間可將問題及解法的先驗知識進行適當的編碼。必須對追尋的目標有所假設才能讓模型有效的從資料中學習。

好的模型架構：能加速學習、縮小搜尋空間

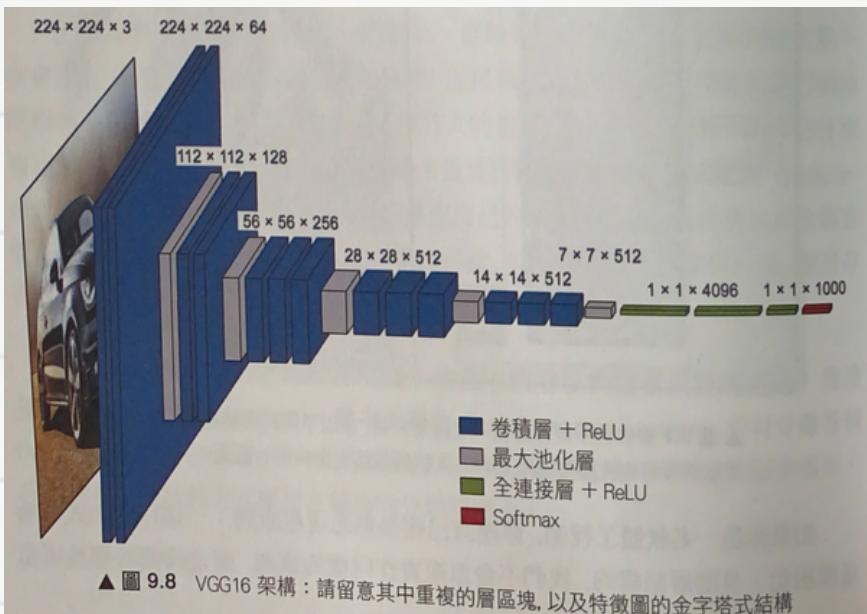
模組化、階層化、及重複使用

將複雜內容解構成模組→組織模組的階層結構→重複使用相同的模組



模組化、階層化、及重複使用

熱門的卷積神經網路架構都不僅是被架構成層，而是架構成重複的、一整組的層，大多是金字塔式的結構



▲ 圖 9.8 VGG16 架構：請留意其中重複的層區塊，以及特徵圖的金字塔式結構

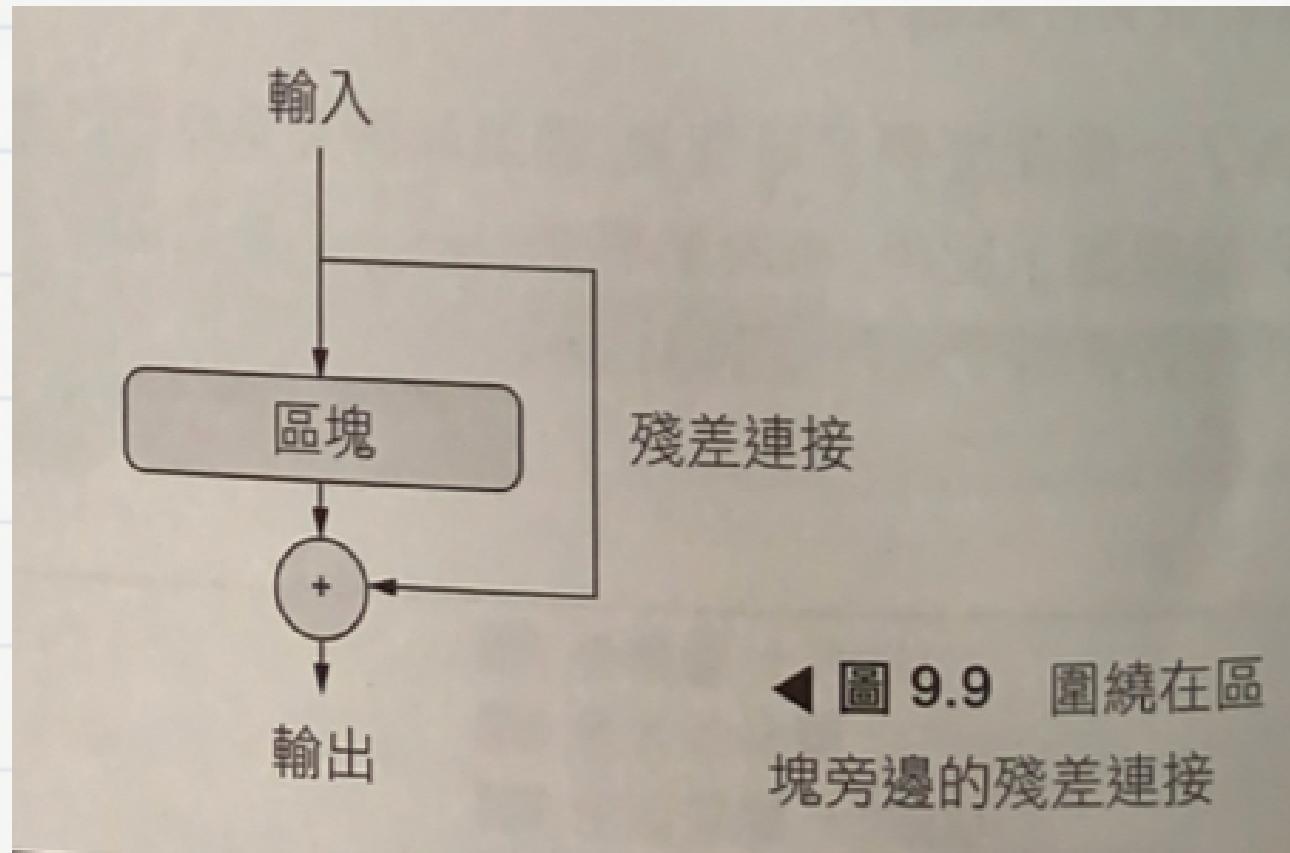
越深的階層結構越能促進特徵的重複使用
但因為梯度消失，堆疊的深度會受限制，因此要使用到殘差連接

殘差連接

梯度消失：序列式深度學習模型的反向傳播失去效用

$$y = f_4(f_3(f_2(f_1(x))))$$

解決方式：強迫每個函式都變得不具資訊破壞性



批次正規化

正規化：讓機器學習模型所看到的不同樣本能更相似

- 在訓練期間，批次正規化會用當前批次資料的平均值和變異數來進行正規化。另外還會以指數移動平均的算法，用每一批次的平均值和變異數來持續更新移動平均值及移動變異數，以供推論時期使用
- 在推論期間，由於無法取得足夠大的批次資料來計算具整體代表性的平均值與變異數，所以會改用訓練時所算出、具整體代表性的移動平均值及移動變異數來做正規化。

程式 9.4 不建議的批次正規化用法

```
x = layers.Conv2D(32, 3, activation="relu")(x)
x = layers.BatchNormalization()(x)
```

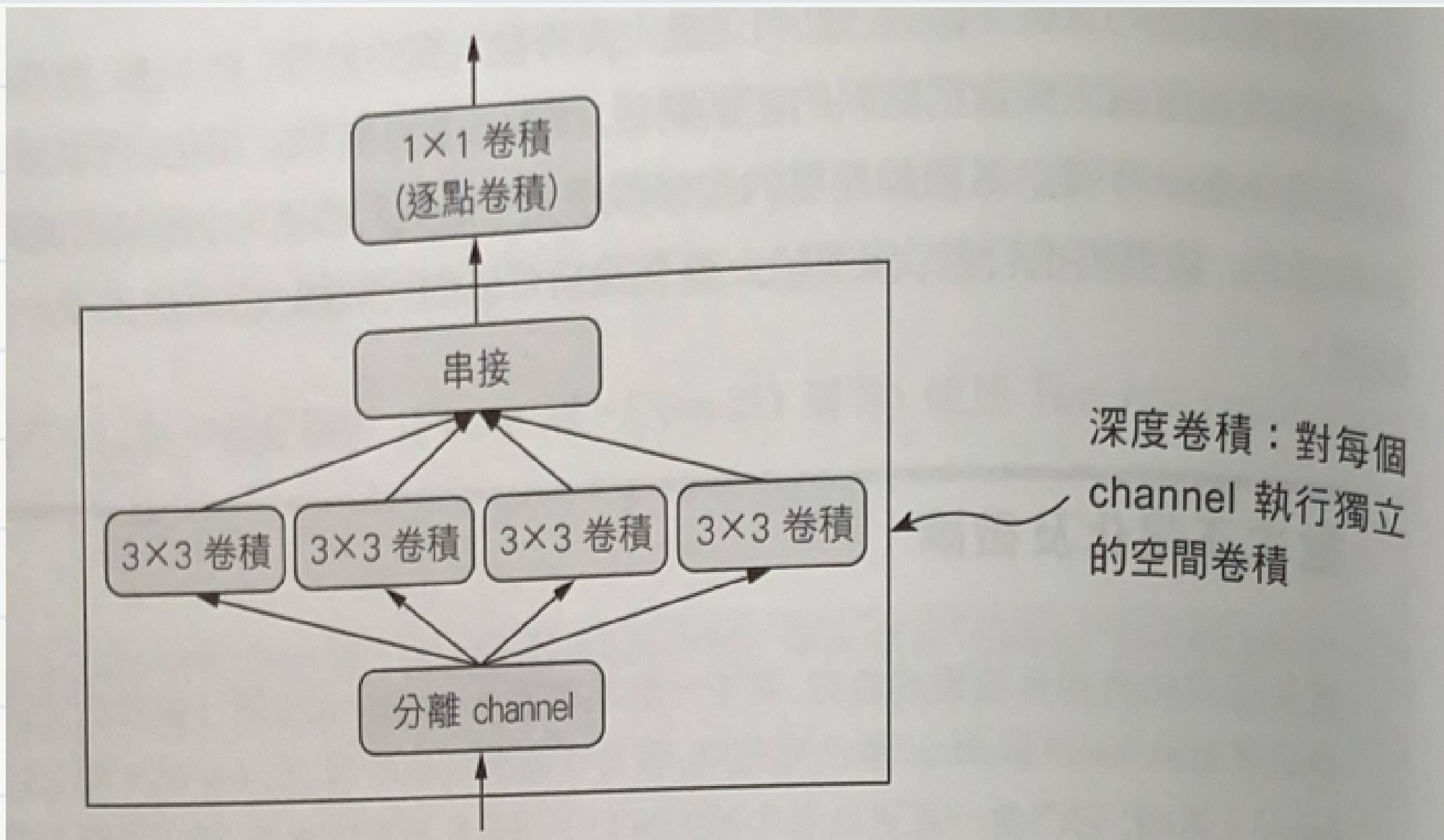
程式 9.5 建議的批次正規化用法

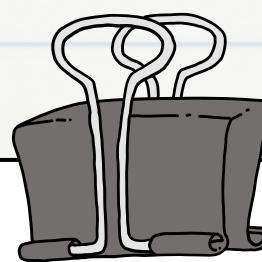
```
x = layers.Conv2D(32, 3, use_bias=False)(x) ← 此處先不使用激活函數
x = layers.BatchNormalization()(x)
x = layers.Activation("relu")(x) ← 經過正規化後，再使用激活函數
```

深度可分離卷積

取代Conv2D層，使模型更小，更加精簡，並且讓模型的表現提升

透過逐點卷積將輸出Channel混合之前，獨立地對輸入的每個channel 執行
空間卷積

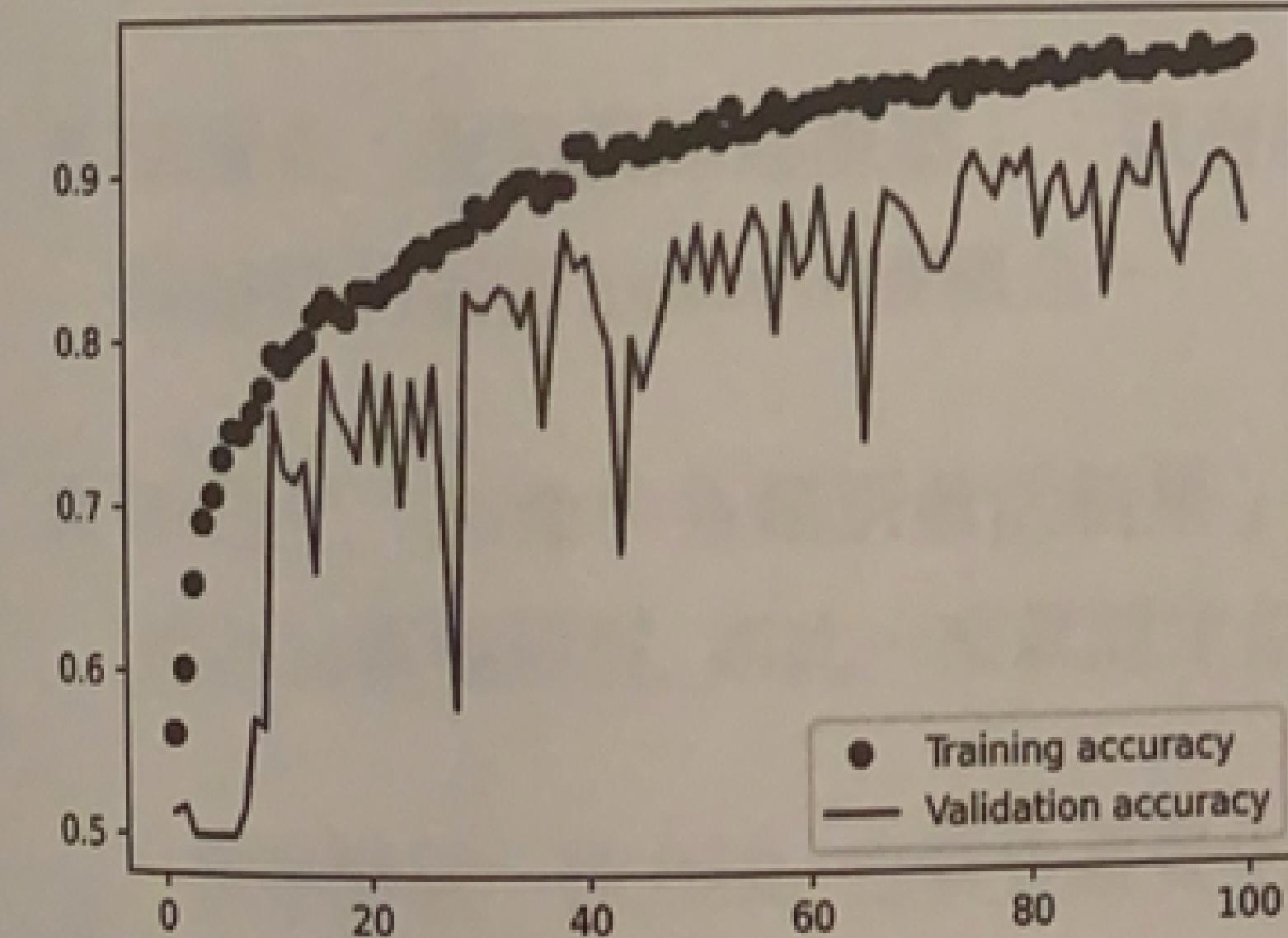




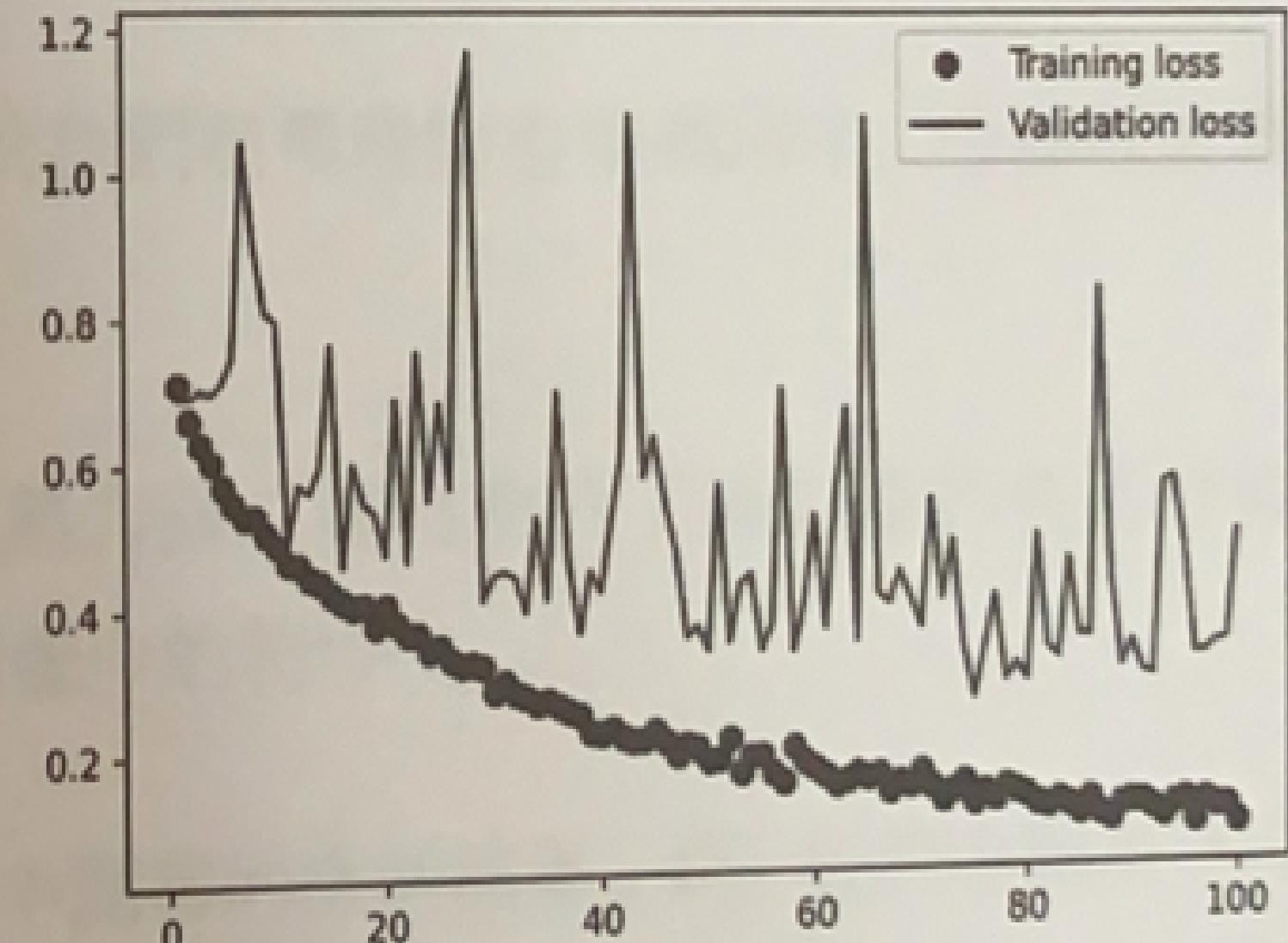
卷積神經網路原則

- 模型應該由重複的區塊組成，這些區塊通常包含多個卷積層和一個最大池化層。
- 層中的過濾器數量應該逐漸增加，而空間特徵圖的尺寸則逐漸變小。
- 深又窄的神經網路，比淺又寬的好。
- 在區塊旁邊使用殘差連接，有助於訓練更深的網路。
- 在卷積層之後使用批次正規化。
- 使用 SeparableConv2D 層來取代 Conv2D 層是有利的，因為其參數使用效率較高。

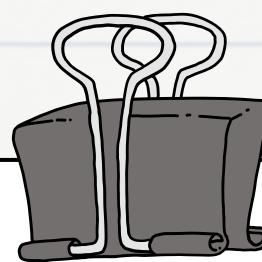
訓練和驗證準確度



訓練和驗證損失



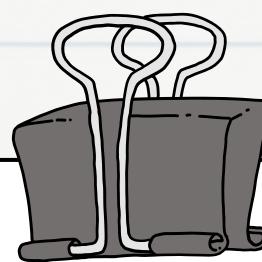
▲ 圖 9.11 Xception 架構的訓練和驗證評量指標



9-4 卷積神經網路學到了什麼？

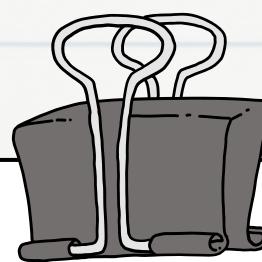
將卷積網路學習到的內容視覺化技術：

- 視覺化「卷積神經網路的中間輸出」
- 視覺化「卷積網路過濾器」
- 視覺化「影像中各類別激活結果」的熱力圖



視覺化中間層的激活結果

呈現模型中各卷積層和池化層在給定輸入後的傳回值，進而讓我們看到輸入是如何被分解到不同的過濾器做處理



視覺化卷積神經網路的過濾器

透過在輸入空間應用梯度上升，呈現每個過濾器會對什麼視覺pattern做出反應



區塊 2



區塊 4



區塊 8

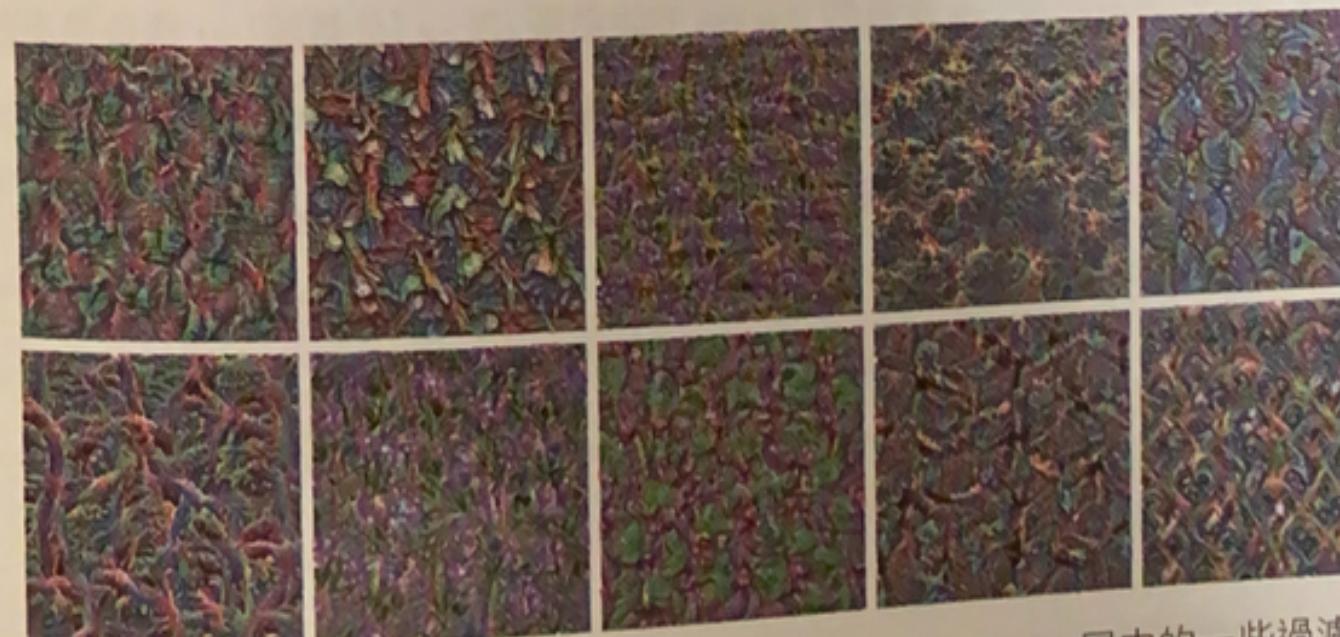


圖 9.17 block2_sepconv1 層、block4_sepconv1 層及 block8_sepconv1 層中的一些過濾器 pattern



視覺化類別激活熱力圖

可在輸入影像上生成類別激活結果的熱力圖。類別激活熱力圖是一個與特定輸出類別有關的2D分數網格，它會針對輸入影像中的每個位置做計算，並指出每個位置對特定類別的重要性。



◀ 圖 9.18
非洲象的測試圖片