

## 應用深度卷積神經網路建立電子商業服裝類型之分類系統

組別/編號:工業工程與管理組\_B1-3 學生: 王語喬 劉宇珊 余軒承 指導老師:鄭春生 教授

### 研究動機與目的

由於電子商務越發的流行，我們希望顧客能透過以圖搜圖的方式，來尋找到相似或相同款式的衣物。我們將利用網路查詢實驗資料，並參考過去相關研究，進一步探討如何改善缺點。在最後我們預期的成果為機器可以成功辨識圖像中衣物所代表的種類及屬性，並檢出相關圖片，且能夠應用在電子商務等，如 Fig 1 所示。

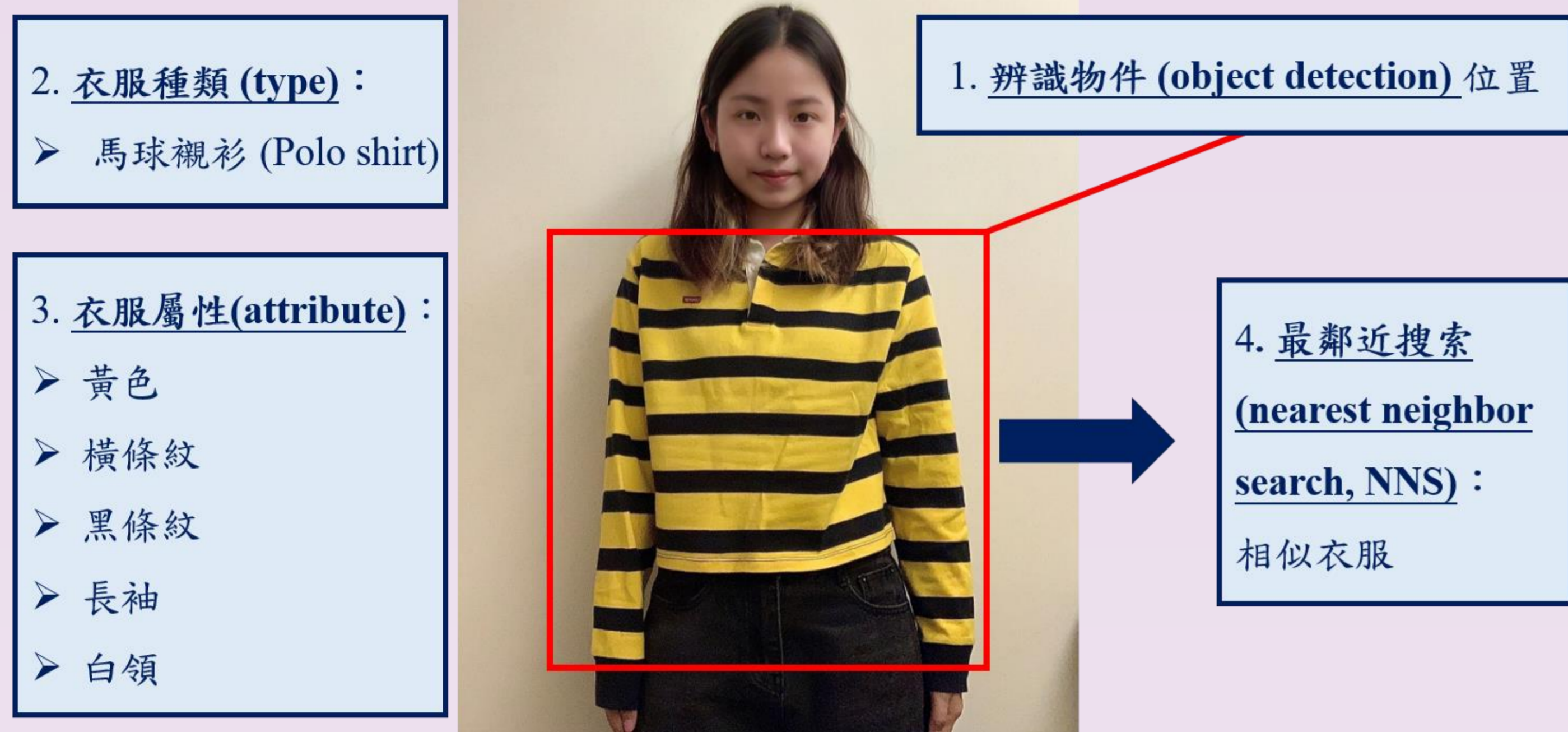


Fig1. 檢測圖像示意圖

### 研究方法

#### ● 卷積神經網路 (Convolutional Neural Network)

卷積神經網路為一種前饋神經網路，在處理各項大型圖像有非常良好的表現，在圖像方面主要有 4 個技術應用：物體定位、物體識別、目標分割、關鍵點檢測。而它的結構是由一個或多個卷積層 (convolutional layer)、池化層 (pooling layer)、激發函數 (activation function) 和全連接層 (fully-connected layer) 所組成，架構如 Fig 2 所示。

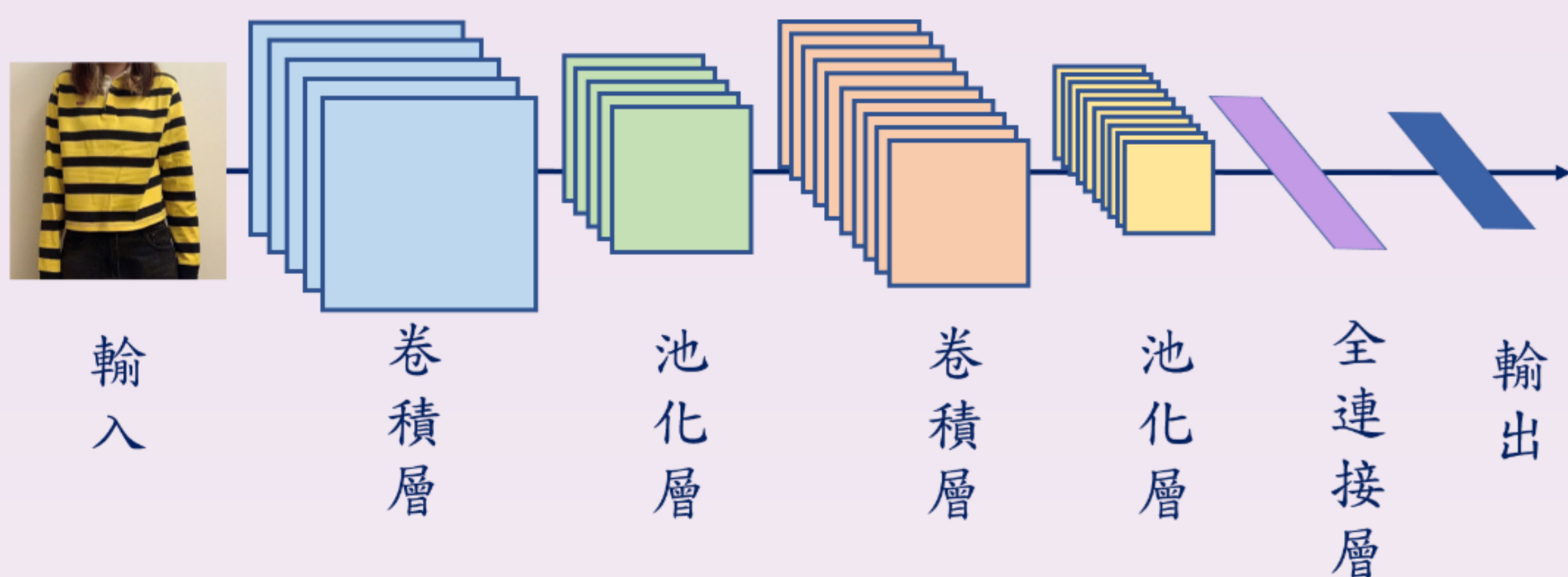


Fig2. 卷積神經網路基礎架構

#### ● 遷移學習

架構如 Fig 3 所示，卷積基底 (convolutional base) 其功用是將輸入的圖片資料進行特徵的擷取訓練；而分類器 (classifier) 將卷積基底訓練後的特徵進行分類。經由資料集大小與相似性的矩陣進行分類並進行微調，微調決策圖如 Fig 4 所示。

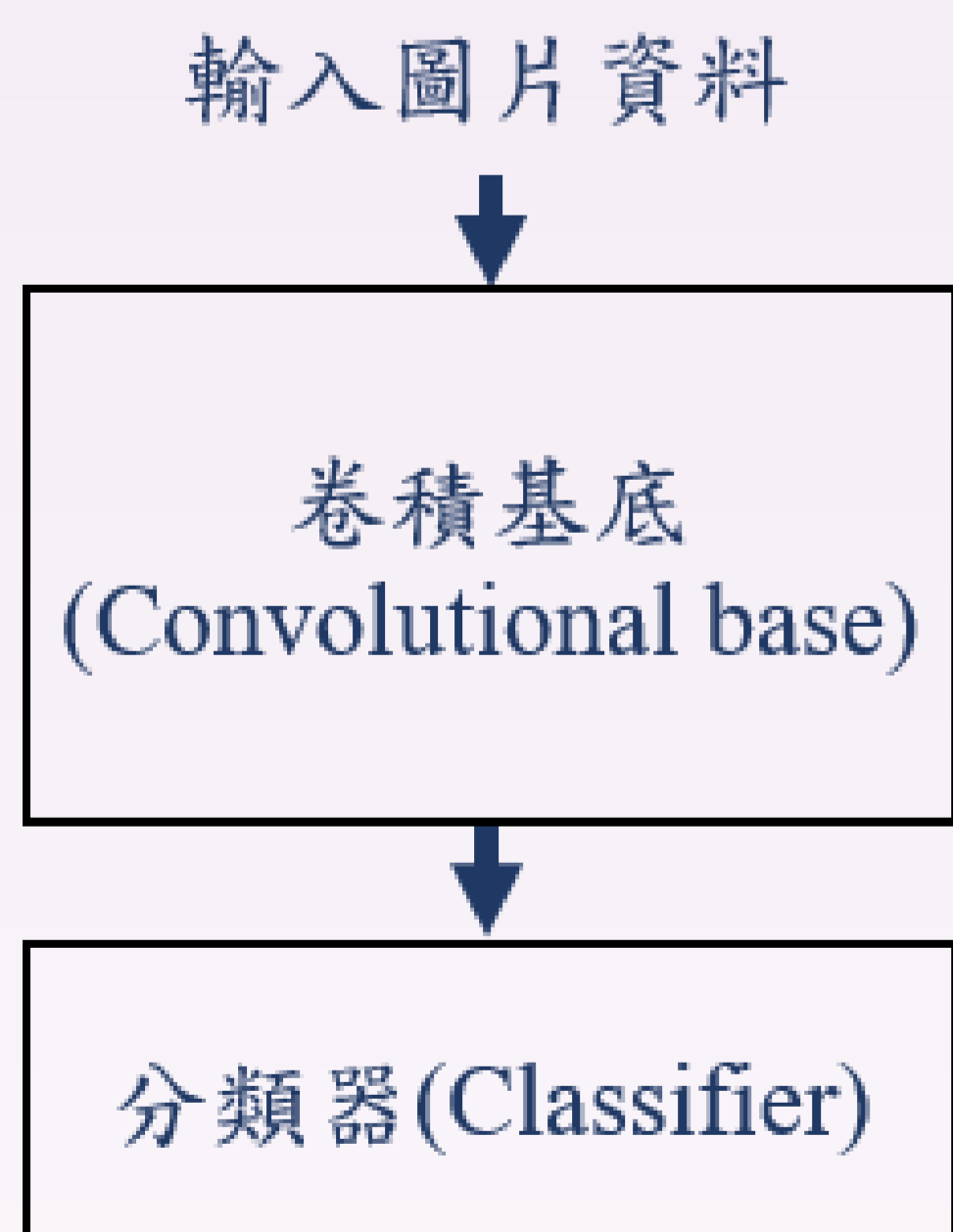


Fig3. 遷移學習架構圖

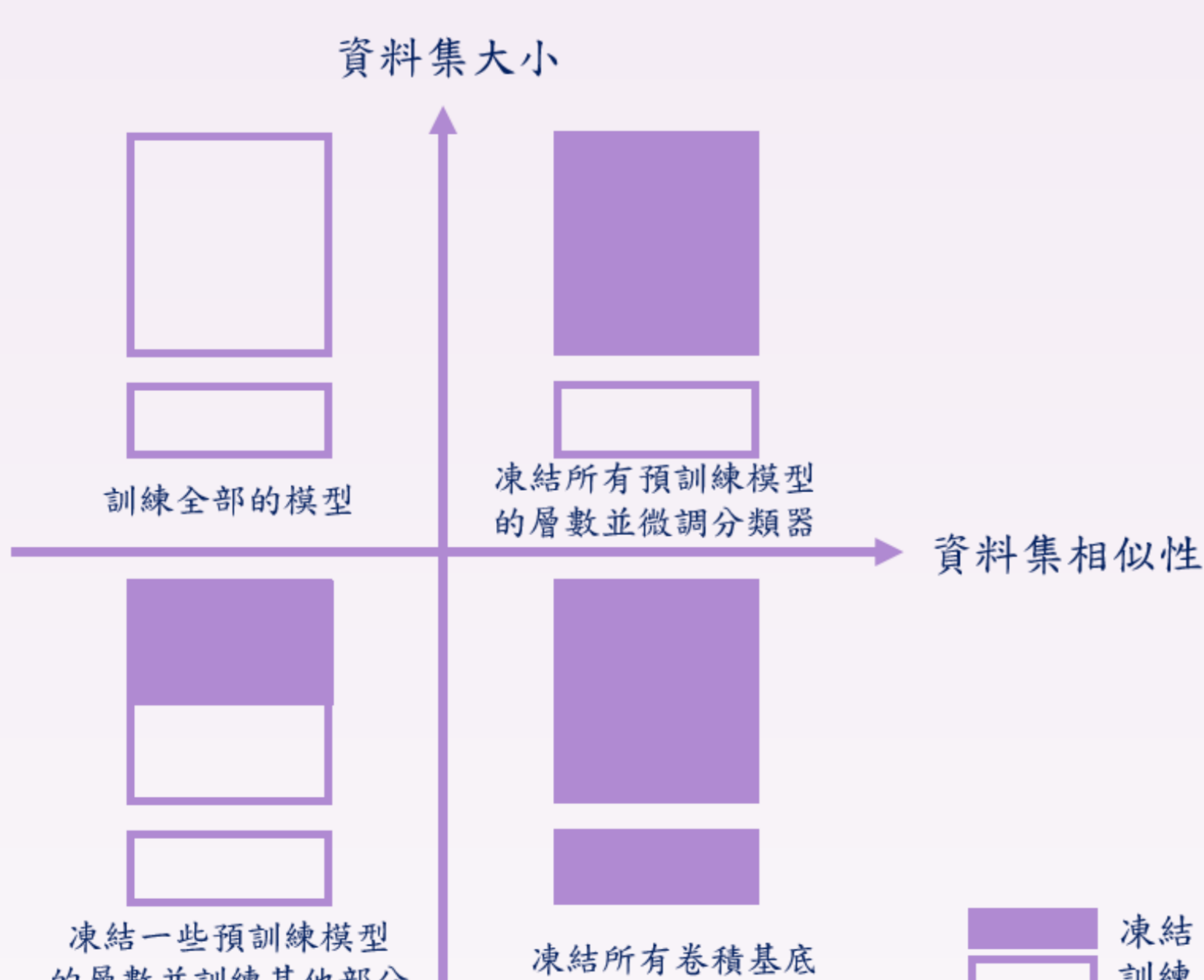


Fig4. 微調決策圖

#### ● Xception

Xception 利用深度可分卷積 (depth-wise separable convolution) 技術，先進行 1x1 卷積後再進行 3x3 卷積操作，其架構如 Fig 5 所示。

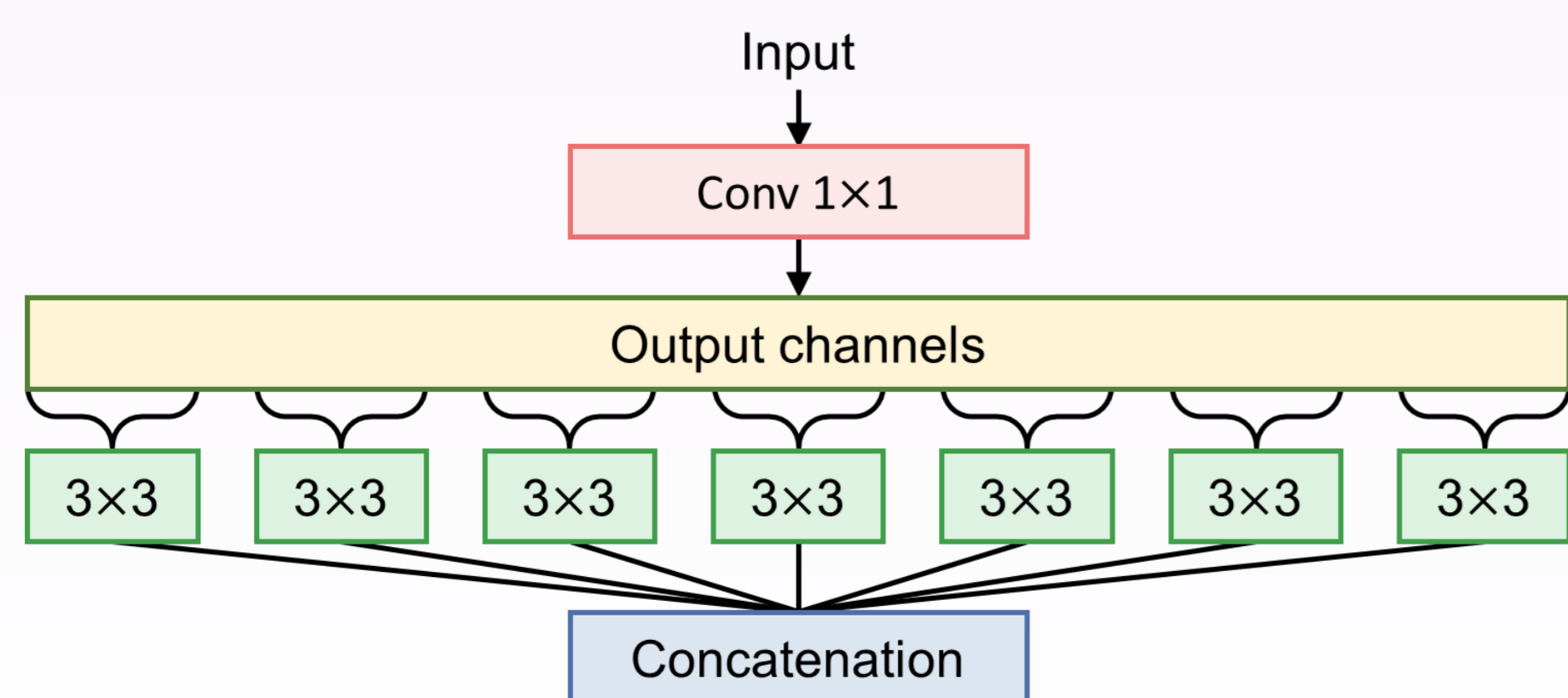
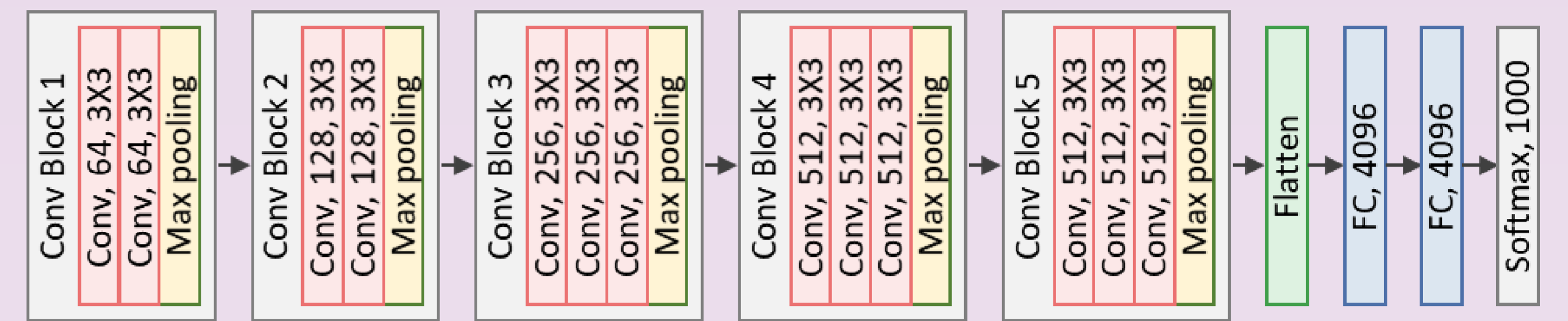


Fig5. Xception 架構圖

#### ● VGG-16

VGG-16 由 13 個卷積層和 3 個全連接層組成，並重複地結合許多 3x3 的卷積層和 2x2 的池化層，另外在每個池化層後都接著 ReLU 激發函數，過濾器的個數則隨著模型深度遞增，VGG-16 之架構圖如 Fig 6 所示。



Conv, a, b x b: 卷積層，過濾器個數為 a，過濾器大小為 b x b

Fig6. VGG-16 架構圖

### 研究績效評估

我們利用混淆矩陣作為分類績效的指標，其中 Tab 1 分類結果的數量，可以進一步計算不同的績效指標，較具代表性的指標為正確率 (accuracy)。

Tab1. 混淆矩陣

	預測為真	預測為假
實際為真	True Positive (TP)	False Negative (FN)
實際為假	False Positive (FP)	True Negative (TN)

$$\text{正確率: } accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$



Fig7. 各種衣物類別之範例

### 研究績效與結果

Tab2. 各類遷移學習模型之分類正確率 (%)

Statistics\Model	CNN	VGG-16	Xception
Maximum	85.48	94.95	96.04
Minimum	83.06	92.79	95.14
Average(%)	84.52	93.64	95.50
StdDev	0.70	0.69	0.28

從上表中我們可以看出 VGG-16 及 Xception 之平均分類正確率皆明顯高於 CNN，分別提高了 9.12% 及 10.98%，而 Xception 又相比 VGG-16 提高了 1.86%。

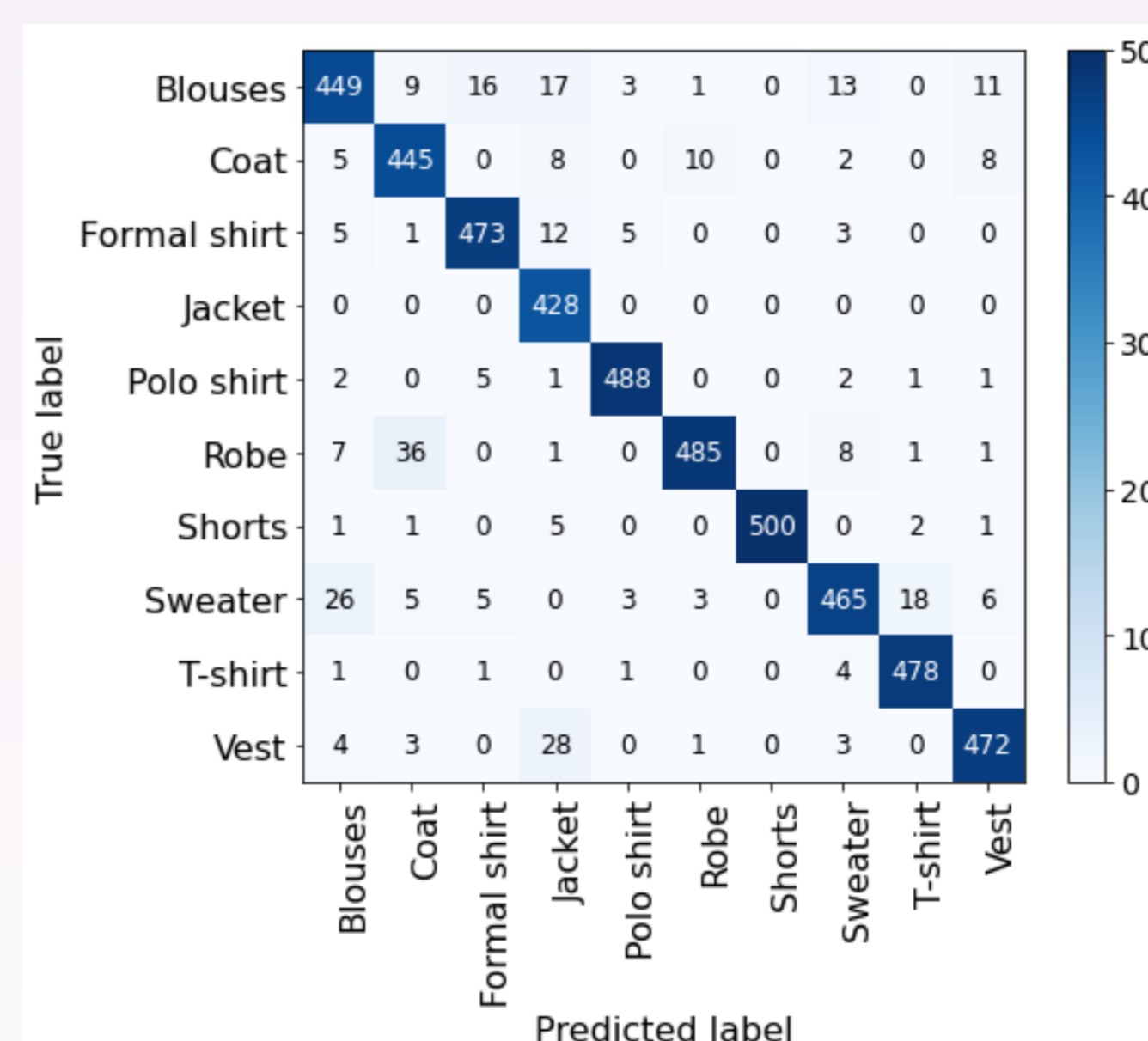


Fig7. VGG-16 之混淆矩陣

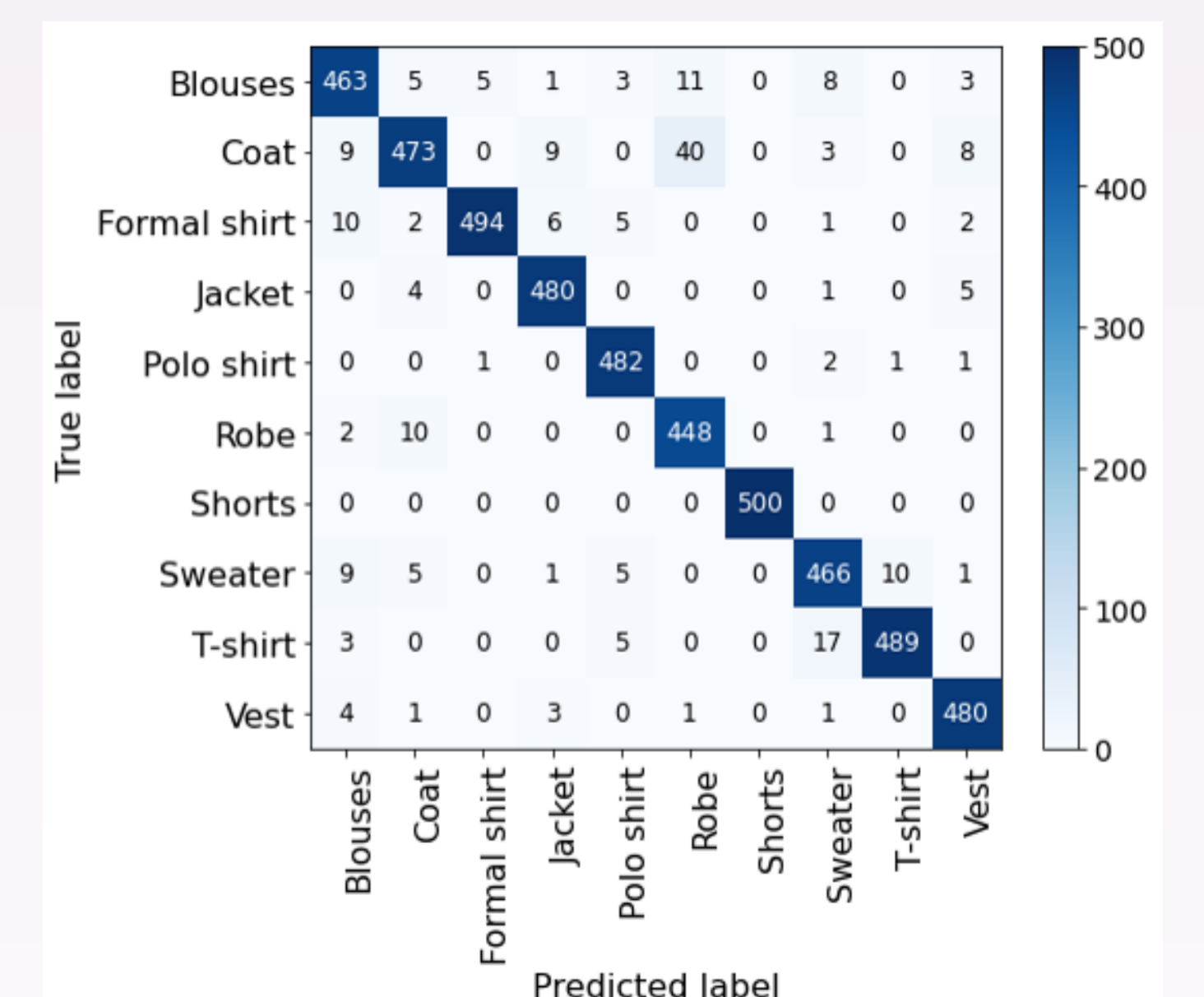


Fig8. Xception 之混淆矩陣

從 Fig 7 與 Fig 8 的混淆矩陣結果可以得知，Xception 除了對長袍類別的分類能力較弱，其他類別之正確率則與 VGG-16 相當或者更高，且從 Tab 2 中可以看到 Xception 平均正確率高達 95.5%，因此這項技術是可以運用在電子商務及社群媒體等行業，並用來預測消費者行為並推薦適當產品，讓顧客能夠更快速取得想要的資訊。未來研究中也可以將深度卷積神經網路應用於圖像資料中的屬性分類或檢索，讓衣物圖像分類的用途更加廣泛。