

# Deep Learning - Assignment I

RE6121029 張立勳

**關鍵字**—HoG, ORB, BRISK, K-Means, K-NN, SVM, CatBoost.

## I. 介紹

本次作業為一機器學習模型應用於圖像多分類之實作結果，其流程主要可分為三部分：1. 資料前處理 2. 特徵提取 3. 模型訓練。第一部分讀取圖像、調整圖像大小、轉換成灰階圖像；第二部分使用 HoG, ORB, BRISK 三種特徵提取工具並利用 K-Means 分群將所有特徵向量調整成相同長度；第三部分使用第二部分所得之新特徵向量訓練 K-NN, SVM, CatBoost 三種分類模型並比較模型準確度和 F1-score。

資料來源：<https://www.dropbox.com/t/tpNFrucmj4GQQV6d>

## II. 使用方法

### A. HoG

在一圖像中，局部目標的表象和形狀能夠被梯度或邊緣的方向密度分布很好地描述。具體的實現方法是：首先將圖像分成小的連通區域，我們把它叫細胞單元。然後採集細胞單元中各像素點的梯度的或邊緣的方向直方圖。最後把這些直方圖組合起來就可以構成特徵描述器。

### B. ORB

這是一種在圖像處理和電腦視覺領域常見的特徵提取演算法。它結合了快速特徵檢測 (FAST) 和二進制魯棒獨立特徵 (BRIEF) 兩種技術。ORB 演算法能夠快速地檢測圖像中的關鍵點，並生成描述這些關鍵點的特徵向量，用於識別、匹配以及追蹤目標。這種演算法具有較高的計算效率和穩定性，適用於各種實時圖像處理應用。

### C. BRISK

這是一種用於圖像處理的特徵提取演算法，能夠快速檢測圖像中的關鍵點並生成描述符。BRISK 利用尺度空間和方向來檢測關鍵點，並使用二進制描述符來描述這些關鍵點的局部特徵。它具有計算效率高、穩定性強的特點，適用於各種圖像識別和匹配應用。

### D. K-Means

這是一種分群演算法，用於將資料集分成 K 個不同的群。演算法根據資料點之間的距離將它們分配到最接近的群中心，然後更新每個群中心，直到收斂為止。

### E. K-NN

這是一種基於資料的監督式學習分類器。它通過測量新資料點與訓練資料集中已標記資料點之間的距離來進行分類。KNN 會找出與新資料點最相似的 K 個鄰居，然後利用它們的標籤進行預測。

### F. SVM

這是一種監督式學習分類器。它通過找到將不同類別分開的最佳超平面來進行分類。SVM 的目標是找到能夠使間隔最大化的超平面，從而增強模型的泛化能力。

### G. CatBoost

這是一種 Gradient Descent Tree 模型，適用於處理類別特徵的機器學習問題。它能夠自動處理類別特徵的編碼，無需預處理，並具有出色的結果和穩定性。CatBoost 採用 Symmetric Tree and Symmetric Leaf Splitting 技術，有效地減少了過擬合問題。同時，CatBoost 具有 Learning Rate Adjustment 和 Random Permutation 等功能，進一步提高了模型的準確性。

## III. 實作結果

本次作業為多元分類問題實作，目標是準確區分 5 類圖像。原始的訓練資料有 4980 筆、測試資料有 10 筆，經特徵提取 (HoG, ORB, BRISK) 並清除無法提取特徵的資料後，轉換後訓練資料以特徵提取工具為基準分別為 4980、4967、4952 筆，測試資料皆為 10 筆。經特徵提取後，利用 K-Means 分群方法將原本長度不一的特徵向量轉換成固定長度為 50 的特徵向量，作為新訓練資料。

### A. 模型結果

首先，讓所有模型用訓練集訓練，並觀察模型在測試集的表現。

#### 1) HoG

Model	Metrics	
	Accuracy	F1-score
K-NN Classifier	0.000	0.000
Support Vector Machine	0.005	$5.051 \times 10^{-5}$
CatBoost Classifier	0.005	$4.975 \times 10^{-5}$

Table 1. 各模型預測結果

#### 2) ORB

Model	Metrics	
	Accuracy	F1-score
K-NN Classifier	0.010	0.008
Support Vector Machine	0.005	0.005
CatBoost Classifier	0.005	0.005

Table 2. 各模型預測結果

#### 3) BRISK

Model	Metrics	
	Accuracy	F1-score
K-NN Classifier	0.015	0.008
Support Vector Machine	0.010	0.010
CatBoost Classifier	0.015	0.012

Table 3. 各模型預測結果

由 Table 1, 2, 3 可知，這三種模型都不能準確預測圖像。在這三種特徵提取工具之中，HoG 對於此資料集來說並不能有效提取對模型而言有用的資訊，相對地 BRISK 能提供給模型更有用的資訊。而以模型表現而言，CatBoost 是整體表現較好的。

## IV. 結論

Identify applicable funding agency here. If none, delete this text box.

考量到這三種模型表現都不好，而且不同的特徵提取工具在模型表現上也沒顯示出明顯的差異，也許可以嘗試其他模型或是增加 K-Means 的群數以增加 feature 數量，試圖找出更多有效資訊。