

Machine Learning 2020 - Homework 10 Report

學號：b08902100, 系級：資工一, 姓名：江昱勳

1. 任取一個 baseline model (sample code 裡定義的 fcnn, cnn, vae) 與你在 kaggle leaderboard 上表現最好的單純 autoencoder 架構的 model (如果表現最好的 model 就是 sample code 裡定義的 model 的話就再任選一個, e.g. 如果 fcnn 最好那就再選 fcnn), 對各自重建的 testing data 的 image 中選出與原圖 mse 最大的兩張加上最小的兩張並畫出來。(假設有五張圖, 每張圖經由 autoencoder A 重建的圖片與原圖的 MSE 分別為 $[25.4, 33.6, 15, 39, 54.8]$, 則 MSE 最大的兩張是圖 4、5 而最小的是圖 1、3)。須同時附上原圖與經 autoencoder 重建的圖片。(圖片總數: $(\text{原圖} + \text{重建}) \times (\text{兩顆 model}) \times (\text{mse 最大兩張} + \text{mse 最小兩張}) = 16 \text{ 張}$)

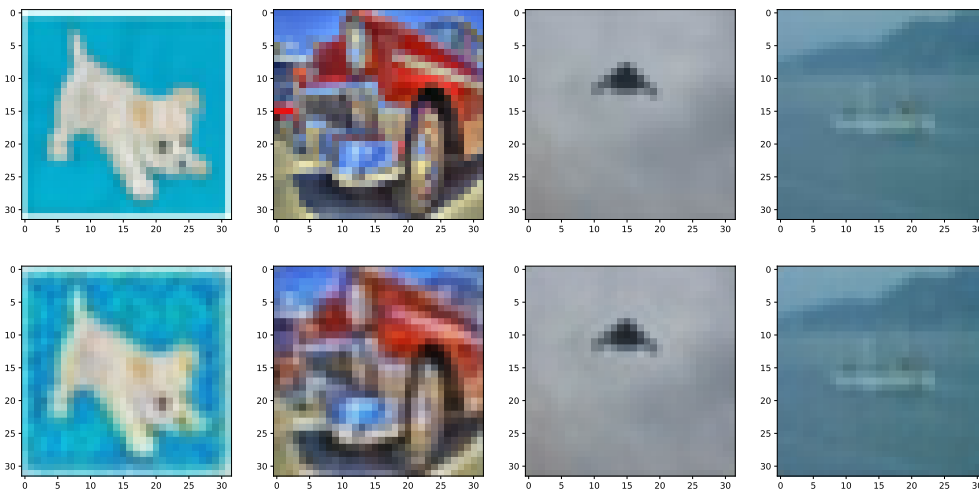


Figure 1: cnn 的 baseline model

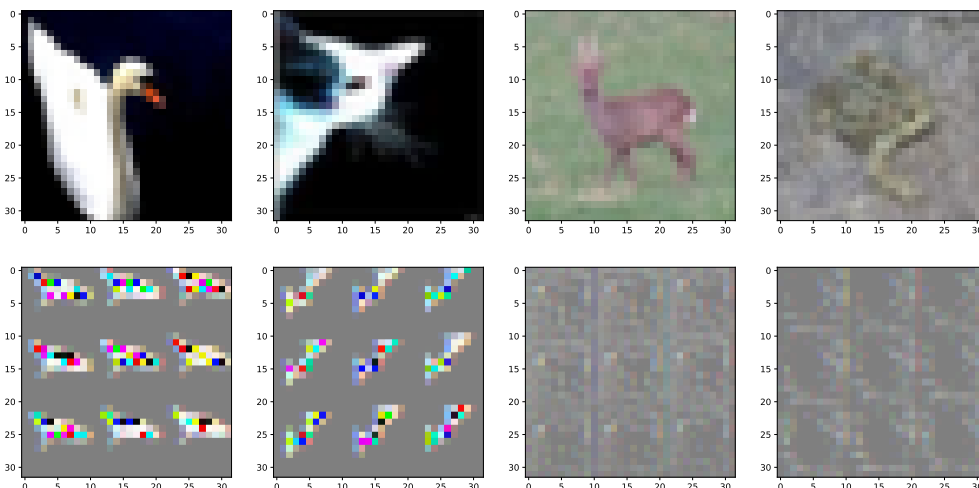


Figure 2: 最好的 autoencoder

上面的圖片前兩者皆是最差的兩個，後兩者皆是最好的兩個，而上方是原本的圖片，下面是 autoencoder 的輸出

2. 嘗試把 sample code 中的 K-means 與 PCA 分別做在 autoencoder 的 encoder output 上，並回報兩者的 auc score 以及本來 model 的 auc。autoencoder 不限。不論分數與本來的 model 相比有上升還是下降，請同學簡述原因。

在不使用 PCA 與 KNN 的情況下的 score 為 0.62262，若使用 PCA 降維成 16 維則 score 為 0.66144，使用 K-Means 分成 6 群的話 score 為 0.69602。可以發現將 PCA 應用再 autoencoder 的 encoder output 上可以有一定幅度的提升，可以猜測是因為 anomaly 與正常的東西的組成不太一樣，所以使用 PCA 投影回去時會有明顯的不同，另外從 K-Means 的結果我們更可以知道，這種不同更可能是某種距離上的不同，所以直接使用 K-Means 來找出比較遠的點的時候會有更好的效果。

3. 如 hw9，使用 PCA 或 T-sne 將 testing data 投影在 2 維平面上，並將 testing data 經第 1 題的兩顆 model 的 encoder 降維後的 output 投影在 2 維平面上，觀察經 encoder 降維後是否分成兩群的情況更明顯。(因未給定 testing label，所以點不須著色)

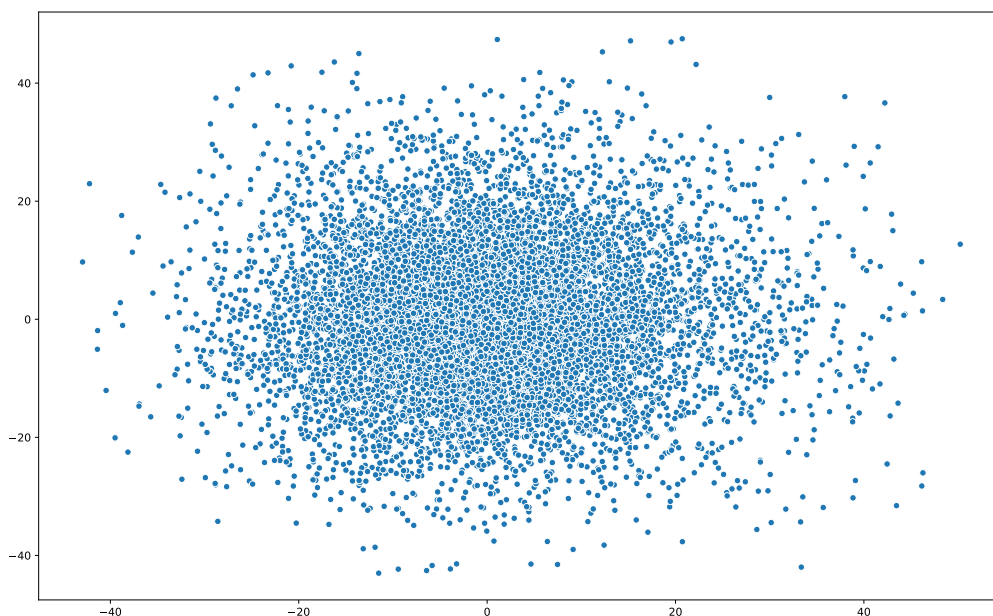


Figure 3: PCA on original image

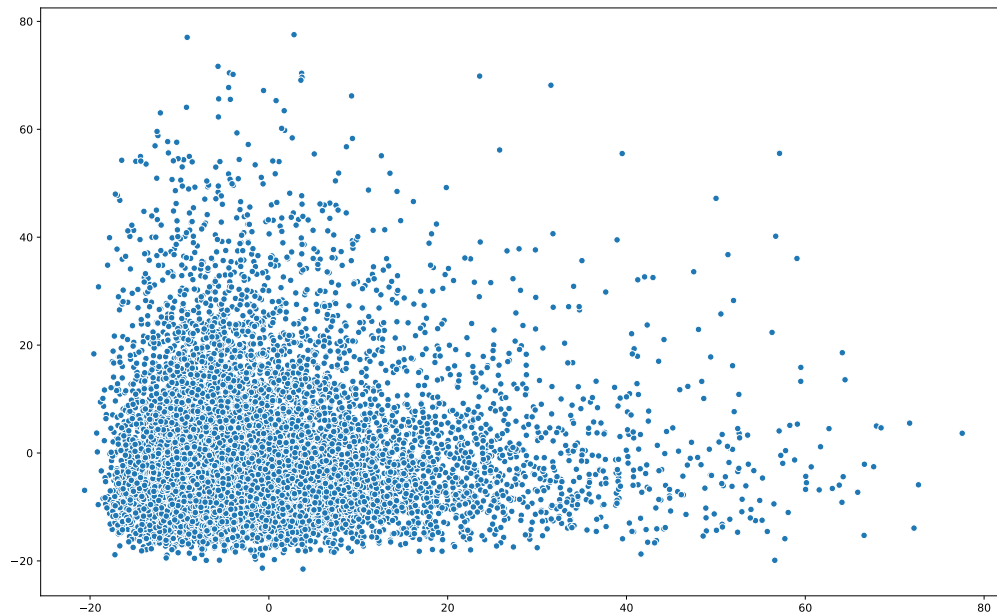


Figure 4: PCA on baseline model

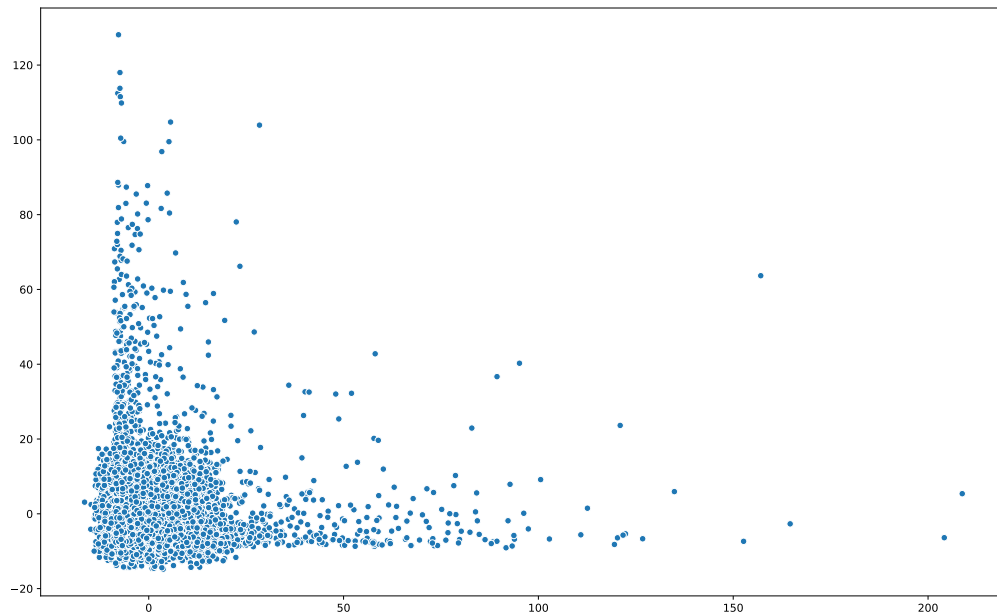


Figure 5: PCA on best model

可以發現越好的 model 會在降為後的圖形上有越集中於某一塊的取向。

4. 說明為何使用 auc score 來衡量而非 binary classification 常用的 f1 score。如果使用 f1 score 會有什麼不便之處？

因為助教有提到這次的 anomaly 其實不多，而若使用 f1 score 的話，很有可能直接輸出皆不是 anomaly 會有很好的效果，但是我們希望的是能偵測出 anomaly，也就是我們更偏向「寧可錯殺而不要放過一

個」的模式，因此採用 **auc score** 才比較容易達到如此的目標。

另一方面若使用 **auc score** 可以比較全面的評估一個模型的靈敏度 (Sensitivity) 以及特異度 (Specificity)，因為他比較像是各種參數下的 **F score** 的綜合。