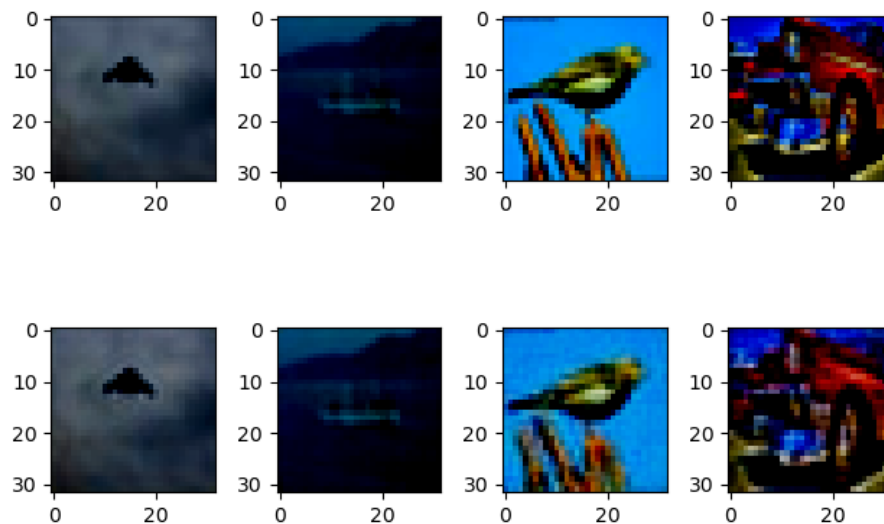


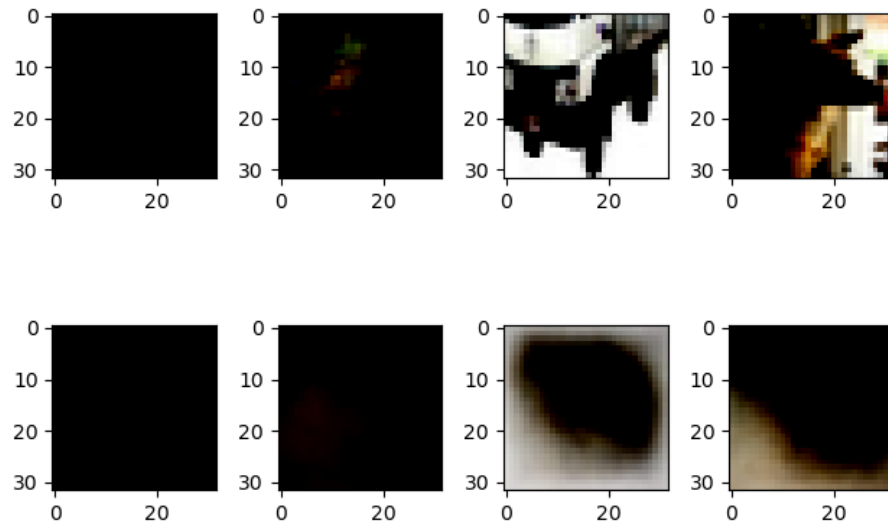
Machine Learning HW10

學號: B06902060 系級: 資工三 姓名: 鄒宗霖

1. (2%) 任取一個 baseline model (sample code 裡定義的 fcn, cnn, vae) 與你在 kaggle leaderboard 上表現最好的 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 cnn 最好那就再選 fcn) · 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: (原圖 + 重建) x (兩類 model) x (mse 最大兩張 + mse 最小兩張) = 16 張)



上圖為原始的 testing data 以及經由 (cnn baseline model's) autoencoder 重建後的圖片, 左邊兩列為 mse 最小的兩張圖、右邊兩列為 mse 最大的兩張圖, 我們可以看到儘管是 mse 最大的兩張重建的圖片, 被還原的程度仍然相當高, 也就是說不管是 normal 還是 anomaly image 都能被 autoencoder 重建 (mse 都很小), 因此這個 model 不適合拿來實作 anomaly detection。



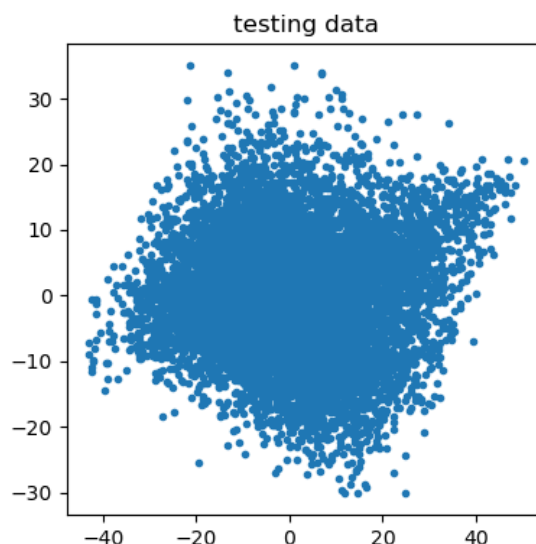
上圖為原始的 testing data 以及經由 (best model's) autoencoder 重建後的圖片，左邊兩列為 mse 最小的兩張圖、右邊兩列為 mse 最大的兩張圖，我們可以看到 mse 最小的兩張重建的圖片與原始圖片幾乎一樣 (都是接近全黑)；然而 mse 最大的兩張重建的圖片與原始圖片差距甚大，也就是說 autoencoder 不是每張圖片都能重建 (能重建的幾乎都是 normal image)，因此這個 model 比較適合拿來實作 anomaly detection。

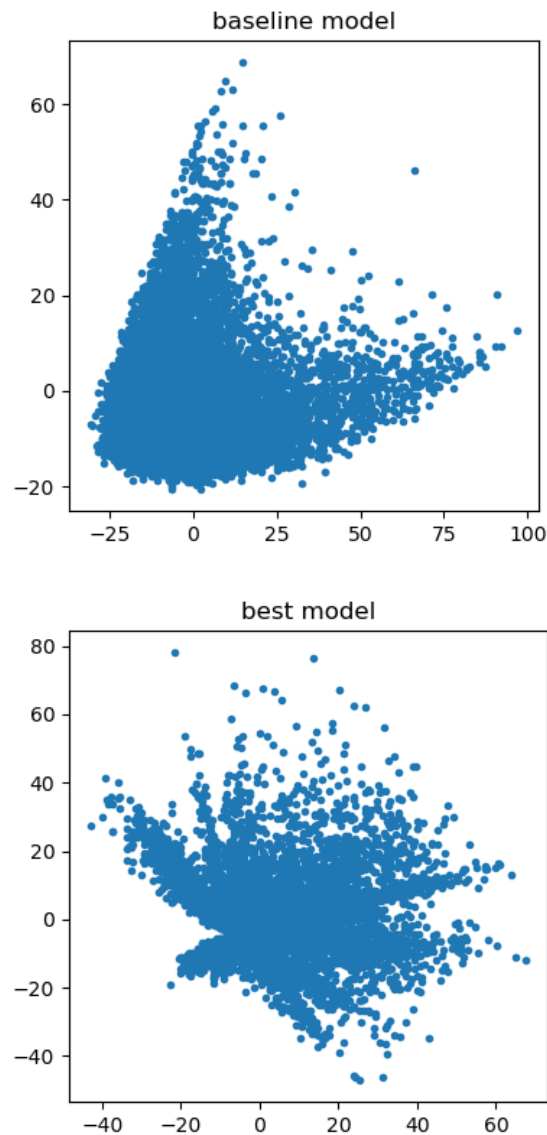
2. (1%) 嘗試把 sample code 中的 KNN 與 PCA 做在 autoencoder 的 encoder output 上，並回報兩者的 auc score

	KNN	PCA	baseline
auc score	0.62884	0.56228	0.55311

上表為把 KNN, PCA 實作在 autoencoder's encoder output 上的結果，可以看到實作後的 auc score 都比 baseline model 還高，可能的原因是 normal image's latent code 比較能代表那張圖 (圖片被 decoder 還原的程度較高)，且這些圖片來自於 training data 的某幾個類別，因此在降維後的空間中比較接近。

3. (1%) 如 hw9，使用 PCA 或 T-sne 將 testing data 投影在二維平面上，並將 testing data 經第一題的兩類 model 的 encoder 降維後的 output 投影在二維平面上，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)





上面三張圖分別為利用 PCA 將 testing data, baseline model's encoder output, best model's encoder output 投影在二維平面上的結果，可以看到從上到下資料點漸漸分成兩群。

4. (2%) 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？

$$f1\ score = 2 * \frac{precision * recall}{precision + recall}$$

$$presicion = \frac{true\ positives}{true\ positives + false\ positives}$$

$$recall = \frac{true\ positives}{true\ positives + false\ negatives}$$

上面幾條數學式子為 f1 score 的計算方式，在求出 f1 score 之前我們必須知道 presicion, recall 分別是多少，這就代表我們必須決定一個 threshold 來判定哪些預測結果為 anomaly、哪些預測結果為 normal；然而 auc score 的計算方式不需要決定 threshold 來區分 anomaly 以及 normal，他的橫軸代表 FPR (False Positive Rate)、縱軸代表 TPR (True Positive Rate)，只需要畫出 auc-roc curve 並計算曲線底下的面積即可。