

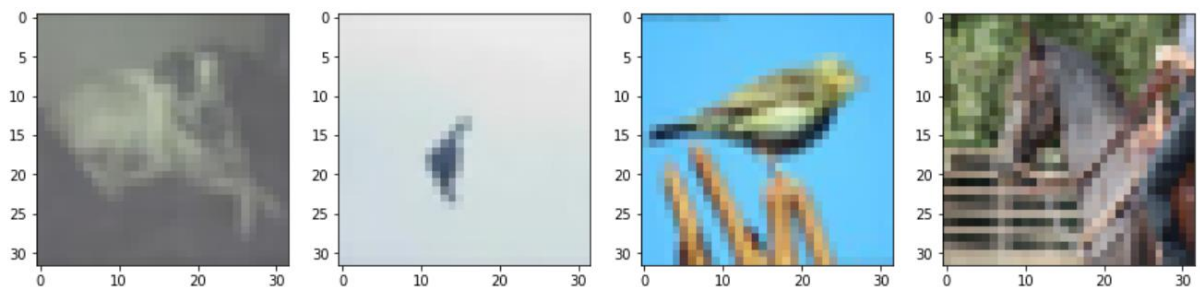
學號： b05901070 系級：電機四

姓名： 蔡昌廷

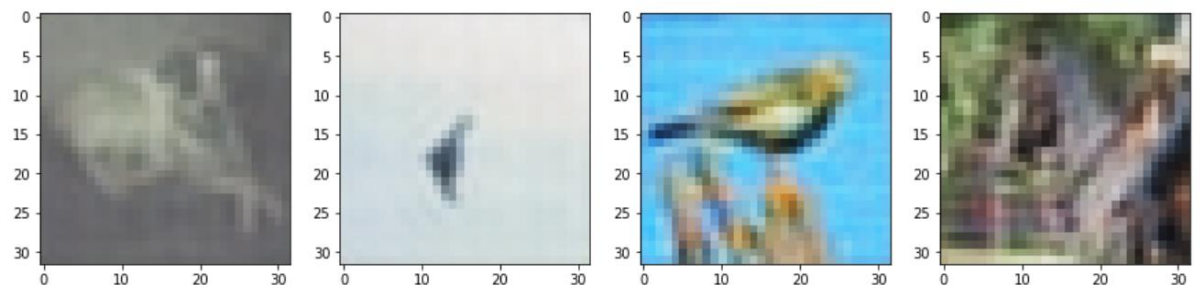
1. (2%) 任取一個 baseline model (sample code 裡定義的 fcn , cnn , vae) 與你在 kaggle leaderboard 上表現最好的單純 autoencoder 架構的 model (如果表現最好的 model 就是 sample code 裡定義的 model 的話就再任選一個, e.g. 如果 cnn 最好那就再選 fcn), 對各自重建的 testing data 的 image 中選出與原圖 mse 最大的兩張加上最小的兩張並畫出來。(假設有五張圖, 每張圖經由 autoencoder A 重建的圖片與原圖的 MSE 分別為 [25.4, 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)*(兩顆 model)*(mse 最大兩張+mse 最小兩張) = 16 張)

baseline CNN : (MSE : 1.59873、1.60379、74.9526、76.2217)

原圖

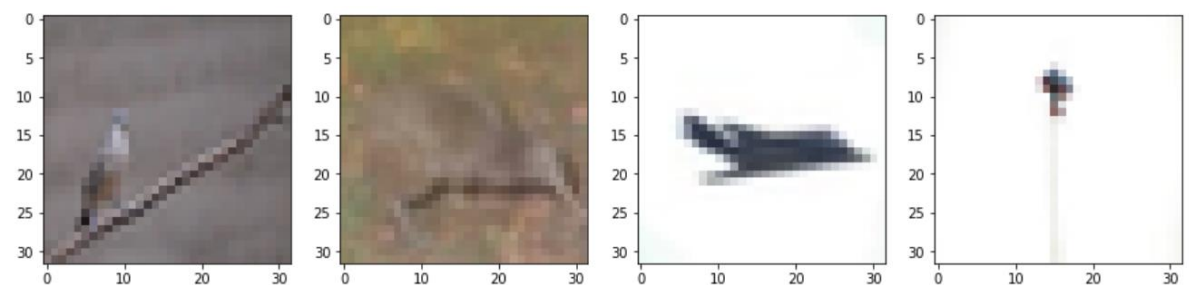


重建之後的圖

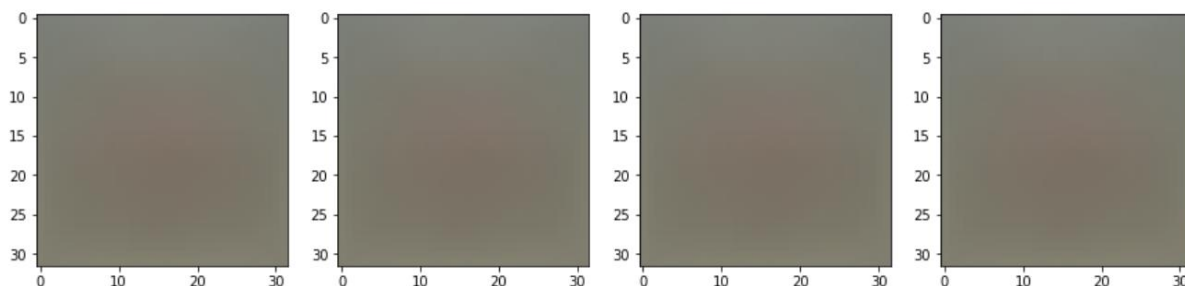


Best model : (MSE : 49.82854、59.5902、3064.3826、3429.0576)

原圖



重構後的圖



2. (1%) 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。

使用之前：0.61007

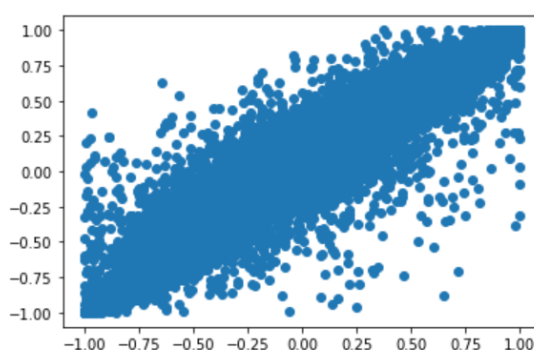
使用 K-means 之後：0.50688

使用 PCA 後：0.50817

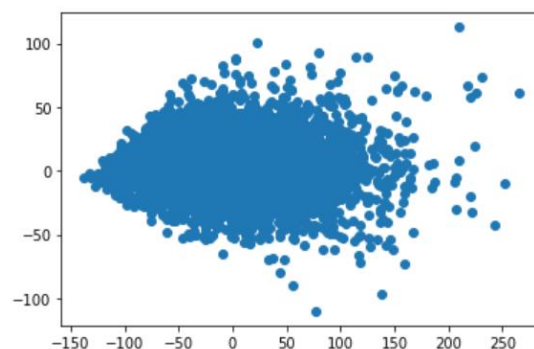
在此處我實驗過很多次，也用過不同種的方法會發現 K-means 跟 PCA 的效果弄起來都沒有比較好，我認為可能原因是因為圖片經過 encoder 之後已經降維成 100 維，因此對此 100 維再進行 PCA 或是 K-means 可能效果上並不會有顯著的改善。

3. (1%) 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）

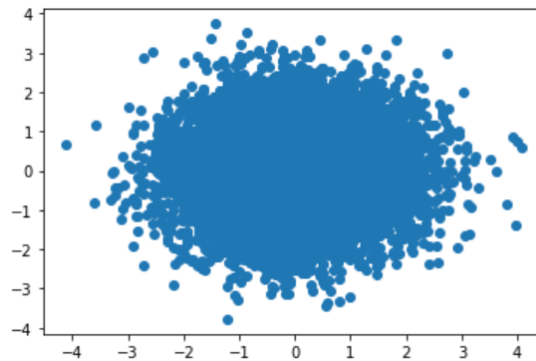
直接將 testing data 使用 PCA 降到 2 維



將經過 baseline CNN encode 後經過 PCA 降到 2 維



將經過 best model encode 後經過 PCA 降到 2 維



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？
- 此處使用 AUC score 最大的原因是因為 AUC 可以衡量此次分類問題是否能夠將兩種不同種類的資料區分的最開，AUC 會根據不同的 threshold 去計算其在 ROC 以下的面積，這種方法更適合去評估此次異常偵測的問題，如果是使用 f1 score 的有可能會導致兩個 model 一個 precision 很高，但 recall 很低，而另一個 precision 很低，但 recall 很高輸出一樣的 f1 score，這樣就比較難去評估哪一種 model 比較好。