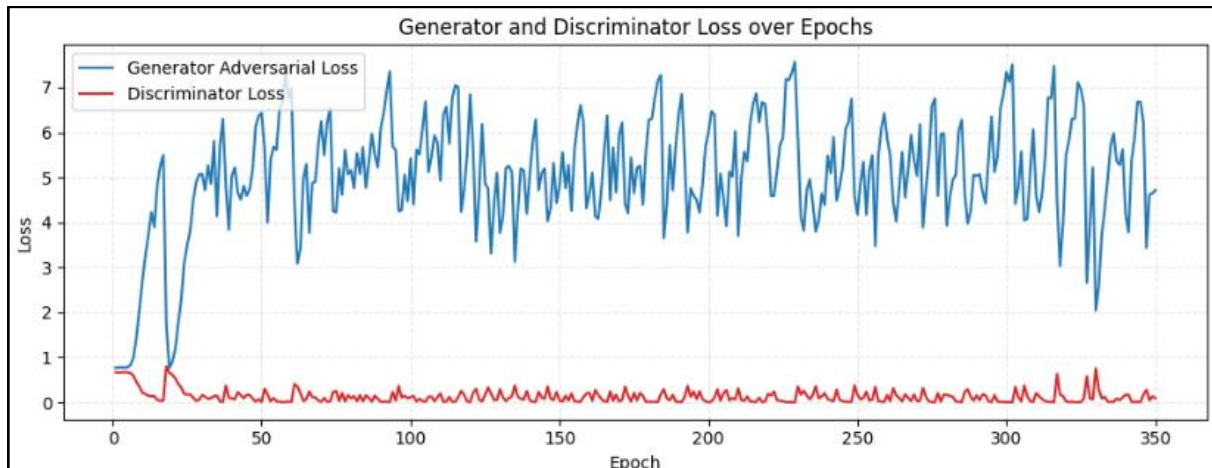


## HW2: GAN\_Pix2Pix

314581006 林鼎翔

### 1. Print the generator and discriminator training curve (epoch vs loss).



discriminator 一開始的平均 loss 由 0.56 到 0.094，顯示 discriminator 很快收斂並維持穩定。generator adversarial loss 由 1.20 到 5.20，代表在 discriminator 變強後 generator 需要付出更高對抗成本。Loss curve 顯示兩者後續訓練皆趨於平穩，無明顯的梯度爆炸且看起來有很好的在做對抗訓練。

### 2. Quantitative test result. Print PSNR SSIM, and LPIPS, and explain the representative meaning of PSNR SSIM, and LPIPS.

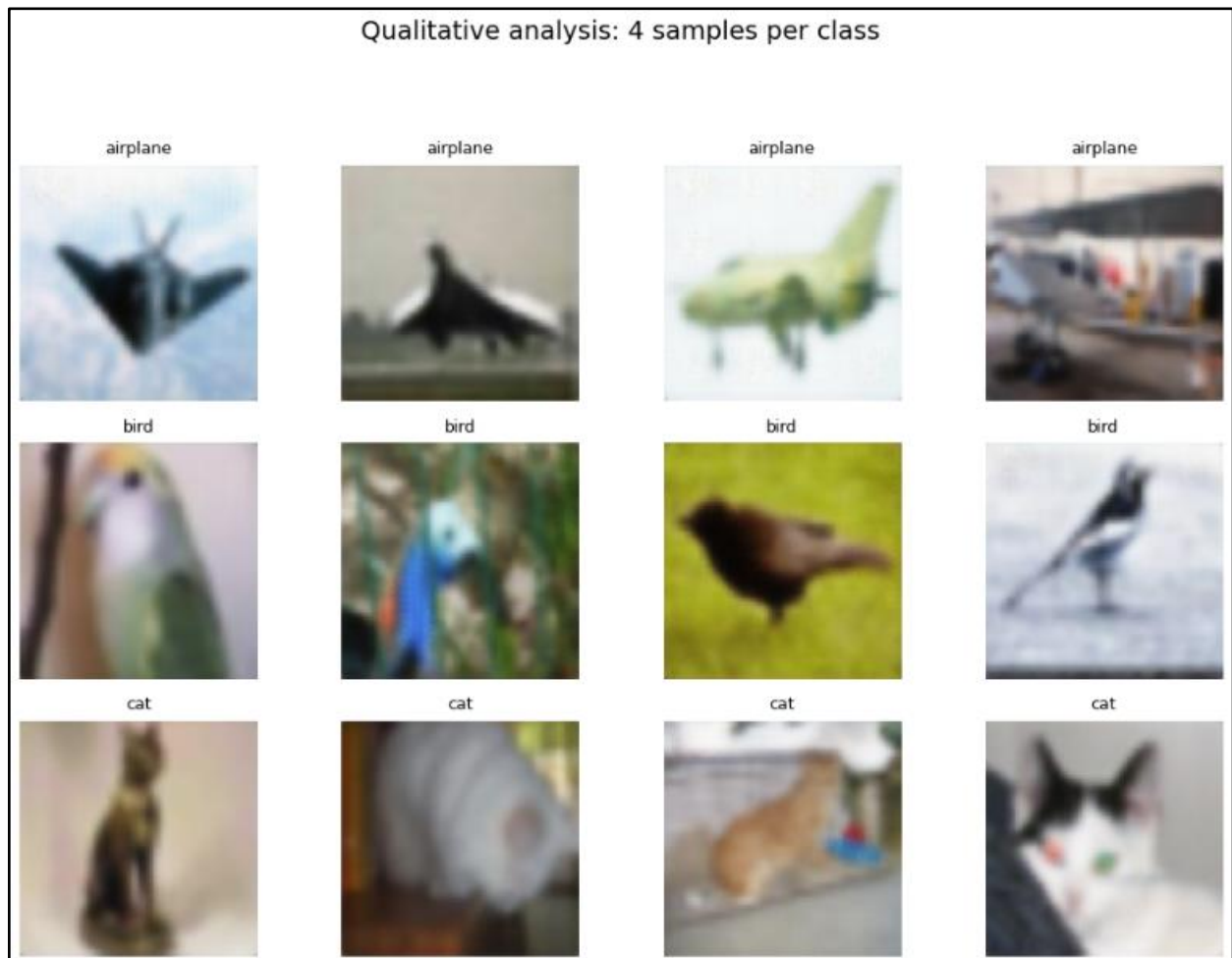
```
# Epoch = 350
evaluate()

Evaluation on 100 samples:
PSNR : 40.0898
SSIM : 0.9843
LPIPS : 0.0308
```

- PSNR (Peak Signal-to-Noise Ratio): 衡量重建影像與參考影像在像素層級的差異，數值愈高表示像素誤差愈小，常用來評估低階影像復原或壓縮品質。
- SSIM (Structural Similarity Index Measure): 著重亮度、對比與結構三個分量，模擬人眼對局部視窗結構的敏感度；值域 0~1，越接近 1 代表在結構與對比上越接近真實影像。
- LPIPS (Learned Perceptual Image Patch Similarity): 評估感知上的差異，值越小視覺上越相似；能捕捉人類主觀上更敏感的高階語意與紋理變化。

### 3. Qualitative result. Print 4 images of each inference from 3 categories.

下圖每列對應 airplane、bird、cat，每列 4 張樣本，展示訓練後能維持類別特徵且避免顏色崩壞。



- airplane：機體輪廓、翼尖和尾翼細節都可以很好分辨。
- bird：羽毛色塊平滑，可以分辨出胸腹部，鳥和背景也能明顯區別出來。
- cat：貓的形狀能明顯被區別出來。

### 4. Discuss the pros and cons of Pix2Pix compared with older methods.

Pix2Pix 和 Auto-encoder 比較的話：

- 優點：Pix2Pix 將訓練目標直接對應輸入與目標影像，可在生成過程中保留空間細節，對比傳統 auto-encoder 僅透過重建 loss，Pix2Pix 靠對抗訓練與 skip connections 讓輸出更銳利、能控制輸出風格；同時可以對不同條件（例如類別或語意圖）產生對應影像。

- 缺點：需要成對的資料集，收集成本高。對抗訓練較不穩定，需要平衡 G/D，像是我在做這份作業的時候，就發現在對抗訓練時，discriminator 太強，GAN 很難訓練起來，generator 產生不出很像的 data，最後我調整 loss function 和模型架構後，才成功訓練起來。而且模型參數量與計算量高於單純 auto-encoder，推論成本也較大；若 paired data 品質不一致，也容易造成 artifact 或 mode collapse，而 auto-encoder 在無監督情境下則是可以利用更多未標記的資料。