



GAN大(四)事

403410623 劉冠林
403410771 王柏翰
403410169 張軒齊



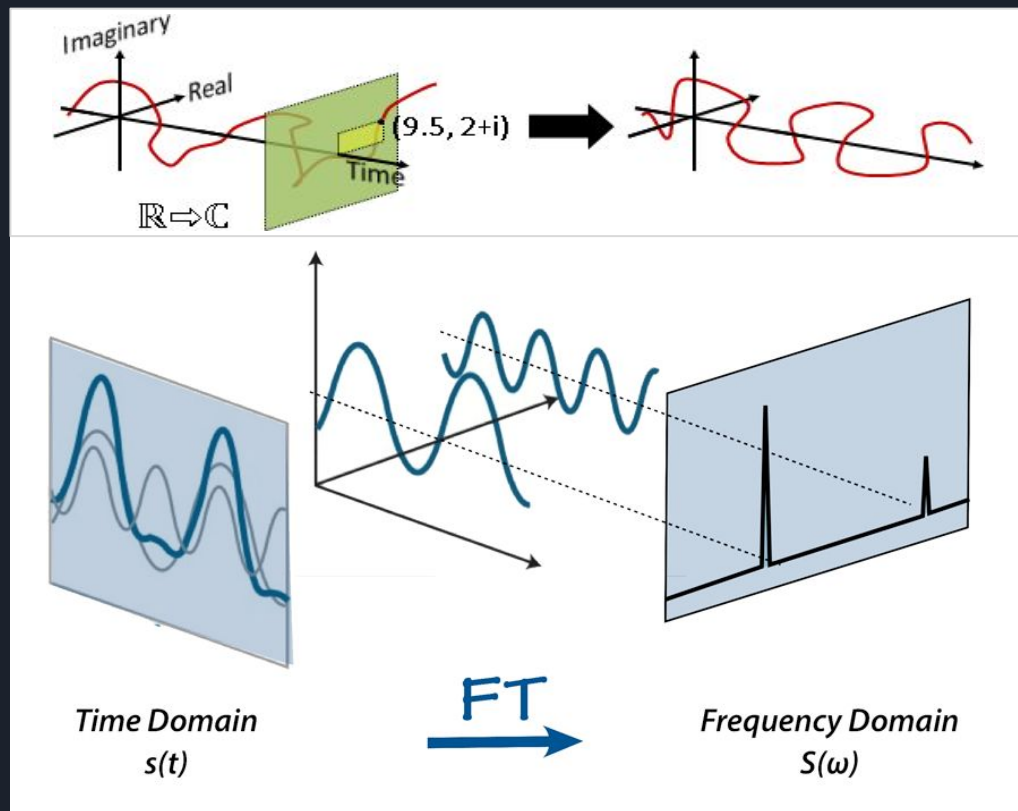
GAN聲音

主旨:

我們給他一個聲音，讓機器學習並模仿



Input: sound to 2d array (frequency, time)





音訊檔

=> 切成一單位440的資料100份($100 * 440$)

=> 傅立葉

=> 保留虛數部分

=> 正規化(位移,縮放)

=> 取正整數

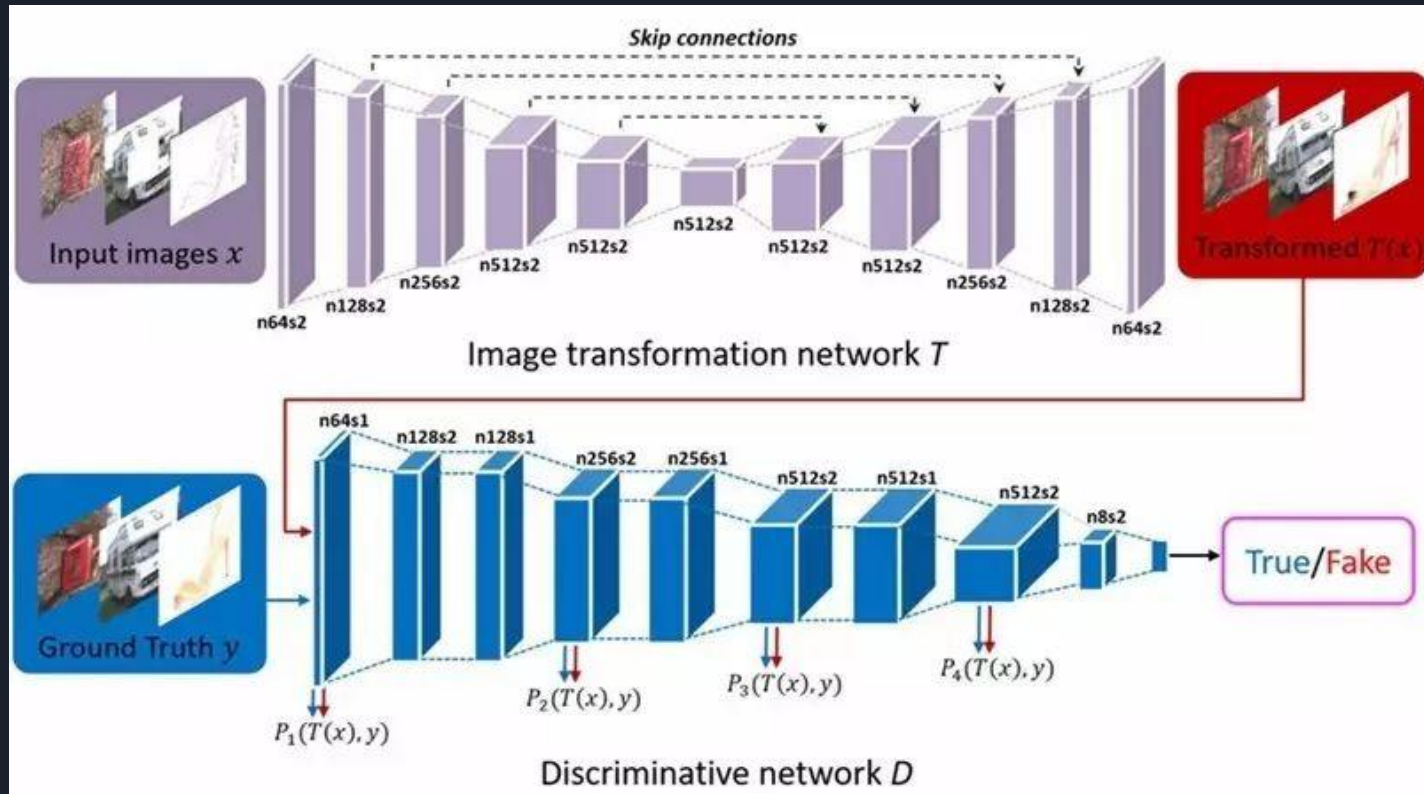
=> 訓練資料

murmur...

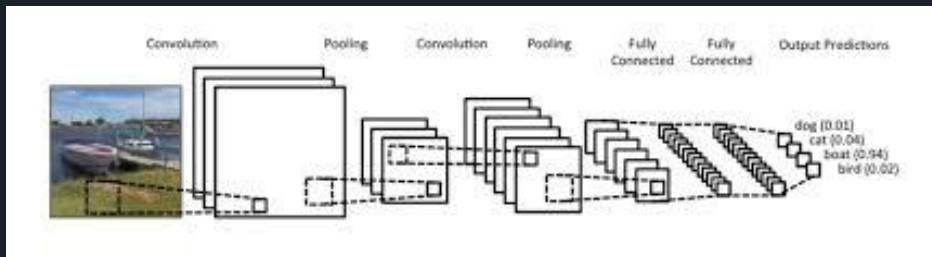
其實問題出在我們的 GAN 模型問題, 因為採用最基本的 CNN 本身就是拿來跑圖的運算, 因此我們只好轉成圖片格式。

如果時程夠多, 應該可以來改模型來做變化

GAN: 2 CNNs model to training data

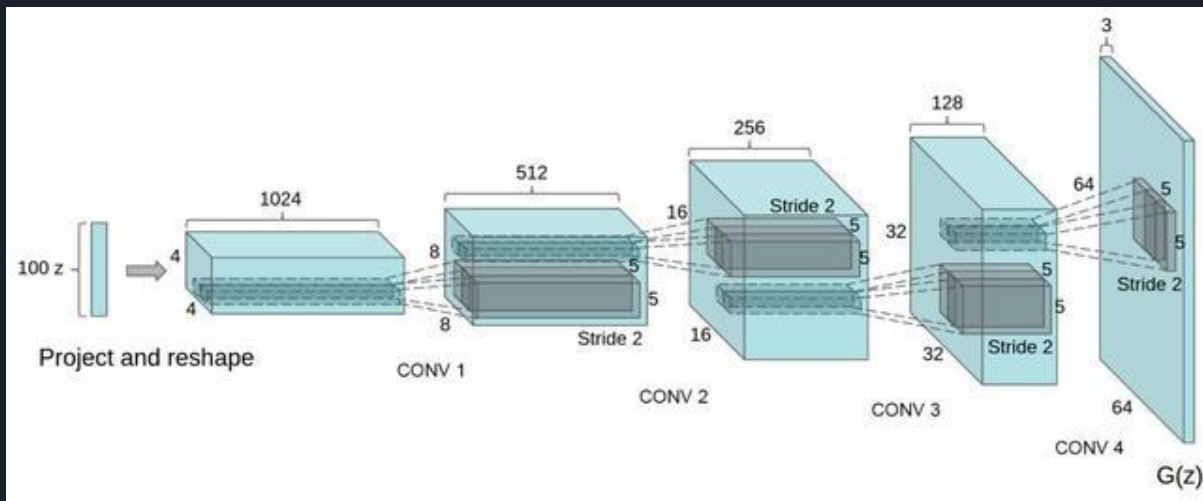


CNN



=> 偵探
(discriminator)

Inv CNN



=> 工匠
(generator)



Results



Discussion

1. 由於這次模型選用最基本的GAN，沒有參數的設置，所以只會學習越來越像的圖而已，而且我們沒有太深入了解GAN的運作，以至於沒有改變模型結構
2. 由於聲音轉圖片，會有失真與轉換圖片問題，聲音的取樣、區間、正規化都還得重新規劃
3. 沒想到要訓練這麼久QAQ



END