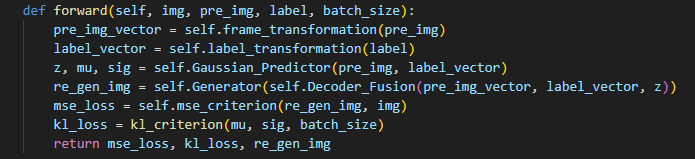
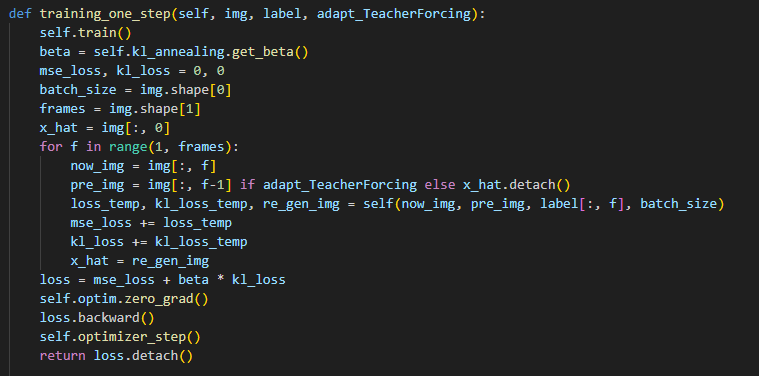
1. Introduction

本次實驗實作CVAE，和一般VAE不同的是，CVAE可以給decoder其他額外資訊來生成特定圖片，本次實驗及透過給decoder過去的資訊來生成現在的圖片。而在訓練上則有Techer forcing和KL annealing等方法來實現，並以PSNR score當作評分標準。

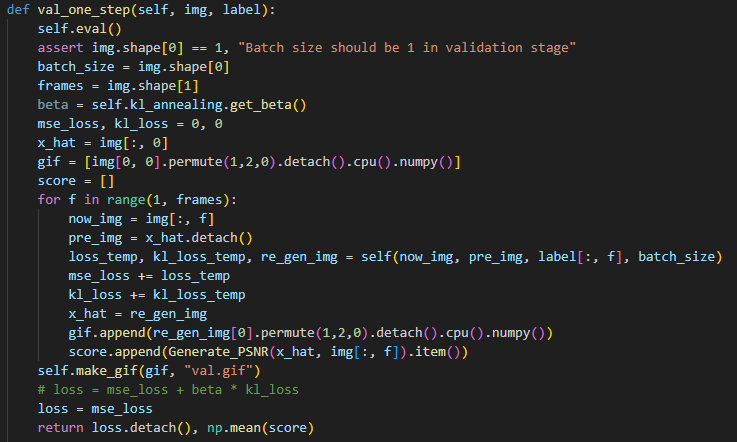
1. Implementation detail
   1. training/testing protocol



forward當中先將input轉換並丟進Gaussian Predictor，得到這些input在latent space的分布情況(mu, sig)，接著只要在這個分布當中抽一個點，並加上condition(pre\_img\_vector, label\_vector)丟給decoder和Generator，產生模型輸出，而mse\_loss是模型輸出和實際的差異，kl\_loss則是latent space分布和常態分佈的差異。

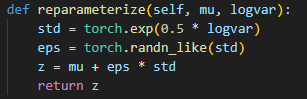


training當中以第0幀當起點，丟給模型forward回傳mse\_loss和kl\_loss以及第1幀的predict，反覆循環直到所有幀被預測完，之後再做反向傳遞，Techer forcing是判斷下一幀的輸入是否要是模型上一幀的輸出還是實際幀。



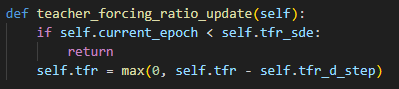
testing當中和training類似，差異在為了評估模型的能力，需要強制使用模型上一幀的輸出當作下一幀的輸入，而圖片中loss的計算只有mse loss，其意思代表評估方式最終看重的是模型輸出的預測能力而非考慮latent space的分布。

* 1. reparameterization tricks



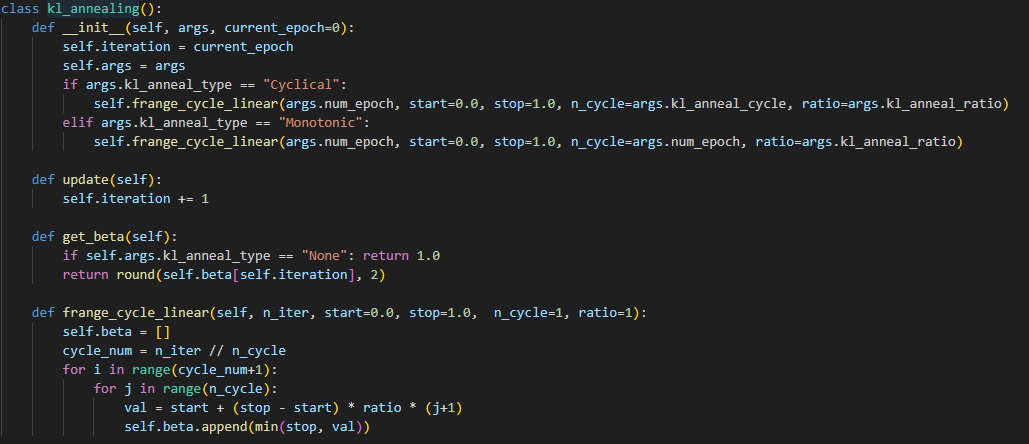
先前提到要在latent space中抽一個點，但如果真的隨機抽樣的話會無法做反向傳播(back propagation會斷在此處)，因此透過mu + sig \* eps的方式，eps是一個常態分布，整體代表要選一個點而這個點會偏離以mu為中心eps的標準差，這就和在N(mu, sig)中抽樣行為類似，且此方法也可微分做反向傳播。

* 1. teacher forcing strategy



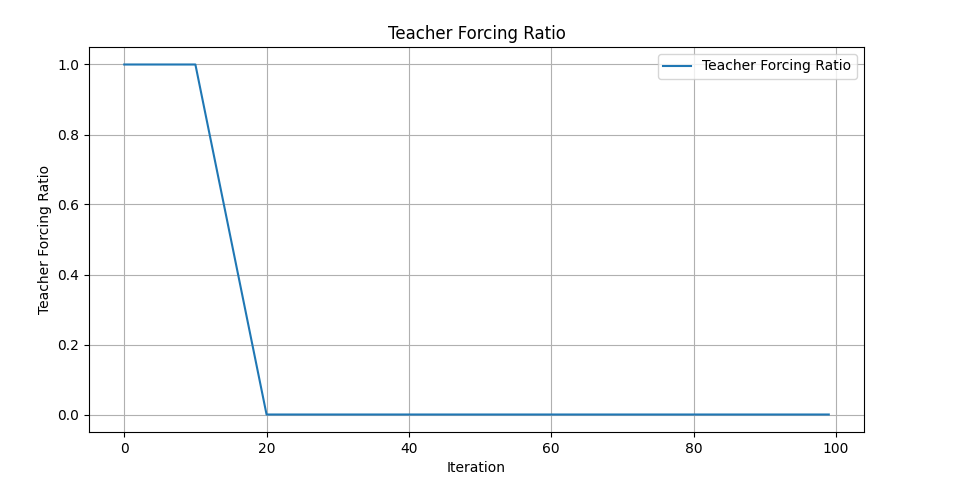
在一定epoch之前保證tfr為預設值(1.0)，之後每一次epoch衰減一定值。

* 1. kl annealing ratio

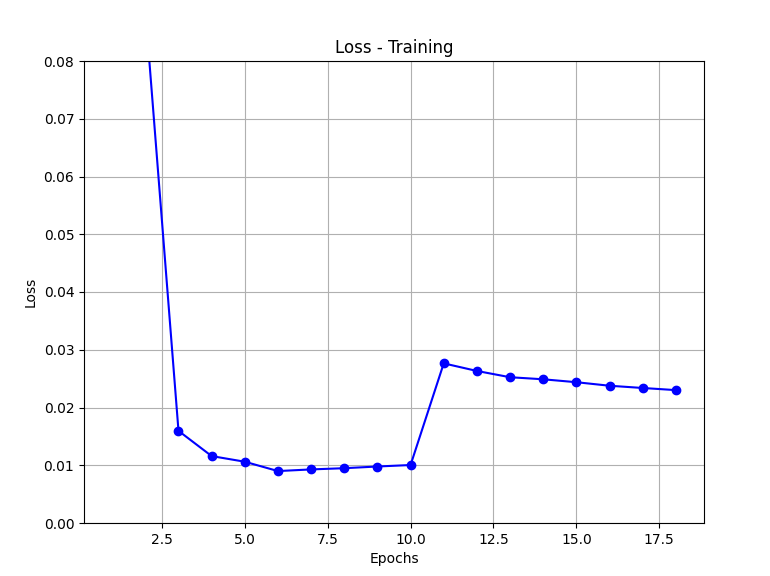


如果是Cyclical，則beta週期性的從start->stop，如果是Monotonic，則週期為1的版本，如果beta在遞增的途中超過stop，則會停留在stop不動。

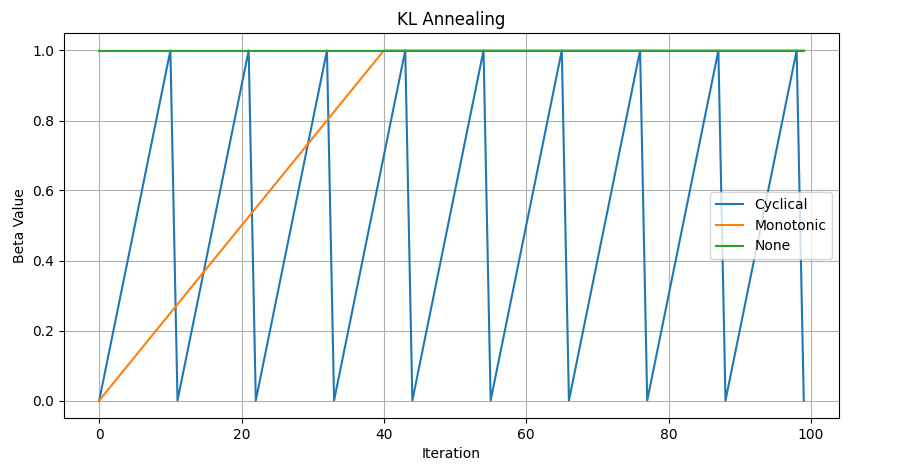
1. Analysis and Discussion
   1. teacher forcing ratio



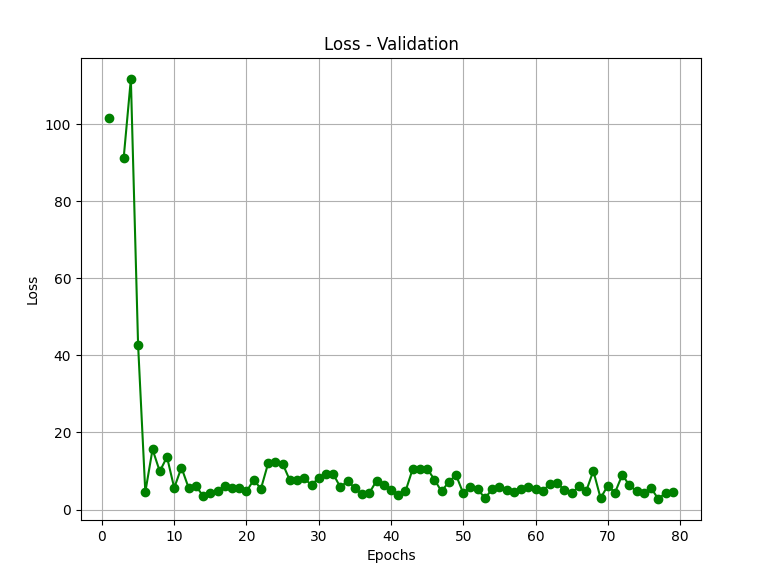
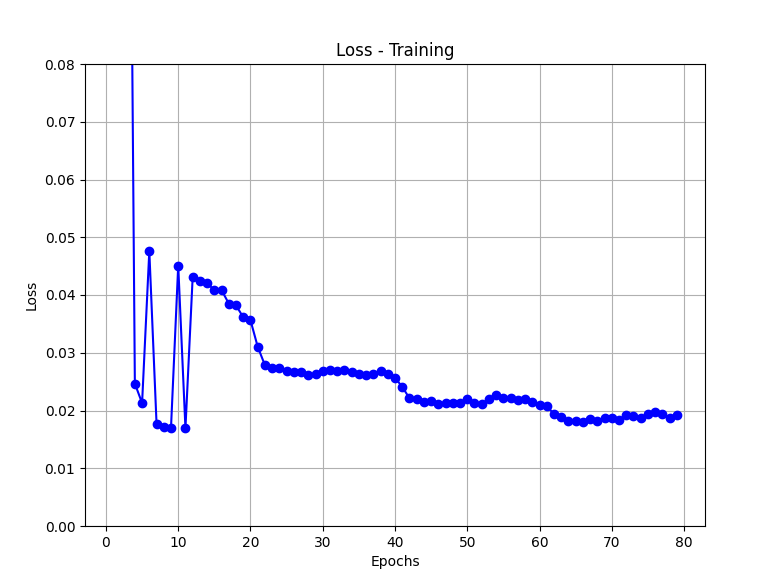
tfr 的幫助能夠讓模型很好的學習下一幀的圖片，但是仍有缺點，如下圖所示，training loss在第10個epochs的時候突然上升，這說明即使對模型學習下一幀圖面的生成能力有幫助，但卻沒有在每一幀跌代時將部分錯誤導正的能力，也就是如果模型在某一幀開始偏離ground truth，模型接下來沒有能力彌補錯誤。



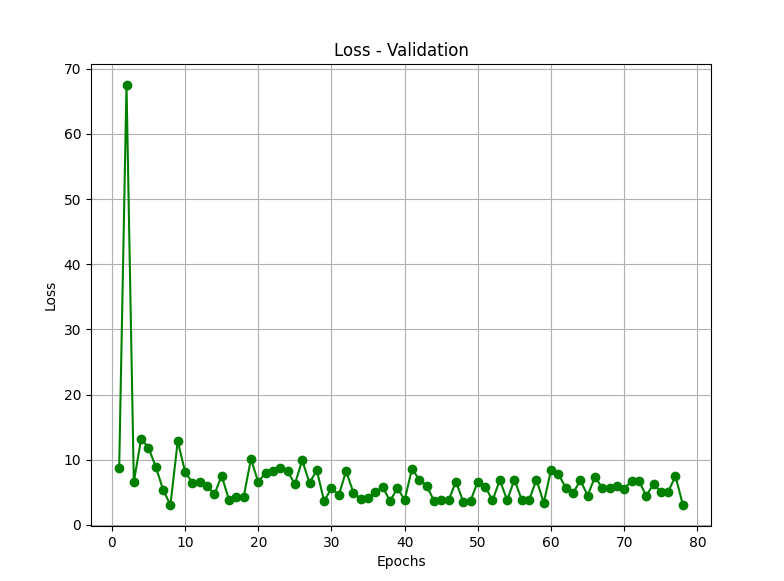
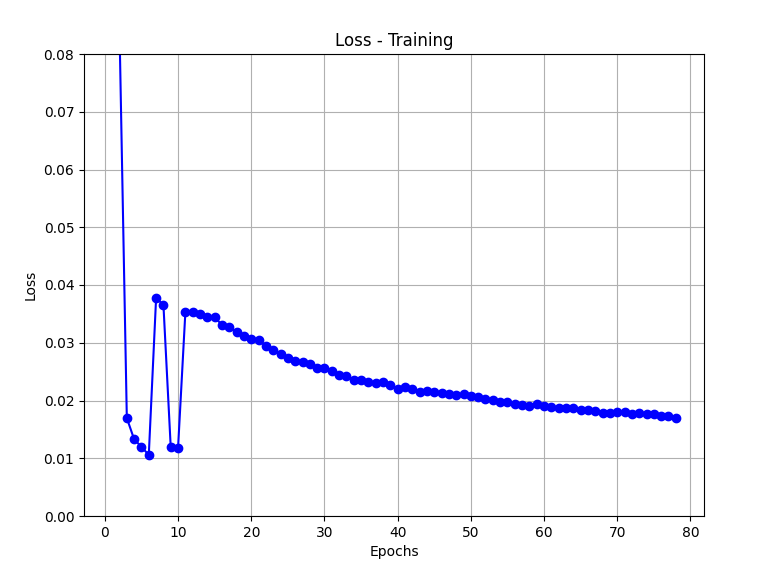
* 1. KL annealing



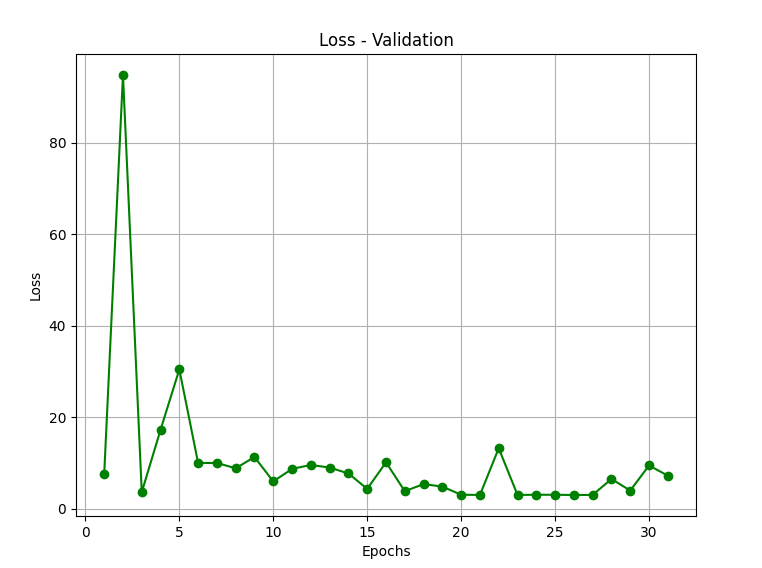
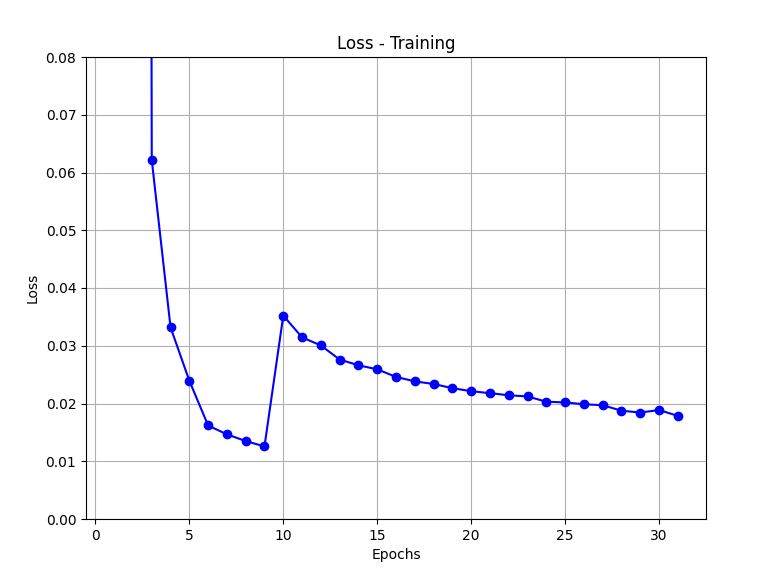
* + 1. Cyclical



* + 1. Monotonic



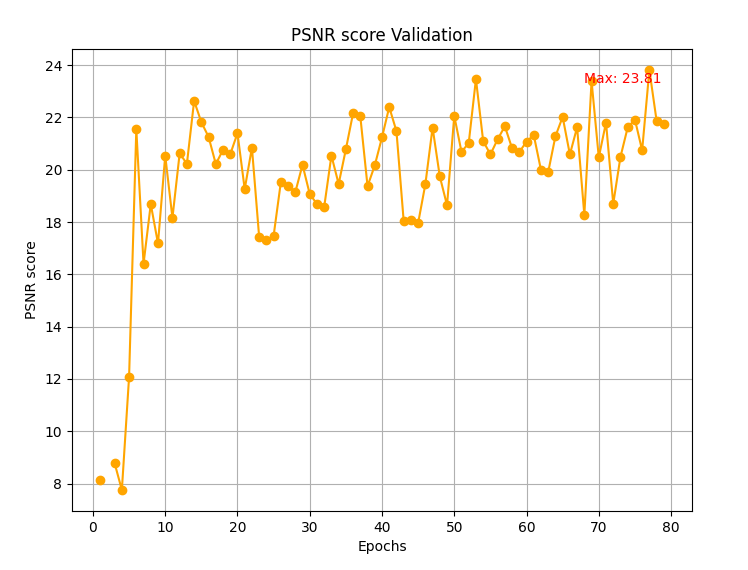
* + 1. None



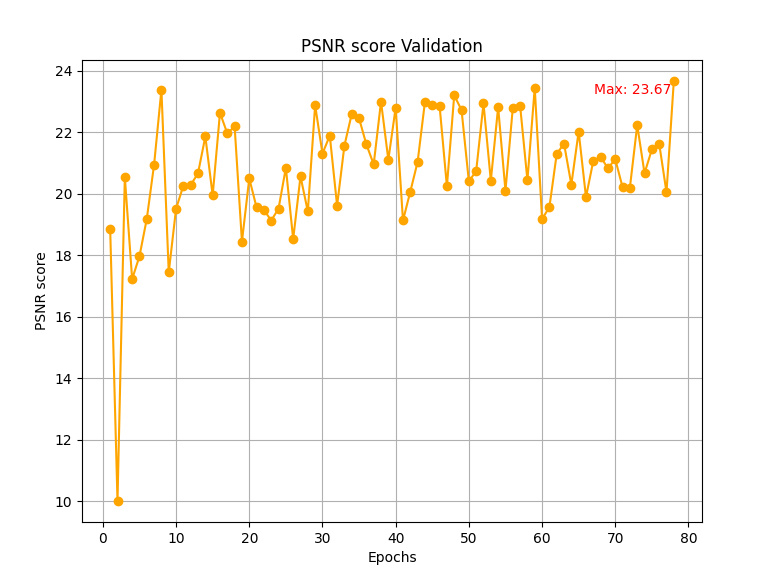
* + 1. Little Conclusion

實驗顯示，使用不同方式的KL annealing，對於最終訓練成果的影響不大，但是相比之下，使用Cyclical和Monotonic的loss曲線相對穩定。

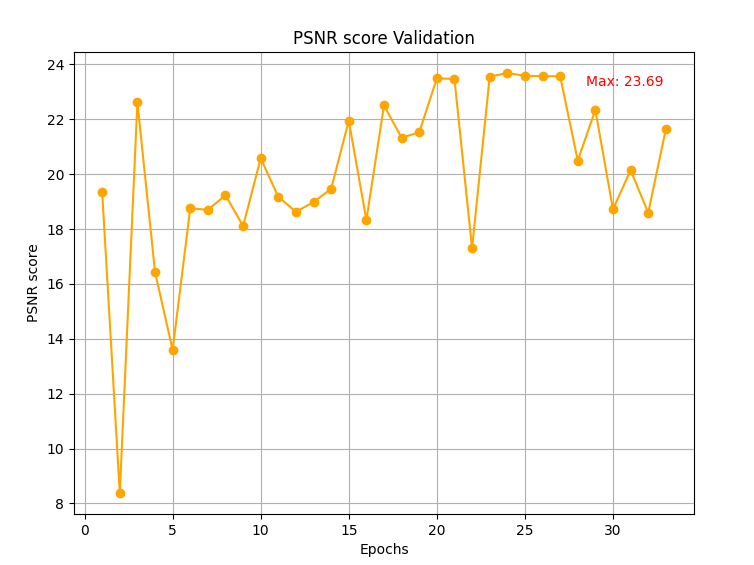
* 1. PSNR
     1. Cyclical



* + 1. Monotonic



* + 1. None

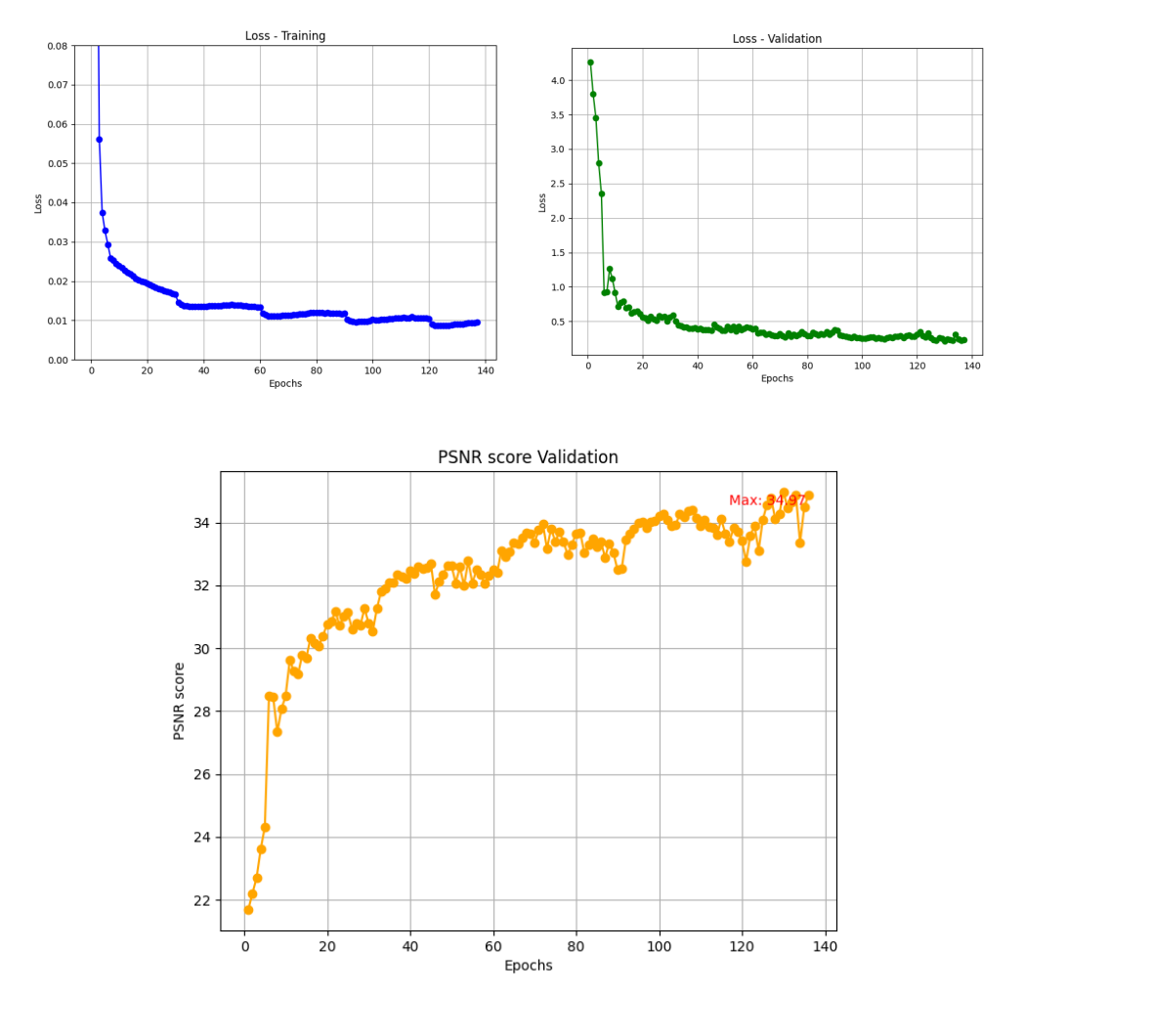


* 1. Other training strategy analysis
     1. 將 tfr 設為 0

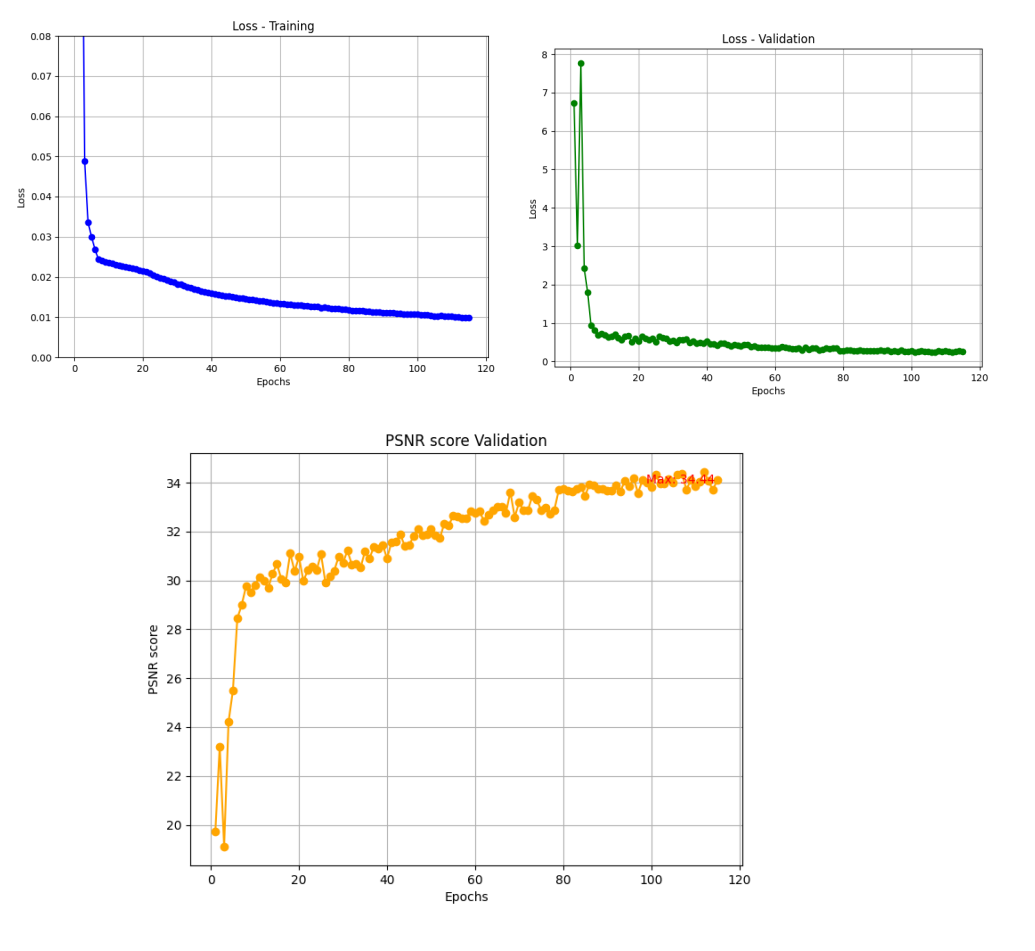
既然 tfr 有可能會讓模型失去導正錯誤的能力，

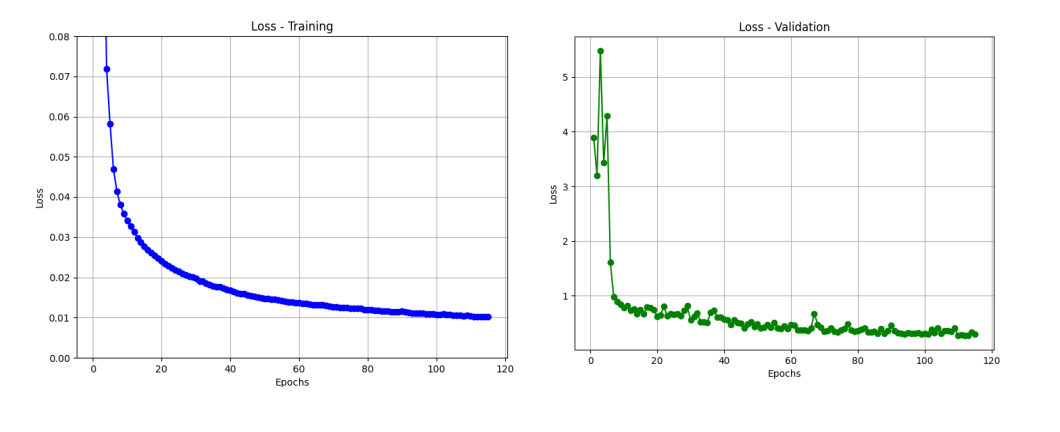
在此就不使用 tfr 的方法。結果顯示，如果將 tfr移除，則不管使用哪一種 KL annealing 的方式，所有的模型成效皆會**大幅提升**。

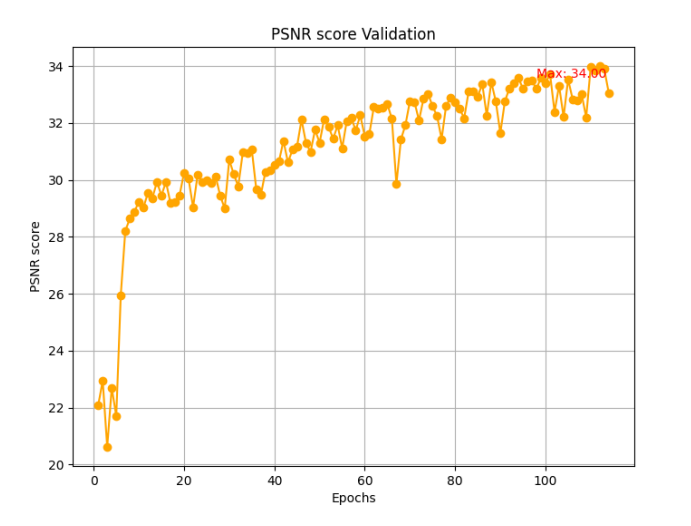
Cyclical:



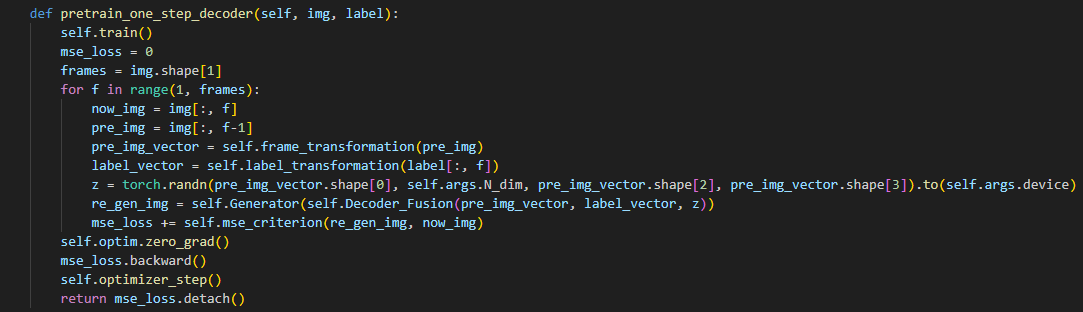
Monotonic:



None:



* + 1. Pretrain



這邊使用了其他方法，為了避免在訓練初期，encoder所產出的output品質不好，連帶影響到decoder，因此先把decoder訓練好，在pretrain當中假設encoder已經是有能力區分不同圖片的latent space分布，且其分布接近標準常態分布的情況下去訓練的。

1. execute code
   1. Training

程式會把所模型和所有分析存在save\_root

python ./Lab4\_template/Trainer.py --save\_root ./save\_path

* 1. Testing

python ./Lab4\_template/Tester.py --save\_root ./save\_path --ckpt\_path ./save\_path/best\_score.ckpt