

進捗報告

1 DAMSM と GAN の同期学習

従来手法では DAMSM は事前学習するモデルであり、GAN モデルの訓練時には text encoder や image encoder のパラメーターは固定されていましたが、先々週のゼミでの助言を受け、DAMSM と GAN を同期的に学習するようにモデルを変更してみました。

DAMSM の段階では実画像しか取り扱っていませんでしたが、GAN では生成画像も取り扱えるため、real と fake それぞれの word loss, sentence loss を DAMSM に backward するようにしています。その比率として、real での loss を 0.999, fake での loss を 0.001 かけて backward しています。

使用画像を 25 種類の鳥から 64 種類の鳥に増加させ、200epoch 程動かしていました。それに伴い DAMSM の学習時のバッチサイズを 64 にしていますが、GAN の学習や画像生成時のバッチサイズはメモリの都合上変更なしです。

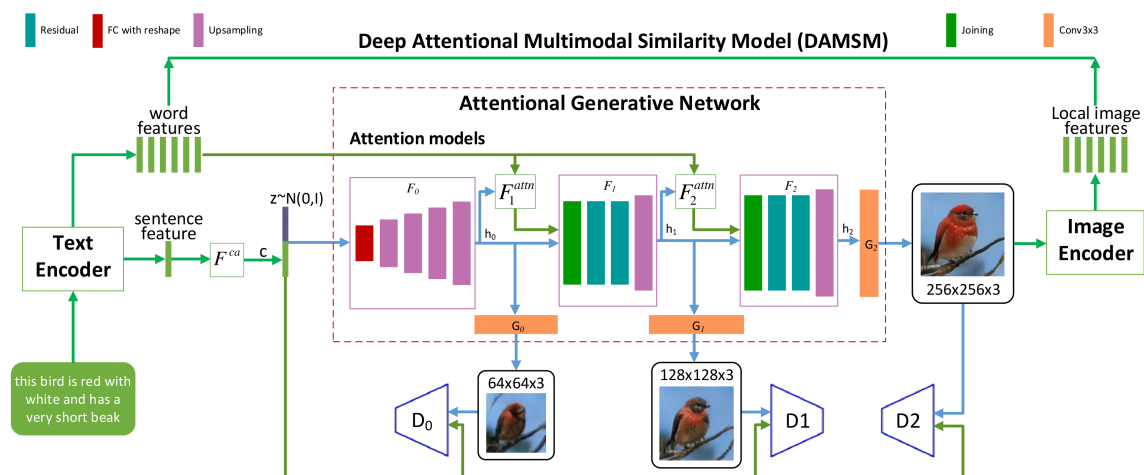


図 1: AttnGAN のモデル概要図