

深度學習應用——作業四報告

B03902072 江廷睿

1 Model Description

1.1 模型架構

圖片標籤的前處理

圖片部份的前處理有：把原圖縮小到 64×64 ，並把 $[0, 255]$ 的顏色線性對應到 $[-1, +1]$ 。標籤則是只取頭髮顏色與眼睛顏色相關的標籤，並把頭髮顏色跟眼睛顏色各自 one-hot encode 後連接起來作為 generator 與 discriminator 的 condition。

Generator

把 condition 的向量跟 100 維從標準常態分佈取樣出來的雜訊相接，通過 filter 個數為 512、256、128、64、3，尺寸為 4×4 ，stride 為 2 的 deconvolution layer，除了最後一層以外，每一層都使用 Batch Normalization 與激發函數 RuLU。最後一層則是使用 tanh 作為激發函數，產生 64×64 ，通道數為 3 的圖像。

Discriminator

把輸入通過 filter 個數分別為 32、64、128、256，尺寸為 4×4 ，stride 為 2，的 convolution layer，得到 4×4 ，通道數為 256 的張量。接著把 condition 的向量複製 16 份，接到 convolution layer 產生的張量後。接著再把得到的張量通過個數為 128、128，尺寸為 1×1 與 4×4 的 convolution layer。其中除了最後一層以外，在每一層後都使用 Leaky ReLU 作為激發函數。

1.2 Generator 跟 Discriminator 的目標函數

令 x 為隨機取樣的真實圖片， c 為那張圖片的 condition， c' 為隨機取樣的錯誤 condition， G 為 generator 函數、 D 為 discriminator 函數 α 為從 $[0, 1)$ 的均勻分布取樣的隨機數。generator 的目標函數為

$$\max D(G(c), c)$$

discriminator 的目標函數為

$$\max D(x, c) - \max\{D(G(c), c), D(x, c')\} - \lambda(\|\nabla D(\alpha x + (1 - \alpha)G(c))\|_2 - 1)^2$$

這裡 λ 使用 10。

1.3 其餘超參數

- Optimizer: RMSprop
- Learning Rate: 0.0002
- 此外每更新 5 次 discriminator 才更新一次 generator 。

2 How do you improve your performance

2.1 使用 Improved W-GAN

相較於 W-GAN，Improved W-GAN 不對 discriminator 的參數的最 weight-clipping，而是在產生的圖片與真的圖片之間隨機內插一點，並在目標函數中減少 generator 在該點的梯度的 L2-norm。

2.2 Discriminator 的目標函數

原先使用的損失函數為

$$\max D(x, c) - 0.5D(G(c), c) - 0.5D(x, c') - \lambda(\|\nabla D(\alpha x + (1 - \alpha)G(c))\|_2 - 1)^2$$

但有可能是因為 discriminator 可以把 $D(G(c), c)$ 降得很低，以至於忽略了 $D(G(c), c)$ ，所以產生的圖像不太考慮 condition。因此若是把目標函數改成

$$\max D(x, c) - \max\{D(G(c), c), D(x, c')\} - \lambda(\|\nabla D(\alpha x + (1 - \alpha)G(c))\|_2 - 1)^2$$

就可以強迫 discriminator 必須要同時能夠分辨出假的圖像還要能分辨出 condition 才能增加整體目標函數的數值。

3 Experiment settings and observation

3.1 使用 Improved W-GAN

使用 improved W-GAN 相較於一般的 GAN，可以產生更清晰的圖片。

3.2 Discriminator 的目標函數

在使用更改後的目標函數後，產生出來的圖片顯然更與 condition 相關。