Generative Adversarial Network (GAN)實作:

基於MNIST資料集的手寫字體生成

B0829039王語堂

B0829027鄭星逸

B0829005黃淞蜂

B0829016邱弘璽

題目介紹

生成對抗網路（英語：Generative Adversarial Network，簡稱GAN）是非監督式學習的一種方法，透過兩個神經網路相互博弈的方式進行學習。該方法由伊恩·古德費洛等人於2014年提出。生成對抗網路由一個生成網路與一個判別網路組成。生成網路從潛在空間（latent space）中隨機取樣作為輸入，其輸出結果需要盡量模仿訓練集中的真實樣本。判別網路的輸入則為真實樣本或生成網路的輸出，其目的是將生成網路的輸出從真實樣本中盡可能分辨出來。而生成網路則要盡可能地欺騙判別網路。兩個網路相互對抗、不斷調整參數，最終目的是使判別網路無法判斷生成網路的輸出結果是否真實。

由GAN deepfake生成的人臉生成對抗網路常用於生成以假亂真的圖片。此外，該方法還被用於生成影片、三維物體模型等。

生成器的任務是從隨機噪聲中生成數據樣本，試圖使其看起來與真實數據相似。判別器則試圖區分生成器生成的數據和真實數據。生成器和判別器通過對抗訓練的方式進行互相競爭和學習。

GAN的訓練過程如下：首先，生成器通過隨機噪聲生成一批假數據樣本。然後，判別器將真實數據和生成器生成的數據混合在一起，並嘗試區分它們。接下來，根據判別器的結果，生成器進行調整，以生成更逼真的數據樣本，使得判別器更難區分。這個過程反復進行，直到生成器能夠生成具有高度逼真性的數據，而判別器難以區分真假數據。

GAN的優點是能夠生成高度逼真的數據樣本，並且不需要大量標記的訓練數據。它可以應用於圖像生成、圖像修復、風格轉換等許多應用領域。然而，GAN的訓練過程相對複雜且不穩定，需要適當的超參數調整和技巧來獲得好的結果。

而在這個報告中，我們將以實作的方式，以MNIST手寫數字資料集作為我們的訓練資料，深入了解生成對抗網絡(GAN)。

MNIST手寫數字資料集是一種常用於深度學習研究的資料集，包含了70000張28x28像素的手寫數字圖像，每張圖像都有相對應的標籤，表示該圖像表示的數字。在這個報告中，我們將以MNIST資料集為基礎，訓練一個GAN模型，生成新的手寫數字圖像。

動機

1. 為何選擇GAN：生成對抗網絡（GAN）是深度學習領域中非常重要的模型。它具有生成高度逼真數據的能力，並在許多應用領域中取得了突破性的成果。通過研究GAN，我們可以了解到這一強大模型的工作原理和應用。

2.為何選擇MNIST數據集：MNIST數據集是機器學習領域中非常著名且廣泛使用的數據集之一。它包含了大量的手寫數字圖像，是進行圖像分類和生成的理想數據集。選擇使用MNIST手寫字體數據集，可以讓我們在具體的實例中探討GAN的應用，並更好地理解GAN在圖像生成方面的能力。

綜上所述，選擇GAN和MNIST手寫字體作為報告的主題，能夠讓我們深入研究和理解GAN模型的運作機制，並在實際應用中應用到一個廣泛使用的數據集上，用GAN來生成逼真的手寫字體數字。同時，這個主題也讓我們能夠探索優化GAN性能和生成高品質手寫字體圖像的方法。

研究方法

我們的研究方法將遵循以下步驟：

1. 資料準備：我們將使用MNIST數據集，這是一個公開可用的手寫數字數據集，包含60,000個訓練樣本和10,000個測試樣本。每個樣本都是一個28x28的灰度圖像，代表從0到9的一個數字。

2. GAN模型建構：我們將建構一個GAN模型，其中包括一個生成器和一個判別器。生成器的任務是生成看起來像真實MNIST圖像的假圖像，而判別器的任務是分辨一個給定的圖像是真實的MNIST圖像還是生成器生成的假圖像。

3. 模型訓練：我們將使用反向傳播的技術來訓練GAN模型。首先使用生成器生成假的MNIST圖像，然後將這些假圖像和真實的MNIST圖像一起提供給判別器進行訓練。根據判別器的性能，我們將更新生成器和判別器的權重。

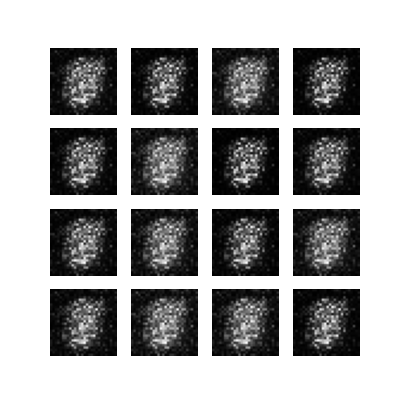
4. 模型評估：在訓練過程中和訓練結束後，我們將檢查生成的圖像的品質，評估我們的GAN模型的性能。

5. 結果分析和討論：最後，我們將分析和討論我們的實驗結果，包括模型的訓練過程、生成圖像的質量以及模型性能的評估結果。我們也將討論任何觀察到的問題，並提出可能的解決方案。

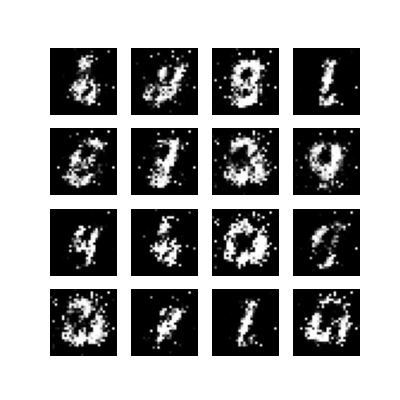
實驗結果

訓練過程

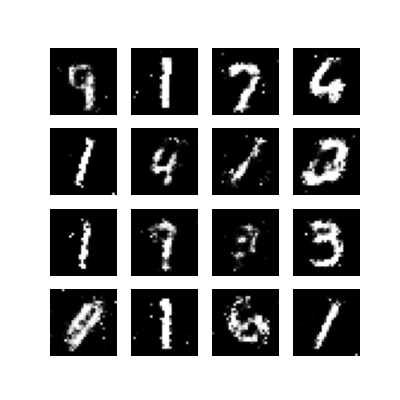
Epoch1



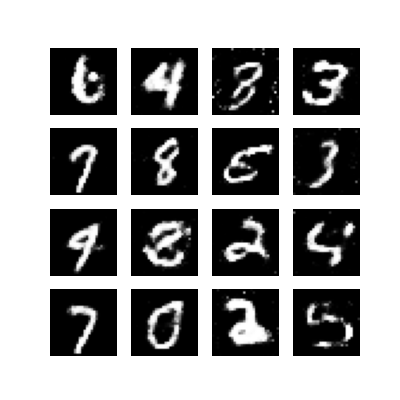
Epoch 5



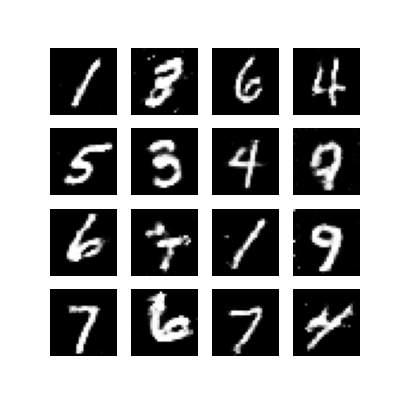
Epoch 10



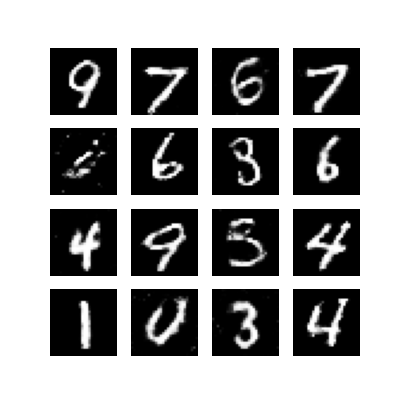
Epoch 50



Epoch 100

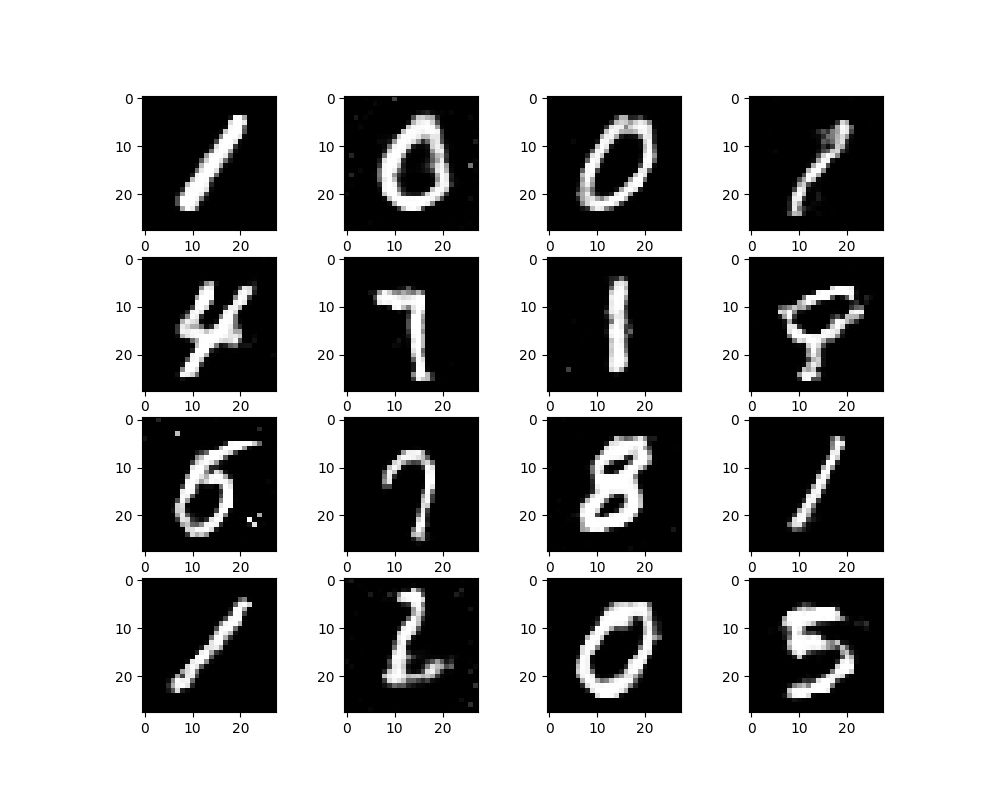


Epoch 200



可以看的出來隨著訓練的Epoch的增加，圖像的精度也在提升，可見本實驗中的GAN之架構是有效的。

最後結果



我們的GAN模型成功地生成了新的手寫數字圖像。這些生成的圖像在視覺上與MNIST數據集中的真實圖像非常相似，顯示出GAN模型強大的生成能力。我們從生成的圖像中可以清楚地識別出0到9的所有數字。

困難與解決方法

1.損失函數的選擇：GAN使用的損失函數通常是非常特殊的，選擇不適當的損失函數可能導致訓練困難或品質下降。

解決辦法:試驗不同的損失函數，並觀察它們對訓練的影響。

2.超參數調整：模型的超參數，如學習率、批次大小、優化器的參數等，對訓練的效果和速度有很大的影響。不同的超參數組合可能會產生不同的結果。

解決辦法:進行超參數的調整和優化，以找到最佳的訓練設定。

3.模型評估：GAN的訓練並沒有明確的標準評估方式。評估和比較不同模型的品質相對有挑戰性的。

解決辦法:以視覺進行人工的評斷，以確保生成的圖像的品質。

結果討論

儘管我們的模型表現良好，但在某些情況下，生成的圖像仍然存在一些問題，例如數字的形狀偶爾會出現扭曲或不清晰。在未來的工作中，我們將探索更進一步改進模型性能的方法，例如使用更精良的網路架構，或者對參數進行調整。

參考文獻

Goodfellow, Ian J.; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua. Generative Adversarial Networks. 2014.

<https://arxiv.org/abs/1406.2661>

專題分工與貢獻

B0829039王語堂(25%程式撰寫與編輯報告內容)

B0829027鄭星逸(25%程式撰寫與編輯報告內容)

B0829005黃淞蜂(25%程式撰寫與編輯報告內容)

B0829016邱弘璽(25%程式撰寫與編輯報告內容)