

主題一：GAN 模型實際操作生成圖片

一、模型介紹：

網站名稱：AI Gahaku

網站連結：<https://ai-art.tokyo/>

網站描述：利用生成對抗網絡（GAN）將上傳的照片變成各種經典畫作的風格，特別是將人像照片變成肖像畫風格。

模型技術：pix2pix（cGAN）

技術簡介：條件生成對抗網絡（cGAN，Conditional Generative Adversarial Networks）是一種生成對抗網絡（GAN）的變體，其中生成圖像的過程是「有條件」的，生成時不僅利用隨機的 noise 來創造圖像，還會根據給定的條件（如輸入圖像或特定的標籤）來生成對應的圖像，完成各種圖像的轉換。

二、生成結果：

本文透過上傳多位於 Google 搜尋後的隨機現任總統照片，觀察到 AI Gahaku 能夠在 5 秒內生成出近 40 種的經典畫作風格照片，圖 1 至圖 4 為 4 組生成前後的比較範例。生成的結果可以滿有效的去除人像外的背景，但生成出的結果有些不符合原圖像、不自然或缺少細節的描繪，例如：圖 1 賴清德膚色為黃種人膚色，但生成結果卻偏向白種人的膚色；圖 2 及圖 3 中川普的髮絲生成出一片白色並未有紋理感，普丁額頭上的紋路也並未被描繪出；圖 4 的雙眼扭曲且牙齒生成的結果十分不自然。此部分推測可能受限於經典畫作多參考於 19 世紀的歐洲畫作，因此川普及普丁兩位生成的結果十分還原原本的樣貌，而非白種人的膚色生成後顏色偏亮，加上經典畫作中很少露齒笑的樣貌，再加上紋路細節的辨識也不尚精確，因此造成圖 4 生成圖像的不自然。

由上述可見，GAN 模型本身的特點也可能對生成品質產生影響。GAN 由兩個部分組成：生成器和判別器，兩者相互對抗訓練。生成器的目標是創建真實的圖像，而判別器則試圖辨別這些圖像是否真實。儘管這樣的對抗過程能夠促進生成器創造出高品質的圖像，但如果訓練數據集中缺乏現代人物的多樣性，生成的結果可能仍會偏離真實的樣貌。



圖 1 賴清德生成前(左)後(右)比較



圖 2 唐納·川普生成前(左)後(右)比較



圖 3 弗拉迪米爾·普丁生成前(左)後(右)比較



圖 4 西里爾·拉瑪佛沙生成前(左)後(右)比較

三、現今 GAN 較少人使用的原因：

依據 Google 搜尋後上普遍的說法，主要是因為隨著新技術如擴散模型（Stable Diffusion、DALL·E 2）出現，這些方法能提供更穩定、高質量且易於控制的結果，如 GAN 的訓練過程不穩定，需要大量的計算資源而容易導致模型崩潰，而 Stable Diffusion 訓練過程穩定；Stable Diffusion 相較於 GAN 更能夠生成更清晰或真實的圖像，進而取代了 GAN 的應用。