

運用生成對抗網路產製超解析衛星影像之先期評估

張庭韶¹ 蔡富安^{2*}

摘要

本研究探討使用生成對抗網路 (GAN) 模型提升衛星影像的空間解析度，以解決因拍攝角度、天氣狀況及感測器限制導致的解析度下降問題。研究使用中、高及超高分辨率影像進行降解析處理，並透過 GAN 進行訓練生成超解析影像。GAN 訓練過程中，生成器負責將低解析度影像重建為高解析度影像，判別器則區分生成影像與真實影像的差異。研究亦使用 VGG-19 預訓練模型進行特徵提取，提升生成影像品質。結果顯示，隨訓練次數增加，影像細節變得更為清晰及銳利化，且在結構相似性指標上優於傳統方法。然而，影像生成過程中出現色彩偏移及偽影現象。為改善此問題，建議進行更深層次訓練或使用後處理技術，並優化模型架構，如移除 Batch Normalization。綜上所述，GAN 模型具有提升衛星影像解析度的潛力，惟仍尚有影像色彩偏移及偽影問題，未來可針對模型穩定性和影像後處理進行優化。

關鍵詞：超解析、衛星影像、生成對抗網路、深度學習

1. 前言

衛星影像判讀可應用於追蹤特定目標及災損評估等議題，對於特定事件從事精準判釋及分析而言，除須仰賴分析師自身經驗及熟悉事件相關背景知識外，關鍵在於是否能於短時間內獲取多幅衛星影像，影像是否可清晰呈現目標細節。對於超高分辨率衛星影像而言 (例如：Maxar WorldView 系列影像、Airbus Pleiades 影像)，影像最佳解析度均小於 1 m，惟影像於偵照時，可能受限於拍攝傾角過大或天氣狀況不佳等因素，影像空間解析度可能會降低至 1 m (或以上)，導致無法清晰呈現微小目標之特徵細節。另囿於衛星軌道設計，上述星種均於太陽同步軌道運行 (Boain, 2004)，因此重返週期較長，單一星種需經過 1-2 天後方可針對相同地點執行偵照，無法於短時間內獲取大量影像。在另一方面，部分新興微衛星群 (例如：Planet Labs 公司所屬 SkySat 衛星、Dove 衛星)，衛星在軌數量較多，因此重返週期短，可於短時間內重複偵照同一目標 (Kim *et al.*, 2022)，但影像空間解析度通常相較於超

高分辨率影像略顯遜色。部分衛星影像應用需於特定日期偵照特定目標 (例如：監控非法走私船艦)，並進行目標判釋 (Aiello *et al.*, 2019)，因此若能提升原始衛星影像解析度，將可幫助分析師正確判釋目標。

目前深度學習技術已廣泛運用遙測領域，物件偵測 (Object Detection) 及影像分類 (Classification) 應用實例屢見不鮮 (Li *et al.*, 2017)。現今亦有多項神經網路模型應用於影像超解析 (Super resolution)。本研究嘗試運用生成對抗網路 (Generative Adversarial Network) 深度學習技術 (Goodfellow *et al.*, 2014)，透過生成器 (Generator) 盡可能模仿訓練集中真實樣本之各種特徵，生成和高分辨率影像非常相似之超解析影像。研究中透過判別器 (Discriminator) 區分超解析及高分辨率影像，並將鑑別結果回饋給生成器，在兩組神經網路不斷對抗情形下，透過持續訓練機制，當判別器將超解析影像判定為高分辨率影像時，表示生成器已有足夠能力產製高品質超解析影像，所產製之超解

¹ 國立中央大學土木工程學系 碩士

² 國立中央大學太空及遙測研究中心 教授

* 通訊作者，電話：03-4227151 轉 57619, E-mail: ftsai@csrnr.ncu.edu.tw

收到日期：民國 113 年 04 月 22 日

修改日期：民國 113 年 06 月 26 日

接受日期：民國 113 年 09 月 25 日

析衛星影像於視覺上將會更銳利化，目標細節亦可更被清楚呈現。

2. 相關研究

傳統超解析方法主要運用內插技術提升影像解析度 (Khaledyan *et al.*, 2020)，例如：雙三次插值 (bicubic interpolation)。惟此方法用於重建影像複雜紋理及邊緣細節時，效果表現不佳。近年來，伴隨深度學習技術蓬勃發展，影像超解析技術有顯著的突破，生成對抗網路即為一重要方法，透過生成器及判別器相互對抗，並交替訓練，可依據原始影像紋理特徵，有效重建並生成更高解析度影像，且該影像具備更多的細節及真實感。儘管生成對抗網路模型可有效提升影像解析度，但該模型仍存有一些問題，例如：訓練過程不穩定、生成影像有偽影產生、生成影像過於平滑及遺失部分紋理特徵等。因此，為解決前述問題，近期有許多研究針對生成對抗網路模型進行調整，並探討訓練資料的多樣性及模型架構是否與生成圖片的品質有關。期間經不斷修改模型架構及調整參數後，研究者們陸續提出改進版本，其中 ESRGAN 及 Real ESRGAN 係為以 SRGAN 為基礎研改之模型，與 SRGAN 模型相比，前述 2 個模型均將生成對抗網路內之生成器架構由殘差網路 (ResNet) 改為密集卷積網路 (Dense Net)，以提升模型學習影像特徵能力，另於密集卷積網路內移除 Batch Normalization，解決影像內隨機產生偽影情形 (Wang *et al.*, 2018、Wang *et al.*, 2021)，後續更有研究者依據 Real ESRGAN 模型進行修改，提出 StarSRGAN (Vo & Bui, 2023)，針對模型內的判別器架構由 U-Net 改為多尺度注意力 U-Net，其中多尺度可提升重建影像紋理及邊緣細節能力，注意力機制則可使模型聚焦於影像基本特徵，以整體強化影像超解析能力。

除了生成對抗網路外，其他基於卷積神經網路 (CNNs) 架構的深度學習方法亦於影像超解析上有顯著卓越的成果，例如：SRCNN、EDSR 及 VDSR 等 (Dong *et al.*, 2016、Lim *et al.*, 2017、Kim *et al.*, 2016)。這些方法主要探討如何運用殘差網路

(ResNet) 改善深度學習梯度消失問題，並可使模型構建深層的神經網路，以獲取影像內的複雜特徵，進而獲取超解析的成果 (He *et al.*, 2016)。

3. 研究方法

本研究主要使用三組衛星影像資料，第一組為中解析度衛星影像，空間解析度約 3 m；第二組為超高解析度衛星影像，空間解析度約 31 cm；第三組資料包含高解析度及超高解析度衛星影像，空間解析度介於 31 cm 至 6 m。主要流程首先為運用衛星影像訓練生成對抗網路模型，先將原始解析度之衛星影像進行 4 倍降解析 (Downsampling)，後續將 80%降解析度影像及其對應之原始解析度影像做為訓練資料，匯入模型從事訓練；待模型訓練完成後，匯出權重，並將剩餘 20%降解析影像資料匯入模型內，產製超解析 (Super resolution) 衛星影像。後續運用峰值訊噪比 (Peak signal-to-noise ratio, PSNR) 及結構相似性 (Structural Similarity Index, SSIM) 針對超解析影像及原始解析度影像進行比對，以評估超解析影像品質，本研究整體流程圖如圖 1 所示。

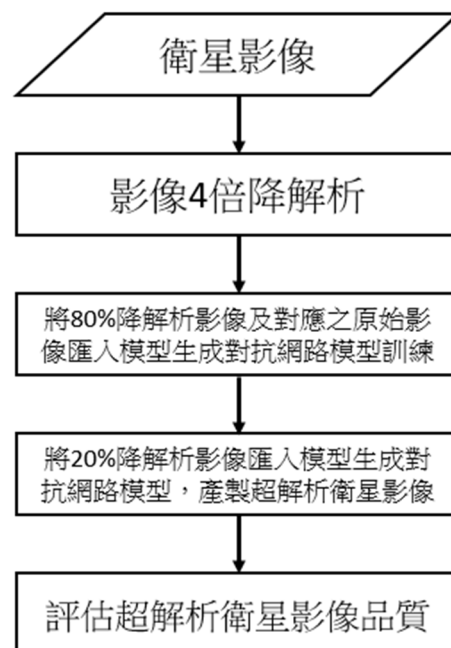


圖 1 整體研究流程

3.1 深度學習模型

利用深度學習模型進行分析主要目的為提升衛星影像空間解析度，以強化目標特徵細節。在將訓練資料輸入模型前，需對原始影像進行前處理，處理完成後，方可匯入模型訓練。本研究運用生成對抗網路提升衛星影像解析度流程包含 4 個步驟，分別為影像前處理、訓練生成對抗網路模型、產製超解析影像及超解析衛星影像品質評估，詳細說明分述如后。

3.1.1 影像前處理

(1) 裁切影像圖塊 (Image Patches)

由於深度學習模型的訓練和推理過程涉及大量的數學運算，包括矩陣乘法、卷積、激活函數等，現今常見深度學習模型框架，例如：Tensorflow 及 Pytorch，均支援在圖形處理單元 GPU 上執行深度學習模型，以從事加速運算。然而在處理大型模型和資料情況下，GPU 可能會有記憶體不足問題。因此，本研究若直接將整幅衛星影像輸入模型，將會導致記憶體不足而無法訓練，因此必須先行將整張影像裁切為多張影像圖塊，再匯入模型進行訓練，期可更順利處理大型、高解析度或不統一尺寸的影像，同時提升訓練效率和模型性能。

(2) 影像降解析 (Downsampling)

「降解析」為將原始高解析度影像轉換為低解析度版本，而「超解析」則是試圖從低解析度影像中重建出近似高解析度的影像，因此本研究透過先將原始高解析影像從事降解析處理，以獲取低解析度影像，並匯入生成對抗網路模型進行訓練。藉由影像內局部上下文 (Relative global contextual information) 關係，模型可根據低解析度影像中的結構和模式，更有效地預測高解析度影像中的細節，並在超解析過程中嘗試恢復丟失的細節。影像前處理流程，如圖 2 所示。

3.1.2 訓練生成對抗網路模型

生成對抗網路模型架構主要分為生成器 (Generator) 及判別器 (Discriminator) 組成 (Xiong

et al., 2020)，以下將模型各模組功能及訓練方式簡述如后。

(1) 生成器 (Generator)

生成器負責將低解析度影像轉換為高解析度影像，該模組使用卷積神經網路學習如何生成逼近於原始高解析度影像的圖像。

(2) 判別器 (Discriminator)

判別器是生成對抗網路中的另一個模組，亦由卷積神經網路構成，目標為區分原始高解析度影像及生成的超解析影像。

(3) VGG-19 預訓練模型

為訓練判別器，本研究運用預訓練的 VGG-19 網路萃取影像的特徵。這些特徵可用於衡量生成影像和原始影像間的差異，以利指導生成器生成更加真實的影像。判別器藉由該模型獲取影像特徵後，可計算生成的超解析影像和原始高解析度影像之間的感知損失 (Perceptual Loss)，故能更加客觀地評估 2 張影像於人眼視覺上的差異 (Ledig *et al.*, 2017)。此感知損失在訓練過程中將逐漸被最小化，並不斷回饋給生成器，將有助於生成器生成更逼近原始影像的超解析影像。

(4) 模型訓練方式

A. 訓練生成器

首先將低解析度影像輸入生成器，獲得生成的超解析影像，後續運用預訓練好的 VGG-19 網路提取生成的超解析影像和原始高解析影像的特徵，以計算感知損失，並衡量生成的超解析影像和原始高解析影像間的差異。本研究以特徵之間的均方誤差 (Mean Square Error, MSE) 作為感知損失，並在逐漸最小化感知損失的過程中，持續更新生成器的參數，以利生成器生成更逼近原始影像的超解析影像。

B. 訓練判別器

接續運用原始高解析度影像和生成的超解析影像訓練判別器，並計算對抗損失 (Adversarial Loss)；此損失目的為使判別器分辨生成的超解析影像和原始高解析度影像的特徵。本研究運用二元交叉熵 (binary cross-entropy) 衡量對抗損失，在訓練過程中，嘗試逐漸最小化對抗

損失，並持續更新判別器參數，以利判別器能夠更好地區分生成的超解析影像和原始高解析度影像。

C.交替訓練：

交替訓練生成器和判別器，於每個訓練步驟中，先更新生成器，再更新判別器，依循上

述方法重複訓練，並即時調整超參數 (Hyperparameter)，例如：學習率 (Learning Rate)，以利模型收斂並獲取更佳品質之超解析影像。簡而言之，生成器負責產製超解析衛星影像，判別器負責區分高解析度影像與超解析影像。訓練生成對抗網路模型流程，如圖 3 所示。

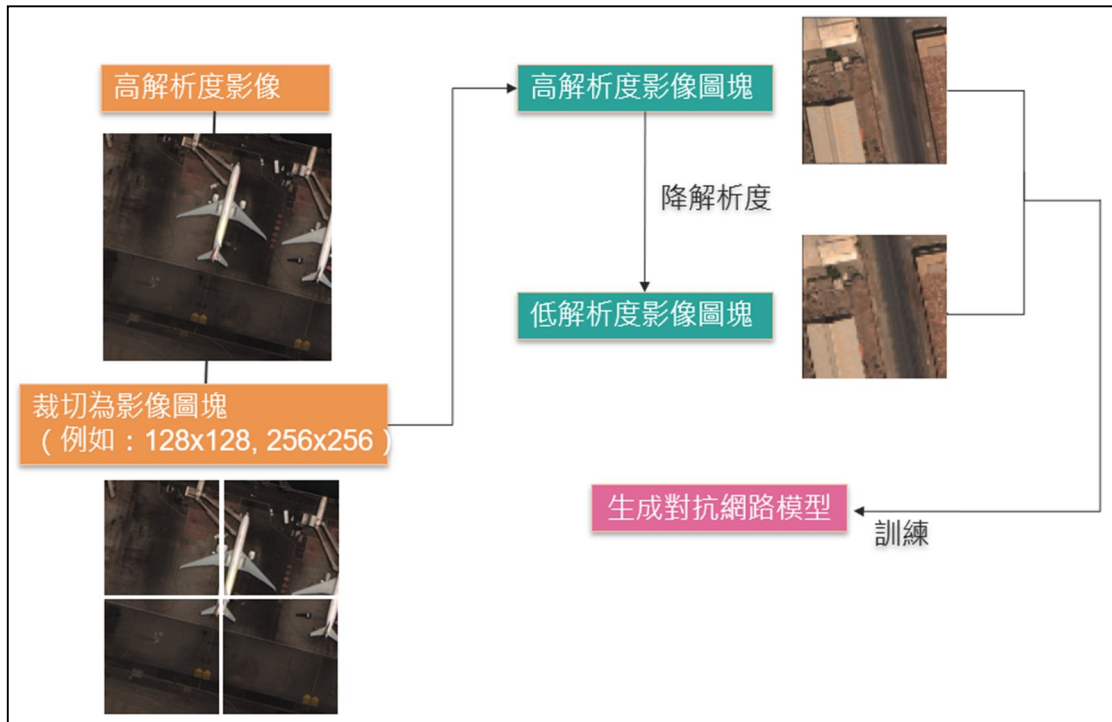


圖 2 影像前處理

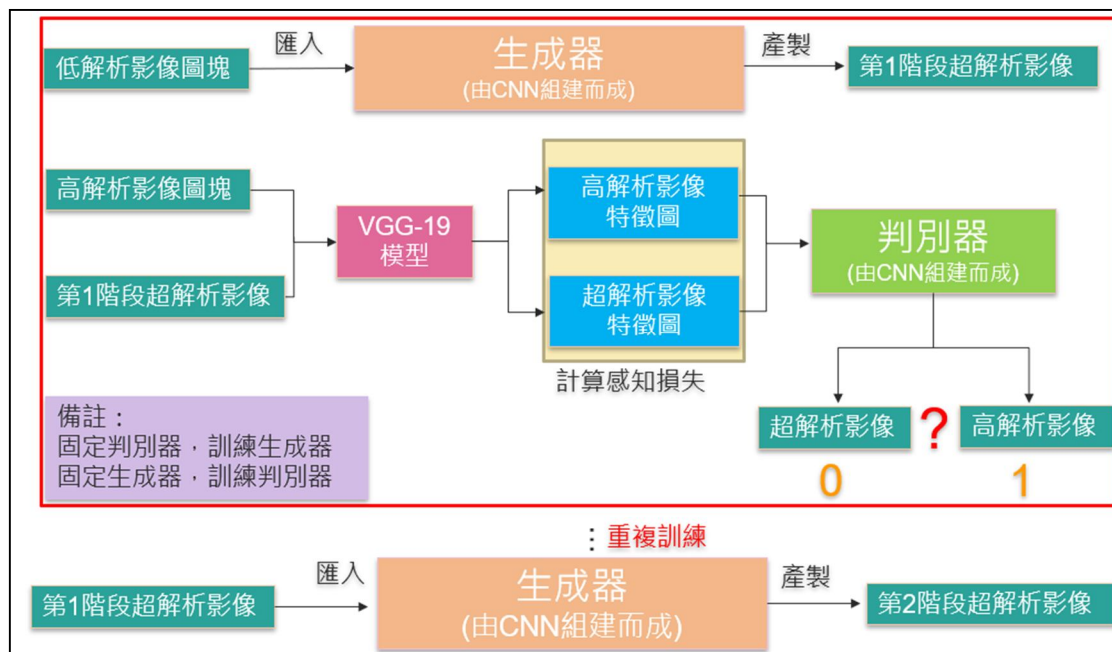


圖 3 訓練生成對抗網路模型流程

3.1.3 超解析影像品質評估

本研究採用 2 種量化分析指標評估超解析影像品質，分別為峰值訊噪比 (PSNR) 及結構相似性 (SSIM)，以下將此 2 種評估指標運算方式及評估方法 (Alain & Djemel, 2010) 簡述如后。

(1) 峰值訊噪比

為一常用於衡量影像品質的指標，用來量化原始影像和重建影像之間的差異，該指標計算方式基於影像的均方誤差 (MSE)，即重建影像和原始影像之間每個像素的差異平方的平均值，單位為分貝 (dB)，值愈高表示重建影像與原始影像之間的相似度越高，惟該指標存在部分局限性。例如：無法完全捕捉到人眼針對影像的失真型態。

(2) 結構相似性

主要考慮影像的亮度、對比度和結構等特徵，以更貼近人眼感知的方式評估影像的相似度，該指標值的範圍在 -1 到 1 之間，當值越接近 1，表示兩張影像越相似，也就是說，較高的 SSIM 值意味著重建影像與原始影像的結構、對比度和亮度等方面更相似。

3.1.4 神經網路模型

本研究運用的神經網路模型為 SRGAN (Super Resolution Generative Adversarial Network)，該模型主要由生成器及判別器建構而成，其中本研究所使用之生成器由 16 個殘差塊 (Residual Block)、2 個向上採樣模塊 (Upsampling Block)、卷積層 (Convolution Layers) 及輸出層組成；每個殘差塊內包含各 2 個卷積層及 Batch Normalization 層、1 個激活函數。殘差塊內的卷積層用於提取影像的特徵，Batch Normalization 層則可將每一層輸入的數據進行常態化，使數據分佈控制於特定範圍內，以有效改善梯度消失或爆炸問題，並加速模型收斂 (Shibani *et al.*, 2018)。每個殘差塊間均使用跳過連接 (skip connections) 保留低解析度影像的細節，如此一來有助於生成器學習低解析度到高解析度的映射，同時保留影像的結構。另向上採樣模塊運用 Pixel Shuffle 方法，透過重新排列像素，將低解析度

的特徵圖轉換為高解析度的影像。至於判別器組成，本研究先行構建判別器模塊 (discriminator block)，每個判別器模塊由卷積層、Batch Normalization 層及 LeakyReLU 激活函數 (Activation Function) 組成。判別器整體由 8 個判別器模塊、各 2 個全連接層 (Dense layer) 及激活函數組成 (第一個為 LeakyReLU、第二個為 Sigmoid)，其中卷積層用於提取影像特徵，激活函數放在卷積層後對影像特徵執行非線性處理，以利判別器學習更複雜的特徵。另全連接層則位於判別器的倒數第二層，主要用於將最終特徵表示映射為輸出分數，透過該分數判定輸入之影像是生成的超解析影像還是真實高解析度影像。SRGAN 生成器及判別器架構，如圖 4。

4. 研究成果及分析

4.1 研究資料

4.1.1 深度學習資料集訓練設定

本研究分別運用中解析度、超高解析度及混合高解析度及超高解析之衛星影像圖塊作為訓練資料集，以測試生成對抗網路模型對於不同資料來源之通用性，以下謹將各資料集詳細資訊分述如后 (詳如附表 1)。

(1) 中解析度衛星影像

本資料集係採用中解析度衛星影像 4000 張做為訓練及測試資料集，偵照區域位於美國舊金山灣 (San Francisco Bay) 及聖佩德羅灣 (San Pedro Bay)，每張影像大小均為 80×80 。影像均已透過融合演算法處理 (Pansharpened)，其波段包含紅光、綠光及藍光等 3 個波段，空間解析約 3 m，其中 1000 張影像內包含各式小型船隻，另 3000 張影像則未包含任何船隻。本研究先行將上述 4000 張影像尺寸調整為 128×128 ，後續再對該批影像進 4 倍降解析處理，故每張低解析度影像之尺寸為 32×32 。本研究隨機挑選 3600 張低解析度影像做為訓練樣本 (Training Set)，其餘 400 張做為測試樣本 (Testing Set)，每次迭代取 6 張影像 (6 Batch Size)，整個訓練過程將進行 30 輪 (30 Epochs) 計算。

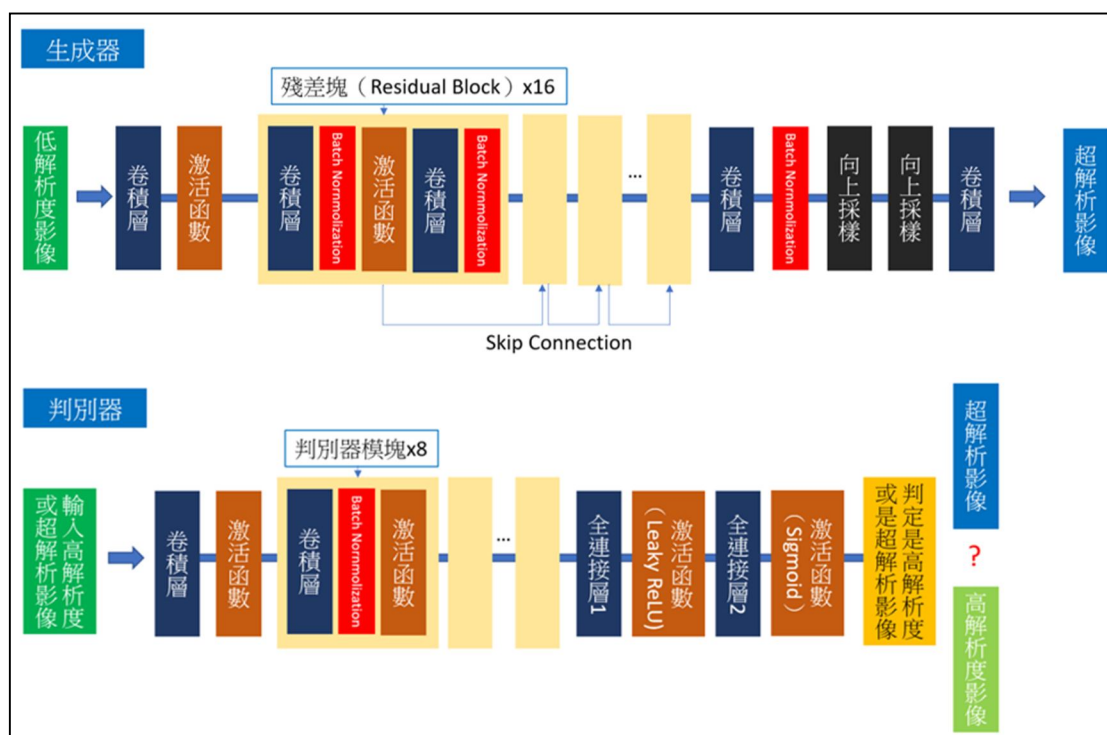


圖 4 SRGAN 生成器及判別器架構

表 1 影像資料集

	中解析度衛星影像	超高解析度衛星影像	高解析/超高解析度衛星影像
空間解析度	約 3 m	約 31 cm	介於 31 cm 至 6 m
影像張數	4000 張	2719 張	1275 張
原始解析度之影像大小	128 × 128	256 × 256	512 × 512
低解析度之影像大小	32 × 32	64 × 64	128 × 128
影像訓練張數	3600 張	2447 張	1147 張
影像測試張數	400 張	272 張	128 張
訓練次數	30 epochs	50 epochs	70 epochs

(2) 超高解析衛星影像

本資料係採用超高解析度衛星影像 2719 張做為訓練及測試資料集，空間解析度均約 31 cm，影像均已透過融合演算法處理，其波段包含紅光、綠光及藍光等 3 個波段。此影像資料集內包含多種地景地貌特徵，例如：機場、城市、山地、港口及海岸等，另本研究亦調整訓練影像之尺寸，將上述 2719 張影像尺寸調整為 256 × 256，後續再對該批影像進 4 倍行降解析處理，故每張低解析度影像之影像尺寸為 64 × 64。本研究隨機挑選 2447 張低解析度影像做為訓練樣本 (Training Set)，其餘 272 張做為測試樣本 (Testing Set)，每次迭代取 1 張影像 (1 Batch Size，受限於硬體因素影響)，整個訓練過

程將進行 50 輪 (50 Epochs) 計算。

(3) 混合高解析及超高解析度衛星影像

本資料係採用混合高解析超高解析度衛星影像共 1275 張做為訓練及測試資料集，影像空間解析度介於 31 cm 至 6 m，影像波段包含紅光、綠光及藍光等 3 個波段。本研究先行將上述 1275 張影像尺寸調整為 512 × 512，後續再對該批影像進 4 倍降解析處理，故每張低解析度影像之影像尺寸為 128 × 128。本研究隨機挑選 1147 張低解析度影像做為訓練樣本 (Training Set)，其餘 128 張做為測試樣本 (Testing Set)，每次迭代取 1 張影像 (1 Batch Size，受限於硬體因素影響)，整個訓練過程將進行 70 輪 (70 Epochs) 計算。

4.2 成果分析與討論

4.2.1 成果分析

對於中解析度衛星影像，本研究將 3600 張影像匯入生成對抗網路模型訓練共計 30 輪，並分別輸出訓練 10 輪、20 輪及 30 輪之模型權重，後續分別將上述權重及測試樣本 400 張衛星影像匯入生成對抗網路模型。對於超高解析度衛星影像，本研究將 2447 張影像匯入生成對抗網路模型訓練共計 50 輪，並分別輸出訓練 10 輪、30 輪及 50 輪之模型權重，後續分別將上述權重及測試樣本 272 張衛星影像匯入相同模型。另對於混合高解析度及超高解析度衛星影像資料集，本研究將 1147 張影像匯入生成對抗網路模型訓練共計 70 輪，並分別輸出訓練 10 輪、30 輪、50 輪及 70 輪之模型權重，後續分別將上述權重及做為測試樣本之 128 張衛星影像匯入相同模型，以產製超解析衛星影像。

實驗結果顯示不論運用前述任何一種資料集做為訓練資料，生成對抗網路模型產製之超解析影像，於人眼視覺效果上，均能夠更加清晰呈現船艦或建物之特徵細節，且伴隨訓練次數增加，對於影像內的目標，其細節亦會更加清晰及銳利。惟目前遭逢問題為超解析影像均有色調偏移 (Color Drift) 現象；另以混合高解析度及超高解析度衛星影像作為訓練資料，所產製之超解析影像，影像內偶含有雜訊或偽影。上述問題須透過影像後處理或調整模型架構等方法進行改善，於問題討論章節將進行更細部探討。除運用人眼視覺直觀評估影像品質外，本研究亦以運用量化分析指標「峰值訊噪比」及「結構相似性」2 種指標評估超解析影像重建品質。實驗結果發現訓練愈多輪，結構相似性數值愈高，代表模型透過更深層學習，超解析影像與原始高解析度影像在結構上之相似程度會逐漸提升。惟在峰值訊噪比方面，發現訓練至特定輪數後，超解析影像之數值可能不升反降，顯示峰值訊噪比可能無法完全客觀評估超解析影像重建品質，另相比於運用 Bicubic Interpolation 所產製之超解析影像，於「結構相似性」指標方面，發現運用生成對抗網路產製之超解析影像效果較佳，惟於「峰值訊噪比」部分，

則發現 Bicubic Interpolation 產製之超解析影像的品質比生成對抗網路產製之超解析影像效果佳，造成前述 2 個指標的差異可能係為「峰值訊噪比」係用來衡量原始影像與失真影像之間的相似度，主要評估像素強度的差異，故當影像色調發生偏移時，像素的色彩分佈會與原始影像有所不同，導致兩者的像素值差異增大，因此「峰值訊噪比」的值會變小；「結構相似性」則以亮度 (luminance)、對比度 (contrast) 和結構 (structure) 評估影像相似性，當影像色調發生偏移時，可能會影響亮度或對比度，但對於影像的結構 (如邊緣或紋理) 沒有明顯變化，故色調偏移對「結構相似性」指標的影響較小，綜上所述，相較於 Bicubic Interpolation 產製之超解析影像，生成對抗網路產製之超解析影像於結構上會與原始影像更相近，故影像的邊緣及紋理可於視覺上更清楚呈現。

圖 5 所示為中解析度衛星影像產製之超解析影像比較；圖 6 則為超高解析度衛星影像產製之超解析影像比較；圖 7 為混合高解析度及超高解析度衛星影像產製之超解析影像比較。經綜整上述實驗結果，顯示生成對抗網路模型產製之超解析影像可明顯提高影像的細節，增進判讀成效，惟影像色彩偏移及偶有雜訊或偽影產生，仍為須持續改善議題。

4.2.2 問題討論

(1) 影像色調偏移

相比於原始影像，本研究生成之超解析影像均有色調偏移現象。經檢視模型架構，發現生成器內有向上採樣模塊，當影像進行向上採樣時，由於影像尺寸擴大，需要透過內插生成新的像素值，而此插值方法可能導致像素值間產生微小變化而影響影像整體色彩。實驗結果發現，若將模型進行更深層訓練，將有助於減輕影像色調偏移情形，其可能係因模型可從原始高解析度影像中學習到更深層次特徵，故對於低解析影像進行超解析處理時，可更佳保留原始低解析影像色調及紋理細節，所產製超解析影像之色調及特徵細節，將會更接近原始高解析影像。除前述方法外，透過影像後處理亦可減輕色調偏移現象，可將原始低解析影像及超解析影

像進行直方圖匹配 (Histogram Matching), 使超解析影像色調可接近原低解析影像色調, 縱然無法完全一致, 但仍可有效改善色調偏移情形, 圖 8 為比較低解析度影像、原始超解析影像及經直方圖匹配處理後之超解析影像, 圖 9 為前述影像於紅光、綠光及藍光波段之影像直方圖分布。

(2) 改善影像內含雜訊或偽影相關方法

本研究混合高解析度及超高解析度影像作為訓練資料, 每張影像大小為 512 × 512, 實驗結果發現生成之超解析影像內偶有大量雜訊或偽影, 以下為改善影像內含雜訊或偽影相關方法。

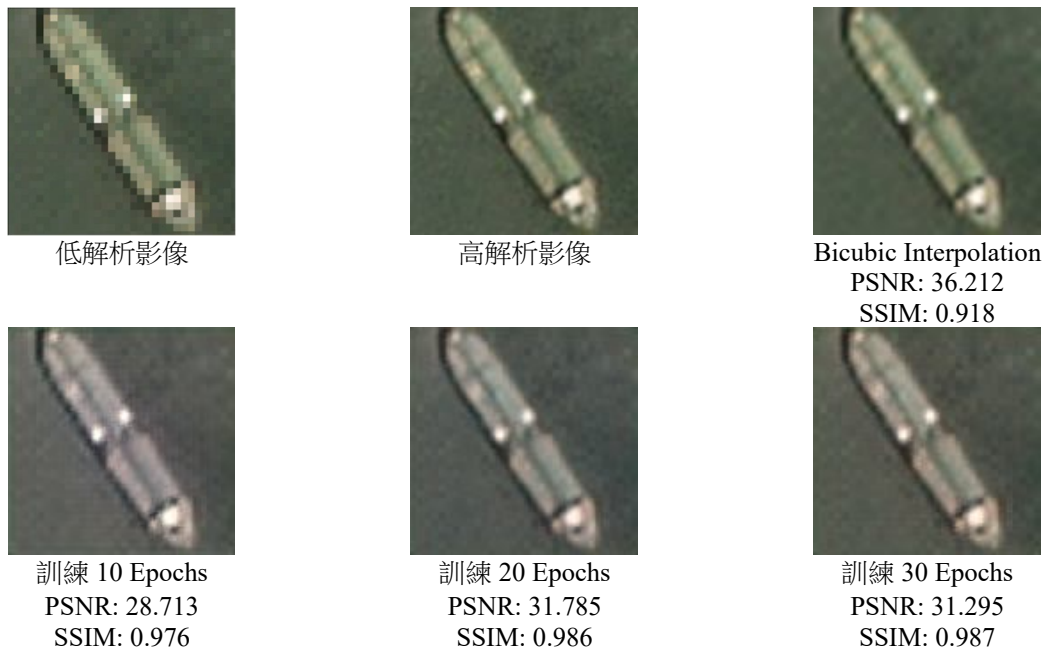


圖 5 中解析度衛星影像產製之超解析影像比較

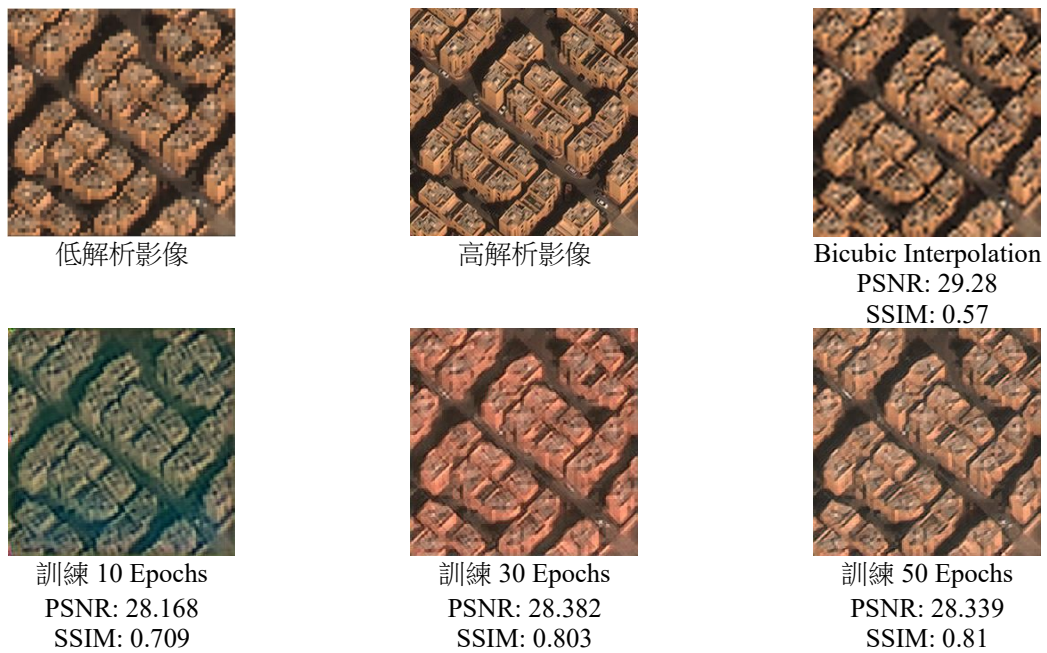


圖 6 超高解析度衛星影像產製之超解析影像比較

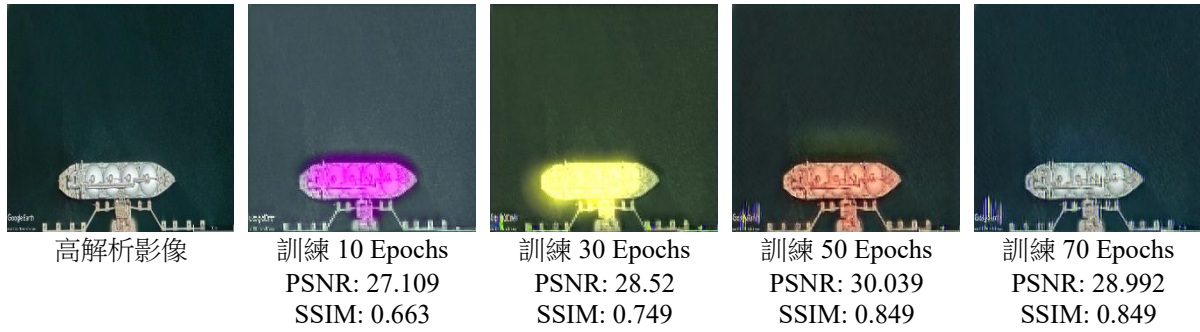


圖 7 混合高解析度及超高解析度衛星影像產製之超解析影像比較

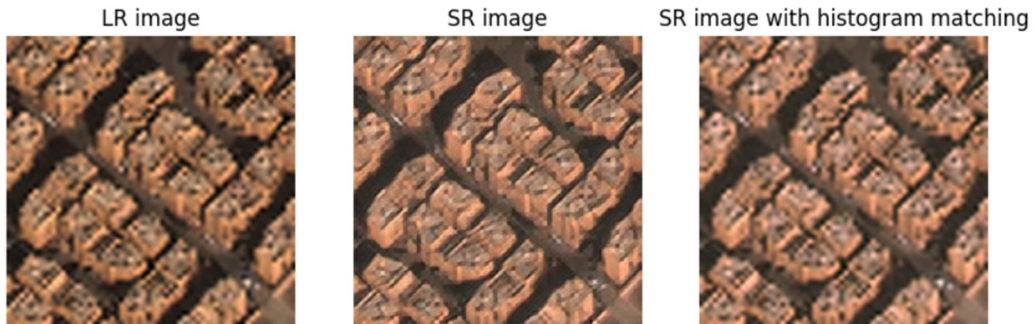


圖 8 比較低解析度影像、原始超解析影像及經直方圖匹配處理後之超解析影像

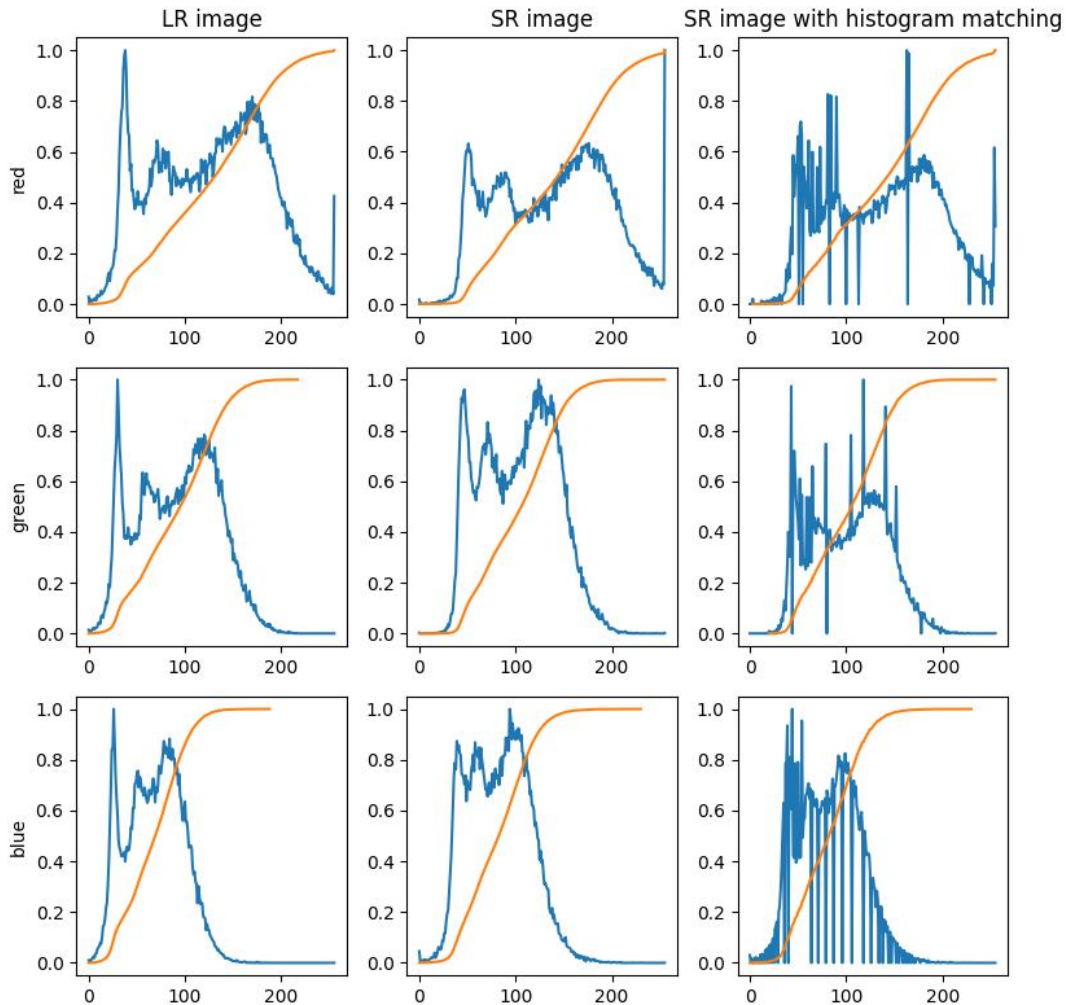


圖 9 低解析度影像、原始超解析影像及經直方圖匹配之超解析影像於紅光、綠光及藍光波段之影像直方圖分布

A. 縮小訓練資料尺寸，減少模型運算資源

相比於中解析度及超高解析度衛星影像資料集，每張影像大小分別為 128×128 及 256×256 ，運用該資料集所生成的超解析影像未有大量雜訊或偽影。因此研判係因較大尺寸影像需消耗更多計算資源和記憶體，將可能導致模型無法有效地學習並生成完整超解析影像，故建議於模型訓練前，需考量硬體設備情況，彈性調整訓練資料之影像大小。

B. 調整生成對抗網路模型架構

為改善影像上隨機出現偽影情況，首先可針對模型架構進行調整，ESRGAN (係由 SRGAN 做為基底改進之模型) 於 2018 年被提出 (Wang *et al.*, 2018)，文獻指出 SRGAN 模型內係使用 Batch Normalization 於訓練中利用批量數據的均值和方差來對影像特徵進行歸一化處理，而在測試階段則使用整個訓練數據集的估計均值和方差。惟當訓練數據和測試數據間的統計特性差異較大時，特別於生成對抗網路 (GAN) 模型框架下，Batch Normalization 更容易產生不穩定的偽影，且偽影在不同的訓練迭代和設置中會隨機出現，導致訓練過程不穩定，限制模型的泛化能力，故於文獻中作者選移除 Batch Normalization，以提升模型訓練穩定度及提升泛化能力，另移除 Batch Normalization 可有效減少模型計算複雜度，有效降低顯卡內存使用量，實驗結果發現運用 SRGAN 模型產製之超解析影像確實於影像內偶有雜訊或偽影，經參考上述方法，未來可嘗試於模型內移除 Batch Normalization 作為後續研究方向，驗證是否可有效改善影像上隨機出現雜訊或偽影問題。

C. 使用單一解析度影像作為訓練資料

本研究使用三種不同組合的訓練資料，主要嘗試運用不同解析度的影像，測試是否能讓不同解析度的影像達到超解析效果。例如：使用解析度 30 cm 的 WorldView-3 超高解析度衛星影像作為訓練資料，期可提升高解析度影像 (如福衛 5 號影像) 的解析度；本研究亦探討將高解析度和超高解析度衛星影像混合作為訓練資料，以期進

一步提升中解析度影像的解析度 (如 Planetscope Dove 影像)；此外亦將解析度 3 m 之中解析影像作為訓練資料，期可提升中低解析度影像的解析度 (如 Sentinel-2 影像)，惟本次僅將測試資料進行降解析，再依模型訓練後的權重，執行超解析處理，以利運用量化指標評估超解析的效果，實驗結果顯示，發現若使用單一解析度影像執行訓練，測試影像上不會產生偽影或雜訊，惟若將混合超高解析度/高解析度影像匯入模型進行訓練，於測試影像發現會隨機產生偽影或雜訊，其可能原因包含本身模型架構造造成訓練不穩定，再者訓練樣本的解析度差異較大，造成訓練時梯度不穩定，模型在嘗試將這些不同解析度的影像學習到相同的超解析度邏輯時，可能會難以統一學習特徵，導致模型不易收斂或收斂速度變慢，最終造成生成結果不穩定，故為改善此問題，建議將訓練樣本的解析度調整更相近，或根據解析度差異進行分層訓練，從而幫助模型更好地學習不同解析度下的特徵。

5. 結論

本研究希望能透過運用生成對抗網路模型，有效提升多種衛星影像之空間解析度，以利後續目標偵測、辨識、和影像判讀等應用。實驗結果顯示，模型透過三個不同的資料集上進行訓練，所生成的超解析影像已經有效地強化目標物的特徵細節。此外，在視覺上，整體影像呈現出更明顯的清晰度和銳利度。綜上顯示，該模型呈現出色的通用性、適應性和泛化能力，使其能夠適用於多種不同來源的影像，進行高品質的超解析處理。然而，所生成的超解析影像均顯示出色調偏移情形，研判起因於生成對抗網路模型在生成超解析影像時需要進行插值和估算，惟為提升影像解析度，向上採樣係為必要的步驟，故導致色調偏移無法避免。儘管如此，本研究可透過更深層的神經網路訓練或進行影像後處理，嘗試改善色調偏移的情況。總體而言，生成式對抗網路深度學習可有效達成衛星影像的超解析處理，提高影像的空間細節，有利於目標偵測、辨識及影像判讀等應用；惟超解析影像的色彩偏移

則不利於光譜相關的處理與分析。

參考文獻

- Aiello, M., Vezzoli, R., and Gianinetta, M., 2019. Object-based image analysis approach for vessel detection on optical and radar images, *Journal of Applied Remote Sensing*, 13(01): 014502, DOI: 10.1117/1.JRS.13.014502.
- Alain, H., and Djemel, Z., 2010. Image quality metrics: PSNR vs. SSIM, *Proceedings of the 20th International Conference on Pattern Recognition*, Istanbul, Turkey., pp.2366-2369, DOI: 10.1109/ICPR.2010.579.
- Boain, R.J., 2004. A-B-Cs of sun-synchronous orbit mission design, *Proceedings of the AAS/AIAA Space Flight Mechanics Conference*, AAS04-108, Hawaii, USA.
- Dong, C., Loy, C.C., He, K., and Tang, X., 2016. Image super resolution using deep convolutional networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2): 295-307, DOI: 10.1109/TPAMI.2015.2439281.
- Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., 2014. Generative adversarial nets, *Proceedings of the Advances in Neural Information Processing Systems 27 (NIPS 2014)*.
- He, K., Zhang, X., Ren, S., and Sun, J., 2016. Deep residual learning for image recognition, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, pp.770-778, DOI: 10.1109/CVPR.2016.90.
- Khaledyan, D., Amirany, A., Jafari, K., Moaiyeri, M.H., Khuzani, A.Z., and Mashhadi, N., 2020. Low-cost implementation of bilinear and bicubic image interpolation for real-time image super resolution, *Proceedings of the IEEE Global Humanitarian Technology Conference (GHTC)*, Seattle, USA, DOI: 10.1109/GHTC46280.2020.9342625.
- Kim, J., Lee, J.K., and Lee, K.M., 2016. Accurate image super resolution using very deep convolutional networks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, pp. 1646-1654, DOI: 10.1109/CVPR.2016.182.
- Kim, M., Park, S., Sampath, A., Anderson, C., and Stensaas, G.L., 2022. System characterization report on Planet SkySat, Chapter E of System Characterization of Earth Observation Sensors, U.S. Geological Survey Open-File Report 2021-1030-E, DOI: 10.3133/ofr20211030E.
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., and Shi, W., 2017. Photo-realistic single image super resolution using a generative adversarial network, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, pp. 105-114, DOI: 10.1109/CVPR.2017.19.
- Li, J., Qu, C., and Shao, J., 2017. Ship detection in SAR images is based on an improved faster R-CNN, *Proceedings of the SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA)*, Beijing, China, DOI: 10.1109/BIGSAR DATA.2017.8124934.
- Lim, B., Son, S., Kim, H., Nah, S., and Lee, K., 2017. Enhanced deep residual networks for single image super resolution, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA, pp.1132-1140, DOI: 10.1109/CVPRW.2017.151.
- Shibani, S., Dimitris, T., Andrew, I., and Aleksander, M., 2018. How does batch normalization help optimization?, *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS 2018)*, Montréal, Canada, DOI: 10.48550/arXiv.1805.11604.

- Vo, K.D., and Bui, L.T., 2023. StarSRGAN: Improving real-world blind super-resolution, Proceedings of the International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2023 (WSCG 2023), Czech Republic, pp.62-72, DOI: 10.24132/CSRN.3301.9.
- Wang, X., Xie, L., Dong, C., and Shan, Y., 2021. Real-ESRGAN: Training real-world blind super resolution with pure synthetic data, Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, pp. 1905-1914, DOI: 10.1109/ICCVW54120.2021.00217.
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., and Loy, C.C., 2018. ESRGAN: Enhanced super-resolution generative adversarial networks, Proceedings of the European Conference on Computer Vision (ECCV) Workshops, pp 63-79, DOI: 10.1007/978-3-030-11021-5_5.
- Xiong, L.F., Guo, S.X., Chen, J.S., Deng, X.P., Sun, L.Y., Zheng, X.R., and Xu, W.N., 2020. Improved SRGAN for remote sensing image super resolution across locations and sensors, Remote Sensing, 12(8): 1263, DOI: 10.3390/rs12081263.

The Preliminary Evaluation of Generating Super-Resolution Satellite Images Using Generative Adversarial Networks

Ting-Shao Chang¹ Fuan Tsai^{2*}

Abstract

This study explores the use of Generative Adversarial Networks (GAN) to enhance the spatial resolution of satellite images, addressing resolution degradation caused by factors such as off-nadir angles, weather conditions, and sensor limitations. The research utilizes medium, high, and very-high-resolution images, applying downsampling and training the GAN to generate super-resolution images. During GAN training, the generator reconstructs low-resolution images into super-resolution ones, while the discriminator distinguishes between generated and real images. The study also utilizes the VGG-19 pre-trained model for feature extraction to improve image quality. Experimental results show that image details become sharper as training progresses, and the GAN outperforms traditional methods in terms of structural similarity. However, issues like color shifts and artifacts emerged during image generation. To address these problems, the study recommends deeper training, post-processing techniques, and model optimizations such as removing Batch Normalization. Overall, while GAN models show potential for enhancing satellite image resolution, further improvements are needed to resolve color shifts and artifacts, focusing on model stability and post-processing.

Keywords: Super Resolution, Satellite Imagery, Generative Adversarial Network, Deep Learning

¹ Master, Department of Civil Engineering, National Central University

² Professor, Center for Space and Remote Sensing Research, National Central University

* Corresponding Author, Tel: 886-3-4227151 ext. 57619, E-mail: ftsai@csrsr.ncu.edu.tw

Received Date: Apr. 22, 2024

Revised Date: Jun. 26, 2024

Accepted Date: Sep.25, 2024