

Super-resolución de imágenes satelitales mediante GAN: Un enfoque basado en entrenamiento con imágenes aéreas

Magda Alexandra Trujillo-Jiménez^{1,5}[0000-0001-5506-3496], Francisco Iaconis²[0000-0002-2373-5793],
Debora Pollicelli³[0009-0006-8771-2538], Gisela Noelia Revollo Sarmiento⁴[0000-0002-1532-5428]

¹ Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur - CONICET, Bahía Blanca, Argentina.

² Instituto de Física del Sur, Departamento de Física, Universidad Nacional del Sur - CONICET, Bahía Blanca, Argentina.

³ Laboratorio de Investigación en Informática, Departamento de Informática, Facultad de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco, Puerto Madryn, Argentina.

⁴ Instituto de Ecorregiones Andinas (INECOA - CONICET), Facultad de Ingeniería, Universidad Nacional de Jujuy, Jujuy, Argentina.

⁵ Instituto Patagónico de Ciencias Sociales y Humanas, Centro Nacional Patagónico CCT CENPAT CONICET, Puerto Madryn, Argentina.
mtrujillo@cenpat-conicet.gob.ar

Resumen. Las imágenes satelitales suelen presentar limitaciones en su resolución espacial y alto costo, lo que dificulta su uso en aplicaciones como el monitoreo urbano o de fauna. En este trabajo, se propone un enfoque novedoso que utiliza imágenes aéreas de alta resolución para entrenar un modelo de super-resolución basado en Redes Generativas Adversarias (GANs). En particular, se adaptó el modelo ESRGAN (Enhanced Super-Resolution Generative Adversarial Network), optimizando sus parámetros para reducir el tiempo de entrenamiento y mejorar su eficiencia. Los resultados preliminares muestran una mejora visual significativa en la calidad de las imágenes satelitales, demostrando la viabilidad de transferir la capacidad de super-resolución desde imágenes aéreas a imágenes satelitales. Este trabajo sienta las bases para futuras investigaciones orientadas a desarrollar modelos especializados en la super-resolución de imágenes.

Palabras clave: Redes Generativas Adversarias, Imágenes Satelitales, Imágenes Aéreas.

Super-resolution of satellite imagery using GAN: An approach based on training with aerial imagery.

Abstract. Satellite images usually present limitations in their spatial resolution and high cost, which hinders their use in applications such as urban or wildlife monitoring. In this work, a novel approach is proposed that uses high-resolution aerial imagery to train a super-resolution model based on Generative Adversarial Networks (GANs). In particular, the ESRGAN (Enhanced Super-Resolution Generative Adversarial Network) model was adapted, optimizing its parameters to reduce the training time and improve its efficiency. Preliminary results show a significant visual improvement in the quality of satellite images, demonstrating the feasibility of transferring super-resolution capability from aerial images to satellite images. This work lays the foundation for future research aimed at developing specialized models for super-resolution imaging.

Keywords: Adversarial Generative Networks, Satellite Imagery, Aerial Imagery.

1 Introducción

La resolución espacial de las imágenes satelitales es un factor crítico en aplicaciones como el monitoreo urbano, la gestión de recursos naturales y la conservación de fauna (Liu et al., 2015, Wang et al., 2019). Sin embargo, debido a limitaciones técnicas y costos asociados, muchas de estas imágenes presentan una resolución insuficiente para abordar tareas que requieren un alto nivel de detalle (Shan et al., 2018). Esto ha llevado a la exploración de técnicas de super-resolución (Yang et al., 2019, Karwowska et al., 2022), que buscan mejorar la calidad de las imágenes mediante el procesamiento computacional. Entre estas técnicas, las Redes Generativas Adversarias (GANs) han demostrado un gran potencial debido a su capacidad para generar detalles de alta calidad a partir de imágenes de baja resolución (Ledig et al., 2017, Karwowska et al., 2022).

En este contexto, el modelo ESRGAN (Enhanced Super-Resolution Generative Adversarial Network) se ha posicionado como una de las arquitecturas más efectivas para tareas de super-resolución en imágenes generales (Wang et al., 2018, 2021). No obstante, su aplicación en dominios específicos, como las imágenes satelitales, requiere adaptaciones y entrenamientos especializados. Este trabajo propone un enfoque novedoso que utiliza imágenes aéreas de drones de alta resolución, más precisamente ortofotos para entrenar este modelo, con el objetivo de transferir esta capacidad de super-resolución a imágenes satelitales de baja resolución de la misma zona geográfica. Las imágenes aéreas, al tener una resolución significativamente mayor, permiten generar pares de imágenes de alta y baja resolución ideales para el entrenamiento del modelo. Los resultados preliminares, muestran una mejora significativa en la calidad de las imágenes aéreas y satelitales, lo que valida el potencial de esta metodología.

2 Materiales y Métodos

2.1 Conjunto de datos y pre-procesamiento

Se utilizó un conjunto de datos compuesto por 2 imágenes aéreas tipo ortofotos de la ciudad de Lleida. Las imágenes tienen una resolución espacial de 25 cm el pixel y fueron descargadas en formato RGB (banda roja, verde y azul), utilizando el complemento “Open ICGC” (ICGC, 2024), del Sistema de Información Geográfico QGIS (<https://qgis.org/>). Las ortofotos son imágenes aéreas verticales corregidas geométricamente para mantener una escala uniforme en toda su extensión, lo que permite representar con precisión la superficie terrestre. Una ortofoto en color captura información del espectro visible, combinando las bandas RGB (Rojo, Verde y Azul) para generar una imagen en "color natural", que refleja la apariencia real del paisaje. Las imágenes originales, de tamaños de 15715x12261 y 13480x14564 píxeles, fueron procesadas para generar 1418 recortes de alta y baja resolución. Para ello, se extrajeron recortes de 512x512 píxeles como referencia de alta resolución. Luego, cada recorte fue escalado a una versión de baja resolución de 128x128 píxeles, utilizando el método de interpolación por área, que preserva mejor los detalles al reducir el tamaño de la imagen. Esta reducción se realizó dividiendo las dimensiones de la imagen por un factor de 4.

Para el proceso de prueba, se utilizaron imágenes multiespectrales SuperView-1 con aproximadamente un 10% de coberturas de nubes. Las bandas utilizadas fueron azul (450–520 nm), verde (520–590 nm) y rojo (630–690 nm) con una resolución espacial de 2 m. Se realizaron recortes en RGB de 615 x 615 píxeles.

Es fundamental que el entrenamiento se realice con una gran cantidad y diversidad de imágenes, de modo que incluso los patrones menos frecuentes estén estadísticamente representados. En caso contrario, por ejemplo en contextos donde existan particularidades como cultivos experimentales con líneas de siembra circulares en lugar de longitudinales, el modelo tenderá a reconstruir patrones mayoritarios (e.g., líneas longitudinales), introduciendo sesgos. Cabe destacar que, dado que la escala de interpolación de píxeles es menor que los patrones espaciales de interés, esta limitación no representa un obstáculo crítico para la calidad general de los resultados. El entrenamiento se realizó utilizando una estación de trabajo equipada con un procesador Intel Core i9-13900KF y una GPU NVIDIA RTX 3080 Ti de 12 GB de VRAM GDDR6 (ASUS TUF).

2.2 Entrenamiento del modelo

Para mejorar la resolución de las imágenes satelitales, se utilizó ESRGAN (Enhanced Super-Resolution Generative Adversarial Network) (Wang et al., 2018, 2021), un modelo basado en Redes Generativas Adversarias (GANs) diseñado para la super-resolución de imágenes. ESRGAN se basa en una arquitectura con bloques residuales densos (RRDB), los cuales permiten mejorar la reconstrucción de detalles finos y reducir artefactos no deseados en la imagen generada. A diferencia de modelos previos, ESRGAN incorpora una función de pérdida perceptual, que optimiza la similitud visual entre la imagen generada y la imagen real de alta resolución, logrando resultados más realistas y detallados. Dado que ESRGAN ha sido desarrollado y

optimizado para imágenes generales, en este trabajo se realizaron modificaciones en su entrenamiento para adaptarlo al dominio específico de imágenes aéreas y satelitales.

Para optimizar el entrenamiento y el uso de recursos computacionales, se ajustaron varios parámetros del modelo. El tamaño del batch se configuró en 2 imágenes en la primera fase y en 1 en la fase optimizada, reduciendo el consumo de memoria. Se habilitó el uso de GPU para acelerar el proceso de entrenamiento mediante paralelización, y se fijaron 1000 iteraciones para balancear tiempo de cómputo y calidad de resultados. Además, se emplearon pesos pre-entrenados (ESRGAN-V1) para reducir el tiempo de convergencia. El modelo utilizó funciones de pérdida combinadas (L1, perceptual y adversaria), mejorando tanto la reconstrucción estructural como la calidad visual de las imágenes generadas.

3 Resultados Preliminares

Los primeros resultados demuestran que el modelo entrenado con imágenes aéreas de drones logra una mejora visual significativa en la resolución y calidad de las imágenes satelitales. Las imágenes generadas por el modelo presentan detalles más nítidos y una mejor definición en comparación con las versiones originales de baja resolución. En particular, se observa una mayor claridad en estructuras urbanas, como edificios y calles, así como en detalles del terreno, como vegetación y cuerpos de agua (ver Fig.1). Estos resultados sugieren que el enfoque propuesto tiene potencial para aplicaciones prácticas en tareas de monitoreo urbano y análisis de cambios en entornos específicos.

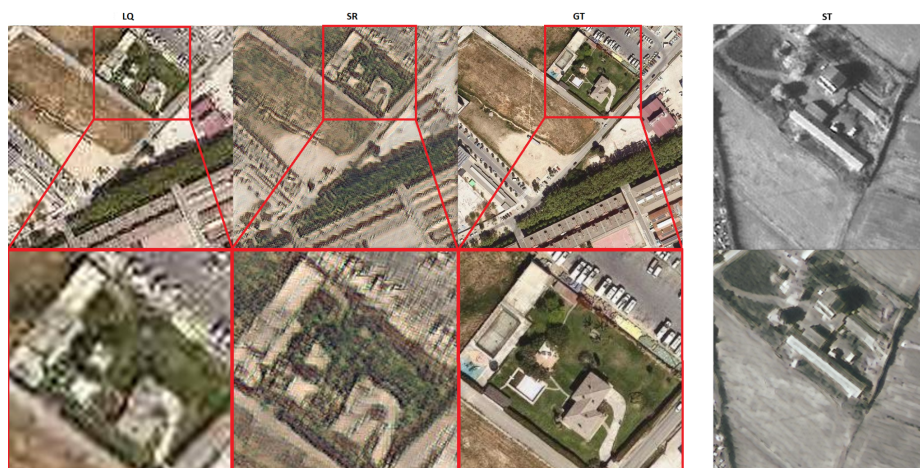


Fig. 1. Resultados de super-resolución utilizando la red ESRGAN. Comparación visual entre imágenes de baja calidad (LQ), imágenes mejoradas mediante super-resolución (SR) y las imágenes de referencia de alta calidad (GT). A la derecha, ejemplo de su aplicación en imágenes satelitales (ST). Los resultados ilustran la capacidad del modelo para recuperar detalles finos y mejorar la percepción visual en imágenes de baja resolución.

4 Conclusiones y Trabajo Futuro

Este estudio demuestra la viabilidad de transferir la capacidad de un modelo de super-resolución entrenado con imágenes de drones a imágenes satelitales de baja resolución. Sin embargo, se identifican algunos desafíos, como la necesidad de ajustar el modelo para manejar las diferencias espectrales y geométricas entre imágenes aéreas y satélites. Para abordar estas limitaciones se propone como trabajos futuros: (1) ampliar el conjunto de datos, incluyendo imágenes de diversas regiones geográficas y condiciones ambientales, con el fin de mejorar la generalización del modelo y reducir su dependencia de datos específicos, (2) explorar adaptaciones específicas, como el uso de otros modelos arquitectónicos (3) incorporar métricas cuantitativas robustas (p. ej. Peak Signal-to-Noise Ratio, Structural Similarity Index, Learned Perceptual Image Patch Similarity, Mean Squared Error y Multi-Scale Structural Similarity Index) para evaluar de manera más precisa el rendimiento del modelo y poder comparar los modelos con enfoques alternativos (Arabboev et al., 2024). Este trabajo sienta las bases para el desarrollo de modelos especializados en la super-resolución de imágenes satelitales, a bajo costo, con aplicaciones potenciales en monitoreo urbano y la conservación de la fauna.

Bibliografía

- Liu, T., & Yang, X. (2015). Monitoring land changes in an urban area using satellite imagery, GIS and landscape metrics. *Applied geography*, 56, 42-54.
- Wang, D., Shao, Q., & Yue, H. (2019). Surveying wild animals from satellites, manned aircraft and unmanned aerial systems (UASs): A review. *Remote Sensing*, 11(11), 1308.
- Shan, J., Weng, Q., Ehlers, M., Quattrochi, D. A., Zhou, G., Stilla, U., ... & Anderson, S. (2018). *Urban remote sensing*. CRC press.
- Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J. H., & Liao, Q. (2019). Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia*, 21(12), 3106-3121.
- K. Karwowska and D. Wierzbicki, "Using Super-Resolution Algorithms for Small Satellite Imagery: A Systematic Review," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 3292-3312, 2022, doi: 10.1109/JSTARS.2022.3167646.
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4681-4690).
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., ... & Change Loy, C. (2018). Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops* (pp. 0-0).
- Wang, X., Xie, L., Dong, C., & Shan, Y. (2021). Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 1905-1914).
- Institut Cartogràfic i Geològic de Catalunya ICGC. (s.f.). Ortofoto 25 cm v7r0 color. Infraestructura de Dades Espacials de Catalunya (IDE).
- Arabboev, M., Begmatov, S., Rikhsivoev, M., Nosirov, K., & Saydiakbarov, S. (2024). A comprehensive review of image super-resolution metrics: classical and AI-based approaches. *Acta IMEKO*, 13(1), 1-8.