**UNIVERSIDAD REAL, MAYOR Y PONTIFICIA DE SAN FRANCISCO XAVIER DE CHUQUISACA**

Facultad de ciencias de la tecnología

Ing. En Ciencias de la Computación



PSNR Mejora de resolución en imágenes

Vargas Paredes Jherson Efrain

Sucre – Bolivia

2025

Índice

[1.1 Antecedentes 1](#_Toc209735868)

[1.2 Tema 1](#_Toc209735869)

[1.2.1 ¿Qué es un agente de IA? 1](#_Toc209735870)

[1.2.2 Evolución y contexto del auge de los agentes IA 1](#_Toc209735871)

[1.3 AutoGPT 2](#_Toc209735872)

[1.3.1 ¿Qué es AutoGPT? 2](#_Toc209735873)

[1.3.2 Arquitectura 2](#_Toc209735874)

[1.4 OpenAgents 4](#_Toc209735875)

[1.4.1 ¿Qué es OpenAgents? 4](#_Toc209735876)

[1.4.2 Arquitectura 4](#_Toc209735877)

[1.5 Comparación 6](#_Toc209735878)

[1.6 Casos de uso 7](#_Toc209735879)

[1.6.1 AutoGPT 7](#_Toc209735880)

[1.6.2 OpenAgents 7](#_Toc209735881)

[1.7 Ventajas y limitaciones 8](#_Toc209735882)

[1.7.1 AutoGPT 8](#_Toc209735883)

[1.7.2 Enfoque 9](#_Toc209735884)

[1.8 Herramientas y frameworks relacionados, alternativas y tendencias (2025) 9](#_Toc209735885)

[1.9 Conclusiones y Recomendaciones 10](#_Toc209735886)

[Bibliografía 10](#_Toc209735887)

# Antecedentes

El campo de la superresolución de imágenes ha experimentado una revolución sustancial en la última década, motivada por avances en el aprendizaje profundo y la evolución de arquitecturas dedicadas al procesamiento visual. Originalmente, la superresolución de imágenes (SR) se abordaba mediante técnicas tradicionales como interpolaciones (bilineal, bicúbica) y métodos basados en patches, que resultaban eficientes computacionalmente, pero mediocres en la reconstrucción de detalles y texturas. La llegada de redes neuronales profundas, especialmente las CNNs, permitió mapear de manera no lineal imágenes de baja resolución (LR) a sus contrapartes de alta resolución (HR), aprendiendo representaciones y patrones que los enfoques clásicos no podían modelar.

A partir de 2014, con la introducción de SRCNN, se sentaron las bases para una ola de modelos más sofisticados y profundos. VDSR, FSRCNN y ERSGAN supusieron avances considerables, dando como resultados con una mejora en resolución y nitidez sorprendentes para la época, sin embargo, la capacidad de “alucinar” en estos modelos a pesar de permitir texturas y bordes definidos muchas veces suponía un problema ya en imágenes con poca información podía tergiversar los valores de la imagen y crear algo totalmente diferente a la imagen original.

Bajo ese contexto surge PSNR más que proponer una arquitectura novedosa nos propone mejorar la métrica de valoración respecto a la imagen real. Esto quiere decir que ahora la red tratara de alucinar a un valor de manera que a nivel de valores matriciales el resultado sea lo más parecido a la imagen real pero también tratando de agregar detalles para mejorar la imagen.

La transición de la SR basada en loss pixel a loss perceptuales y adversariales marcó un punto de inflexión. Es importante destacar que esta evolución estuvo motivada tanto por los avances científicos como por las crecientes exigencias de aplicaciones reales: desde la restauración de material audiovisual, recuperación de capturas médicas hasta la mejora de videos de vigilancia y satelitales.

# Tema

Este informe profundiza en PSNR (Peak Signal-to-Noise Ratio o Relación señal-ruido máxima), analizando sus fundamentos, arquitectura, evolución, aplicaciones y relevancia en el ecosistema de la inteligencia artificial aplicada a la superresolución de imágenes. PSNR representa un hito en la reconstrucción perceptual, perfeccionando caminos iniciados por SRGAN y abordando limitaciones como artefactos y detalles poco naturales.

## Arquitectura y funcionamiento

PSNR surge como sucesor directo de ERSGAN, resolviendo importantes debilidades en la generación de detalles y la aparición de artefactos indeseables. Este avance es el resultado de tres innovaciones clave:

***a) Residual-in-Residual Dense Blocks (RRDB):*** PSNR reemplaza los bloques residuales estándar de SRGAN por bloques densos residuales en residuales. Los RRDB eliminan la normalización por lotes, permitiendo una mayor profundidad y mejor transmisión de gradientes, lo que se traduce en una mejor preservación de texturas y detalles de alta frecuencia. Su estructura interna utiliza múltiples conexiones densas y residuales, facilitando la reutilización de características y la mejora perceptual sin el costo de artefactos introducidos por normalización excesiva.

***b) Discriminador Relativista (RaGAN):*** La arquitectura del discriminador también evoluciona al emplear una versión relativista, que aprende no solo a distinguir entre imágenes reales y generadas, sino a juzgar la “realidad relativa” de una imagen respecto al conjunto. Esto refuerza la presión adversarial sobre el generador, promoviendo salidas más cercanas a la apariencia natural.

***c) Pérdida Perceptual Mejorada:*** En lugar de comparar características tras las capas de activación (como en SRGAN), PSNR utiliza mapas de características antes de la activación. Este cambio brinda una supervisión más fuerte de brillo y texturas, conduciendo a imágenes coherentes perceptualmente. Además, se incorpora una pérdida de contenido (L1 o L2), pérdida adversarial basada en RaGAN y, en variantes como Real-ESRGAN, adaptación a degradaciones del mundo real.

## Proceso de Entrenamiento y Estrategias Adicionales

El entrenamiento de PSNR se realiza tradicionalmente en dos fases: primero, una pre-entrenamiento enfocado en PSNR (minimizando MSE/L1 con dataset de imágenes HR y LR), seguido por un ajuste fino adversarial con pérdidas perceptuales y adversariales para mejorar la calidad visual. Además, técnicas como el residual scaling y la inicialización pequeña permiten profundizar la red sin provocar divergencias o explosión de gradientes.

Una técnica característica es la interpolación de redes: se combinan modelos orientados a PSNR puro y modelos adversariales, permitiendo ajustar el balance entre fidelidad cuantitativa y percepción visual sin reentrenar modelos desde cero.

## Problemas y Tecnologías similares

El PSNR es rápido de calcular, reproducible y ampliamente adoptado, sobre todo en contextos controlados y para compresión de imágenes técnicas (gráficos, diagramas, logotipos) donde los errores localizados pueden ser críticos. Sin embargo, PSNR no considera la percepción visual humana (no distingue alteraciones imperceptibles localizadas ni es sensible a la estructura, brillo o contraste) y puede asignar altos valores numéricos a imágenes que, en la práctica, lucen borrosas, artificiales.

Esta limitación ha motivado el desarrollo y adopción de modelos más alineadas con la visión humana, como SSIM, LPIPS, DLSS y Transformers que incorporan elementos que se perciben mejor en el sistema visual humano.

# 1.6 Casos de uso

La superresolución basada en PSNR encontraron aplicaciones en una amplia diversidad de industrias y contextos críticos, entre los que destacan:

* **Restauración de material audiovisual y fotografías antiguas:** PSNR permite resucitar detalles y texturas en imágenes históricas o deterioradas donde la eliminación de desenfoques y artefactos es esencial para preservación cultural.
* **Imágenes médicas:** El aumento de claridad y la supresión de ruido en radiografías, resonancias y tomografías facilita diagnósticos más precisos y un mejor seguimiento clínico.
* **Sensores remotos, teledetección y vigilancia:** La capacidad de PSNR para recuperar información de imágenes satelitales o de drones incrementa el valor estratégico en agricultura de precisión, monitoreo ambiental y seguridad nacional.
* **Videojuegos y animación:** Aplicaciones para la ampliación y mejora de texturas de bajo presupuesto sin perder coherencia visual ni generar artefactos ajenos al estilo original.
* **Reconstrucción facial y redes sociales:** Algunas variantes de PSNR muestra una mejora evidente en tareas de aumento de calidad en imágenes de rostros, preservando texturas y detalles sutiles (ojos, piel, cabello) fundamentales para la percepción humana y biométrica.
* **Optimización de contenido multimedia para la web:** El uso de PSNR permite aumentar la resolución y la calidad de imágenes en sitios de comercio electrónico, marketing y publicación digital, mejorando la experiencia visual y el ratio de conversión de clientes.

No obstante, el éxito de estos modelos depende de la adecuada selección de métricas de evaluación y de la capacidad de adaptación a los propios conjuntos de datos, degradaciones y requisitos específicos del dominio.

# 1.7 Ventajas y limitaciones

**a) Ventajas de PSNR**

* **Calidad cuantitativa superior:** Incluyendo detalles de alta frecuencia y texturas naturales, esenciales para aplicaciones culturales, médicas o publicitarias.
* **Arquitectura flexible:** Al estar basada en ERSGAN y ser parte de RRDBs permite una red más profunda y fácilmente reentrenable, facilitando la transferencia a nuevas tareas y degradaciones.
* **Robustez a condiciones complejas:** PSNR, en particular, mejora la resiliencia ante ruidos, artefactos y distorsiones reales, haciéndolo viable para imágenes del mundo real, no solo datasets sintéticos.
* **Versatilidad en aplicaciones:** Los avances permiten desde la restauración de imágenes históricas hasta la potenciación de video digital, imágenes médicas o vigilancia remota.

**b) Limitaciones**

* **Costo computacional elevado:** Los modelos GAN, especialmente PSNR y variantes profundas, requieren mucho tiempo de cómputo y hardware especializado para su entrenamiento y ejecución.
* **Propagación de artefactos:** Sin un ajuste cuidadoso de hiperparámetros, es posible que se generen detalles artificiales o poco naturales, especialmente en contextos con datos insuficientes o ruidos atípicos.
* **Dependencia de la evaluación subjetiva:** Aunque PSNR ofrecen guía cuantitativa, la verdadera calidad es, en última instancia, percibida por humanos, lo que complica la comparación directa y la optimización algorítmica.
* **Requerimiento de grandes datasets:** El rendimiento óptimo de PSNR y sus variantes depende de la disponibilidad de grandes cantidades de imágenes pareadas HR/LR y de la diversidad de degradaciones presentes en los datos de entrenamiento.
* **Limitaciones de PSNR:** En resultados, falla en capturar la similitud perceptual; no detecta bien errores estructurales ni localizados, y puede sobrevalorar restauraciones borrosas, pero numéricamente estables.

# 1.9 Conclusiones y Recomendaciones

La combinación de modelos GAN, con métricas objetivas y subjetivas, marca el estado del arte en la superresolución de imágenes. PSNR demuestra que es posible trascender el mero ajuste numérico y avanzar hacia una calidad visual que responde a la percepción humana, lo que ha sido refrendado por la adopción masiva en aplicaciones industriales, médicas y culturales.

Sin embargo, este salto cualitativo viene acompañado de desafíos: la necesidad de recursos computacionales sustanciales, la elevada sensibilidad a los hiperparámetros y la dificultad de evaluar y comparar métodos de forma universal. La investigación muestra que, mientras PSNR sigue resultando útil para comparaciones automatizadas y control de calidad técnica, debe ser complementada (nunca sustituida) por métricas perceptuales especialmente en dominios donde la decisión final la toma el ojo humano.

Como tendencia emergente, variantes como AESRGAN muestran la viabilidad de modelos híbridos y ajustados a degradaciones del mundo real, así como el potencial de mecanismos de atención para la preservación de estructuras semánticas críticas.

# Bibliografía

1. Ledig, C., Theis, L., Huszár, F., Caballero, J., Aitken, A. P., Tejani, A., et al. (2017). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. arXiv preprint arXiv:1609.04802v5. <https://arxiv.org/pdf/1609.04802v5>
2. Rubii. (2025, marzo 12). Real-ESRGAN: Tutorial completo para mejorar la calidad de imágenes y videos. AI News ES. <https://github.com/eroge69/RealesrganGUI>
3. Unified Image Tools. (2025). Métricas de Calidad de Imagen SSIM/PSNR/Butteraugli Guía Práctica 2025. <https://unifiedimagetools.com/es/articles/image-quality-metrics-ssim-psnr-butteraugli-2025>
4. Wikipedia. (2025). PSNR. <https://es.wikipedia.org/wiki/PSNR>
5. IBM. (2025). Casos de uso, ejemplos y aplicaciones generativas. <https://www.ibm.com/es-es/think/topics/generative-ai-use-cases>
6. Wikipedia. (2025). Peak Signal-to-Noise Ratio. <https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio>
7. Patsnap Eureka. (2025, julio 10). Perceptual Metrics Face-Off: LPIPS vs SSIM vs PSNR. <https://eureka.patsnap.com/article/perceptual-metrics-face-off-lpips-vs-ssim-vs-psnr>