



Modelado de lentes gravitacionales a través de redes neuronales convolucionales

Modeling of gravitational lenses through convolutional
neural networks

Juan Jordi Ancona Flores*
Alberto Hernández Almada
Octavio Cornejo Pérez

Universidad Autónoma de Querétaro,
Querétaro, México.

*jancona16@alumnos.uaq.mx

04

¿CÓMO CITAR ESTE ARTÍCULO?

J. J. Ancona Flores, A. Hernández Almada, y O. Cornejo Pérez, "Modelado de lentes gravitacionales a través de redes neuronales convolucionales", *Perspectivas de la Ciencia y la Tecnología*, vol. 9, núm. 16, pp. 54-74, 2026.



Resumen

El presente estudio aborda el modelado de lentes gravitacionales fuertes (SGL) mediante redes neuronales convolucionales (CNN), con el objetivo de mejorar la inferencia de parámetros físicos asociados al modelo elipsoide isotérmico singular (SIE). Las SGL constituyen herramientas fundamentales para investigar la materia oscura y la energía oscura; y del mismo modo, para obtener información acerca de la estructura interna de las galaxias. Pero el volumen creciente de datos astronómicos compromete la eficiencia de los métodos tradicionales de modelado. Se generó un conjunto sintético de 76 396 imágenes de sistemas de lentes galaxia-galaxia, derivadas del catálogo del Telescopio Espacial Chino (CSST) utilizando la paquetería Lenstronomy. El modelo CNN se diseñó a partir de la arquitectura Alex-

Net, incorporando modificaciones como estratos adicionales de convolución y normalización por lotes, con el fin de optimizar la extracción de características y reducir el costo computacional. La red predice cuatro parámetros del modelo SIE: el radio de Einstein (θ_E), la relación axial (f) y las componentes de la elipticidad (ϵ_x , ϵ_y). Los resultados muestran una elevada capacidad predictiva, cuyos coeficientes de determinación R^2 van de 0.95 y 0.97, y los errores relativos promedio entre 2.6 % y 5.1 %. Las imágenes reconstruidas alcanzaron una fidelidad PSNR de hasta 37 dB. Estos hallazgos avalan el potencial de las arquitecturas ligeras de CNN orientadas al modelado automático de lentes gravitacionales en grandes volúmenes de datos astronómicos.

Palabras clave: AlexNet, aprendizaje profundo, lentes gravitacionales fuertes, radio de einstein, reconstrucción de Imágenes, redes neuronales convolucionales.

Abstract

This study addresses the modeling of strong gravitational lenses (SGL) through convolutional neural networks (CNN), aiming to improve the inference of physical parameters associated with the singular isothermal ellipsoid (SIE) model. SGL are essential tools for probing dark matter and dark energy, as well as for obtaining insights into the internal structure of galaxies. However, the increasing volume of astronomical data challenges the efficiency of traditional modeling methods. A synthetic dataset of 76 396 galaxy-galaxy lensing images was generated from the Chinese Space Station Telescope (CSST) catalog using the Lenstronomy package. The CNN model was based on the AlexNet architecture,

with modifications including additional convolutional layers and batch normalization to optimize feature extraction and reduce computational cost. The network predicts four parameters of the SIE model: Einstein radius (θ_E), axis ratio (f), and ellipticity components (ϵ_x , ϵ_y). Results demonstrate high predictive performance, with determination coefficients (R^2) ranging from 0.95 to 0.97 and average relative errors between 2.6% and 5.1%. Furthermore, reconstructed images achieved a fidelity of 37 dB in PSNR. These findings confirm the effectiveness of lightweight CNN architectures for the automatic modeling of gravitational lenses across massive astronomical datasets.

Keywords: AlexNet, deep learning, strong gravitational lenses, Einstein's radio, image reconstruction, convolutional neural networks.



Introducción

Una lente gravitacional es un fenómeno astronómico producido por la curvatura del espacio-tiempo alrededor de un objeto masivo, como una galaxia, un cúmulo de galaxias o un cuásar, cuya influencia gravitacional altera la trayectoria de la luz procedente de cuerpos más distantes. Cuando la masa proyectada, que es la acumulada a lo largo de la línea de visión, dentro de la región característica conocida como *radio de Einstein* alcanza una magnitud suficiente para producir múltiples imágenes, el fenómeno resultante recibe el nombre de *lente gravitacional fuerte* (SGL, por sus siglas en inglés). Más aún, cuando la alineación observador-lente-fuente es perfecta, y la fuente alcanza una masa suficientemente grande, se proyecta una circunferencia completa denominada *anillo de Einstein* [1], [2], [3], [4].

La SGL funciona como un telescopio de alta resolución que permite detectar fuentes difíciles de observar directamente. Además, permite sondear la materia oscura dentro de las regiones centrales de los halos, así como determinar el gradiente del perfil de densidad de masa interna de las galaxias y las restricciones en la densidad de energía oscura dentro del Universo [5], [6]. Gracias a las SGL, los telescopios espaciales Euclid [7], [8], el Gran telescopio VLT [9] y el James Webb [10], [11] ya están proporcionando datos sin precedentes, y se espera que las instalaciones de las próximas generaciones, como el telescopio Simonyi en el Observatorio Vera C. Rubin [12], el Extremely Large Telescope [13], [14] y el Telescopio de la Estación Espacial de China (CSST) [15] capturen en pocos años bases de datos en el orden de 10^5 imágenes [16].

Los catálogos conformados hasta ahora incluyen el Cosmic Lens All-Sky Survey [17], [18], el COSMOS [19] y el Sloan Lens ACS [20], [21]. A partir de estos datos, las técnicas convencionales de modelado de lentes intentan abordar el problema inverso no lineal de reconstruir el brillo de las fuentes de lentes, mientras modelan simultáneamente el potencial gravitacional de las lentes. Sin embargo, el volumen ingente (y creciente) de información recopilada impone un reto para los métodos tradicionales analíticos en cuanto a tiempo y recursos.

Tras la detección de una posible lente, resulta fundamental desarrollar un modelo que represente con precisión la distribución total de la masa proyectada, a fin de mitigar la aparición de falsos positivos. Las técnicas de Cadenas de Markov de Monte Carlo se han implementado para el modelado de estos sistemas; sin embargo, su eficiencia ha decaído con el incremento exponencial en el volumen de datos [22], [23]. Para atenuar este problema, las herramientas de inteligencia artificial, como las redes neuronales convolucionales (CNN), han mostrado éxito en la cla-



sificación de imágenes estelares, consolidándose como la técnica predilecta para la búsqueda e identificación de SGL.

Zhang [24] introdujo estas aplicaciones en 1988, al proponer una CNN bidimensional conocida como *red neuronal artificial invariante*. La utilidad principal de las CNN en la clasificación astronómica inició con la caracterización del espectro en el Sloan Digital Sky Survey [25]. Asimismo, Petrillo *et al.* [26] fueron vanguardia en el uso de un método de clasificación morfológica basado en CNN para identificar SGL del catálogo Kilo Degree. Hezaveh *et al.* [27] mostraron la eficiencia de las CNN en la inferencia de valores paramétricos de modelos de masa a partir de imágenes preprocesadas de SGL en alta resolución.

El equipo Highly Optimized Lensing Investigations of Supernovas, Microlensing Objects, and Kinematics of Ellipticals and Spirals (HOLISMOKES) desarrolló una CNN destinada a modelar imágenes de galaxias a partir de la información extraída de SGL [28] y predecir cinco parámetros del modelo de masa elipsoide isotérmico singular (SIE). Adicionalmente, las redes neuronales bayesianas han permitido extraer información altamente abstracta a partir de imágenes complejas. Esta capacidad se ejemplifica en el trabajo de Ji Won Park *et al.* [29], quienes, mediante retrasos de tiempo de lente simulados, determinaron la constante de Hubble; de modo similar, Pearson *et al.* [30] entrenaron una CNN bayesiana aproximada para predecir los parámetros del perfil de masa.

Más recientemente, la Colaboración Euclid [31] diseñó una red neuronal bayesiana para modelar lentes gravitacionales de forma automática y eficiente. La investigación de Parlange *et al.* [32] se sumerge en modelos de transformadores para detectar SGL a partir de conjuntos de datos de HOLISMOKES VI [33] y SuGOHI X [34], los cuales se espera captar con el telescopio Subaru, que emplea una cámara Hyper Suprime-Cam [35].

El presente estudio muestra los resultados de un examen integral del rendimiento de las CNN en la inferencia de los parámetros físicos asociados con SGL. Este análisis detallado implica comparar el conjunto completo de imágenes en busca de similitudes o discrepancias entre las imágenes simuladas y las modeladas con base en los valores inferidos por la CNN. El conjunto de datos sintético comprende 76 396 imágenes de sistemas simulados de lentes de galaxias derivados del catálogo del CSST y generados en un entorno libre de ruido; se apartaron 70 000 imágenes para el proceso de validación cruzada *k-fold*: en cada pliegue se dedicó el 75 % de las imágenes al entrenamiento y el 25 % a la validación. Las 6396 imágenes restantes se reservaron para fines de prueba.



Para este trabajo se creó un modelo CNN basado en la arquitectura AlexNet [36], debido a que las arquitecturas profundas reales exigen un costo computacional excesivo para el procesamiento de millones de imágenes. Las modificaciones realizadas en la red permiten mantener una arquitectura más ligera en comparación con las actuales, facilitando las inferencias rápidas a un bajo costo computacional. La adición de una capa convolucional al bloque central mejora la profundidad de extracción de datos y mitiga la linealidad. El sistema predice cuatro parámetros del modelo de lente SIE: las componentes de elipticidad del lente (ϵ_x y ϵ_y), la proporción axial (f) y el radio de Einstein (θ_E) (ver Sección 2 para sus definiciones). La capacidad predictiva del modelo se evalúa midiendo el error relativo entre los valores reales e inferidos, junto con otras métricas de evaluación. El objetivo de esta investigación es avanzar la interpretabilidad y eficacia de los métodos de aprendizaje profundo en el modelado de lentes gravitacionales a través del uso de una arquitectura ligera diseñada para procesar volúmenes masivos de datos.

El trabajo está organizado de la siguiente manera: en la *Fundamentación teórica* se realiza una breve descripción de los sistemas de lentes gravitacionales fuertes de galaxia-galaxia, así como del trasfondo de las redes neuronales convolucionales. En *Materiales y métodos* se detallan los datos sintéticos utilizados, mostrando los criterios para realizar las simulaciones; también se introduce la arquitectura específica del modelo modificado de CNN. El apartado de *Resultados* muestra las inferencias generadas por el modelo, así como un análisis de los errores relativos. Por último, se exponen las conclusiones derivadas de las pruebas realizadas con el modelo de CNN, así como las recomendaciones y propuestas para futuros trabajos.

Fundamentación teórica

Ecuación de la lente

En la Figura 1 se observa un esquema representativo de un sistema conformado por una concentración de masa (llamada *lente gravitacional*) localizada a una distancia DL del observador, la cual desvía los rayos de luz provenientes de una fuente ubicada a una distancia DS respecto al observador, asimismo, la distancia entre la fuente y la lente se denota por DLS . En este trabajo se adopta el paradigma de la *pantalla delgada*, es decir, se asume que las dimensiones tanto de la fuente como de la lente y el observador son despreciables comparadas con las distancias que los separan [1]; esto nos permite localizar a la lente



y a la fuente en espacios bidimensionales paralelos, llamados respectivamente *plano de la fuente* y *plano de la lente*. Entonces, un haz de luz emitido desde una ubicación en el plano de la fuente $\vec{\eta} = \beta \vec{D}_S$ se desvía en un ángulo $\vec{\alpha}$ en la posición $\vec{\xi} = \vec{\theta} D_L$ al atravesar el plano de la lente, donde $\vec{\xi}$ denota la distancia de dicha posición a la lente (ver Figura 1 para las definiciones de $\vec{\beta}$ y $\vec{\theta}$). Como resultado de esta desviación, el observador percibe imágenes de la fuente en las posiciones angulares aparentes $\vec{\theta}_i$ (el índice i designa cada imagen individual). La ecuación de la lente relaciona los ángulos $\vec{\theta}$, $\vec{\beta}$ y $\vec{\alpha}$ con las distancias D_L , D_S y D_{LS} como se expresa a continuación:

$$\vec{\theta} D_S = \vec{\beta} D_S + \vec{\alpha} D_{LS} \quad (1)$$

Para simplificar la ecuación, se define el ángulo de desviación reducido:

$$\vec{\alpha}(\vec{\theta}) = \frac{D_{LS}}{D_S} \vec{\alpha} \quad (2)$$

Simplificando la ecuación de la lente:

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}) \quad (3)$$

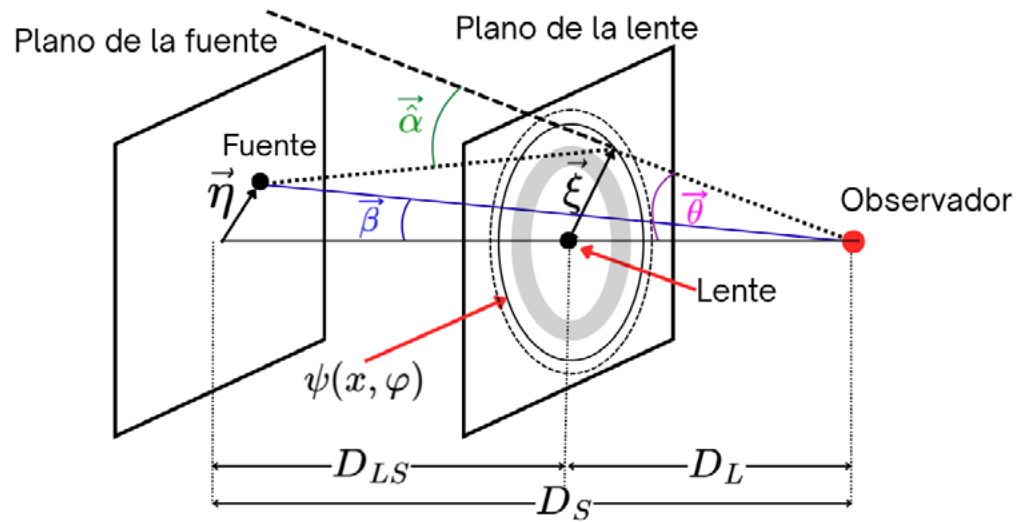


FIGURA 1.
Esquema de un sistema de lente gravitacional. Fuente: adaptado de [37].

Modelo elipsoide isotérmico singular (SIE)

La modelación realista de lentes gravitacionales generadas por los sistemas galaxia-galaxia requiere calcular las distribuciones de masa definidas por contornos de densidad elíptica. Esta investigación se centra en la densidad de superficie según el modelo SIE, predominante en el estudio de lentes gravitacionales fuertes a escala galáctica porque proporciona una descripción efectiva de la distribución total de masa proyectada en el plano de la lente. Es decir, permite representar de manera realista y con simplicidad computacional la



morfología de las galaxias que constituyen una SGL [1], [2], [3], [4], [5], [38]. El modelo SIE puede entenderse como una refinación del modelo esférico isotérmico singular (SIS), que asume simetría esférica y una distribución de velocidad isotérmica. En el modelo SIS, la densidad de masa ($\Sigma_{SIS}(\xi)$) en el plano de la lente está dada por:

$$\Sigma_{SIS}(\xi) = \frac{\sigma_v^2}{2G\xi} \quad (4)$$

donde σ_v^2 representa la velocidad de dispersión de la galaxia lente y ξ es la distancia desde el centro de la lente hasta la posición en el plano de la lente donde ocurre la desviación de la luz (Figura 1).

Para generar las variaciones en la densidad de masa con respecto a la simetría esférica, el SIE introduce elipticidad en la distribución de la masa proyectada, reemplazando la coordenada radial circular en el plano de la lente por una elíptica. Siguiendo [38], esto se logra con la transformación:

$$\xi \rightarrow \sqrt{\xi_1^2 + f^2 \xi_2^2} \quad (5)$$

donde ξ_1 y ξ_2 son las coordenadas de la posición de la fuente en el plano de la lente y f denota la relación axial de la elipse, definida en $0 < f \leq 1$ [39]. Bajo esta premisa, el perfil de densidad de masa para el modelo SIE es:

$$\Sigma_{SIE}(\xi) = \frac{\sigma_v^2}{2G\xi} \frac{\sqrt{f}}{\sqrt{\xi_1^2 + f^2 \xi_2^2}} \quad (6)$$

El potencial gravitacional asociado a la Ecuación (6) se expresa en coordenadas polares como:

$$\psi(x, \varphi) = \frac{\sqrt{fx}}{f'} \left[\sin(\varphi) \arcsin(f' \sin \varphi) + \cos(\varphi) \sinh^{-1}\left(\frac{f'}{f} \cos \varphi\right) \right] \quad (7)$$

Donde:

φ ángulo de posición medido entre el semieje mayor de la elipse de masa de la lente y el eje x del plano de la lente

$$f' = \sqrt{1 - f^2}$$

$$\vec{x} = \vec{\xi}/\xi_0$$

ξ_0 longitud de escala característica de la lente

$$x = |\vec{x}|$$

La relación entre este potencial y el ángulo de deflexión está dada por:

$$\vec{\alpha} = \nabla \psi \quad (8)$$



Cuando existe alineación entre la fuente, la lente y el observador, se obtiene un efecto conocido como *anillo de Einstein*; su radio, conocido como *ángulo de Einstein*, está definido por:

$$\theta_E = 4\pi \frac{\sigma_v^2}{c^2} \frac{D_{LS}}{D_S} \quad (9)$$

donde c denota la velocidad de la luz [38]. Finalmente, la elipticidad está típicamente parametrizada en función de φ como:

$$\varepsilon_x = \frac{1-f}{1+f} \cos(2\varphi) \quad (10)$$

$$\varepsilon_y = \frac{1-f}{1+f} \sin(2\varphi) \quad (11)$$

Ambas componentes están definidas en el intervalo $[-1,1]$.

Redes neuronales convolucionales

Yann LeCun [39] fue pionero en la aplicación efectiva del método de retropropagación para CNN dirigidas al reconocimiento de código escrito a mano, asignando al proceso el término *convolución*. Las CNN representan una evolución de los modelos *feedforward* estándar, y sirven para identificar jerarquías de características mediante mecanismos evolutivos que reducen el costo computacional. Frente a las redes convencionales, las CNN ofrecen ventajas porque optimizan la convergencia al procesar la información de entrada: limitan la conexión de cada capa neuronal a subgrupos específicos de la capa previa y disminuyen la carga de parámetros mediante la gestión del peso, la reducción de dimensionalidad o el *downsampling*.

Una arquitectura CNN típica posee cuatro componentes principales: capas convolucionales, capas de agrupamiento (*pooling*), funciones de activación y capas totalmente conectadas (*fully connected*), como se muestra en la Figura 2. Las capas convolucionales aplican filtros en los datos de entrada para identificar patrones. Las capas de agrupación realizan un muestreo descendente a lo largo de dimensiones espaciales, simplificando así las cargas de parámetros en el proceso de activación; su función principal es reducir las dimensiones espaciales de los mapas de activación sin omitir las características críticas, lo que reduce las demandas computacionales, controla el sobreajuste y logra representaciones invariantes de traducciones pequeñas. Las funciones de activación son indispensables para la expresión de características complejas y actúan de manera similar al modelo neuronal humano al determinar qué información se propaga a cada neurona sucesora; entre funciones de activación se utiliza predominantemente la unidad lineal rectificada (ReLU) [39], que ayuda al mo-



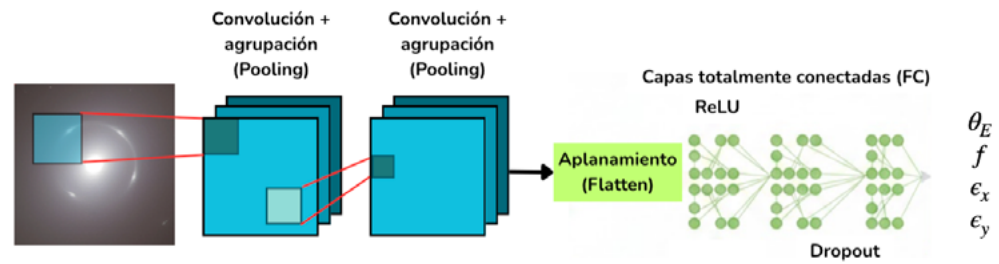
delo a discernir patrones no lineales de imagen. La función ReLU se escribe como:

$$\text{ReLU}(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases} \quad (12)$$

Si una entrada x es negativa, su valor de función es 0; si es positiva, su valor se mantiene intacto.

Además, como en otras redes neuronales, el entrenamiento de CNN requiere algoritmos de aprendizaje automático y regularización; uno de ellos es el *dropout*, que funciona omitiendo neuronas de forma aleatoria para prevenir la coadaptación de detectores de características y reducir el sobreajuste. Esta técnica permite entrenar múltiples configuraciones de red de manera eficiente y en tiempos razonables [40], [41], [42], [43]. Normalmente, la tasa de selección del *dropout* está entre el 20 y 50 %.

FIGURA 2.
Arquitectura ilustrativa de un modelo de CNN. Fuente: elaboración propia.



Materiales y métodos

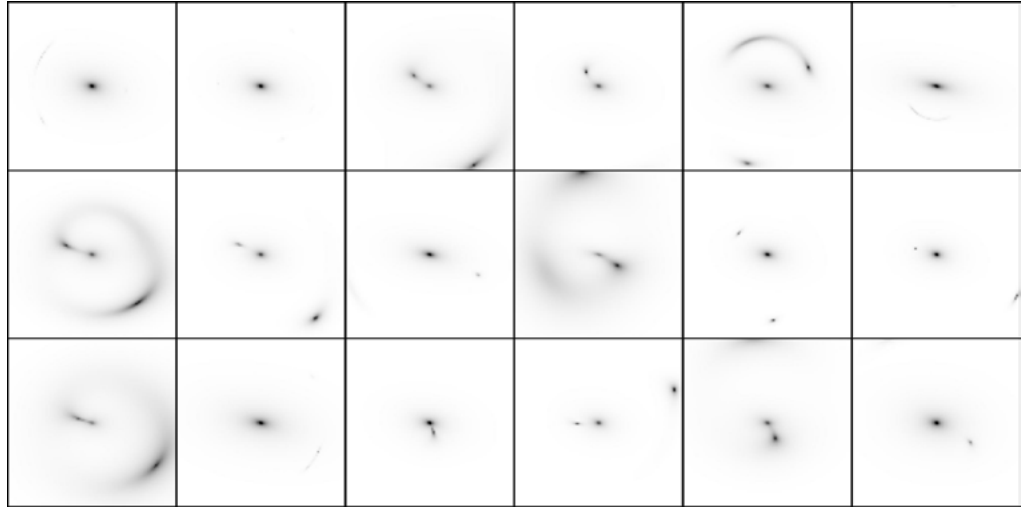
Base de datos simulada del catálogo CSST

El conjunto de datos utilizado para entrenar, validar y probar el modelo CNN consta de 76 396 imágenes de 100×100 píxeles, 1 canal de color y resolución angular de 0.06 segundos de arco/píxel, generadas según las configuraciones de sistemas de lentes fuertes reportadas y validadas por el telescopio espacial CSST [15]. Dicha muestra corresponde a las capturas que se esperarían en barridos angulares amplios del cielo en exposiciones individuales.

La Figura 3 muestra varios ejemplos aleatorios de imágenes sintéticas utilizadas en el análisis. Las imágenes se generaron mediante el paquete *Lenstronomy*, en el entorno Python (3.11.10), que modela el perfil de luminosidad de la fuente mediante el método de Sérsic [44]. S. Mukherjee *et al.* [45] determinaron que la selección meticulosa de los parámetros de origen para el modelado de lentes posee una relevancia limitada, y no introduce sesgos sustanciales en el proce-



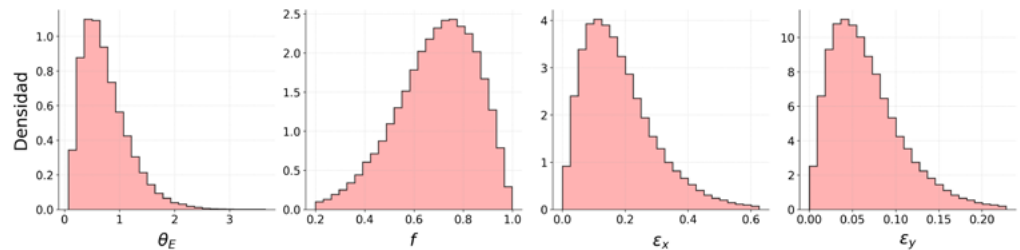
FIGURA 3.
Ejemplos aleatorios
de imágenes sintéticas
de sistemas de lentes
galaxia-galaxia
utilizados en el
entrenamiento.
Fuente: elaboración
propia.



Como se exhibe en las Ecuaciones 9 y 10, las componentes de elipticidad se calcularon en función de la proporción axial de las configuraciones de la muestra de CSST [15]. Para este trabajo, el ángulo de posición utilizado para calcular las componentes de elipticidad para la lente ($\varepsilon_x, \varepsilon_y$) se estableció arbitrariamente en 10° relativo al norte, mientras que el ángulo de posición para computar las componentes de elipticidad de la fuente se extrajo directamente de la muestra.

La distribución de masa de la lente se caracteriza por medio del modelo SIE; y las distribuciones de brillo con el modelo Sérsic a $n_s = 4$. La Figura 4 muestra la distribución paramétrica SIE de la base de datos del CSST. Los registros de θ_E cubren una región entre $[0.07, 3.65]$. Para f , los valores se encuentran entre $[0.2, 1.0]$. Las componentes de elipticidad ε_x y ε_y cubren los intervalos $[0.0, 0.6]$ y $[0.0, 0.2]$, respectivamente.

FIGURA 4.
Distribuciones
de los datos de los 4
parámetros. Fuente:
elaboración propia.



Modelo de CNN

El objetivo principal es entrenar una CNN modificada a partir de AlexNet para predecir los cuatro parámetros característicos del modelo SIE ($\theta_E, f, \varepsilon_x, \varepsilon_y$) en

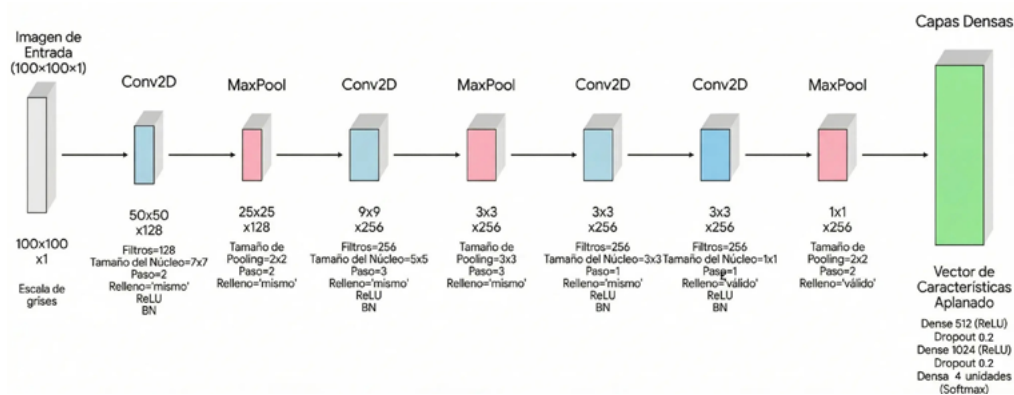


el modelado de lentes gravitacionales. Como se indicó previamente, el conjunto de datos se subdividió en 70 000 imágenes para efectuar un esquema *k-fold* de cuatro iteraciones: 75 % de las imágenes se destinaron al entrenamiento y 25 % a la validación; las 6396 imágenes restantes se reservaron para poner a prueba la capacidad predictiva de la CNN tras el entrenamiento.

Un elemento crucial del conjunto de datos es el formato binario TFRecords optimizado con TensorFlow para almacenar imágenes. Este formato es ideal para grandes conjuntos de datos, ya que permite un almacenamiento compacto y acelera la lectura de datos durante el entrenamiento del modelo, incluso si los registros no caben completamente en la memoria. Asimismo, facilita el etiquetado y la organización sistemática de las imágenes, lo que permite que un solo archivo se divida en muestras de entrenamiento, validación y prueba. Además, el soporte del formato para imágenes en escala de grises agiliza los procedimientos de preprocesamiento.

AlexNet fue introducida por Alex Krizhevskii *et al.* en 2012 y ganó el Desafío de Reconocimiento Visual a Gran Escala de ImageNet de 2012, competencia en la cual se evaluó el rendimiento de algoritmos en detección de objetos y categorización de imágenes. Esta arquitectura logró una tasa de error del 15.3 % en la clasificación de imágenes a gran escala, al ampliar los conceptos fundamentales de LeNet y aplicar los principios básicos de las CNN a una red más extensa y profunda. El sistema integra la función de activación ReLU, el *dropout* y la normalización de respuesta local (LRN) por primera vez dentro de una CNN. El modelo de CNN implementado en esta investigación incorpora modificaciones en AlexNet tanto en el número de capas como de neuronas. En la Figura 5, los bloques de convolución están indicados por el color azul, los de agrupamiento por el color rosa, y las capas densas está representada por el bloque verde.

FIGURA 5.
Esquema de la CNN
utilizada en este
trabajo. Fuente:
elaboración propia.



El modelo comprende seis bloques que preceden a las capas completamente conectadas. Los bloques iniciales consisten en una capa convolucional, seguida de la normalización por lotes y las capas de agrupación máxima. La prime-



ra capa convolucional aplica 128 filtros de tamaño 7×7 con una zancada de 2, reduciendo así la resolución espacial. La segunda capa convolucional utiliza 256 filtros de tamaño 5×5 con una zancada de 3, facilitando la extracción de características a escala media. El bloque convolucional profundo, que consiste en cuatro capas convolucionales con 256 filtros de tamaño 1×1 , captura representaciones jerárquicas y mejora la profundidad de la red sin aumentar considerablemente el número de parámetros. Este bloque concluye con una capa *max-pooling* para reducir aún más la dimensionalidad espacial. Los mapas de características aplanados son procesados por dos capas densas que consisten en 512 y 1024 neuronas, respectivamente. En última instancia, el modelo concluye con una capa densa que posee activación lineal y que produce los cuatro resultados correspondientes a los parámetros de interés.

La mejora primordial que presenta el modelo utilizado en este trabajo son las capas *BatchNormalization*, que estabilizan el entrenamiento, permiten tasas de aprendizaje más altas y reducen la sensibilidad a la inicialización. Otra ventaja es la inclusión de una capa convolucional adicional que mejora la capacidad del modelo para extraer detalles de las imágenes. Además, al implementar la factorización del filtro dentro del bloque de capa convolucional 1×1 se representan mejor las características sin elevar el costo computacional.

Debido a que las componentes de elipticidad producen variaciones más débiles en las imágenes que otros parámetros del modelo de lente, la red tendía a subajustarlas. Para mitigar este efecto, se aplicó un esquema de ponderación en la función de pérdida para asignar pesos más altos a los términos asociados con ε_x y ε_y ; se trató de un proceso iterativo de ajuste empírico según la contribución de cada componente al fenómeno, evaluando su efecto sobre la convergencia del modelo y el error de predicción en el conjunto de validación. En específico, los pesos para el radio de Einstein y la relación del eje se definieron en 1.0, y para las componentes de elipticidad en 3.0, respectivamente; de esa manera, la influencia de estos parámetros sobre el gradiente fue más fuerte durante el entrenamiento, estimulando la sensibilidad del modelo a las particularidades de las firmas morfológicas.

Para el refinamiento del peso se utilizó el NAdam, una adaptación del optimizador Adam, que incorpora a su vez el impulso Nesterov. El gradiente acelerado de Nesterov (NAG) es una técnica de optimización de primer orden que en ciertos contextos eleva y acelera la tasa de convergencia respecto del método tradicional de descenso de gradiente estocástico (SGD), y fomenta la precisión para minimizar la función de pérdida. El NAG se distingue del SGD porque da pri-



mero un paso sustancial en la dirección del gradiente acumulado, y luego evalúa la posición terminal del gradiente para realizar los ajustes necesarios.

En cuanto a la mitigación del sobreajuste, se integró en el modelo un *callback* de *ReduceLROnPlateau* para regular sistemáticamente la tasa de aprendizaje y refinar así las métricas de entrenamiento. En este trabajo, el factor λ para *ReduceLROnPlateau* se fijó en 0.5. La tasa de aprendizaje ϵ mínima elegida fue 10^{-8} . Por último, se evaluó el rendimiento del modelo con base en las 6396 imágenes reservadas para este fin.

Métricas de evaluación

Para evaluar y comparar el rendimiento del modelo de manera cuantitativa, se seleccionó el error cuadrático medio ponderado (WMSE) como función de pérdida, mientras que, como métrica de evaluación, sirvió el error absoluto medio (MAE). Como se mencionó anteriormente, los pesos para el radio de Einstein y la relación axial se establecieron en 1.0, y respecto a las componentes de elipticidad de la lente en 3.0, siguiendo el enfoque de aprendizaje multitarea propuesto por Caruana [46]. Cabe mencionar que se asignó un mayor peso a las componentes de la elipticidad en la función de pérdida debido a la alta complejidad asociada a su inferencia, derivada de las degeneraciones existentes entre estos parámetros y la relación axial (f). El valor de 3.0 fue seleccionado empíricamente mediante un proceso iterativo, porque mostró el mejor desempeño predictivo global durante la validación del modelo. Además, se calculó para cada parámetro el coeficiente de determinación R^2 , que indica en qué medida la varianza en los valores observados puede explicarse por las inferencias del modelo. Estas métricas se exponen en las siguientes expresiones,

$$WMSE = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^k w_j (y_{i,j} - \hat{y}_{i,j})^2 \quad (13)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^k |y_{i,j} - \hat{y}_{i,j}| \quad (14)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (15)$$

Donde:

y_i el valor inferido actual

$\hat{y}_i$ el valor real actual

$\bar{y}$ la media de los valores reales

Σ_j la suma sobre el número de parámetros



El error relativo ϵ_r de cada valor inferido y refiere a la divergencia proporcional de éste respecto del valor real $\hat{y}$. La Ecuación 16 ilustra la expresión correspondiente a dicho cálculo:

$$\epsilon_r = \left| \frac{\hat{y} - y}{\hat{y}} \right| \quad (16)$$

Asimismo, para medir la desviación y precisión de las imágenes se calcula el sesgo entre los valores inferidos respecto de los originales mediante la siguiente métrica:

$$\mu = \text{mediana}(\vec{\hat{y}} - \vec{y}) \quad (17)$$

Donde:

$\vec{\hat{y}}$ vector de los valores inferidos

$\vec{y}$ vector de los valores originales

Adicionalmente, para reforzar los hallazgos es crucial un análisis comparativo entre imágenes simuladas a partir de los resultados del modelo CNN y las imágenes de referencia, usando métricas objetivas como la proporción máxima de señal a ruido (PSNR) y el error cuadrático medio (MSE). La PSNR se expresa en decibelios: una reconstrucción de las imágenes se considera buena en el rango de 30 a 40 dB. Para una imagen de referencia X y una de prueba Y de tamaños $M \times N$, la proporción de ruido se calcula mediante:

$$PSNR(X, Y) = 10 \log_{10} (MAX_{img}^2 / MSE_{img}(X, Y)) \quad (18)$$

Donde:

$$MSE_{img} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (X_{ij} - Y_{ij})^2 \quad (19)$$

MAX_{img} = valor máximo que un píxel puede tener en la imagen

Como complemento, se calculó el índice de similitud estructural (SSIM), que mide la semejanza perceptiva en función de las diferencias estructurales entre una imagen original y una imagen reconstruida; se calcula como una combinación de la luminancia (l), el contraste (c) y la estructura (s) [47]:

$$SSIM(X, Y) = f(l(X, Y), c(X, Y), s(X, Y)) \quad (20)$$

Un SSIM de 1.0 indica un nivel de fidelidad óptimo, mientras que los procesos de reconstrucción cuyo SSIM tiende asintóticamente a 0 se consideran deficientes.



Resultados

Esta sección está dedicada a analizar el rendimiento del modelo CNN en la inferencia de cuatro parámetros del modelo SIE. Se observó que los resultados reportados corresponden a los valores promedio obtenidos de un procedimiento de validación cruzada que emplea un esquema de partición de 4 pliegues. Específicamente, la muestra de prueba de 6396 imágenes se dividió en cuatro subconjuntos independientes. La Figura 6 ilustra la evolución de las funciones de pérdida y MAE con respecto a cada iteración del modelo de entrenamiento, para las cuales se aprecia un comportamiento consistente. La Tabla 1 resume los resultados que se discuten en esta sección.

FIGURA 6.
Métricas empleadas en el entrenamiento y validación del modelo.
Fuente: elaboración propia.

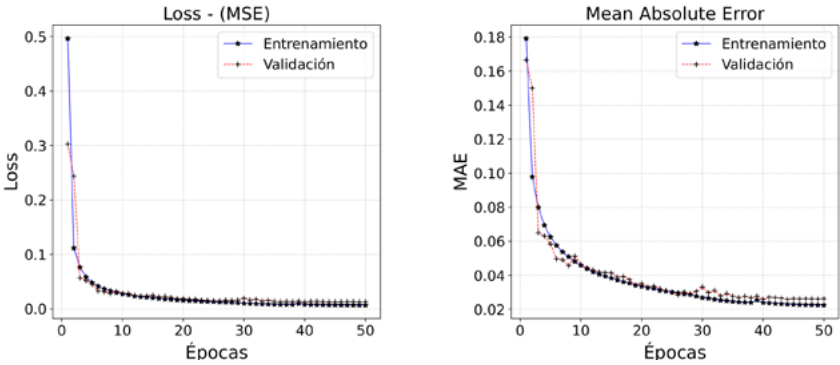


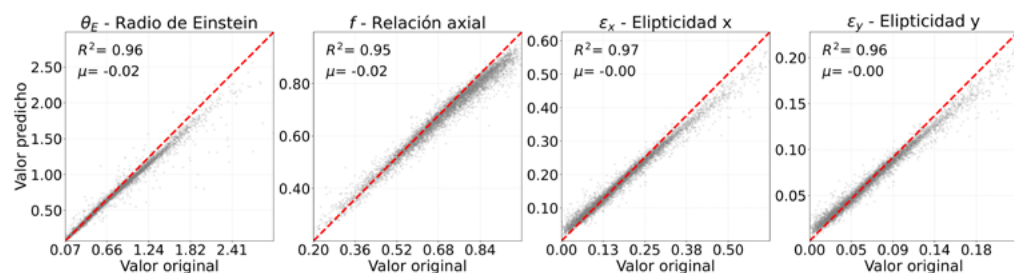
TABLA 1.
Resultados del coeficiente de determinación del error relativo de cada parámetro para cada modelo.

MÉTRICAS		f		f		ε_y		ε_y	
Loss	MAE	R^2	R^2 (%)	R^2	ϵ_r (%)	ϵ_r	ϵ_r (%)	ϵ_r	ϵ_r (%)
0.006	0.022	0.96	4.00	0.95	2.44	0.97	4.86	0.96	5.11

Las distribuciones de dispersión entre los parámetros SIE verdaderos e inferidos (Figura 7) muestran una fuerte tendencia a la relación uno a uno. Las inferencias son más precisas a valores más bajos del radio de Einstein, mientras que se hace evidente un sesgo sistemático para $\theta_E \gtrsim 1.2$, consistente con la distribución de las muestras de entrenamiento. Se observa una tendencia similar para el parámetro de relación de ejes f , donde el modelo subestima sistemáticamente los valores por encima de $f \gtrsim 0.8$. Y las componentes de elipticidad ε_x y ε_y exhiben un sesgo creciente a valores altos, particularmente cuando superan 0.4 y 0.12, respectivamente.


FIGURA 7.

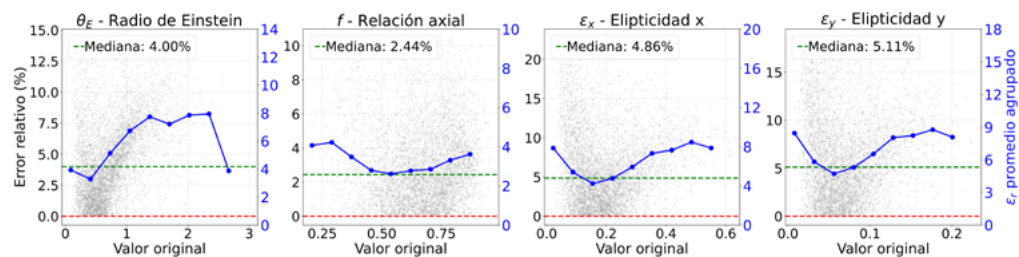
Valores inferidos frente a valores originales para los parámetros de lentes estimados. Fuente: elaboración propia.



Estas diferencias se pueden apreciar en la Figura 8, donde se agrega el MAE agrupado para visualizar con mayor detalle el patrón de los residuos. Los errores relativos de los parámetros SIE se presentan en la Figura 9. El error relativo medio más bajo se obtiene para el parámetro f , con valores de alrededor del 2.6 %, mientras que los parámetros de elipticidad muestran las tasas más altas, alcanzando valores relativos medios de hasta el 5.1 %. En general, el modelo logra una alta capacidad predictiva para los parámetros SIE, con R^2 en el rango de 0.95 a 0.97.

FIGURA 8.

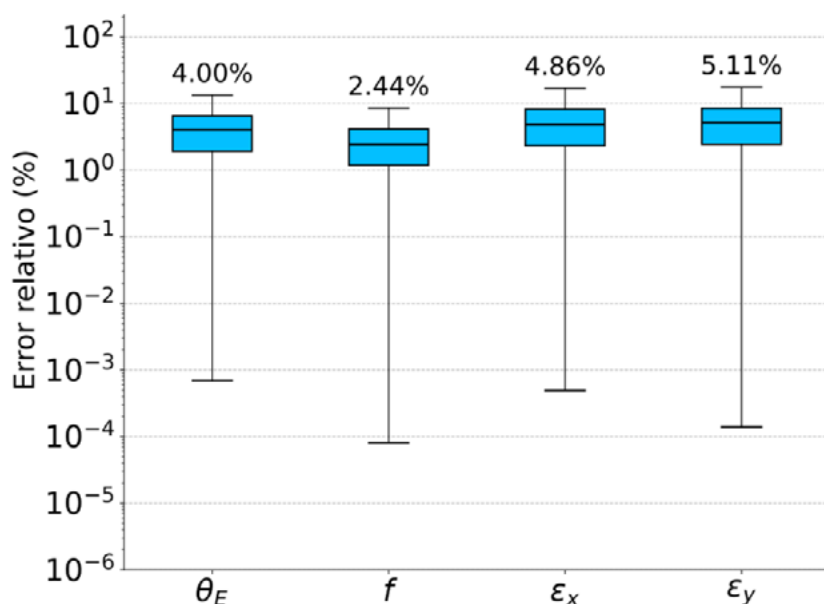
Valores originales frente al error relativo para los parámetros de lentes estimados. Fuente: elaboración propia.



En cuanto a la PSNR, las imágenes de origen reconstruidas exhiben un promedio de aproximadamente 37 dB, es decir, dentro del rango de alta fidelidad de reconstrucción.

FIGURA 9.

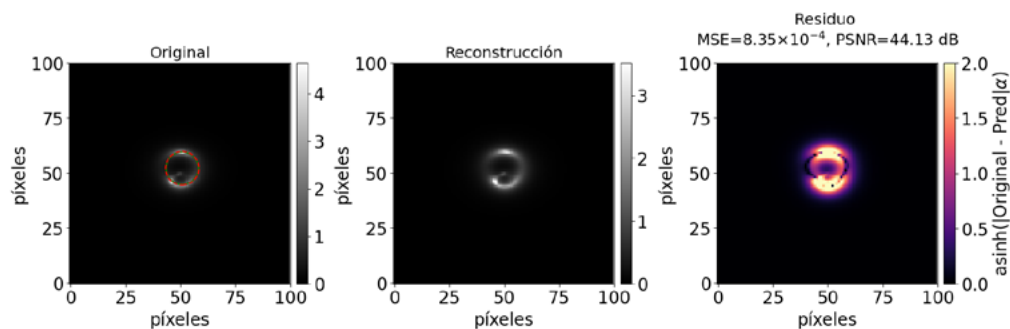
Errores relativos porcentuales por cada parámetro. En la parte superior de las cajas se muestra la mediana de los errores. Fuente: elaboración propia.





Un ejemplo de un sistema de lente galaxia-galaxia reconstruido se muestra en la Figura 10, en la cual la imagen reconstruida coincide estrechamente con la verdadera fuente, con una PSNR de 44.13 dB. Los residuos correspondientes, definidos como la diferencia entre las imágenes verdaderas y reconstruidas, muestran una intensidad muy baja.

FIGURA 10.
Ejemplo aleatorio de un sistema de lente gravitacional fuerte galaxia-galaxia.
Fuente: elaboración propia.



Para mejorar la visibilidad de las características débiles y brillantes, se aplica una transformación asinh de Lupton a los mapas residuales. En la Figura 11 se presentan tres ejemplos representativos de reconstrucciones de SGL, cuyos SSIM próximos a la unidad ratifican la fidelidad de las reconstrucciones respecto de las observaciones originales. Además, al extender la evaluación del SSIM a la totalidad del conjunto de prueba, se obtuvo un valor promedio de 0.99.

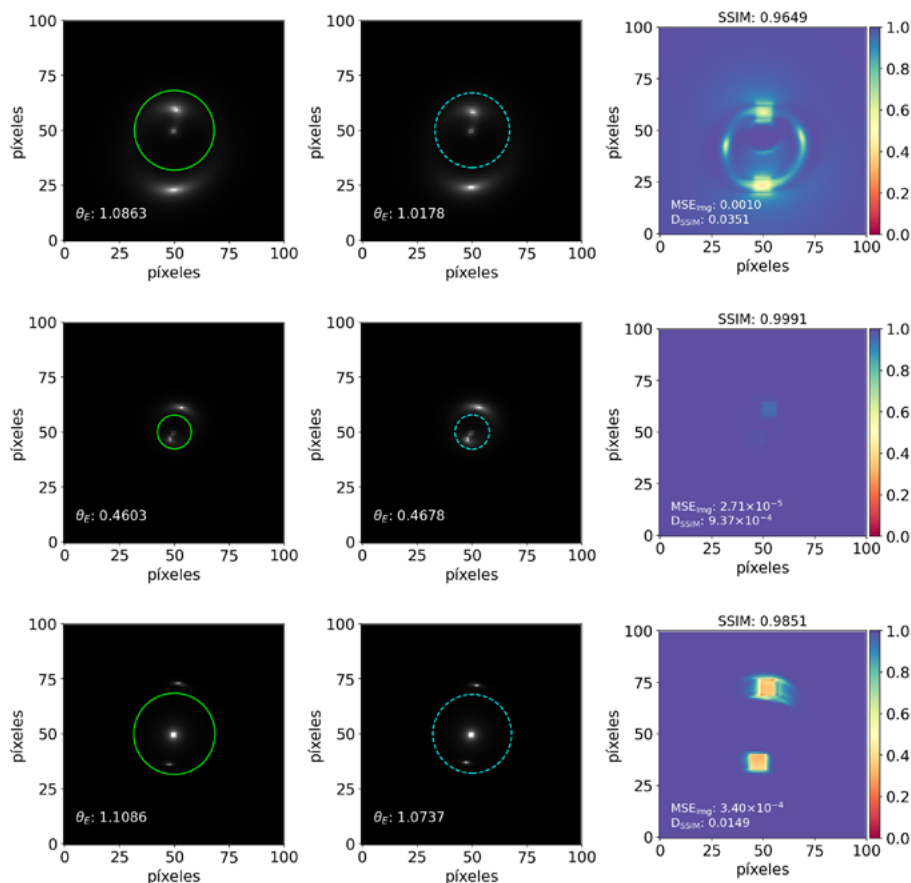


FIGURA 11.
Ejemplos de sistemas reconstruidos de lentes gravitacionales fuertes.
Fuente: elaboración propia.



Discusión y conclusiones

Los resultados muestran que la CNN es capaz de capturar de forma consistente la relación no lineal entre la morfología de los anillos de Einstein y los parámetros morfológicos subyacentes del perfil SIE ($\theta_E, f, \varepsilon_x, \varepsilon_y$). El coeficiente de determinación obtenido para todos los parámetros (R^2 entre 0.95 y 0.97) muestra una elevada capacidad predictiva del modelo, incluso utilizando una arquitectura relativamente ligera en comparación con enfoques más pesados reportados en la literatura.

En particular, el parámetro de relación axial f presentó el menor error relativo promedio ($\sim 2.44\%$), lo cual sugiere que las capas convolucionales capturan con mayor precisión las deformaciones globales asociadas a la elongación de la distribución de masa. Esto puede atribuirse a que dicho parámetro induce cambios geométricos macroscópicos en la forma del anillo, facilitando su extracción a partir de filtros espaciales. Por otro lado, las componentes de elipticidad ($\varepsilon_x, \varepsilon_y$) exhibieron errores relativos mayores ($\sim 5.11\%$), indicando que las variaciones anisotrópicas representan un desafío de mayor envergadura para la red cuando son más sutiles.

El análisis de dispersión entre valores reales e inferidos revela sesgos sistemáticos en regiones específicas del espacio paramétrico, posiblemente a causa de la distribución no uniforme del conjunto de entrenamiento y la incapacidad de generalización a partir de valores menos representados. Puesto de otro modo, la red neuronal presenta un mejor desempeño en los rangos más densamente muestreados de los parámetros, mientras que en regiones menos representadas las predicciones se dispersan, y surge una ligera tendencia a subestimar los valores extremos. En particular, este fenómeno fue más evidente para radios de Einstein superiores a 1.2 y proporciones de elipticidad mayores de 0.8. La incorporación del esquema de ponderación en la función de pérdida WMSE demostró ser un elemento clave para equilibrar el aprendizaje multitarea. Al asignar pesos mayores a las componentes de elipticidad, se incrementó la sensibilidad del gradiente respecto a estos parámetros cuya firma física es tenue.

Respecto a la fidelidad de reconstrucción de imágenes, el PSNR promedio de aproximadamente 37 dB, junto con ejemplos individuales que alcanzan 44.13 dB, indica que las imágenes reconstruidas simulan adecuadamente la intensidad y distribución global del brillo. Por otro lado, la similitud estructural observada sugiere que la red replica fielmente los parámetros numéricos conservando la geometría de las curvas críticas proyectadas en el plano de imagen. Este resultado es importante porque, en sistemas de lentes gravitacionales, pequeñas



variaciones en parámetros como la elipticidad o el radio de Einstein producen deformaciones coherentes en la morfología de los arcos. En este contexto, un SSIM promedio de 0.99 avala la capacidad del modelo para aprender a relacionar consistentemente la distribución de masa con la formación de imágenes, más allá de una simple aproximación estadística basada en regresión.

En conclusión, las CNN constituyen una herramienta robusta y escalable para el modelado de SGL. La arquitectura compacta propuesta permitió inferir con alta precisión los parámetros del modelo SIE, alcanzando errores relativos entre 2.44 y 5.11 % bajo condiciones ideales a bajo costo computacional durante entrenamiento e inferencia. Estos resultados constituyen un punto de partida y referencia para evitar de manera sistemática que, en trabajos futuros, se degrade la precisión al incorporar parámetros adicionales, como el ruido instrumental, el fondo astronómico, el *shear* externo o incluso perfiles de densidad como el Navarro-Frenk-White, correspondiente a la materia oscura [48]. Finalmente, la integración de funciones de pérdida informadas por la física puede mejorar la interpretabilidad y permitir que el modelo, además de ajustar parámetros observacionales, respete directamente las leyes físicas que gobiernan el fenómeno de las lentes gravitacionales.

Referencias

- [1] P. Schneider, C. Kochanek, y J. Wambsganss, *Gravitational Lensing: Strong, Weak and Micro*. Alemania: Springer, 2006, DOI: [10.1007/978-3-540-30310-7](https://doi.org/10.1007/978-3-540-30310-7)
- [2] P. Schneider, *Extragalactic Astronomy and Cosmology: An Introduction*. Alemania: Springer, 2015, DOI: [10.1007/978-3-642-54083-7](https://doi.org/10.1007/978-3-642-54083-7)
- [3] Y. D. Hezaveh *et al.*, "Detection of lensing substructure using ALMA observations of the dusty galaxy SDP.81", arXiv:1601.01388, 2016, DOI: [10.48550/arXiv.1601.01388](https://doi.org/10.48550/arXiv.1601.01388)
- [4] M. Meneghetti, *Introduction to Gravitational Lensing With Python Examples*, Suiza: Springer, 2021, DOI: [10.1007/978-3-030-73582-1](https://doi.org/10.1007/978-3-030-73582-1)
- [5] A. J. Shajib *et al.*, "Strong Lensing by Galaxies", *Space Science Reviews*, vol. 220, núm. 87, 2024, DOI: [10.1007/s11214-024-01105-x](https://doi.org/10.1007/s11214-024-01105-x)
- [6] T. Treu, "Strong Lensing by Galaxies", *Annual Review of Astronomy and Astrophysics*, vol. 48, pp. 87-125, 2010, DOI: [10.1146/annurev-astro-081309-130924](https://doi.org/10.1146/annurev-astro-081309-130924)
- [7] Euclid Collaboration, "Euclid: I. Overview of the Euclid Mission", *Astronomy & Astrophysics*, vol. 697, art. A1, 2025, DOI: [10.1051/0004-6361/202450810](https://doi.org/10.1051/0004-6361/202450810)
- [8] R. Laureijs *et al.*, "Euclid Definition Study Report", arXiv:1110.3193, 2011, DOI: [10.48550/arXiv.1110.3193](https://doi.org/10.48550/arXiv.1110.3193)
- [9] D. Enard, "The European Southern Observatory Very Large Telescope", *Journal of Optics*, vol. 22, núm. 2, pp. 33-50, 1991, DOI: [10.1088/0150-536X/22/2/001](https://doi.org/10.1088/0150-536X/22/2/001)
- [10] M. W. McElwain *et al.*, "The James Webb Space Telescope Mission: Optical Telescope Element Design, Development, and Performance", *Publications of the Astronomical Society of the Pacific*, vol. 135, art. 058001, 2023, DOI: [10.1088/1538-3873/acada0](https://doi.org/10.1088/1538-3873/acada0)
- [11] J. P. Gardner *et al.*, "The James Webb Space Telescope", *Space Science Reviews*, vol. 123, pp. 485-606, 2006, DOI: [10.1007/s11214-006-8315-7](https://doi.org/10.1007/s11214-006-8315-7)
- [12] F. B. Bianco *et al.*, "Optimization of the Observing Cadence for the Rubin Observatory Legacy Survey of Space and Time: A Pioneering Process of Community-focused Experimental Design", *The Astrophysical Journal Supplement Series*, vol. 258, núm. 1, 2021, DOI: <https://doi.org/10.3847/1538-4365/ac3e72>



- [13] I. Hook, "The Science Case for the European ELT", en *Science with the VLT in the ELT Era*, A. Moorwood, Ed., 2009, pp. 225-232, DOI: [10.1007/978-1-4020-9190-2_38](https://doi.org/10.1007/978-1-4020-9190-2_38)
- [14] E. Palle *et al.*, "Ground-breaking exoplanet science with the ANDES spectrograph at the ELT", *Experimental Astronomy*, vol. 59, núm. 29, 2025, DOI: [10.1007/s10686-025-10000-4](https://doi.org/10.1007/s10686-025-10000-4)
- [15] Y. Gong *et al.*, "Introduction to the Chinese Space Station Survey Telescope (CSST)", *Science China Physics, Mechanics & Astronomy*, vol. 69, art. 239501, 2026, DOI: [10.1007/s11433-025-2809-00](https://doi.org/10.1007/s11433-025-2809-00)
- [16] T. E. Collett, "The Population of Galaxy-Galaxy Strong Lenses in Forthcoming Optical Imaging Surveys", *The Astrophysical Journal*, vol. 811, núm. 1, 2015, DOI: [10.1088/0004-637X/811/1/20](https://doi.org/10.1088/0004-637X/811/1/20)
- [17] S. T. Myers *et al.*, "The Cosmic Lens All-Sky Survey - I. Source selection and observations", *Monthly Notices of the Royal Astronomical Society*, vol. 341, núm. 1, pp. 1-12, 2003, DOI: [10.1046/j.1365-8711.2003.06256.x](https://doi.org/10.1046/j.1365-8711.2003.06256.x)
- [18] I. W. A. Browne *et al.*, "The Cosmic Lens All-Sky Survey - II. Gravitational lens candidate selection and follow-up", *Monthly Notices of the Royal Astronomical Society*, vol. 341, núm. 1, pp. 13-32, 2003, DOI: [10.1046/j.1365-8711.2003.06257.x](https://doi.org/10.1046/j.1365-8711.2003.06257.x)
- [19] N. Scoville *et al.*, "The Cosmic Evolution Survey (COSMOS): Overview", *The Astrophysical Journal Supplement Series*, vol. 172, 2007.
- [20] A. S. Bolton *et al.*, "The Sloan Lens ACS Survey. I. A Large Spectroscopically Selected Sample of Massive Early-Type Lens Galaxies", *The Astrophysical Journal*, vol. 638, núm. 2, pp. 703-724, 2006, DOI: [10.1086/498884](https://doi.org/10.1086/498884)
- [21] A. S. Bolton *et al.*, "The Sloan Lens ACS Survey. V. The Full ACS Strong-Lens Sample", *The Astrophysical Journal*, vol. 682, núm. 2, pp. 964-984, 2008, DOI: [10.1086/589327](https://doi.org/10.1086/589327)
- [22] F. Sciortino *et al.*, "Inference of experimental radial impurity transport on alcator C-Mod: Bayesian parameter estimation and model selection", *Nuclear Fusion*, vol. 60, núm. 12, 2020, DOI: [10.1088/1741-4326/abae85](https://doi.org/10.1088/1741-4326/abae85)
- [23] A. Fowlie, W. Handley, y L. Su, "Nested sampling cross-checks using order statistics", *Monthly Notices of the Royal Astronomical Society*, vol. 497, núm. 4, pp. 5256-5263, 2020, DOI: [10.1093/mnras/staa2345](https://doi.org/10.1093/mnras/staa2345)
- [24] W. Zhang *et al.*, "Shift-invariant neural network for image processing: learning and generalization", *Applications of Artificial Neural Networks III*, vol. 1709, pp. 257-268, 1992, DOI: [10.1117/12.140004](https://doi.org/10.1117/12.140004)
- [25] P. Hála, "Spectral classification using convolutional neural networks", arXiv:1412.8341, 2014, DOI: [10.48550/arXiv.1412.8341](https://doi.org/10.48550/arXiv.1412.8341)
- [26] C. E. Petrillo *et al.*, "Finding strong gravitational lenses in the Kilo Degree Survey with Convolutional Neural Networks", *Monthly Notices of the Royal Astronomical Society*, vol. 472, núm. 1, pp. 1129-1150, 2017, DOI: [10.1093/mnras/stx2052](https://doi.org/10.1093/mnras/stx2052)
- [27] Y. D. Hezaveh, L. P. Levasseur, y P. J. Marshall, "Fast automated analysis of strong gravitational lenses with convolutional neural networks", *Nature*, vol. 548, pp. 555-557, 2017, DOI: [10.1038/nature23463](https://doi.org/10.1038/nature23463)
- [28] S. Schuldt *et al.*, "HOLISMOKES IV. Efficient mass modeling of strong lenses through deep learning", *Astronomy & Astrophysics*, vol. 646, art. A126, 2021, DOI: [10.1051/0004-6361/202039574](https://doi.org/10.1051/0004-6361/202039574)
- [29] J. W. Park *et al.*, "Large-scale Gravitational Lens Modeling with Bayesian Neural Networks for Accurate and Precise Inference of the Hubble Constant", *The Astrophysical Journal*, vol. 910, núm. 1, 2021, DOI: [10.3847/1538-4357/abdfc4](https://doi.org/10.3847/1538-4357/abdfc4)
- [30] J. Pearson, J. Maresca, N. Li, y S. Dye, "Strong lens modelling: comparing and combining Bayesian neural networks and parametric profile fitting", *Monthly Notices of the Royal Astronomical Society*, vol. 505, núm. 3, pp. 4362-4382, 2021, DOI: [10.1093/mnras/stab1547](https://doi.org/10.1093/mnras/stab1547)
- [31] V. Busillo *et al.*, "Euclid Quick Data Release (Q1). LEMON - Lens Modelling with Neural networks. Automated and fast modelling of Euclid gravitational lenses with a singular isothermal ellipsoid mass profile", arXiv:2503.15329, 2026, DOI: [10.48550/arXiv.2503.15329](https://doi.org/10.48550/arXiv.2503.15329)
- [32] R. Parlange *et al.*, "GraviT: transfer learning with vision transformers and MLP-Mixer for strong gravitational lens discovery", *Monthly Notices of the Royal Astronomical Society*, vol. 545, núm. 2, 2026, DOI: <https://doi.org/10.1093/mnras/staf1747>
- [33] R. Cañameras *et al.*, "HOLISMOKES VI. New galaxy-scale strong lens candidates from the HSC-SSP imaging survey", *Astronomy & Astrophysics*, vol. 653, art. L6, 2021, DOI: [10.1051/0004-6361/202141758](https://doi.org/10.1051/0004-6361/202141758)
- [34] A. T. Jaelani *et al.*, "Survey of gravitationally lensed objects in HSC imaging (SUGOH) - X. Strong lens finding in the HSC-SSP using convolutional neural networks", *Monthly Notices of the Royal Astronomical Society*, vol. 535, núm. 2, pp. 1625-1639, 2024, DOI: [10.1093/mnras/stae2442](https://doi.org/10.1093/mnras/stae2442)
- [35] L. P. Garate-Nuñez, A. S. G. Robotham, S. Bellstedt, L. J. M. Davies, y C. Martínez-Lombilla, "The Hyper Suprime-Cam extended point spread functions and applications", *Monthly Notices of the Royal Astronomical Society*, vol. 531, núm. 2, pp. 2517-2530, 2024, DOI: [10.1093/mnras/stae1292](https://doi.org/10.1093/mnras/stae1292)
- [36] A. Krizhevsky, I. Sutskever, y G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Communications of the ACM*, vol. 60, núm. 6, pp. 84-90, 2017, DOI: [10.1145/3065386](https://doi.org/10.1145/3065386)
- [37] J. J. Ancona-Flores, A. Hernández-Almada, y V. Motta, "Enhancing Gravitational Lens Study with Deep Learning: A Study on Effects Of Dropout Regularization", *Galaxies*, vol. 14, núm. 2, 2026, DOI: [10.3390/galaxies14020018](https://doi.org/10.3390/galaxies14020018)
- [38] R. Kormann, P. Schneider, y M. Bartelmann, "Isothermal elliptical gravitational lens models", *Astronomy and Astrophysics*, vol.



284, pp. 285-299, 1994. [En línea]. Disponible en: <https://www.semanticscholar.org/paper/Iso-thermal-elliptical-gravitational-lens-models.-Kor-mann-Schneider/ed1caf1f6fdf1866fc6eac34ee-f6a15e780037b2>

- [39] A. Etherington *et al.*, "Automated galaxy-galaxy strong lens modelling: No lens left behind", *Monthly Notices of the Royal Astronomical Society*, vol. 517, núm. 3, pp. 3275-3302, 2022, DOI: [10.1093/mnras/stac2639](https://doi.org/10.1093/mnras/stac2639)
- [40] Y. LeCun *et al.*, "Backpropagation Applied To Handwritten Zip Code Recognition", *Neural Computation*, vol. 1, núm. 4, pp. 541-551, 1989, DOI: [10.1162/neco.1989.1.4.541](https://doi.org/10.1162/neco.1989.1.4.541)
- [41] P. Baldi y P. J. Sadowski, "Understanding Dropout", *Advances in Neural Information Processing Systems*, vol. 26, 2013. [En línea]. Disponible en: https://papers.nips.cc/paper_files/paper/2013/hash/71f6278d140af599e06ad9b-f1ba03cb0-Abstract.html
- [42] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, y R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors", arXiv:1207.0580, 2012, DOI: [10.48550/arXiv.1207.0580](https://doi.org/10.48550/arXiv.1207.0580).
- [43] I. Salehin y D.-K. Kang, "A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain", *Electronics*, vol. 12, núm. 14, 2023, DOI: [10.3390/electronics12143106](https://doi.org/10.3390/electronics12143106)
- [44] J. L. Sérsic y Universidad Nacional de Córdoba Observatorio Astronómico, *Atlas de galaxias australes*. Argentina: Observatorio Astronómico, Universidad Nacional de Córdoba, 1968.
- [45] S. Mukherjee *et al.*, "SEAGLE — I. A pipeline for simulating and modelling strong lenses from cosmological hydrodynamic simulations", *Monthly Notices of the Royal Astronomical Society*, vol. 479, núm. 3, pp. 4108-4125, 2018, DOI: [10.1093/mnras/sty1741](https://doi.org/10.1093/mnras/sty1741)
- [46] R. Caruana, "Multitask Learning", *Machine Learning*, vol. 28, pp. 41-75, 1997, DOI: [10.1023/A:1007379606734](https://doi.org/10.1023/A:1007379606734)
- [47] Z. Wang, A. C. Bovik, H. R. Sheikh, y E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", *IEEE Transactions on Image Processing*, vol. 13, núm. 4, pp. 600-612, 2004, DOI: [10.1109/tip.2003.819861](https://doi.org/10.1109/tip.2003.819861)
- [48] T. Treu y L. V. E. Koopmans, "Massive Dark Matter Halos and Evolution of Early-Type Galaxies to $z \approx 1$ ", *The Astrophysical Journal*, vol. 611, núm. 2, pp. 739-760, 2004, DOI: [10.1086/422245](https://doi.org/10.1086/422245)



PERSPECTIVAS DE LA
CIENCIA Y LA TECNOLOGÍA

¿Quieres publicar en esta revista?

¿Dudas o sugerencias? Escríbenos a:

? perspectivasci@uaq.mx

.....
REVISTA INCLUIDA EN:
.....

VISITA NUESTRO

FISIÓN
PODCAST

Escucha de la voz de
los autores, entrevistas
y comentarios
relacionados a sus
artículos.

Disponible en:

MÁS REVISTAS UAQ EN:



revistas.uaq.mx



ingenieria.uaq.mx

EDICIÓN CUIDADA, DISEÑADA
Y MAQUETADA POR

D | DESPACHO DE
PUBLICACIONES

Visítanos y conoce
las publicaciones que la
FACULTAD DE INGENIERÍA
DE LA UNIVERSIDAD AUTÓNOMA
DE QUERÉTARO
tiene para ti:



UNIVERSIDAD AUTÓNOMA DE QUERÉTARO
FACULTAD DE INGENIERÍA