Análisis estadístico comparativo de métodos clásicos y redes neuronales con arquitecturas Dual y U-Net para recuperación de fase en la técnica de proyección de franjas

Yeison Fabian Mayorga Gamboa y Jose David Silva Solano

Trabajo de Grado para Optar al Título de Ingeniero de Sistemas

Director

Andrés Leonardo González Gómez

Candidato a Doctor en Ciencias de la Computación

Codirector

Jaime Enrique Meneses Fonseca

Doctor en Ciencias para la Ingeniería

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Ingeniería de Sistemas e Informática

Programa de Ingeniería de Sistemas

Bucaramanga

2026

Dedicatoria

A mi padre, Samuel Mayorga Cáceres, por ser mi ejemplo de vida, por enseñarme que los sueños se construyen con esfuerzo, perseverancia y sacrificio. Porque hoy entiendo que muchos de sus sacrificios se convirtieron en mis oportunidades, y que detrás de cada uno de mis logros también están sus esfuerzos y su amor. Gracias por educarme, por sus consejos, por compartir conmigo sus enseñanzas y por mostrarme, a través de su ejemplo, el valor del trabajo y la responsabilidad. Lo que un día soñamos juntos hoy se convierte en realidad, y no hay mayor satisfacción que poder compartir este logro con usted. Esta meta también es suya, porque sin todo lo que ha hecho por mí, llegar hasta aquí no habría sido posible.

A mi madre, Marlene Gamboa Pulido, por enseñarme desde pequeño el verdadero significado de la responsabilidad, el compromiso y el amor incondicional. Gracias por cada sacrificio que hizo para que nunca me faltara lo esencial, por procurar que siempre tuviera un plato de comida en la mesa, un hogar lleno de amor y un lugar donde sentirme seguro. Hoy valoro profundamente cada esfuerzo, incluso aquellos que en su momento no alcanzaba a comprender. Gracias por cuidar de mí, por apoyarme en cada etapa de mi vida y, sobre todo, por creer en mí, incluso en los momentos en los que yo mismo llegué a dudar. Gran parte de lo que soy se lo debo a usted.

A mi hermano, Johan Samuel Mayorga Gamboa, por estar presente a lo largo de toda mi vida, por ser una de esas personas con las que siempre he podido contar y por demostrarme su cariño a través de su apoyo y preocupación constante. Gracias por confiar en mis capacidades, por creer en mí y por acompañarme en los diferentes momentos de este camino. Saber que siempre he tenido a mi hermano a mi lado es algo que valoro más de lo que las palabras pueden expresar.

A mi pareja de vida, Tania Sofia Suarez Carvajal, por su amor, su paciencia y su compañía incondicional. Gracias por nunca dejar de creer en mí, por apoyarme cuando las cosas parecían

difíciles y por recordarme, con sus palabras y sus acciones, que era capaz de alcanzar mis sueños. Gracias por hacerme sentir amado, por ser refugio en los días complicados y alegría en los momentos felices. Tener a alguien con quien compartir mis miedos, mis ilusiones y mis logros ha hecho que este camino sea mucho más bonito. Este logro también lleva un poquito de todo el amor que me ha brindado.

A toda mi familia, porque cada uno, de una u otra manera, ha dejado una huella en mi vida y ha contribuido a la persona que soy hoy. Gracias por sus enseñanzas, sus palabras de aliento, sus consejos, los momentos compartidos y por hacerme sentir que nunca he estado solo. En especial, a mis tíos Rafael Mayorga, Verenice Gamboa, Carlos Gamboa y Susana Mayorga; y a mis primos Daniel Gamboa, Santiago Manrique, Dana Manrique, David Hurtado, Rafael Mayorga, Vanessa Gamboa, Gissela Gamboa, Carolina Gamboa y Camila Gamboa. Gracias por formar parte de mi historia, por los recuerdos que hemos construido juntos y por todo el cariño que me han brindado a lo largo de los años. Cada uno ocupa un lugar especial en mi corazón.

A mis amigos, por su compañía, por las experiencias compartidas, por las risas que hicieron más llevaderos los momentos difíciles y por las palabras de ánimo que muchas veces llegaron cuando más las necesitaba. Gracias a quienes estuvieron presentes durante este proceso y a quienes, desde diferentes etapas de mi vida, aportaron algo valioso a mi crecimiento personal.

Y a todas aquellas personas que, aunque no mencione por su nombre, han formado parte de mi camino, gracias. Cada gesto de cariño, cada consejo, cada enseñanza y cada muestra de apoyo han significado mucho para mí. Llevo conmigo algo de cada persona que me ha ayudado a crecer.

Hoy, al cumplir uno de los sueños más importantes de mi vida, entiendo que este título, aunque lleve mi nombre, también representa el esfuerzo, el amor y los sacrificios de muchas personas que han caminado a mi lado. Nada de esto habría sido igual sin ustedes. Gracias por ser parte de mi vida, por creer en mí y por ayudarme a convertir aquel sueño en una realidad.

A todos ustedes, les dedico este logro con todo mi amor y gratitud. Los amo profundamente.

Yeison Fabian Mayorga Gamboa

A mi madre, Carmen Janeth Solano Rincón, mi todo. Fue mi primera maestra y ha sido, desde siempre, mi guía más fiel. Creyó en mí antes de que yo mismo lo hiciera, y con paciencia, ternura y exigencia me enseñó que el esfuerzo siempre da frutos. Su amor ha estado presente en cada madrugada, en cada cuidado silencioso, en cada preocupación por mi bienestar y en cada palabra que me impulsó a dar lo mejor de mí. No hubo etapa de mi vida en la que no estuviera a mi lado, celebrando mis avances y sosteniéndome en mis caídas. Todo lo que he aprendido lleva algo de ella, y detrás de cada uno de mis logros ha estado siempre su mano. Este título es, ante todo, el fruto de su dedicación y de su amor infinito.

A mi padre, Juan Jose Silva Dueñez, quien convirtió su esfuerzo en mis oportunidades. Con su trabajo incansable y su generosidad sin medida, hizo posible que nunca me faltara nada para seguir adelante. Siempre me dijo que la mejor herencia que podía dejarme era el estudio, y día a día, sin descanso, me la fue entregando con sacrificio y orgullo. Gracias por su apoyo incondicional, por cada consejo que guardo como una lección de vida, por estar pendiente de mí y por enseñarme con su ejemplo el valor del trabajo y la responsabilidad. Gracias por soñar con verme convertido en profesional; hoy ese sueño se hace realidad, y es tan suyo como mío.

A los dos, porque son todo para mí. Todo lo que soy y lo que llegue a ser se lo debo a su amor.

A mis hermanas, Adriana Marcela Silva Solano y Sandra Viviana Silva Solano, por ser mi apoyo incondicional en cada etapa de este camino.

A mi cuñado, Jhon Alberth Montero Pinzón, por su respaldo y sus palabras de ánimo.

A mi sobrina, Alana Gabriela Montero Silva, que con su alegría ilumina mis días. Ojalá algún día, al leer estas líneas, sepa que los sueños se alcanzan con esfuerzo.

A mi compañera de vida, quien sabe que estas líneas son suyas. Fue mi apoyo incondicional a lo largo de este camino y creyó en mí incluso cuando yo dudaba de mis capacidades. Con sus palabras, su paciencia y su amor me sostuvo en los momentos en que más lo necesitaba. Hizo suyos mis sueños como si fueran propios, celebró conmigo cada avance y nunca dejó de impulsarme a seguir adelante. Este título llevará mi nombre, pero una parte de él la construimos juntos, porque a su lado aprendí que, por difícil que sea el camino, nunca se camina solo.

A todos ustedes, gracias por ser parte de mi historia y por hacer posible que hoy uno de mis mayores sueños se convierta en realidad.

Jose David Silva Solano

Agradecimientos

A la Universidad Industrial de Santander, por la formación recibida durante estos años y por disponer de los espacios y recursos que hicieron posible este trabajo.

A nuestro director, Andrés Leonardo González Gómez, por proponer el problema que dio origen a esta investigación, por facilitar el simulador empleado en el segundo experimento y por el acompañamiento sostenido en cada etapa del desarrollo.

A los docentes de la Escuela de Ingeniería de Sistemas e Informática, por la formación que sustenta las decisiones técnicas de este documento, y a los profesores de las demás áreas que acompañaron nuestra formación a lo largo de la carrera y aportaron, cada uno desde su disciplina, a la manera en que hoy abordamos un problema.

Tabla de Contenido

Pág.

Introducción 16
1 Contexto y motivación 16
2 Planteamiento del problema 18
3 Objetivos 19
3.1 Objetivo general 19
3.2 Objetivos específicos 19
4 Alcance y delimitación 20
5 Contribuciones 21
6 Organización del documento 22
7 Fundamentos y antecedentes 23
7.1 Técnica de proyección de franjas 23
7.2 Modelo de la imagen de franjas 25
7.3 Fase continua y fase discontinua 26
7.4 Algoritmos de corrimiento de fase 29
7.5 Desenvolvimiento espacial de la fase 31
7.6 Desenvolvimiento temporal de la fase 33
7.7 Degradaciones radiométricas 35
7.8 Redes neuronales para la estimación de la fase 36
7.8.1 Conceptos de las redes convolucionales 37
7.8.2 Estimación directa: U-Net 38
7.8.3 Estimación indirecta: Dual CNN 38
7.9 Evaluación de la recuperación de fase 39
8 Metodología 42
8.1 Primer experimento: recuperación de fase en condiciones simuladas controladas 43
8.2 Segundo experimento: recuperación de fase sobre geometrías tridimensionales complejas 58
8.3 Evaluación cuantitativa y análisis comparativo 63
8.4 Reproducibilidad y alcance de la evaluación 70
9 Resultados y discusión 72
9.1 Criterios para leer los resultados 72
9.2 Experimento 1: superficies analíticas controladas 74
9.2.1 Fase discontinua 75
9.2.2 Fase continua 77
9.2.3 Costo computacional 79
9.3 Experimento 2: geometrías tridimensionales complejas 81
9.3.1 Fase discontinua 82
9.3.2 Fase continua 84
9.3.3 Costo computacional 86
9.4 Comparación integrada y relación con los objetivos 90

9.5 Limitaciones documentadas 94
10 Conclusiones y trabajo futuro 96
10.1 Conclusiones 96
10.2 Trabajo futuro 98
Referencias Bibliográficas 100
Apéndices 104

Lista de Tablas

Pág.

Tabla 1 Parámetros de las condiciones radiométricas 47
Tabla 2 Configuración principal de entrenamiento 55

Tabla 3 Exp1: error y similaridad de la fase discontinua en Moderate 75
Tabla 4 Exp1: RMSE y MAE continuos en Moderate y VeryStrong 78
Tabla 5 Exp1: costo computacional de cada método 80
Tabla 6 Exp2: error y similaridad de la fase discontinua en Moderate 82
Tabla 7 Exp2: error continuo y cobertura de la fase continua 84
Tabla 8 Exp2: costo computacional de cada método 87
Tabla 9 Efecto de la precisión bf16 en las redes 89
Tabla 10 Control de reproducibilidad de la energía de la Dual CNN 90
Tabla 11 Síntesis por criterio de los dos experimentos 91

Tabla A1 Contrastes globales de Friedman del Exp1 104
Tabla A2 Contrastes globales de Friedman del Exp2 107
Tabla A3 Comparaciones pareadas de Wilcoxon del Exp1 para la fase discontinua 111
Tabla A4 Comparaciones pareadas de Wilcoxon del Exp1 para la fase continua 116
Tabla A5 Comparaciones pareadas de Wilcoxon del Exp2 para la fase discontinua 122
Tabla A6 Comparaciones pareadas de Wilcoxon del Exp2 para la fase continua 125

Lista de Figuras

Pág.

Figura 1 Principio de la técnica de proyección de franjas 24
Figura 2 Modelo de la imagen de franjas 26
Figura 3 Relación entre la fase continua y la fase discontinua 28
Figura 4 Corrimiento de fase en un píxel 30
Figura 5 Desenvolvimiento espacial a lo largo de una fila 32
Figura 6 Desenvolvimiento temporal a lo largo de una fila 34
Figura 7 Degradaciones radiométricas sobre un perfil de franjas 35
Figura 8 Representaciones de salida de las dos redes 39
Figura 9 Error circular de la fase discontinua 40

Figura 10 Organización de los dos experimentos 43
Figura 11 Formación de los datos del primer experimento 45
Figura 12 Transformaciones radiométricas hasta la condición Moderate 48
Figura 13 Entradas y operaciones de los métodos clásicos 51
Figura 14 Representaciones de entrada y salida de las redes 54
Figura 15 Pérdidas de entrenamiento y validación del Exp1 56
Figura 16 Verificación de las salidas de fase del Exp1 57
Figura 17 Formación de una escena del segundo experimento 59
Figura 18 Etapas del desenvolvimiento espacial 61
Figura 19 Máscara válida, región recuperada y píxeles con error de orden 66

Figura 20 Exp1: RMSE circular y SSIM circular de la fase discontinua 76
Figura 21 Exp1: RMSE continuo medio de la fase continua 79
Figura 22 Exp2: RMSE circular y SSIM global de la fase discontinua 83
Figura 23 Exp2: RMSE continuo, error de orden y cobertura correcta 86
Figura 24 Tiempo y energía del procedimiento completo 88

Lista de Apéndices

Pág.

Apéndice A Tablas complementarias de resultados 104

Lista de abreviaturas y siglas

TPF	Técnica de proyección de franjas.
PSA	Algoritmo de corrimiento de fase (*phase-shifting algorithm*).
LTPU	Desenvolvimiento temporal lineal de fase.
CNN	Red neuronal convolucional.
MAE	Error absoluto medio.
RMSE	Raíz del error cuadrático medio.
SSIM	Índice de similitud estructural.

Resumen

Título: Análisis estadístico comparativo de métodos clásicos y redes neuronales con arquitecturas Dual y U-Net para recuperación de fase en la técnica de proyección de franjas[1]

Autores: Yeison Fabian Mayorga Gamboa; Jose David Silva Solano.[2]

Palabras clave: proyección de franjas, recuperación de fase, corrimiento de fase, desenvolvimiento de fase, aprendizaje profundo, U-Net.

Descripción:

La técnica de proyección de franjas codifica la forma de una superficie en la fase de un patrón luminoso. Los algoritmos de corrimiento de fase recuperan esa fase con varias imágenes; algunas redes neuronales profundas la estiman con una sola. Este trabajo comparó, bajo un mismo protocolo, los algoritmos de tres, cuatro y doce imágenes (PSA3, PSA4 y PSA12), el desenvolvimiento temporal lineal (LTPU, solo en el primer experimento) y las redes U-Net y Dual CNN, en la fase discontinua $\varphi(x,y)$ y la fase continua $\Phi(x,y)$. Se realizaron dos experimentos con datos simulados: 800 escenas de superficies analíticas y 1920 escenas generadas a partir de diez modelos tridimensionales. Las redes se entrenaron solo en la condición Moderate y se evaluaron en una Ideal y tres degradadas con ruido gaussiano, no linealidad gamma y saturación especular. La precisión se midió con el RMSE, el MAE y el SSIM y, en la fase continua, con la cobertura y el error de orden; el costo, con el tiempo, la memoria y la energía neta de CPU y GPU. Las diferencias se analizaron con pruebas de Friedman y Wilcoxon con ajuste de Holm. PSA12 obtuvo el menor RMSE circular en la condición de entrenamiento de ambos experimentos (0,0150 rad en el primero) y fue el más robusto en las demás. En esa condición, las redes superaron a PSA3 y PSA4 en las tres métricas del primer experimento; en el segundo, superaron a PSA3 en RMSE y MAE, pero no en SSIM. Las redes presentaron errores altos en Ideal y VeryStrong, alejadas de su dominio de entrenamiento. Aunque requieren una sola imagen, consumieron más energía que los métodos clásicos en el procedimiento completo. La ventaja de cada método depende, entonces, de la condición radiométrica, la métrica y la complejidad de la escena.

[1]Trabajo de grado, modalidad Trabajo de Investigación.

[2]Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Andrés Leonardo González Gómez, PhD(c). Codirector: Jaime Enrique Meneses Fonseca, PhD.

Abstract

Title: Comparative statistical analysis of classical methods and neural networks with Dual and U-Net architectures for phase recovery in the fringe projection technique[1]

Authors: Yeison Fabian Mayorga Gamboa; Jose David Silva Solano.[2]

Keywords: fringe projection, phase retrieval, phase shifting, phase unwrapping, deep learning, U-Net.

Description:

Fringe projection encodes the three-dimensional shape of a surface in the phase of a projected light pattern. Phase-shifting algorithms recover that phase from several images, whereas some deep neural networks estimate it from a single one. This work quantitatively compared, under a common protocol, the three-, four- and twelve-step algorithms (PSA3, PSA4 and PSA12), linear temporal phase unwrapping (LTPU, only in the first experiment), and the U-Net and Dual CNN networks for recovering the discontinuous phase $\varphi(x,y)$ and the continuous phase $\Phi(x,y)$. Two experiments were carried out on simulated data: the first with 800 scenes of analytic surfaces, and the second with 1920 scenes generated from ten three-dimensional models. The networks were trained only under the Moderate condition and evaluated under one Ideal condition and three degraded ones combining Gaussian noise, gamma nonlinearity and specular saturation. Accuracy was measured with RMSE, MAE and SSIM and, for the continuous phase, also with coverage and fringe-order error; cost was measured with time, memory and net CPU and GPU energy. Differences were analyzed with Friedman and Wilcoxon tests with Holm adjustment. PSA12 achieved the lowest circular RMSE under the training condition in both experiments, with a circular RMSE of 0.0150 rad in the first, and was the most robust method under the remaining conditions. Under that condition, the networks outperformed PSA3 and PSA4 in all three discontinuous-phase metrics in the first experiment; in the second, they outperformed PSA3 in RMSE and MAE, but not in SSIM. The networks showed high errors under the Ideal and VeryStrong conditions, far from their training domain. Although they require a single image, they consumed more energy than the classical methods over the complete procedure. It is concluded that the advantage of each method depends on the radiometric condition, the metric and the complexity of the scene.

[1] Undergraduate degree project, research modality.

[2] Faculty of Physical and Mechanical Engineering. School of Systems Engineering and Informatics. Advisor: Andrés Leonardo González Gómez, PhD(c). Co-advisor: Jaime Enrique Meneses Fonseca, PhD.

Introducción

La medición sin contacto de la forma tridimensional de un objeto es una necesidad común en la inspección industrial, la odontología, la metrología facial y las ciencias forenses. La técnica de proyección de franjas responde a esa necesidad codificando la geometría de la superficie en la fase de un patrón luminoso, de modo que la calidad de la reconstrucción depende de la exactitud con la que esa fase se recupera de las imágenes. Este trabajo compara, bajo un mismo protocolo y sobre datos simulados, métodos clásicos de recuperación de fase basados en corrimiento de fase y desenvolvimiento con dos redes neuronales profundas que estiman la fase a partir de una sola imagen. Esta introducción presenta el contexto y la motivación del estudio, el problema que aborda, sus objetivos, su alcance, sus aportes y la organización del documento.

1. Contexto y motivación

En la técnica de proyección de franjas, un proyector ilumina el objeto con un patrón de franjas, generalmente sinusoidal, y una cámara situada en otra dirección registra cómo ese patrón se desplaza y se curva sobre la superficie (Zhang, 2013; González, 2020). Al tratarse de un método óptico sin contacto que mide toda la región iluminada a partir de imágenes, se ha empleado en la reconstrucción de piezas dentales (Múnera et al., 2011), la metrología óptica facial (González et al., 2012), las ciencias forenses (Rincón et al., 2009; González, 2020) y la medición tridimensional en tiempo real (Zhang, 2010).

La altura de cada punto queda codificada en la fase del patrón observado. La fase que se

calcula directamente a partir de las imágenes es una fase discontinua $\varphi(x,y)$, restringida al intervalo $(-\pi,\pi]$, y debe desenvolverse para obtener la fase continua $\Phi(x,y)$, que conserva la correspondencia entre cada píxel de la cámara y una posición del proyector. Las dos etapas condicionan el resultado final: un error en $\varphi(x,y)$ se transmite a $\Phi(x,y)$, y un error de desenvolvimiento introduce saltos de múltiplos de 2π que pueden alterar regiones completas de la superficie (Ghiglia and Pritt, 1998; Zuo et al., 2018).

Los algoritmos de corrimiento de fase constituyen la aproximación clásica para obtener $\varphi(x,y)$: combinan N imágenes del mismo patrón desplazado y calculan la fase de forma independiente en cada píxel (Zuo et al., 2018). Aumentar N reduce la sensibilidad al ruido aditivo (Surrel, 1997) y a los armónicos que introduce la no linealidad entre el proyector y la cámara (Ye et al., 2015; Zuo et al., 2018), pero también aumenta el número de capturas, lo que limita la medición de escenas que cambian durante la adquisición (Zhang, 2010). El desenvolvimiento temporal presenta un compromiso similar: resuelve el orden de franja de cada píxel mediante secuencias de varias frecuencias (Huntley and Saldner, 1993; Saldner and Huntley, 1997), a costa de imágenes adicionales. Los métodos de una sola imagen, como el análisis por transformada de Fourier (Takeda et al., 1982), reducen la adquisición a un único patrón, pero requieren que la información de fase pueda separarse del fondo y de la modulación presentes en esa imagen.

En ese contexto, las redes neuronales profundas se han incorporado al análisis de franjas en tareas como la supresión de ruido (Yan et al., 2019), la extracción de la fase a partir de una sola imagen (Feng et al., 2019; Hu et al., 2021; Kando et al., 2019), el desenvolvimiento (Wang et al., 2019; Li et al., 2020) y la reconstrucción de extremo a extremo (Machineni et al., 2020). Dos formulaciones resultan de especial interés para este trabajo. Hu et al. (2021) emplearon una U-Net para estimar directamente la fase discontinua como una clasificación por píxel, mientras que Feng et al. (2019) propusieron una arquitectura de dos redes, aquí denominada Dual CNN, que estima las componentes del corrimiento de fase y calcula después la fase con la función arcotangente. Ambas prometen la precisión de un método de varias imágenes a partir de un único patrón, lo que motiva

compararlas con los algoritmos clásicos que pretenden sustituir.

2. Planteamiento del problema

Las publicaciones que proponen estas redes las evalúan con sus propios conjuntos de datos, configuraciones de medición, condiciones de entrenamiento y métricas. Esa diversidad dificulta establecer, a partir de la literatura, cómo se comparan entre sí y con los algoritmos de corrimiento de fase cuando todos procesan las mismas escenas. Además, una red aprende la relación entre la imagen y la fase a partir de ejemplos, por lo que su comportamiento ante condiciones radiométricas distintas de las de entrenamiento —ruido, no linealidad gamma, brillos especulares y saturación del sensor— no puede suponerse a partir de su desempeño en la condición de entrenamiento.

La comparación también depende de qué se mida. El error de la fase discontinua no refleja por sí solo los errores de orden que aparecen al desenvolver, y un error bajo sobre una región recuperada pequeña puede ocultar la pérdida de información. A su vez, la ventaja de requerir una sola imagen debe contrastarse con el tiempo, la memoria y la energía del procesamiento completo. Estas consideraciones conducen a la pregunta que orienta el trabajo: *¿en qué medida las redes U-Net y Dual CNN, que estiman la fase discontinua a partir de una sola imagen, alcanzan la precisión y la robustez de los métodos clásicos de corrimiento de fase y desenvolvimiento bajo condiciones radiométricas comunes, y con qué costo computacional?*

3. Objetivos

3.1 Objetivo general

Comparar cuantitativamente el desempeño de métodos clásicos de recuperación de fase en proyección de franjas con el de redes neuronales profundas (*Dual CNN* y *U-Net*), evaluando su precisión (*RMSE*, *MAE*, *SSIM*) y su robustez frente a ruido, saturación y no linealidad, mediante experimentos sobre datos simulados y datos sintéticos basados en modelos 3D complejos.

3.2 Objetivos específicos

- Revisar y sintetizar el estado del arte sobre la recuperación de fase mediante la técnica de proyección de franjas, abarcando métodos clásicos (corrimiento de fase y perfilometría temporal) y arquitecturas basadas en aprendizaje profundo (*Dual CNN* y *U-Net*), identificando sus principales características, limitaciones y criterios de evaluación reportados en la literatura.

- Implementar y validar, en un entorno de simulación controlado, los algoritmos clásicos y las arquitecturas de redes neuronales profundas seleccionadas, garantizando la reproducibilidad de los experimentos mediante control de parámetros, semillas aleatorias y versionado del código y los datos.

- Definir y calcular métricas cuantitativas de desempeño (como el error cuadrático medio —RMSE— y el error absoluto medio —MAE—), así como las métricas de costo computacional (tiempo de inferencia, uso de memoria y eficiencia energética), estableciendo un protocolo estandarizado para el análisis estadístico y comparativo.

- Comparar cuantitativamente los resultados obtenidos por los métodos clásicos y las redes neuronales profundas bajo distintos escenarios simulados, analizando cómo las variaciones

de ruido, desfase y condiciones de iluminación influyen en su precisión y estabilidad.

4. Alcance y delimitación

Los objetivos se desarrollaron mediante dos experimentos complementarios, descritos en el capítulo 8. El primero (Exp1) utiliza 800 escenas de 600×800 píxeles generadas a partir de diez familias de superficies analíticas y compara PSA3, PSA4 y PSA12, el desenvolvimiento temporal lineal (LTPU), la U-Net y la Dual CNN. El segundo (Exp2) utiliza 1920 escenas de 768×1024 píxeles generadas con un simulador suministrado por el director a partir de diez modelos tridimensionales, y conserva los mismos métodos excepto LTPU. El trabajo se delimita de la siguiente manera:

- **Datos.** Todas las imágenes son simuladas. No se realizaron adquisiciones con un sistema físico ni la calibración necesaria para convertir la fase en coordenadas métricas; la comparación se efectúa en el dominio de la fase.
- **Métodos.** Las redes son adaptaciones de las formulaciones de Hu et al. (2021) y Feng et al. (2019), implementadas y entrenadas en el proyecto; no reproducen literalmente las condiciones de esas publicaciones. En cada experimento se entrenaron una sola vez, con semilla fija y únicamente en la condición Moderate.
- **Condiciones radiométricas.** Se evaluaron cuatro condiciones: Ideal, Medium, Moderate y VeryStrong. Las tres últimas combinan ruido gaussiano ($\sigma = 0{,}01$), no linealidad gamma ($\gamma = 2{,}6$) y una contribución especular con recorte en el límite del sensor, que representan respectivamente el ruido, la no linealidad y la saturación del objetivo general y las condiciones de iluminación del objetivo específico 3.2. Estas condiciones no forman un diseño factorial,

por lo que no permiten aislar el efecto de cada mecanismo. Se compararon algoritmos con distintos números de pasos nominales ($N = 3$, 4 y 12), pero no se evaluó su sensibilidad a errores en el desplazamiento de fase. Por tanto, este componente del objetivo específico 3.2 queda fuera del alcance experimental realizado.

- **Desenvolvimiento.** La fase continua de PSA3, PSA4, PSA12 y las redes se obtuvo con desenvolvimiento espacial, y LTPU solo se evaluó en el Exp1. En el Exp2 el desenvolvimiento empleó la máscara del objeto y la línea central del proyector proporcionadas por el simulador como entradas auxiliares.

- **Costo computacional.** Se midieron el número de imágenes, el tiempo y la memoria de inferencia y el tiempo y la energía del procedimiento completo hasta $\widehat{\Phi}(x,y)$ en un único equipo. La energía corresponde al consumo neto de CPU y GPU, no al consumo total en el enchufe, y el número de imágenes no equivale a una medición del tiempo de adquisición.

5. Contribuciones

Los principales aportes del trabajo son los siguientes:

- Un protocolo de comparación que evalúa por separado la fase discontinua, con errores circulares y SSIM, y la fase continua, con errores lineales, cobertura recuperada, error de orden y cobertura correcta. El protocolo incorpora pruebas pareadas por escena con corrección por comparaciones múltiples, tamaños de efecto e intervalos de confianza, y mide el costo sobre el procedimiento completo y no solo sobre la inferencia.

- Evidencia cuantitativa, obtenida en dos escenarios de complejidad geométrica creciente, sobre

el desempeño de tres algoritmos de corrimiento de fase, un método de desenvolvimiento temporal y dos redes neuronales bajo condiciones radiométricas comunes. Esa evidencia muestra que la ventaja de cada método depende de la condición, de la métrica y del experimento, y que la reducción a una sola imagen de las redes no se acompañó de menor energía en el procedimiento completo.

- Una implementación reproducible de los métodos, del entrenamiento y de la evaluación, con particiones fijas por geometría, semilla declarada y registro de configuraciones, manifiestos y código en el repositorio del proyecto, alojado en GitHub (Mayorga Gamboa et al., 2026).

6. Organización del documento

El capítulo 7 presenta los fundamentos necesarios para seguir la metodología: el modelo de la imagen de franjas, la relación entre fase discontinua y fase continua, los algoritmos de corrimiento de fase, el desenvolvimiento espacial y temporal, las degradaciones radiométricas, las redes U-Net y Dual CNN y los criterios de evaluación; con ello responde al objetivo específico 3.2. El capítulo 8 describe los dos experimentos, la generación de los datos, la implementación y el entrenamiento de los métodos y el protocolo de evaluación y de costo, en correspondencia con los objetivos 3.2 y 3.2. El capítulo 9 presenta y discute los resultados de cada experimento, compara ambos y expone las limitaciones del estudio, lo que desarrolla el objetivo 3.2. Finalmente, el capítulo 10 resume las conclusiones según cada objetivo y propone líneas de trabajo futuro. El apéndice A reúne las tablas complementarias de las pruebas estadísticas.

7. Fundamentos y antecedentes

Este capítulo presenta los conceptos necesarios para comprender la metodología del capítulo 8. Su orden sigue la cadena de procesamiento evaluada en el trabajo: la formación de la imagen de franjas, la obtención de la fase discontinua, su desenvolvimiento hasta la fase continua, las degradaciones radiométricas que afectan ese proceso, las redes neuronales comparadas y los criterios con los que se evalúa el resultado. La notación se mantiene en todo el documento: Φ denota la fase continua, φ la fase discontinua y el sombrero identifica una estimación, como en $\widehat{\varphi}$ y $\widehat{\Phi}$. Los mapas de fase se representan en escala de grises. En la fase discontinua, el negro corresponde a $-\pi$ y el blanco a π. En la fase continua, el negro y el blanco corresponden a los límites de la escala utilizada en cada figura, que se indican en su barra de color; cuando se comparan varias imágenes de fase continua, todas emplean los mismos límites, de modo que un mismo tono de gris representa el mismo valor de fase. En las configuraciones simuladas, la fase crece de izquierda a derecha. En las gráficas se mantiene una misma convención de color: la referencia o señal ideal se traza en negro continuo, las estimaciones en azul discontinuo, el error y las señales degradadas en rojo, y la fase discontinua de entrada y los elementos auxiliares en gris. Las marcas sobre las imágenes, como la fila de un perfil, un píxel o una semilla, se dibujan en verde; en los mapas, el exterior de la máscara se muestra en azul claro y los píxeles válidos que quedan sin fase continua, en ámbar.

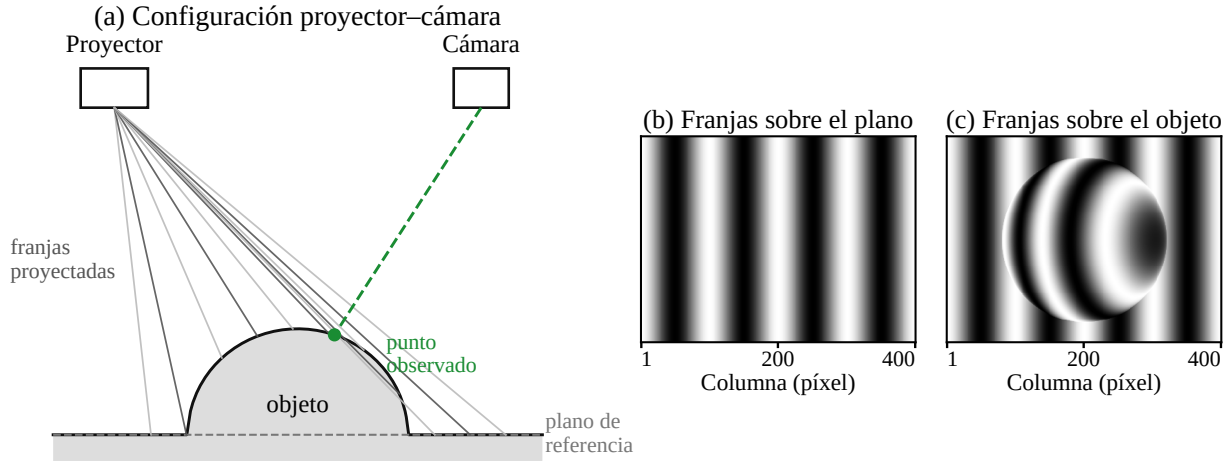
7.1 Técnica de proyección de franjas

La técnica de proyección de franjas es un método de perfilometría óptica sin contacto que permite medir la forma tridimensional de una superficie a partir de imágenes (González, 2020; Zhang, 2013). Un proyector ilumina el objeto con un patrón de franjas, generalmente sinusoidal, y una cámara, situada en una dirección distinta, registra ese patrón sobre la superficie. Como se

observa en la Figura 1, sobre un plano las franjas conservan su forma rectilínea, mientras que sobre un objeto con relieve aparecen desplazadas y curvadas. Ese desplazamiento depende de la altura de cada punto y de la geometría entre el proyector y la cámara.

Figura 1

Principio de la técnica de proyección de franjas



Nota. (a) Esquema lateral: el proyector ilumina el objeto con franjas y la cámara observa la superficie desde otra dirección; la línea discontinua verde indica la observación de un punto del objeto. (b) Franjas sinusoidales registradas sobre un plano de referencia. (c) Las mismas franjas sobre un casquete esférico apoyado en el plano: el relieve desplaza y curva las franjas. Simulación ilustrativa con cuatro periodos en el campo de 400×300 píxeles.

La información de altura queda codificada en la fase del patrón. A lo largo de la imagen proyectada, la fase crece con la columna del proyector, de modo que la fase observada en un píxel de la cámara indica qué columna iluminó el punto correspondiente. Esa identificación requiere, sin embargo, la fase absoluta, es decir, la fase continua referida a un origen conocido. La fase que se calcula directamente a partir de las imágenes solo se conoce dentro de un periodo del patrón, por lo que columnas separadas por un número entero de periodos comparten el mismo valor; la sección 7.3 describe esta ambigüedad y la sección 7.5 la forma de resolverla. La correspondencia con el proyector es la que emplea el segundo experimento para construir su referencia de fase

continua a partir de la coordenada del proyector (ecuación (15a)). Para pasar de la fase a coordenadas métricas se requiere además una calibración del sistema (Gonzalez and Meneses, 2019); el presente trabajo no realiza esa conversión y compara los métodos directamente en el dominio de la fase.

Los métodos para extraer la fase de las imágenes pueden agruparse según el número de imágenes que necesitan. Los métodos de una sola imagen, como el análisis por transformada de Fourier (Takeda et al., 1982) o las redes neuronales estudiadas en este trabajo, estiman la fase a partir de un único patrón. Los algoritmos de corrimiento de fase, en cambio, combinan varias imágenes del mismo patrón desplazado y obtienen la fase píxel a píxel (Zuo et al., 2018). Esta diferencia en el número de imágenes es uno de los criterios de comparación del trabajo.

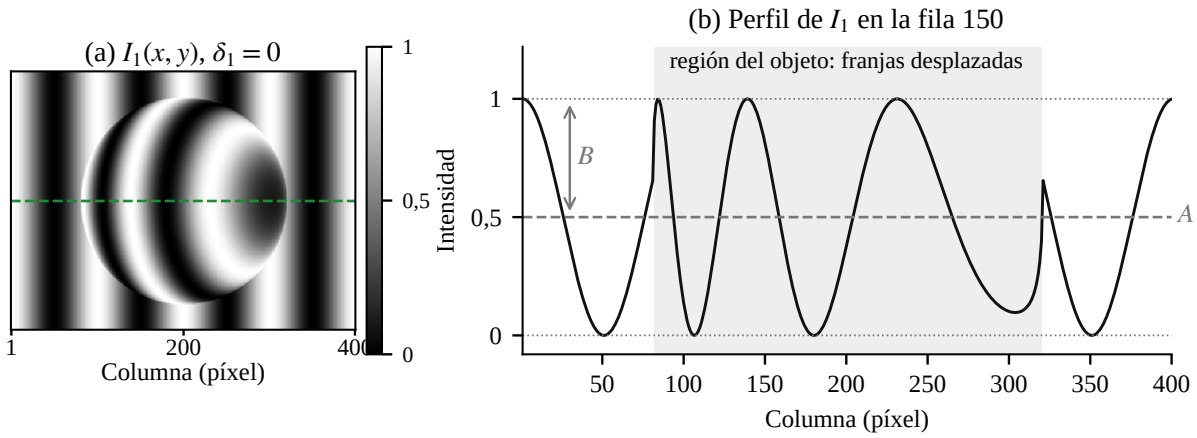
7.2 Modelo de la imagen de franjas

La intensidad registrada en el píxel (x,y) para la n-ésima imagen de una secuencia se modela con la ecuación (1) (Zuo et al., 2018):

$$I_n(x,y) = A(x,y) + B(x,y)\cos\left[\Phi(x,y) + \delta_n\right], \tag{1}$$

donde $A(x,y)$ es la intensidad de fondo, $B(x,y)$ es la amplitud de modulación de las franjas, $\Phi(x,y)$ es la fase continua que codifica la geometría y δ_n es un desplazamiento de fase conocido, introducido por el proyector en cada imagen. El fondo y la modulación dependen de la iluminación y de la reflectancia de la superficie; en el generador del primer experimento ambos se escalan con una respuesta lambertiana normalizada (ecuación (9a)). La Figura 2 muestra el significado de cada término: la intensidad oscila alrededor de A con amplitud B, y la posición de sus máximos y mínimos queda determinada por Φ.

Figura 2
Modelo de la imagen de franjas



Nota. Corresponde a la ecuación (1). (a) Imagen $I_1(x,y)$ con $\delta_1 = 0$ sobre el casquete esférico de la Figura 1; la línea discontinua verde marca la fila 150. (b) Perfil de I_1 en esa fila: la intensidad oscila alrededor del fondo A (línea discontinua gris) con amplitud B (flecha gris); en la región del objeto (fondo gris claro) los máximos se desplazan porque Φ cambia con la altura. Simulación ilustrativa con $A = B = 0{,}5$.

El cociente B/A indica el contraste de las franjas. Donde la modulación es muy baja, por ejemplo en sombras o fuera del objeto, la fase no puede estimarse de forma confiable. Por esa razón las evaluaciones del capítulo 8 se restringen a una máscara de píxeles válidos, que en el primer experimento exige una modulación ideal superior a un umbral.

7.3 Fase continua y fase discontinua

La fase continua $\Phi(x,y)$ es la fase del patrón tal como la determina la geometría, sin restringirla a un intervalo, por lo que puede tomar cualquier valor real. El adjetivo *continua* no significa que varíe suavemente en toda la imagen: un escalón, un borde o una oclusión del objeto pueden producir discontinuidades físicas reales en $\Phi(x,y)$. Lo que la distingue de la fase discontinua es que no contiene los saltos artificiales de 2π que introduce la representación en un intervalo de longitud 2π. Los métodos que estiman la fase a partir de funciones trigonométricas solo pueden devolver su valor principal, comprendido en un intervalo de longitud 2π. Para expresar

esta operación se usa el operador de valor principal de la ecuación (2):

$$\mathscr{W}(\alpha) = \text{atan2}(\sin\alpha, \cos\alpha) \in (-\pi, \pi], \tag{2}$$

donde atan2 es la arcotangente de dos argumentos, que identifica el cuadrante a partir de los signos del seno y del coseno. La fase discontinua es $\varphi(x,y) = \mathscr{W}[\Phi(x,y)]$, y ambas fases se relacionan mediante la ecuación (3):

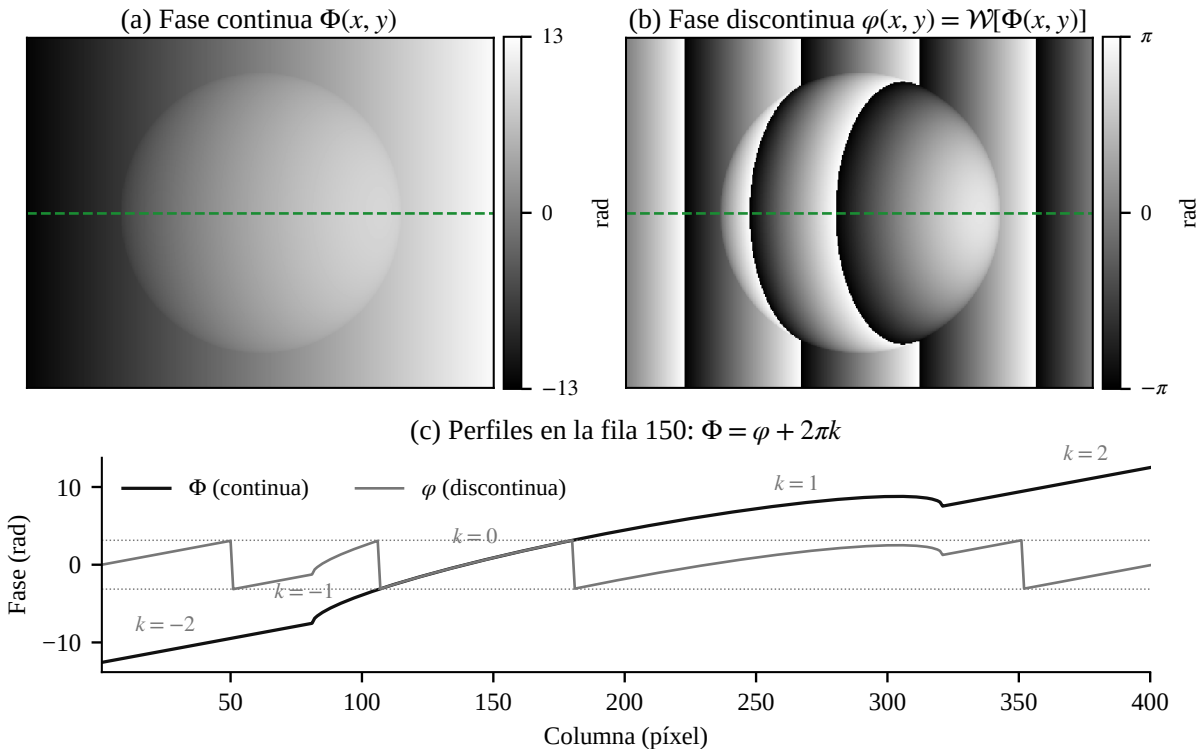
$$\Phi(x,y) = \varphi(x,y) + 2\pi k(x,y), \qquad k(x,y) \in \mathbb{Z}, \tag{3}$$

donde $k(x,y)$ es el orden de franja, es decir, el número entero de ciclos de 2π que separan ambas fases en cada píxel.

La Figura 3 ilustra esta relación. La fase continua crece de izquierda a derecha, desde valores negativos (negro) hasta positivos (blanco). La fase discontinua repite el recorrido de $-\pi$ a π y, cada vez que alcanza π, salta a $-\pi$; en cada tramo entre saltos el orden de franja k es constante. Estos saltos no corresponden a cambios de la superficie, sino a la representación en el intervalo principal.

Figura 3

Relación entre la fase continua y la fase discontinua



Nota. Corresponde al casquete esférico de la Figura 1. (a) Fase continua $\Phi(x,y)$, que crece de izquierda a derecha; escala de grises entre los límites indicados en la barra de color, de -13 rad (negro) a 13 rad (blanco). (b) Fase discontinua $\varphi(x,y) = \mathscr{W}[\Phi(x,y)]$, en escala de grises de $-\pi$ (negro) a π (blanco). (c) Perfiles en la fila 150, marcada en verde en (a) y (b): Φ (negro) y φ (gris); el orden de franja k indicado sobre cada tramo cumple $\Phi = \varphi + 2\pi k$. Simulación ilustrativa.

La recuperación de la fase se divide, por tanto, en dos problemas. El primero consiste en estimar $\widehat{\varphi}$ a partir de las imágenes, mediante corrimiento de fase (sección 7.4) o mediante redes neuronales (sección 7.8). El segundo consiste en estimar el orden de franja para obtener $\widehat{\Phi}$, lo que se denomina desenvolvimiento de fase (secciones 7.5 y 7.6). Un error de $\widehat{\varphi}$ se mide en fracciones de radián, mientras que un error en el orden de franja desplaza la fase continua en múltiplos de 2π; por eso el trabajo evalúa ambas fases por separado.

7.4 Algoritmos de corrimiento de fase

Los algoritmos de corrimiento de fase (PSA, por sus siglas en inglés) registran N imágenes del patrón con desplazamientos conocidos y despejan la fase en cada píxel (Zuo et al., 2018; Rastogi and Hack, 2014). Como la ecuación (1) contiene tres incógnitas por píxel —A, B y Φ—, se requieren al menos tres imágenes. Con desplazamientos equiespaciados $\delta_n = -2\pi(n-1)/N$, $n = 1,\ldots,N$, las sumas ponderadas de las ecuaciones (4a) y (4b) aíslan las componentes en cuadratura de la fase:

$$M = -\frac{2}{N}\sum_{n=1}^{N} I_n \sin\delta_n = B\sin\Phi, \tag{4a}$$

$$D = \frac{2}{N}\sum_{n=1}^{N} I_n \cos\delta_n = B\cos\Phi, \tag{4b}$$

donde se ha omitido (x,y). El resultado se debe a que, para desplazamientos equiespaciados en un ciclo completo, las sumas de $\sin\delta_n$, de $\cos\delta_n$ y de $\sin\delta_n\cos\delta_n$ se anulan, mientras que las de $\sin^2\delta_n$ y $\cos^2\delta_n$ valen $N/2$. Así, el fondo A desaparece y la fase se obtiene con la ecuación (5):

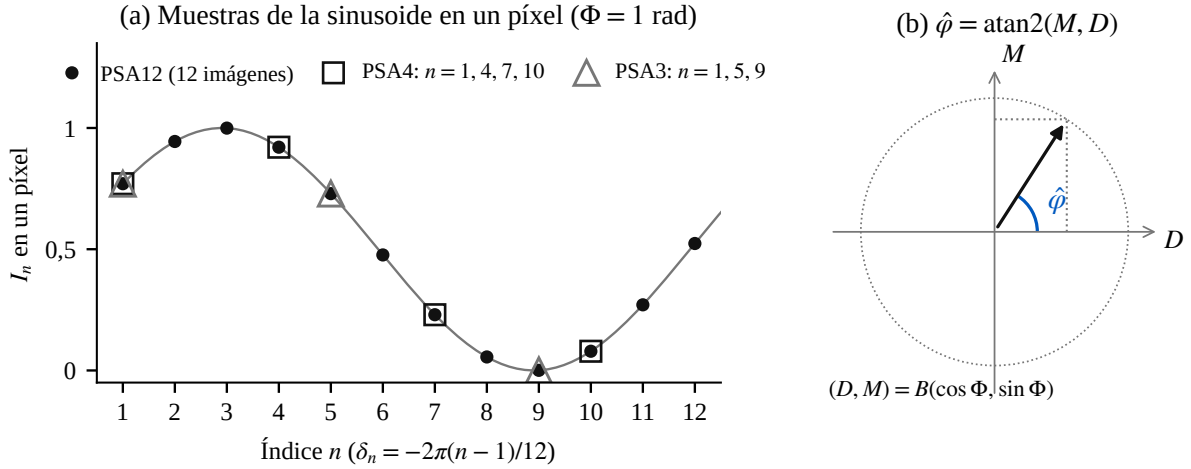
$$\widehat{\varphi} = \operatorname{atan2}(M,D) = \mathscr{W}(\Phi). \tag{5}$$

El uso de atan2 en lugar de la arcotangente de un solo argumento conserva el signo de M y de D y entrega la fase en todo el intervalo $(-\pi,\pi]$. Por ello, el resultado de un PSA es siempre una fase discontinua.

La Figura 4 muestra este procedimiento en un píxel. Las doce intensidades de una secuencia con $N = 12$ muestrean una sinusoide cuya posición depende de Φ. PSA3 y PSA4 emplean subconjuntos equiespaciados de esa misma secuencia, con tres y cuatro muestras por ciclo, que es la configuración utilizada en ambos experimentos (ecuación (11)). El par (D,M) representa un punto de radio B cuyo ángulo es la fase estimada.

Figura 4

Corrimiento de fase en un píxel



Nota. Se representa un píxel con $\Phi = 1$ rad, $A = B = 0{,}5$ y $\delta_n = -2\pi(n-1)/12$. (a) Las doce intensidades I_n (puntos negros) muestrean la sinusoide (línea gris); PSA4 usa las muestras $n = 1, 4, 7, 10$ (cuadrados negros) y PSA3 las muestras $n = 1, 5, 9$ (triángulos grises), todas de la misma secuencia. (b) Las componentes de las ecuaciones (4a) y (4b) forman el punto $(D, M) = B(\cos\Phi, \sin\Phi)$, cuyo ángulo, marcado en azul, es $\widehat{\varphi} = \operatorname{atan2}(M, D)$.

Aumentar el número de imágenes mejora la estimación por dos vías. Frente al ruido aditivo gaussiano de desviación σ, la desviación del error de fase de un algoritmo de N pasos equiespaciados es aproximadamente $\sigma_\varphi \approx \sigma/(B\sqrt{N/2})$ (Surrel, 1997), de modo que disminuye con la raíz del número de imágenes y con la modulación. Además, cuando la sinusoide se deforma y aparecen armónicos de orden superior, un algoritmo de N pasos equiespaciados solo es sensible a los armónicos de orden $jN \pm 1$, con $j = 1, 2, \ldots$ (Zuo et al., 2018). En consecuencia, PSA3 es sensible a los armónicos de orden 2, 4, 5, 7, …; PSA4, a los de orden 3, 5, 7, 9, …; y PSA12, a la familia $12j \pm 1$, cuyos primeros miembros son los armónicos de orden 11 y 13. Como la amplitud de los armónicos suele disminuir con su orden, el primer armónico al que responde cada algoritmo es el que domina su error. Esta propiedad explica la diferencia de comportamiento frente a la no linealidad descrita en la sección 7.7. La contrapartida es el número de imágenes que deben adquirirse: 3, 4 o 12.

7.5 Desenvolvimiento espacial de la fase

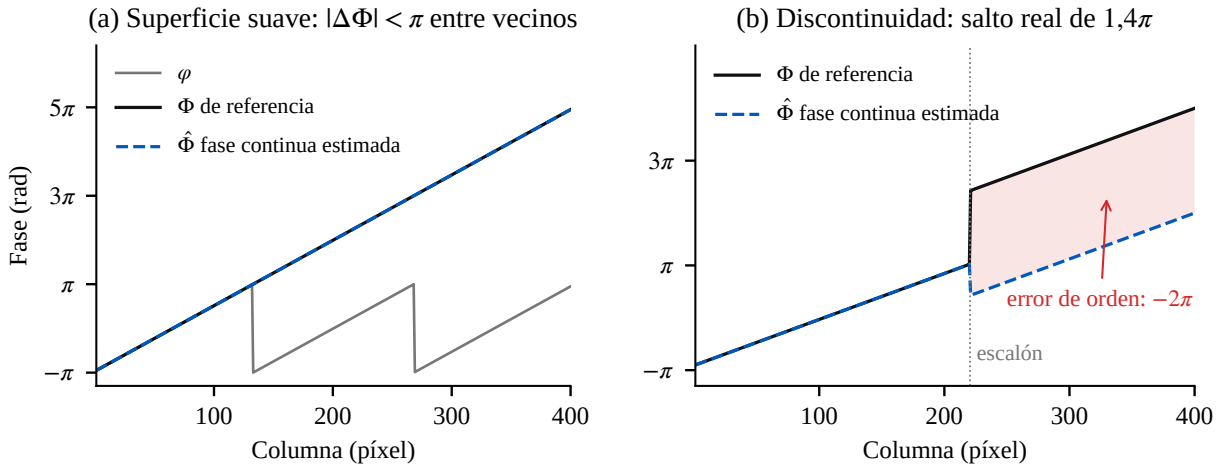
El desenvolvimiento espacial estima el orden de franja comparando cada píxel con sus vecinos (Ghiglia and Pritt, 1998). Si la fase continua cambia menos de π entre dos píxeles vecinos p y q, su diferencia puede recuperarse a partir de la fase discontinua, como indica la ecuación (6):

$$\Phi(q) - \Phi(p) = \mathscr{W}\left[\varphi(q) - \varphi(p)\right], \qquad \text{si } |\Phi(q) - \Phi(p)| < \pi. \tag{6}$$

Partiendo de un píxel semilla y sumando estas diferencias a lo largo de un recorrido se reconstruye la fase continua en toda la región conectada, como muestra la Figura 5(a). El resultado queda determinado salvo un múltiplo entero de 2π común a toda la región, que depende del valor asignado a la semilla. Para obtener una fase absoluta se necesita una referencia externa. En el capítulo 8, el primer experimento parte del centroide de la máscara y su evaluación admite un único desplazamiento global de 2π por escena, mientras que el segundo usa la línea central del proyector, cuya fase es conocida, como origen de la fase.

Figura 5

Desenvolvimiento espacial a lo largo de una fila



Nota. Se aplica la ecuación (6). (a) Superficie suave: la fase continua cambia menos de π entre píxeles vecinos y la suma de las diferencias de φ (gris) recupera la fase continua estimada $\widehat{\Phi}$ (azul discontinuo), que coincide con la referencia Φ (negro). (b) Superficie con un escalón en la columna 220, donde la fase real salta $1{,}4\pi$: la ecuación (6) interpreta el salto como $-0{,}6\pi$, y desde ese punto $\widehat{\Phi}$ (azul discontinuo) queda 2π por debajo de la referencia Φ (negro). La región sombreada en rojo corresponde a un error de orden que se propaga al resto del recorrido. Simulación ilustrativa.

La condición de la ecuación (6) puede incumplirse por ruido, por regiones de baja modulación, por sombras o por discontinuidades reales de la superficie. En esos casos se produce un error de orden: la fase continua queda desplazada en un múltiplo de 2π y el error se propaga a todos los píxeles alcanzados después, como ilustra la Figura 5(b). Estas inconsistencias pueden detectarse mediante residuos: si la suma de las diferencias $\mathscr{W}[\varphi(q) - \varphi(p)]$ alrededor de un ciclo de 2×2 píxeles no es nula, ningún recorrido que rodee ese punto puede integrarse de forma consistente (Ghiglia and Pritt, 1998).

Para limitar la propagación de errores existen dos estrategias complementarias. La primera ordena la propagación según un mapa de calidad, de modo que los píxeles más confiables se desenvuelven primero; entre los mapas de calidad se encuentra la varianza local de las derivadas de la fase (Ghiglia and Pritt, 1998; Cui et al., 2011). La segunda establece barreras alrededor de

residuos o de zonas de baja calidad, que la propagación no atraviesa (Ghiglia and Pritt, 1998; Zhao et al., 2015). Si una región queda aislada por las barreras, puede descartarse en lugar de forzar un orden posiblemente erróneo. Este principio da lugar a dos medidas que se usan en el segundo experimento: la cobertura, que es la fracción de la región válida que se recupera, y la proporción de píxeles recuperados con error de orden.

7.6 Desenvolvimiento temporal de la fase

El desenvolvimiento temporal resuelve el orden de franja en cada píxel de forma independiente, usando una secuencia de patrones con frecuencias crecientes (Huntley and Saldner, 1993; Saldner and Huntley, 1997). Si la frecuencia relativa $f = 1, \ldots, F$ multiplica la fase de un patrón base, la fase continua de cada frecuencia es $\Phi_f = f\,\Phi_1$ y su fase discontinua es $\varphi_f = \mathscr{W}(\Phi_f)$. Cuando la fase del patrón base cumple $|\Phi_1| < \pi$, esta no presenta saltos, y la diferencia entre frecuencias consecutivas, $\Phi_f - \Phi_{f-1} = \Phi_1$, también es menor que π en magnitud. En esas condiciones, el número de vueltas que cambia entre dos frecuencias se obtiene con la ecuación (7a), y la fase continua de la frecuencia más alta, con la ecuación (7b):

$$v_f = \operatorname{round}\left(\frac{\varphi_f - \varphi_{f-1}}{2\pi}\right), \quad f = 2, \ldots, F, \tag{7a}$$

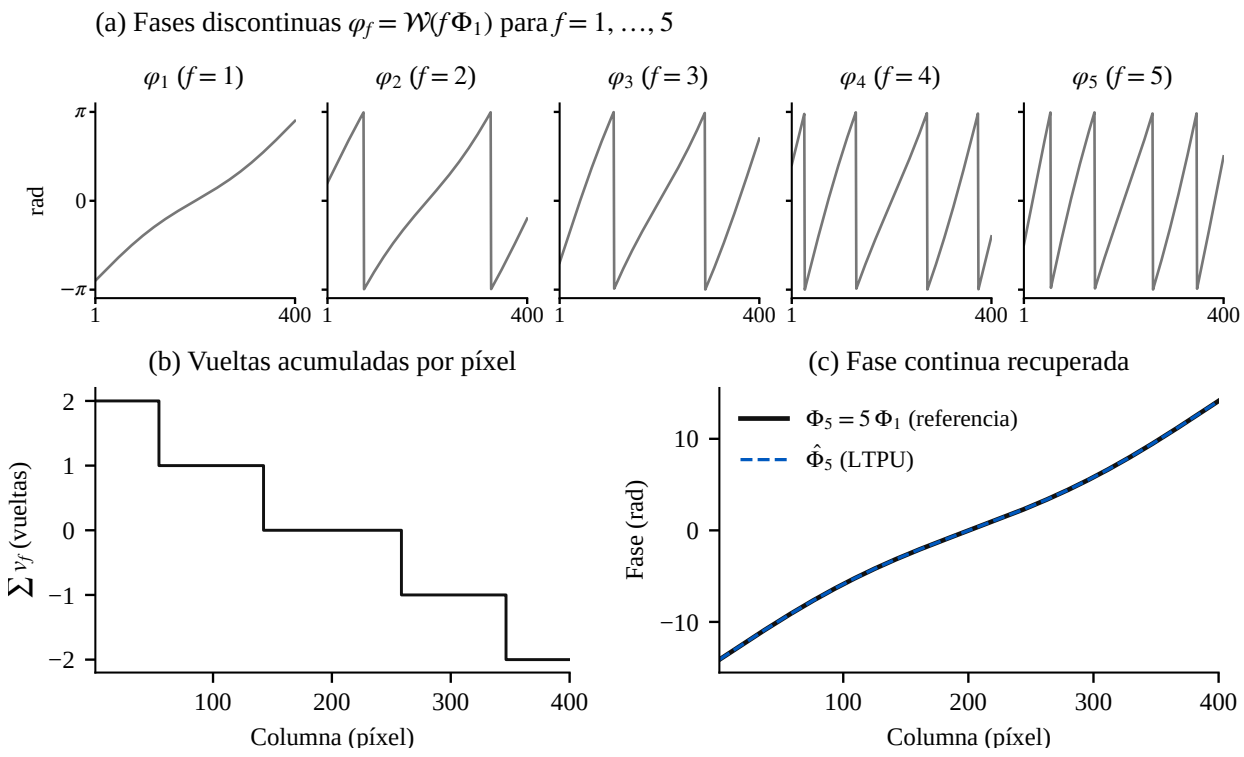
$$\Phi_F = \varphi_F - 2\pi \sum_{f=2}^{F} v_f. \tag{7b}$$

La suma de la ecuación (7b) acumula, píxel a píxel, los cambios de orden de franja entre frecuencias consecutivas, sin consultar los píxeles vecinos. Por eso un error local no se propaga espacialmente y el método puede recuperar superficies con discontinuidades (Saldner and Huntley, 1997). Su costo es el número de imágenes: cada frecuencia requiere su propia secuencia de corrimiento de fase. LTPU, la variante implementada en el primer experimento, emplea cinco frecuencias con cuatro imágenes cada una (ecuación (12)).

La Figura 6 muestra el procedimiento a lo largo de una fila. La fase de baja frecuencia identifica el orden sin ambigüedad, y las frecuencias más altas, que producen más franjas en el mismo campo, conservan esa identificación a medida que se agregan. La condición sobre los incrementos de fase es también su limitación: si el ruido o la degradación hacen que la diferencia entre frecuencias consecutivas supere π, el redondeo de la ecuación (7a) asigna una vuelta incorrecta a ese píxel.

Figura 6

Desenvolvimiento temporal a lo largo de una fila



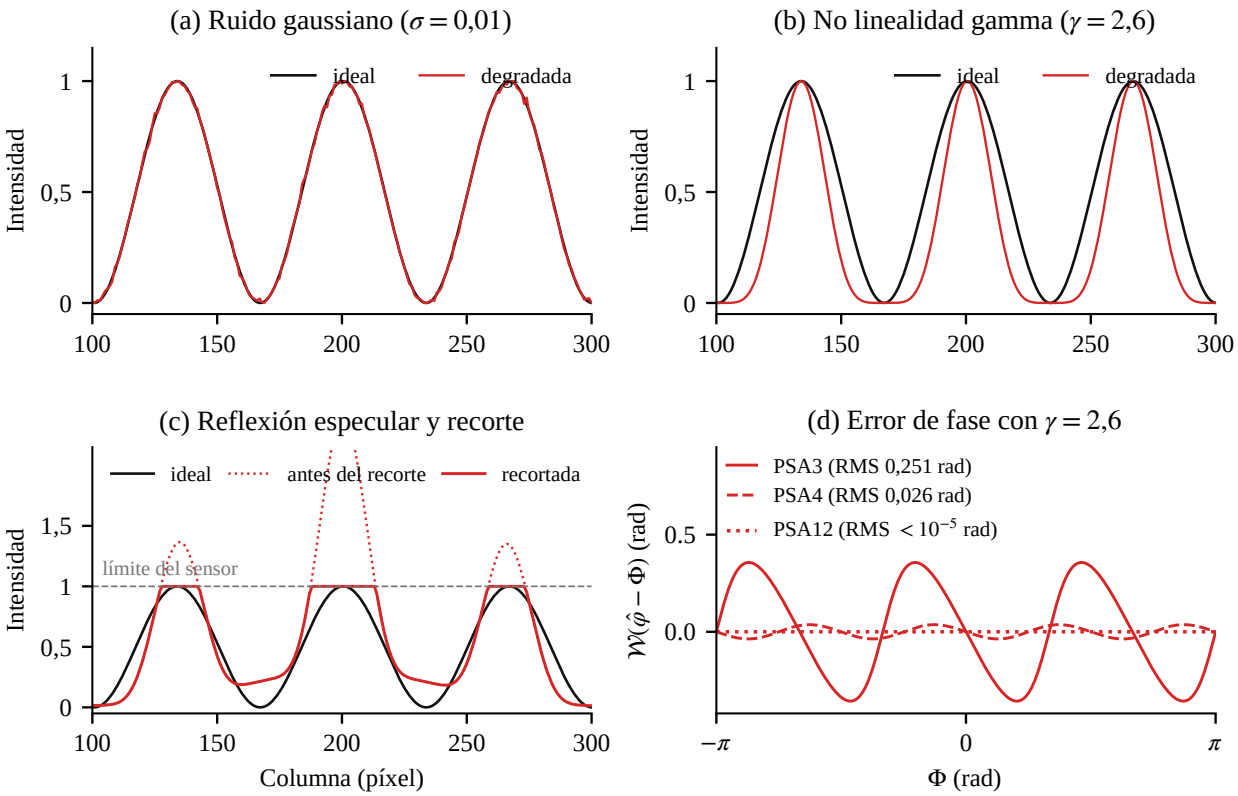
Nota. (a) Fases discontinuas $\varphi_f = \mathcal{W}(f\,\Phi_1)$ para $f = 1, \ldots, 5$: la fase del patrón base ($f = 1$) no presenta saltos y el número de saltos aumenta con la frecuencia. (b) Vueltas acumuladas $\sum v_f$ de la ecuación (7a), calculadas de forma independiente en cada píxel. (c) Fase continua $\widehat{\Phi}_5$ recuperada con la ecuación (7b) (azul discontinuo), superpuesta a la referencia $\Phi_5 = 5\,\Phi_1$ (negro); en esta simulación sin ruido ambas coinciden.

7.7 Degradaciones radiométricas

Las imágenes reales se apartan del modelo de la ecuación (1) por efecto del sensor, del proyector y de la reflexión de la superficie (Zuo et al., 2018). El trabajo considera tres fenómenos, que los experimentos combinan en condiciones compuestas (ecuación (10)). La Figura 7 muestra su efecto sobre un perfil de intensidad y sobre la fase estimada por los PSA.

Figura 7

Degradaciones radiométricas sobre un perfil de franjas



Nota. Se emplea un perfil con $A = B = 0{,}5$; la señal ideal se representa en negro y la degradada en rojo. (a) Ruido gaussiano aditivo con $\sigma = 0{,}01$. (b) No linealidad gamma con $\gamma = 2{,}6$: los máximos se estrechan y los mínimos se ensanchan. (c) Reflexión especular simulada con un lóbulo gaussiano centrado en la columna 200, antes (punteado) y después del recorte en el límite del sensor (línea discontinua gris). (d) Error de fase $\mathcal{W}(\hat{\varphi} - \Phi)$ de PSA3 (línea continua), PSA4 (discontinua) y PSA12 (punteada) cuando solo actúa la no linealidad gamma de (b); entre paréntesis, la raíz del valor cuadrático medio (RMS) del error sobre un ciclo. Los parámetros de ruido y gamma son iguales a los del capítulo 8. Simulación ilustrativa.

El ruido del sensor se modela como una perturbación gaussiana aditiva de media nula, independiente en cada píxel e imagen. Su efecto sobre la fase es aleatorio y, como se indicó en la sección 7.4, disminuye al aumentar el número de imágenes. Con la desviación empleada en los experimentos, $\sigma = 0{,}01$, la perturbación es pequeña frente a la modulación, como se aprecia en la Figura 7(a).

La no linealidad gamma describe la respuesta no lineal entre el nivel enviado al proyector y la intensidad registrada, que puede modelarse como una potencia de la intensidad ideal (Ye et al., 2015; Zuo et al., 2018). Con $\gamma > 1$, la sinusoide se deforma y aparecen armónicos de orden superior, como se observa en la Figura 7(b). A diferencia del ruido, este error es sistemático: se repite con la misma forma en cada ciclo de franja. Su efecto sobre los PSA depende de la sensibilidad a los armónicos descrita en la sección 7.4. En la simulación de la Figura 7(d), con $\gamma = 2{,}6$, el error periódico de PSA3 tiene un valor RMS de 0,251 rad, el de PSA4 de 0,026 rad y el de PSA12 inferior a 10^{-5} rad. Las oscilaciones del error tienen tres y cuatro periodos por ciclo de fase en PSA3 y PSA4, respectivamente.

La reflexión especular concentra la luz en regiones brillantes de la superficie y puede llevar la intensidad por encima del rango del sensor. Los valores que superan ese límite se recortan, como en la Figura 7(c), y la forma sinusoidal se pierde en esas zonas. El recorte no es invertible: en los píxeles saturados de varias imágenes de la secuencia, las sumas de las ecuaciones (4a) y (4b) se calculan con intensidades incorrectas y el error de fase se concentra en las regiones brillantes. En el capítulo 8, este efecto se modela con un mapa especular expandido espacialmente y un recorte en la unidad (ecuación (10c)).

7.8 Redes neuronales para la estimación de la fase

Las redes neuronales profundas se han aplicado a varias etapas del análisis de franjas, como la supresión de ruido (Yan et al., 2019), la extracción de la fase a partir de una sola imagen (Feng et al., 2019; Hu et al., 2021; Kando et al., 2019), el desenvolvimiento (Wang et al., 2019; Li et al.,

2020) y la reconstrucción completa de extremo a extremo (Machineni et al., 2020). Este trabajo se concentra en la extracción de la fase discontinua a partir de una sola imagen, mediante dos formulaciones: la estimación directa de $\widehat{\varphi}$ con una U-Net y la estimación indirecta a través de $\widehat{M}$ y $\widehat{D}$ con una Dual CNN.

7.8.1 Conceptos de las redes convolucionales

Una red neuronal convolucional transforma una imagen mediante capas sucesivas. Cada capa convolucional aplica filtros pequeños, por ejemplo de 3×3 píxeles, cuyos coeficientes se aprenden, y produce varios mapas de características o canales. Entre capas se aplican funciones de activación no lineales, como la ReLU, que anula los valores negativos, y operaciones de normalización por lotes, que estabilizan la escala de las activaciones durante el entrenamiento. Las operaciones de submuestreo, como el *max pooling* de 2×2, reducen la resolución y amplían la región de la imagen que influye en cada salida; las convoluciones transpuestas o el reordenamiento de canales recuperan la resolución. Los bloques residuales suman la entrada de un bloque a su salida, lo que facilita el entrenamiento de redes profundas.

Los parámetros de la red se ajustan minimizando una función de pérdida sobre un conjunto de entrenamiento mediante descenso por gradiente; en este trabajo se emplea el optimizador Adam. Cada recorrido completo por el conjunto de entrenamiento es una época. Un conjunto de validación, separado del de entrenamiento, permite elegir los pesos con menor pérdida de validación y detener el entrenamiento cuando esa pérdida deja de mejorar durante un número fijo de épocas. El desempeño de una red depende de la semejanza entre los datos de entrenamiento y los evaluados; por eso el capítulo 9 distingue la condición de entrenamiento de las demás condiciones radiométricas.

La inferencia puede realizarse con distintas precisiones numéricas. En fp32, cada número ocupa 32 bits, con 8 bits de exponente y 23 de mantisa. En bf16 ocupa 16 bits: conserva los 8 bits de exponente, y por tanto el mismo rango, pero solo 7 bits de mantisa. Reducir la precisión disminuye la memoria y puede reducir el tiempo y la energía, a costa de una menor resolución numérica en los

cálculos intermedios. Por ello el capítulo 8 declara la precisión de cada medición.

7.8.2 Estimación directa: U-Net

La U-Net es una arquitectura de codificador y decodificador propuesta inicialmente para segmentación de imágenes (Ronneberger et al., 2015). El codificador reduce la resolución y aumenta el número de canales para capturar el contexto de la imagen; el decodificador recupera la resolución, y las conexiones de salto concatenan los mapas del codificador con los del decodificador de igual resolución para conservar los detalles finos. Hu et al. (2021) la emplearon para estimar la fase discontinua a partir de un único patrón de franjas, formulando el problema como una clasificación por píxel: el intervalo $(-\pi, \pi]$ se divide en 256 niveles y la red asigna a cada píxel puntuaciones para cada nivel. Las puntuaciones se convierten en probabilidades con la función *softmax* y la clase de mayor puntuación se transforma de nuevo en fase. En la implementación de este trabajo, la red se entrena con la entropía cruzada y la conversión entre clase y fase sigue la ecuación (13).

Esta representación tiene una resolución angular finita: con 256 clases, dos niveles consecutivos están separados por $2\pi/255 \approx 0{,}0246$ rad, como se observa en la Figura 8(b). Además, la clasificación no impone la continuidad entre clases vecinas, por lo que la red debe aprender por sí misma dónde se ubican los saltos de $-\pi$ a π de la fase discontinua.

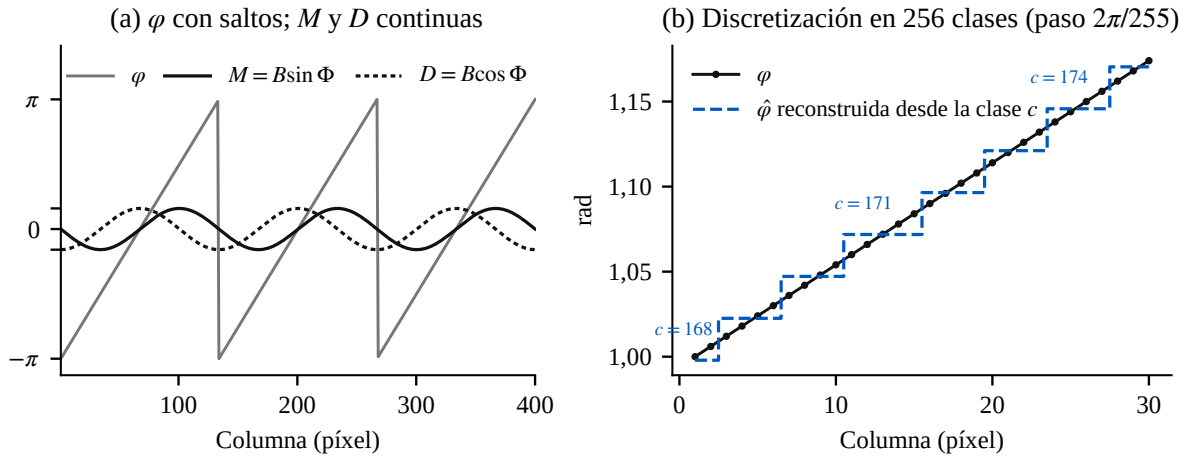
7.8.3 Estimación indirecta: Dual CNN

Feng et al. (2019) propusieron estimar, en lugar de la fase, las cantidades intermedias del corrimiento de fase. Su arquitectura encadena dos redes: la primera estima el fondo $\widehat{A}$ a partir del patrón, y la segunda recibe el patrón junto con ese fondo y estima el numerador $\widehat{M}$ y el denominador $\widehat{D}$ de las ecuaciones (4a) y (4b). La fase se calcula después con $\widehat{\varphi} = \operatorname{atan2}(\widehat{M}, \widehat{D})$, igual que en la ecuación (5). La ventaja de esta formulación se aprecia en la Figura 8(a): $M = B\sin\Phi$ y $D = B\cos\Phi$ son funciones continuas, sin los saltos de 2π de φ, lo que permite plantear su estimación como un problema de regresión; en este trabajo, ambas redes se entrenan con el error cuadrático medio. Los

saltos de la fase discontinua aparecen solo al final, en el cálculo de la arcotangente.

Figura 8

Representaciones de salida de las dos redes



Nota. Perfiles a lo largo de una fila. (a) La fase discontinua φ (gris), estimada directamente por la U-Net, presenta saltos de 2π; las componentes $M = B\sin\Phi$ (negro continuo) y $D = B\cos\Phi$ (negro punteado), estimadas por la Dual CNN, son continuas ($B = 0{,}5$). (b) Discretización de la U-Net en una rampa de fase suave: cada píxel se asigna a una de 256 clases c y la fase reconstruida a partir de la clase, $\widehat{\varphi}$ (azul discontinuo), forma escalones de $2\pi/255$ rad alrededor de la fase original φ (negro). Simulación ilustrativa.

Ambas redes reciben una sola imagen y entregan una fase discontinua, del mismo tipo que la de los PSA. Para obtener la fase continua, su salida se somete al mismo desenvolvimiento espacial que los PSA en cada experimento. Las arquitecturas implementadas, con sus canales y dimensiones, se describen en la Figura 14.

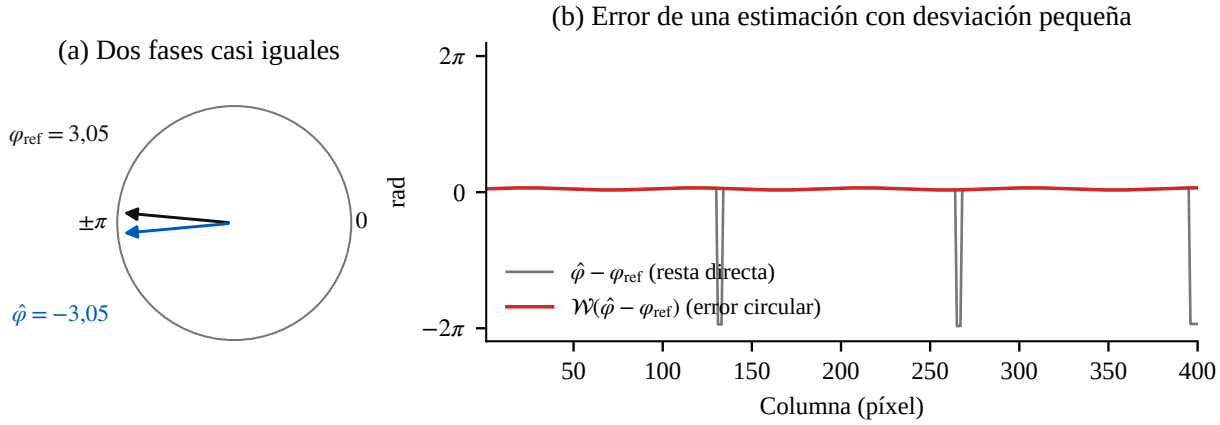
7.9 Evaluación de la recuperación de fase

Comparar los métodos exige medir cuánto se aparta cada estimación de su referencia, y la forma de hacerlo depende del tipo de fase. La fase discontinua es periódica: dos valores cercanos a π y a $-\pi$ representan casi la misma fase, aunque su resta directa sea cercana a 2π. Por eso se usa el error circular $\mathscr{W}(\widehat{\varphi} - \varphi_{\text{ref}})$, que devuelve la diferencia más corta entre ambos ángulos. La Figura 9 ilustra esta diferencia: una estimación con una desviación pequeña produce, con la resta directa,

falsos errores de casi 2π que desaparecen con el error circular.

Figura 9

Error circular de la fase discontinua



Nota. (a) Las fases $\varphi_{\text{ref}} = 3{,}05$ rad y $\widehat{\varphi} = -3{,}05$ rad están separadas por $0{,}18$ rad en el círculo, aunque su resta directa sea $-6{,}10$ rad. (b) Para una estimación desplazada unas décimas de radián respecto a la referencia, la resta directa (gris) presenta picos cercanos a -2π en los saltos de la fase, mientras que el error circular $\mathcal{W}\left(\widehat{\varphi} - \varphi_{\text{ref}}\right)$ (rojo) conserva la desviación real. Simulación ilustrativa.

A partir del error de cada píxel se calculan dos medidas de magnitud. La raíz del error cuadrático medio (RMSE) pondera más las desviaciones grandes, mientras que el error absoluto medio (MAE) resume la magnitud típica del error; cuando el RMSE es mucho mayor que el MAE, el error se concentra en pocos píxeles con desviaciones grandes. En la fase continua, en cambio, el error se calcula con la resta directa, porque allí un salto de 2π sí representa un error: un error mayor que π en valor absoluto indica que el píxel quedó en un orden de franja equivocado. Con ese criterio se definen el error de orden y la cobertura correcta del segundo experimento, que complementan a la cobertura descrita en la sección 7.5. Las definiciones formales de todas estas medidas se presentan en la sección 8.3.

El RMSE y el MAE tratan cada píxel por separado y no indican si la estimación conserva la organización espacial de la fase. Para valorar ese aspecto se emplea el índice de similaridad estructural (SSIM), que compara dos imágenes en ventanas locales a partir de la semejanza de sus

medias, de sus contrastes y de su correlación (Wang et al., 2004). El índice vale 1 cuando ambas imágenes son idénticas y es sensible a la posición de bordes y saltos. Como fue diseñado para imágenes de intensidad, su aplicación a una fase periódica obliga a decidir cómo representarla, y por eso cada experimento declara su protocolo: el primero calcula un SSIM circular sobre el coseno y el seno de la fase, y el segundo un SSIM global sobre la fase normalizada, con una sola ventana.

Como todos los métodos se evalúan sobre las mismas escenas, sus resultados forman medidas pareadas, y las diferencias entre ellos se analizan sin suponer una distribución normal. La prueba de Friedman ordena los métodos dentro de cada escena y contrasta si sus rangos medios difieren; la prueba de rangos con signo de Wilcoxon compara después cada par de métodos, y los valores p se ajustan con el procedimiento de Holm, que controla la probabilidad de cometer al menos un falso positivo dentro de cada familia de contrastes (Demšar, 2006). Un valor p pequeño indica evidencia contra la igualdad de los métodos, pero no informa el tamaño de la diferencia. Por eso se acompaña de la mediana de las diferencias por escena, expresada en la unidad de la métrica, del rango biserial, que indica en qué medida predominan las diferencias en una misma dirección, y de intervalos de confianza obtenidos por remuestreo (*bootstrap*) de escenas.

La precisión se contrasta, por último, con el costo de obtenerla, que se describe con varias magnitudes no equivalentes. El número de imágenes determina el tiempo y la complejidad de la adquisición. El tiempo de cómputo se resume con la mediana de varias repeticiones, después de ejecuciones de calentamiento que estabilizan el equipo. La memoria se registra como el pico de uso durante la ejecución; como en procesadores y tarjetas gráficas se mide con herramientas distintas, sus valores no se suman ni se comparan entre sí. La energía es la integral de la potencia en el tiempo; al restar la potencia del equipo en reposo se obtiene la energía neta atribuible al procedimiento, que es distinta del consumo total medido en el enchufe. El capítulo 8 define el alcance de cada una de estas mediciones en el trabajo.

8. Metodología

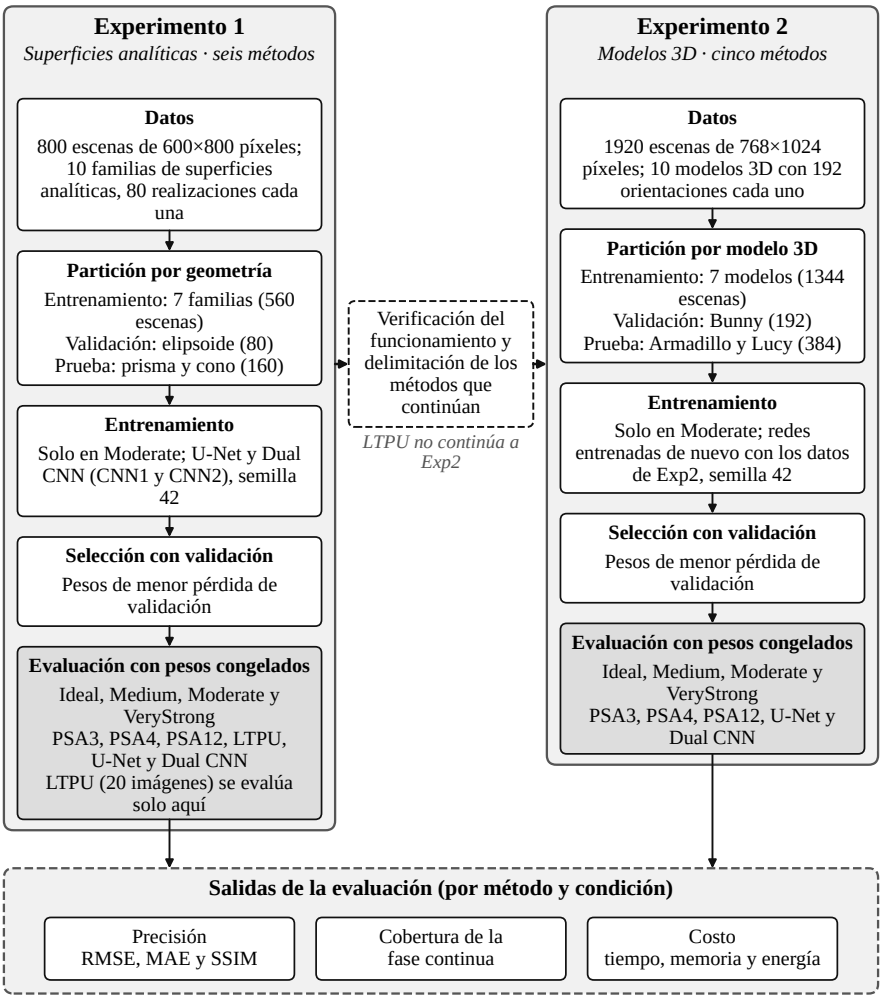
La comparación entre métodos clásicos y redes neuronales se desarrolló mediante dos experimentos complementarios. El primero estableció un entorno controlado para implementar los métodos, entrenar las redes y verificar la recuperación de fase a partir de superficies sintéticas de geometría sencilla. El segundo amplió la evaluación a imágenes generadas a partir de modelos tridimensionales complejos, con variaciones de orientación, visibilidad y respuesta radiométrica. Esta progresión permitió examinar el funcionamiento de las implementaciones antes de estudiar su comportamiento sobre escenas con discontinuidades y regiones de difícil recuperación.

La selección de los métodos se vincula con el primer objetivo específico; su implementación y verificación, con el segundo. El protocolo de precisión y costo computacional desarrolla el tercero, mientras que la aplicación de comparaciones bajo condiciones comunes proporciona la evidencia necesaria para responder al cuarto. Los resultados comparativos y su interpretación se presentan en el capítulo 9. En este capítulo, las salidas de ejemplo se utilizan para explicar las operaciones y sus verificaciones, sin anticipar una superioridad general de algún método.

La figura 10 presenta la relación entre ambos experimentos. Cada uno conserva sus propias particiones, modelos entrenados y referencias; la progresión entre ellos no implica mezclar sus conjuntos de prueba ni trasladar automáticamente sus métricas.

Figura 10

Organización de los dos experimentos



Nota. Se representan las etapas de implementación, entrenamiento, verificación y evaluación.

8.1 Primer experimento: recuperación de fase en condiciones simuladas controladas

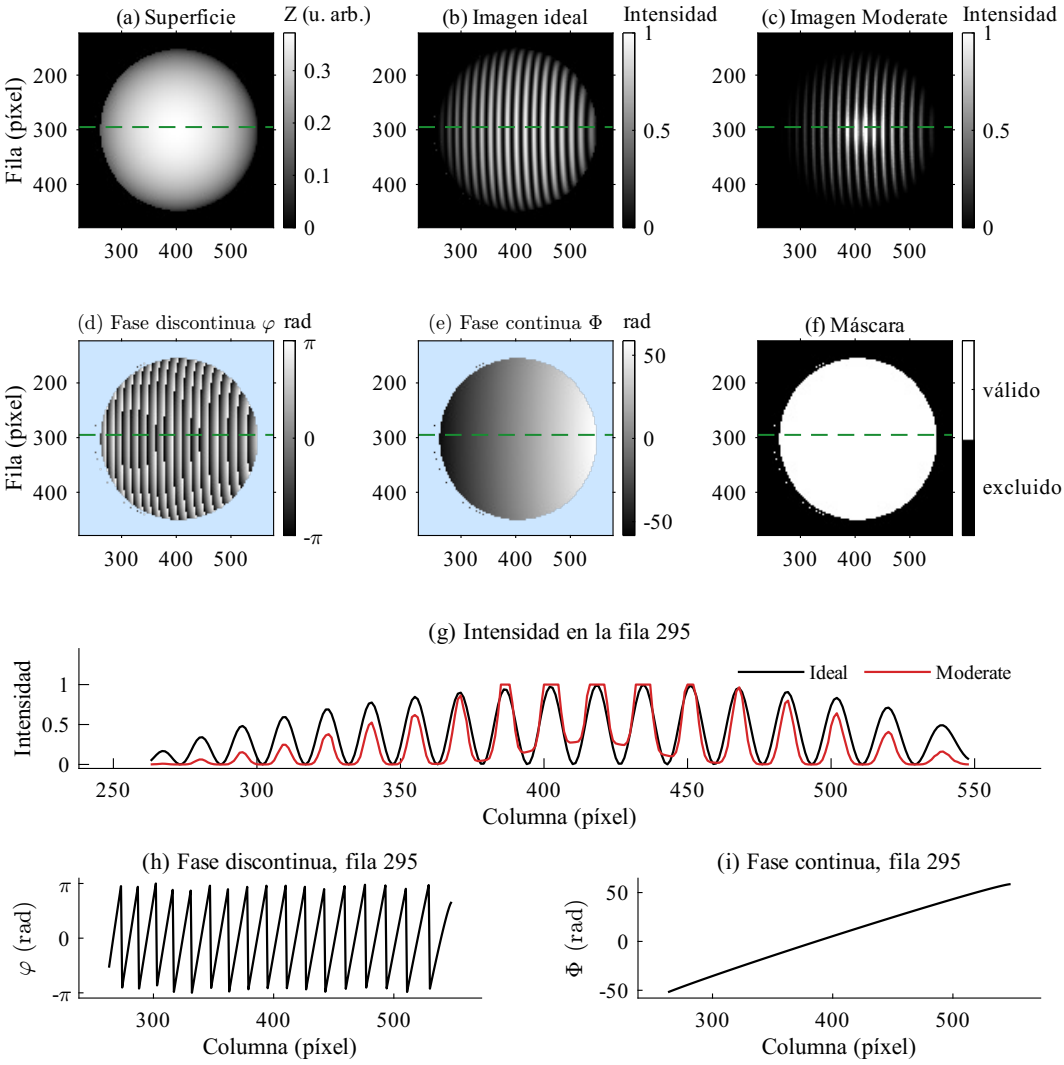
El primer experimento comparó los algoritmos de corrimiento de fase de tres, cuatro y doce imágenes, denominados PSA3, PSA4 y PSA12; el desenvolvimiento temporal lineal, denominado LTPU; y las arquitecturas U-Net y Dual CNN. Esta selección permite contrastar procedimientos que

utilizan distinta cantidad de información de entrada y que recuperan la fase mediante operaciones analíticas o mediante modelos aprendidos. Las redes se implementaron como adaptaciones de los enfoques de clasificación de fase y estimación de componentes en cuadratura descritos, respectivamente, por Hu et al. (2021) y Feng et al. (2019). La implementación utilizada en este trabajo se define a continuación; no se supone que sus condiciones de entrenamiento reproduzcan literalmente las de esas publicaciones.

Se generaron 800 escenas de resolución 600×800 píxeles, distribuidas en diez familias de superficies y 80 realizaciones por familia. Las geometrías comprendieron esfera, plano inclinado, pirámide, paraboloide, semicilindro, tronco de cono, protuberancia gaussiana, elipsoide, prisma y cono. Las variaciones geométricas de cada realización conservaron una referencia de fase conocida, de modo que una diferencia entre la estimación y la referencia pudiera atribuirse al procesamiento bajo las condiciones simuladas. La figura 11 ilustra cómo una superficie origina las imágenes y referencias empleadas en el experimento.

Figura 11

Formación de los datos del primer experimento



Nota. Corresponde a la escena de entrenamiento `Esfera_scene_001`: (a) superficie; (b) imagen en la condición Ideal; (c) imagen en la condición Moderate; (d) fase discontinua de referencia $\varphi(x,y)$; (e) fase continua de referencia $\Phi(x,y)$; (f) máscara de evaluación; (g) intensidad Ideal y Moderate en la fila 295; (h) φ y (i) Φ en la misma fila.

Para describir el generador se utilizan coordenadas normalizadas (x,y) y una superficie

$z(x,y)$. Las ecuaciones (8a) y (8b) establecen la fase base y la fase continua del patrón principal:

$$\Phi_{\text{base}}(x,y) = 2\pi x + \frac{15}{25} z(x,y), \tag{8a}$$

$$\Phi_{\text{ref}}(x,y) = 25\Phi_{\text{base}}(x,y) = 2\pi(25)x + 15z(x,y). \tag{8b}$$

Estas expresiones corresponden al modelo sintético implementado y no constituyen una calibración métrica de un sistema físico. En adelante, Φ denota fase continua, φ fase discontinua y el sombrero identifica una estimación. El subíndice ref identifica la referencia de evaluación. El operador $\mathscr{W}(\alpha) = \text{atan2}(\sin\alpha, \cos\alpha)$ representa el valor principal angular, por lo que la referencia discontinua es $\varphi_{\text{ref}} = \mathscr{W}(\Phi_{\text{ref}})$.

La intensidad ideal y los desplazamientos utilizados en la secuencia de doce imágenes se definen mediante las ecuaciones (9a) y (9b):

$$I_n^{\text{ideal}}(x,y) = A(x,y) + B(x,y)\cos\left[\Phi_{\text{ref}}(x,y) + \delta_n\right], \tag{9a}$$

$$\delta_n = -\frac{2\pi(n-1)}{12}, \qquad n = 1,\ldots,12. \tag{9b}$$

Aquí A es la intensidad de fondo y B la amplitud de modulación. En el generador ambas incorporan la respuesta lambertiana normalizada $r(x,y)$, con $A = B = 0.5r$. Se conserva el signo negativo de los desplazamientos registrado en el código, junto con el signo correspondiente del estimador. La máscara de evaluación se obtiene de la región del objeto y de una modulación ideal superior a 0.05; por tanto, utiliza información auxiliar del simulador y no representa un detector de validez autónomo aplicado a una captura real.

A partir de las imágenes ideales se incorporaron ruido gaussiano, no linealidad gamma y una contribución especular con expansión espacial y recorte. Las ecuaciones (10a), (10b) y (10c)

describen la secuencia aplicada dentro de la región válida:

$$I_n^{\text{ruido}} = \text{clip}(I_n^{\text{ideal}} + \eta_n, 0, 1), \quad \eta_n \sim \mathcal{N}(0, 0.01^2), \tag{10a}$$

$$I_n^{\gamma} = \left(I_n^{\text{ruido}}\right)^{2.6}, \tag{10b}$$

$$I_{n,\ell}^{\text{deg}} = \min\{I_n^{\gamma}(1 + g_\ell S_\ell) + o_\ell S_\ell, 1\}. \tag{10c}$$

Se ha omitido (x, y) para facilitar la lectura. El operador clip restringe las intensidades a $[0, 1]$; S_ℓ es el mapa especular normalizado después de su expansión con un núcleo gaussiano; y g_ℓ y o_ℓ controlan las contribuciones multiplicativa y aditiva. La tabla 1 presenta las configuraciones utilizadas. La condición Ideal conserva las imágenes anteriores a estas degradaciones.

Tabla 1
Parámetros de las condiciones radiométricas

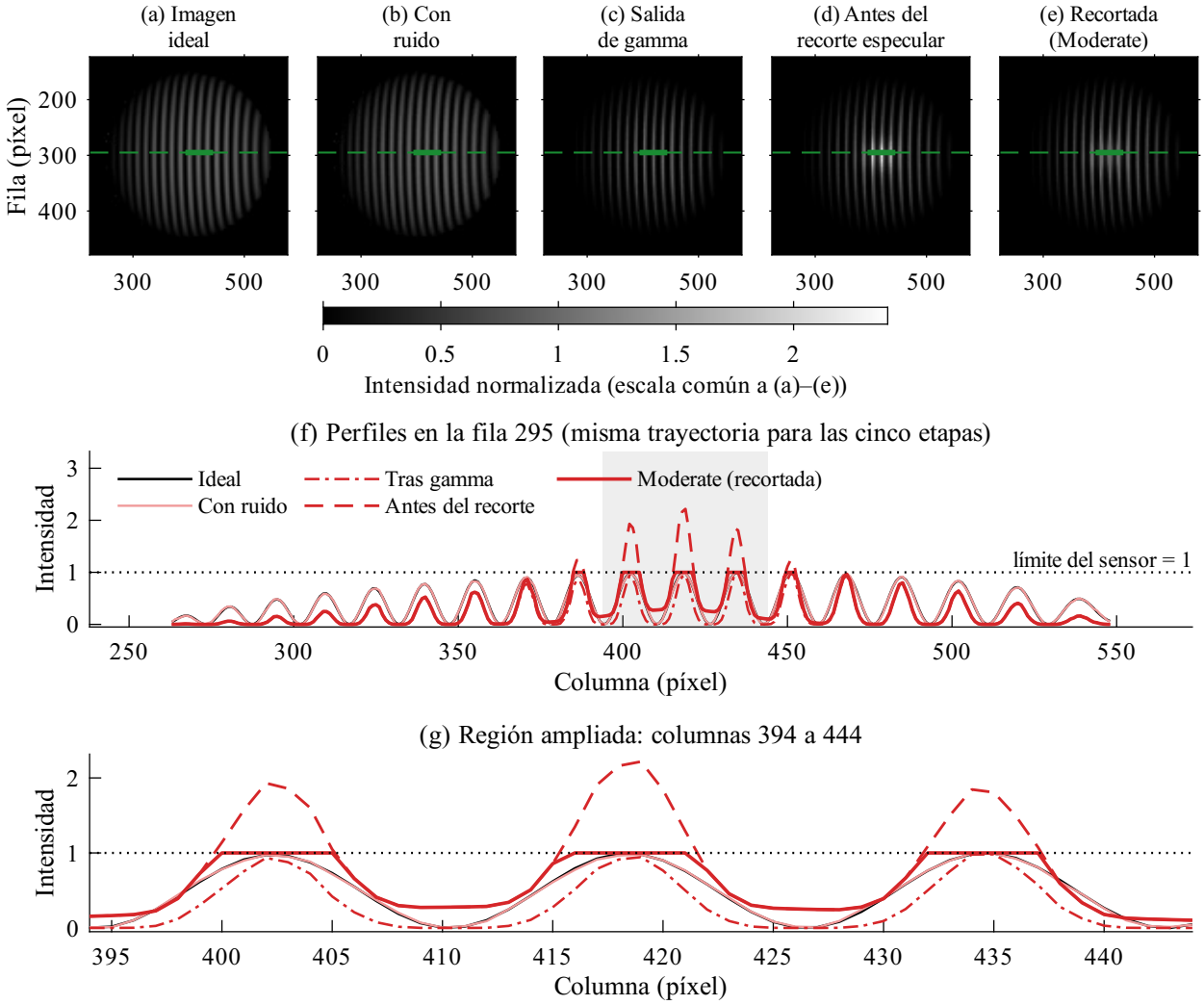
Condición	g_ℓ	o_ℓ	σ_ℓ
Medium	0.75	0.12	6
Moderate	1.10	0.30	9
VeryStrong	1.70	1.00	16

Nota. Las condiciones se utilizaron en ambos experimentos. La desviación espacial se expresa en píxeles.

El mapa especular se calcula a partir de las normales y las direcciones de iluminación y observación, con exponente 60 en el generador. Su expansión reproduce de manera fenomenológica la extensión de las regiones brillantes. Estas condiciones combinan varios mecanismos; no constituyen un diseño factorial que permita atribuir por separado los cambios de desempeño al ruido, a gamma o a la saturación. La figura 12 permite reconocer sus efectos sobre una misma señal antes de analizar el error de fase.

Figura 12

Transformaciones radiométricas hasta la condición Moderate



Nota. Corresponde a la escena de entrenamiento `Esfera_scene_001`: (a) imagen ideal; (b) con ruido gaussiano; (c) tras la no linealidad gamma; (d) antes del recorte especular; (e) imagen recortada (Moderate). (f) Perfiles de las cinco etapas en la fila 295, con el límite del sensor en 1; (g) ampliación de las columnas 394 a 444.

La partición se realizó por familia geométrica: esfera, plano inclinado, pirámide, paraboloide, semicilindro, tronco de cono y protuberancia gaussiana aportaron las 560 escenas de entrenamiento; el elipsoide aportó las 80 de validación; y el prisma y el cono, las 160 de prueba. Todas las condiciones radiométricas de una escena permanecieron en la misma partición. Las redes se

entrenaron únicamente en Moderate y sus parámetros se seleccionaron mediante validación. Los pesos seleccionados se mantuvieron fijos para evaluar Ideal, Medium, Moderate y VeryStrong.

Para PSA3, PSA4 y PSA12 se utilizaron las imágenes con índices $(1,5,9)$, $(1,4,7,10)$ y $(1,\ldots,12)$, respectivamente. Sea $\mathscr{J}_N$ el conjunto de índices correspondiente a $N \in \{3,4,12\}$. Las componentes y la fase estimada se expresan mediante las ecuaciones (11a), (11b) y (11c):

$$M_N = -\frac{2}{N} \sum_{n \in \mathscr{J}_N} I_n \sin \delta_n, \tag{11a}$$

$$D_N = \frac{2}{N} \sum_{n \in \mathscr{J}_N} I_n \cos \delta_n, \tag{11b}$$

$$\widehat{\varphi}_{\mathrm{PSA}N} = \operatorname{atan2}(M_N, D_N). \tag{11c}$$

El factor común $2/N$ no altera el ángulo y se omite en la función de recuperación del repositorio (Mayorga Gamboa et al., 2026). Bajo el modelo ideal, las componentes normalizadas representan $B\sin\Phi_{\mathrm{ref}}$ y $B\cos\Phi_{\mathrm{ref}}$. Esta relación también permite comprobar la consistencia entre las imágenes ideales y las etiquetas de entrenamiento.

LTPU se incluyó como referencia de desenvolvimiento temporal, cuyo principio consiste en resolver las discontinuidades a lo largo de una secuencia de fases en lugar de propagar la solución exclusivamente sobre la imagen (Huntley and Saldner, 1993). La implementación local emplea cinco frecuencias relativas $f = 1,\ldots,5$, con cuatro desplazamientos por frecuencia y un total de veinte imágenes. La fase ideal a cada frecuencia es $f\Phi_{\mathrm{base}}$. Una vez estimadas las fases discontinuas $\widehat{\varphi}_f$, las ecuaciones (12a) y (12b) describen la acumulación temporal implementada y su conversión a la escala del patrón principal:

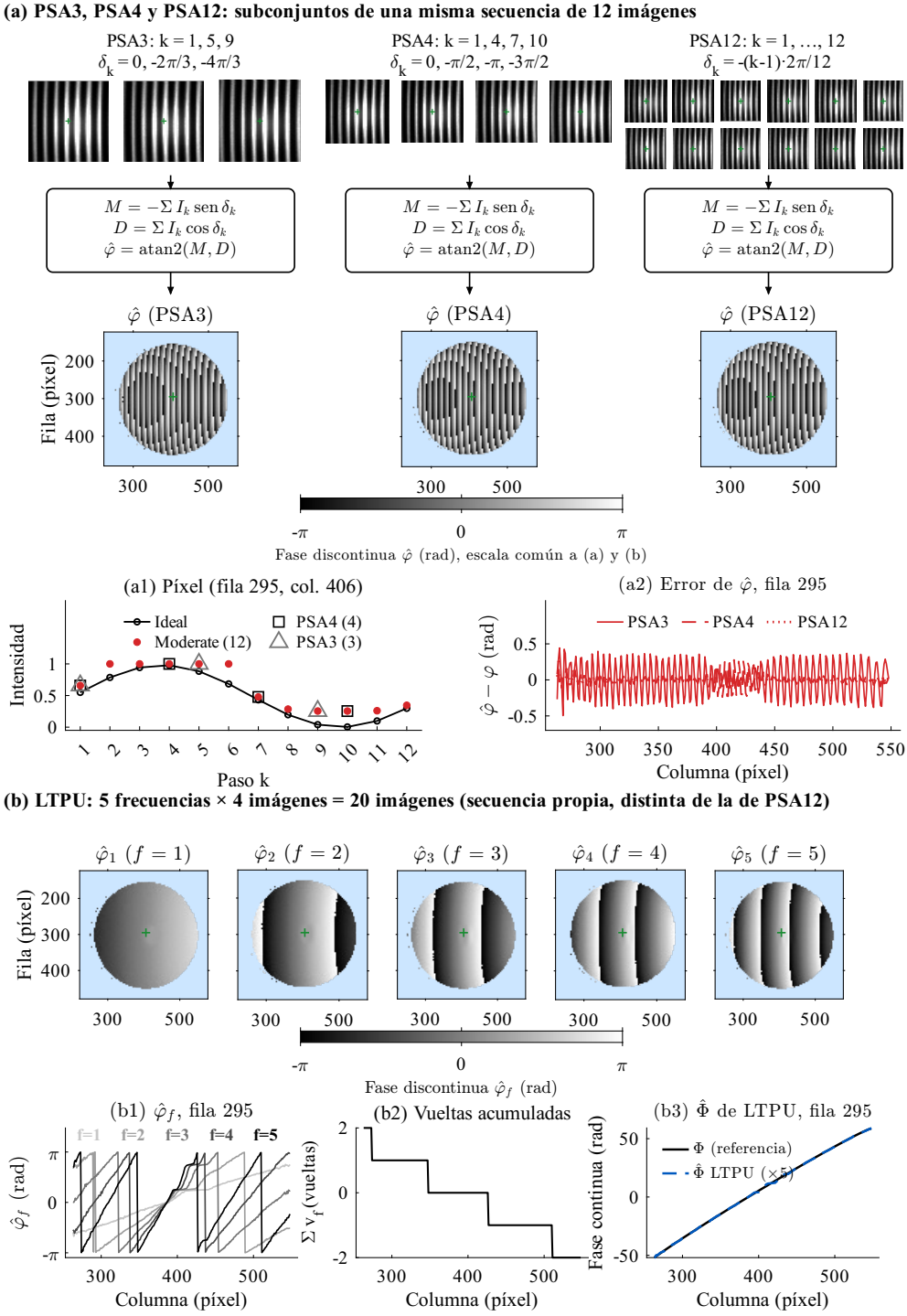
$$v_f = \operatorname{round}\left(\frac{\widehat{\varphi}_f - \widehat{\varphi}_{f-1}}{2\pi}\right), \quad f = 2,\ldots,5, \tag{12a}$$

$$\widehat{\Phi}_{\mathrm{LTPU}} = \frac{25}{5}\left(\widehat{\varphi}_5 - 2\pi \sum_{f=2}^{5} v_f\right). \tag{12b}$$

Esta formulación exige incrementos reales de fase menores que π en magnitud entre frecuencias consecutivas y una referencia inicial compatible con la fase principal. El factor $25/5$ hace comparable la salida con la referencia del Exp1; no se estima a partir del error observado. La figura 13 contrasta las entradas y operaciones de PSA3, PSA4, PSA12 y LTPU.

Figura 13

Entradas y operaciones de los métodos clásicos



Nota. Métodos evaluados en el Exp1. El esquema corresponde a las implementaciones desarrolladas en el proyecto.

Las dos arquitecturas neuronales reciben únicamente la primera imagen de la secuencia, correspondiente a desplazamiento cero. En los archivos preparados esta imagen se identifica como `I2`, aunque coincide con `I01`; esta denominación no indica que la red utilice dos imágenes. U-Net transforma la entrada en 256 puntuaciones por píxel. La codificación de la referencia y la conversión de la clase estimada a fase se presentan en las ecuaciones (13a) y (13b):

$$c_{\mathrm{ref}} = \mathrm{round}\left[255\frac{\varphi_{\mathrm{ref}}+\pi}{2\pi}\right], \tag{13a}$$

$$\widehat{\varphi}_{\mathrm{U}} = \frac{2\pi}{255}\underset{c\in\{0,\ldots,255\}}{\arg\max}\ a_c - \pi, \tag{13b}$$

donde a_c es la puntuación de la clase c. La arquitectura implementada utiliza un codificador con 64, 128, 256 y 512 canales, un nivel central de 1024 canales y un decodificador con conexiones entre niveles de igual resolución. La pérdida corresponde a entropía cruzada promediada dentro de la máscara. La representación introduce una discretización angular que debe distinguirse del error de aprendizaje; su paso no constituye, por sí solo, un límite universal del RMSE.

Dual CNN emplea dos redes sucesivas. La primera estima el fondo y la segunda utiliza ese fondo estimado junto con la imagen para recuperar las componentes de fase. Las ecuaciones (14a), (14b) y (14c) describen este recorrido:

$$\widehat{A} = f_{\theta_1}(I_1), \tag{14a}$$

$$(\widehat{M},\widehat{D}) = g_{\theta_2}(I_1,\widehat{A}), \tag{14b}$$

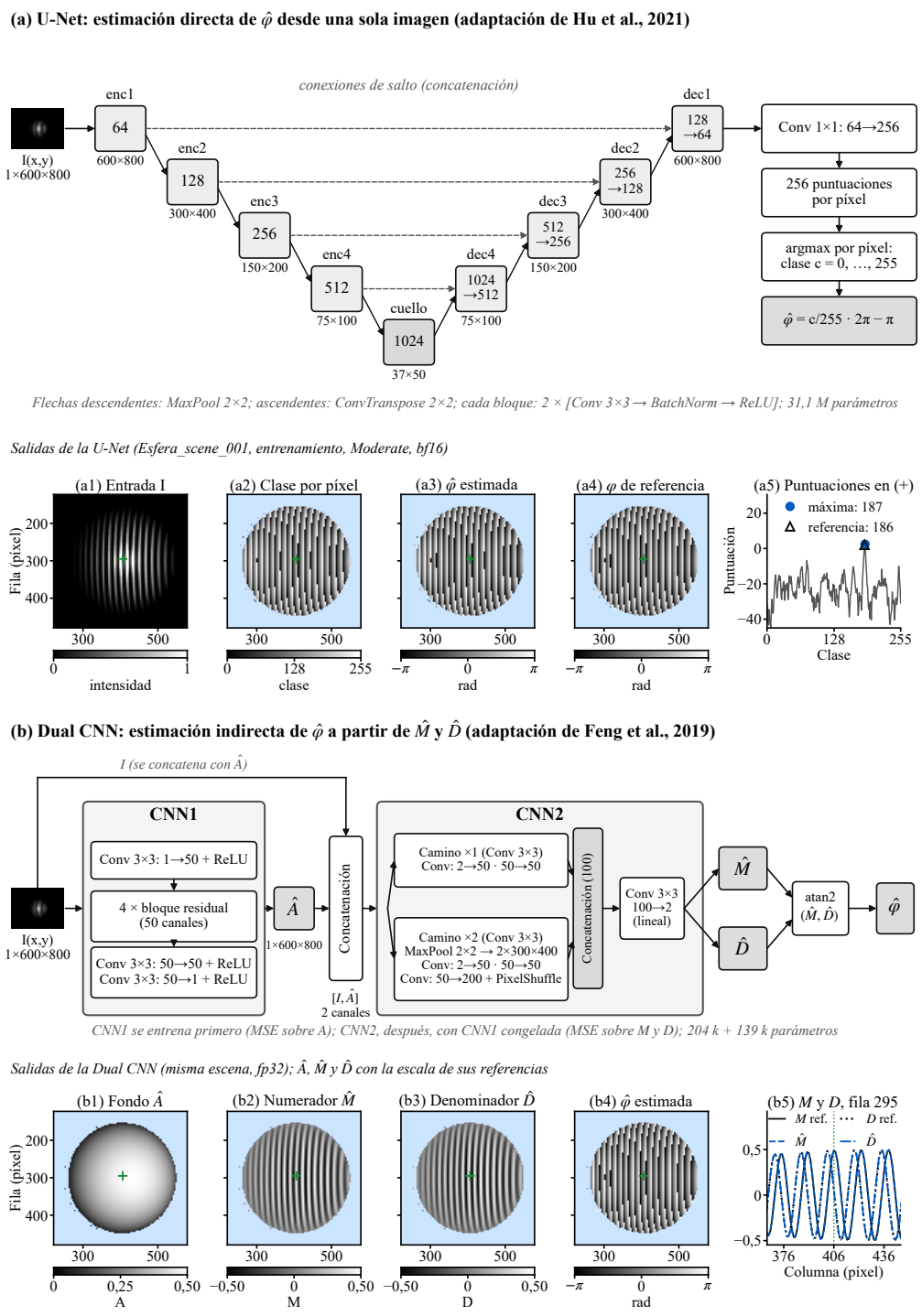
$$\widehat{\varphi}_{\mathrm{D}} = \mathrm{atan2}(\widehat{M},\widehat{D}). \tag{14c}$$

La primera red contiene convoluciones y cuatro bloques residuales; la segunda incorpora dos caminos de procesamiento a distinta escala y una salida lineal de dos canales. Ambas utilizan 50 canales en sus bloques internos. Primero se ajustó CNN1 mediante el error cuadrático del fondo; después se congelaron sus mejores pesos y se entrenó CNN2 con la suma de los errores cuadráticos

de M y D. Las pérdidas se restringieron a la máscara válida. La figura 14 muestra las diferencias entre la clasificación de U-Net y la estimación de componentes de Dual CNN.

Figura 14

Representaciones de entrada y salida de las redes



Nota. Las redes recuperan la fase discontinua. El esquema corresponde a las adaptaciones implementadas en el proyecto.

La corrida definitiva del Exp1 fijó la semilla 42 para el entrenamiento y activó las opciones deterministas disponibles. Esta corrida se conserva separada de los entrenamientos históricos sin semilla. En ambos modelos se utilizó Adam y se seleccionaron los pesos con menor pérdida de validación. La tabla 2 reúne los parámetros principales; la reducción de la tasa de aprendizaje se aplicó con factor 0.5 después de diez épocas sin mejora de validación, y no se empleó aumento de datos.

Tabla 2
Configuración principal de entrenamiento

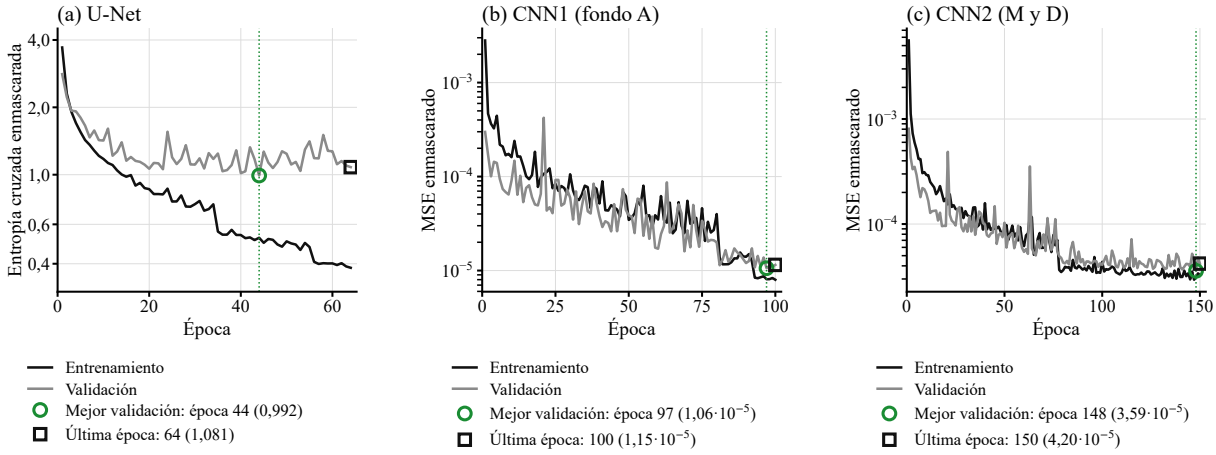
Parámetro	U-Net	CNN1	CNN2
Tasa inicial	10^{-4}	10^{-3}	10^{-3}
Tamaño de lote	1	2	2
Máximo de épocas	150	100	150
Paciencia de parada	20	15	15
Precisión de entrenamiento	bf16	fp32	fp32
Semilla	42	42	42

Nota. La configuración se aplicó por separado en cada experimento. Las épocas son límites máximos, no duraciones efectivas de las corridas.

Los historiales de entrenamiento y validación permiten comprobar la selección de los pesos y detectar divergencias o diferencias persistentes entre ambas pérdidas. La figura 15 ilustra esta verificación. Las curvas no sustituyen la evaluación sobre escenas de prueba ni demuestran, por sí solas, generalización a otras condiciones radiométricas.

Figura 15

Pérdidas de entrenamiento y validación del Exp1



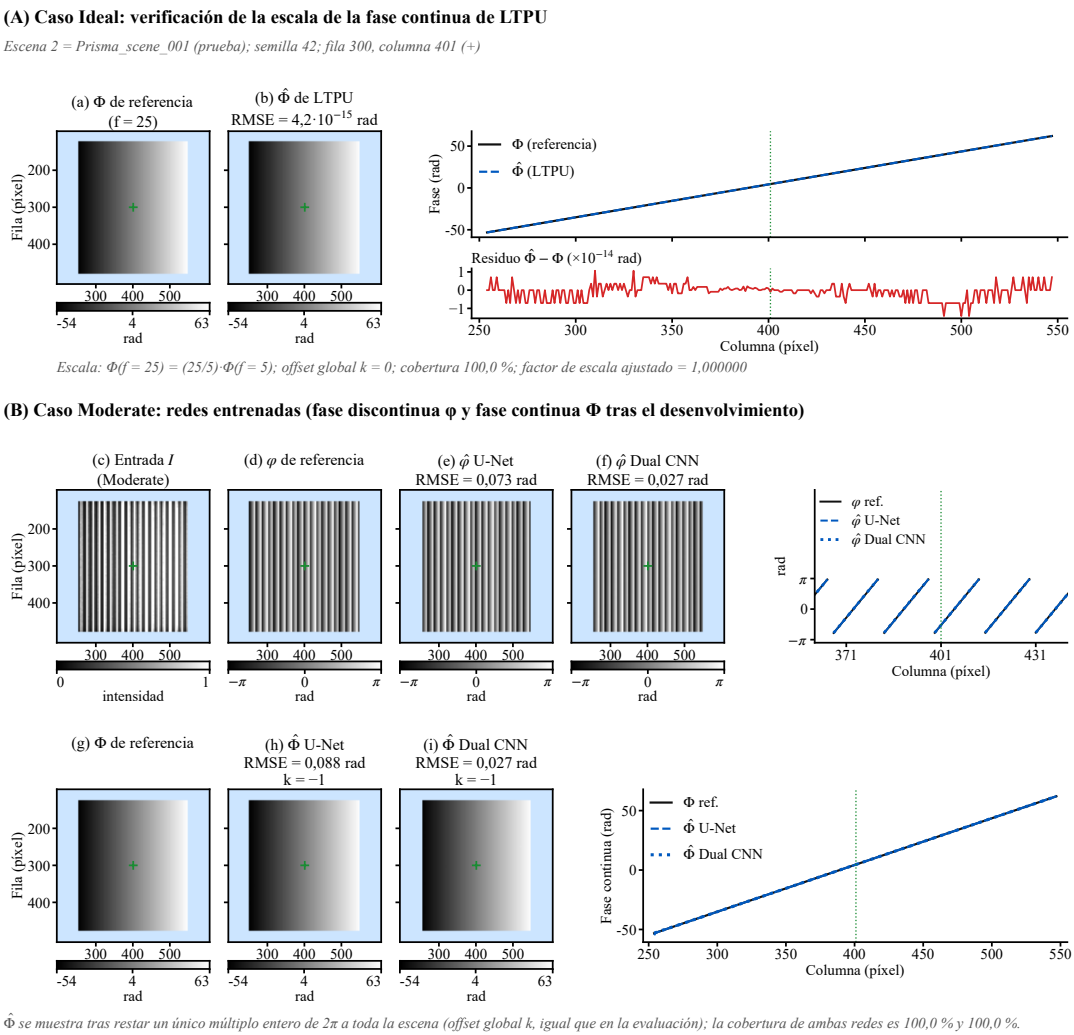
Nota. Corresponde a la corrida definitiva con semilla 42. U-Net se entrenó durante 64 épocas, por parada temprana, y se guardaron los pesos de la época 44; CNN1, durante 100 épocas, con los pesos de la época 97; y CNN2, durante 150 épocas, con los pesos de la época 148. En los tres casos, los pesos guardados corresponden a la menor pérdida de validación.

La verificación de las implementaciones comprendió la consistencia dimensional de entradas, máscaras y salidas, la recuperación de una fase conocida en condiciones ideales y la correspondencia entre la escala de LTPU y la referencia principal. Para PSA3, PSA4, PSA12 y las redes se aplicó el mismo desenvolvimiento espacial clásico con máscara, mientras que LTPU produjo directamente la fase continua mediante su secuencia temporal. La evaluación continua del Exp1 admite un único desplazamiento entero global de 2π por escena para fijar el origen. Este ajuste utiliza la referencia durante la evaluación, no durante la recuperación espacial, y no corrige órdenes de manera independiente en cada píxel.

La figura 16 ilustra las salidas de esta primera prueba experimental. La recuperación de una fase conocida y la disponibilidad de modelos entrenados y salidas consistentes proporcionan la base operativa para continuar al Exp2; la magnitud de las mejoras o deterioros bajo degradación se analiza en el capítulo de resultados.

Figura 16

Verificación de las salidas de fase del Exp1



La continuidad de los métodos se delimitó considerando conjuntamente precisión, cantidad de imágenes y costo de procesamiento. LTPU se mantuvo como referencia temporal en el Exp1, pero no se incorporó al Exp2. La decisión considera su necesidad de veinte imágenes y el mayor tiempo y memoria CPU observados frente a PSA4 en las mediciones del proyecto; no implica que supere en costo a todos los métodos bajo cualquier alcance. En particular, el tiempo hasta fase discontinua y el tiempo del procedimiento hasta fase continua son mediciones diferentes. La

comparación que sustenta esta delimitación se desarrolla en el capítulo 9, conservando también las condiciones en que LTPU ofrece ventajas. Adicionalmente, el conjunto del Exp2 no dispone de la secuencia multifrecuencia requerida. Esta delimitación se explicita como decisión sustentada en la evidencia del primer experimento, sin atribuirle retrospectivamente un carácter de criterio previo.

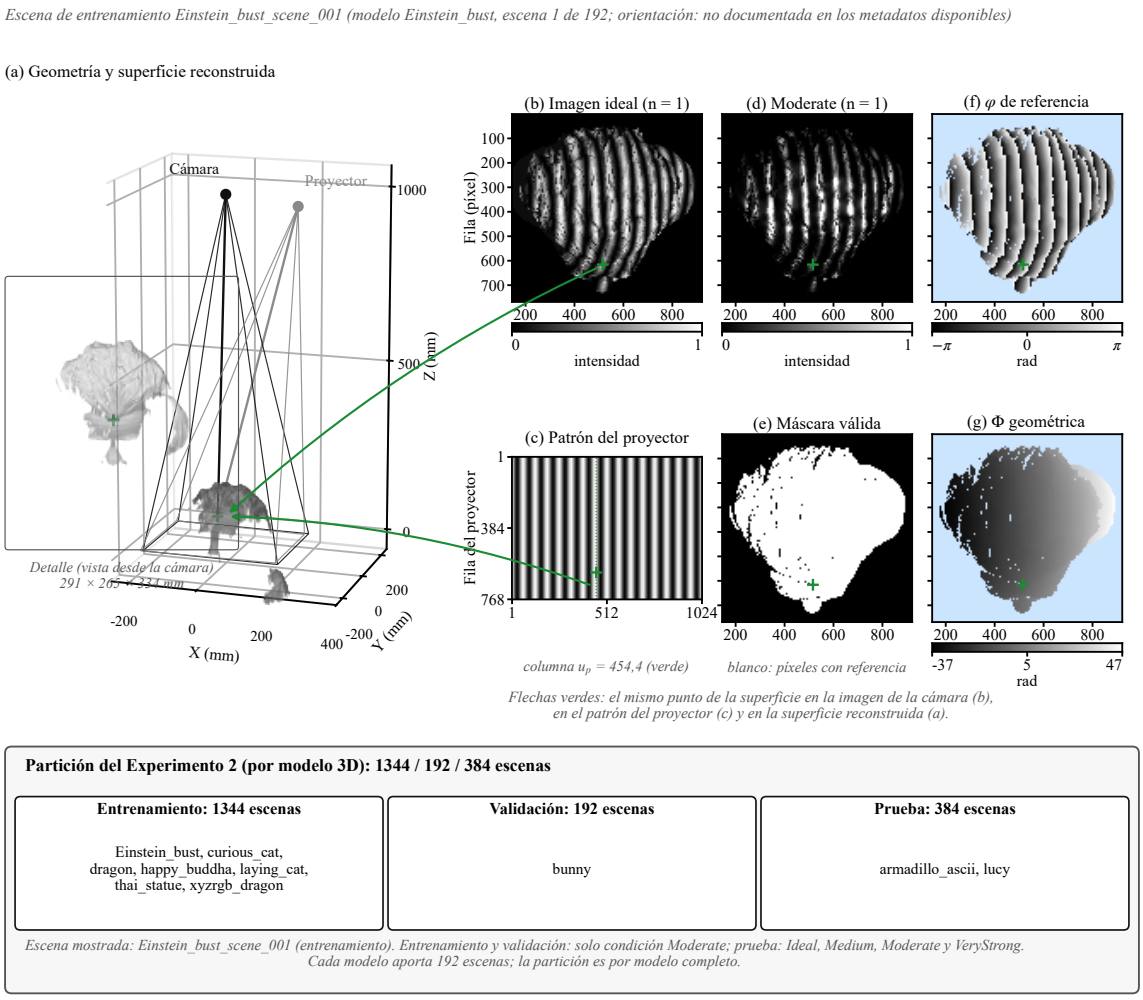
8.2 Segundo experimento: recuperación de fase sobre geometrías tridimensionales complejas

El segundo experimento conservó PSA3, PSA4, PSA12, U-Net y Dual CNN y los evaluó sobre imágenes producidas con `FringePointCloudSimulator`, suministrado por el director. El simulador forma parte del entorno de investigación del que procede el conjunto de datos; su desarrollo no se presenta como una contribución de los estudiantes. A diferencia de las superficies analíticas del Exp1, las mallas tridimensionales incorporan detalles, concavidades y cambios de visibilidad que pueden interrumpir la continuidad espacial de la fase.

El conjunto contiene 1920 escenas de resolución 768×1024 píxeles, correspondientes a diez modelos y 192 orientaciones por modelo. La figura 17 muestra el recorrido desde la malla hasta las imágenes y referencias. La unidad de partición fue el modelo completo: siete modelos aportaron 1344 escenas de entrenamiento, Bunny aportó 192 de validación y Armadillo y Lucy aportaron las 384 escenas de prueba. Los modelos de entrenamiento fueron Einstein Bust, Curious Cat, Dragon, Happy Buddha, Laying Cat, Thai Statue y XYZRGB Dragon. Esta distribución es la consignada en los manifiestos del trabajo de pregrado.

Figura 17

Formación de una escena del segundo experimento



Nota. Se muestra la correspondencia entre imágenes, máscara y referencias de fase. Los datos provienen del simulador suministrado por el director. La superficie tridimensional se presenta como ilustración del entorno simulado; no constituye una validación de reconstrucción métrica mediante los métodos comparados.

Cada escena conserva las doce imágenes de una secuencia de corrimiento de fase y las representaciones ideales utilizadas para generar las etiquetas. Las redes se entrenaron nuevamente con los datos del Exp2, manteniendo la condición Moderate y la selección por pérdida de validación descritas anteriormente. La evaluación se realizó con los pesos propios de este experimento, conge-

lados para las cuatro condiciones Ideal, Medium, Moderate y VeryStrong. No se utilizaron modelos diferentes para cada número de pasos de PSA ni se reentrenó una red por nivel radiométrico.

La referencia continua del Exp2 se obtiene de la correspondencia geométrica con el proyector. Si $u_p(x,y)$ es su coordenada horizontal, $u_{0,p}$ el origen de fase y $T_p = 64$ píxeles el período del patrón, las ecuaciones (15a) y (15b) definen las referencias continua y discontinua:

$$\Phi_{\text{geom}}(x,y) = \frac{2\pi}{T_p}\left[u_p(x,y) - u_{0,p}\right], \tag{15a}$$

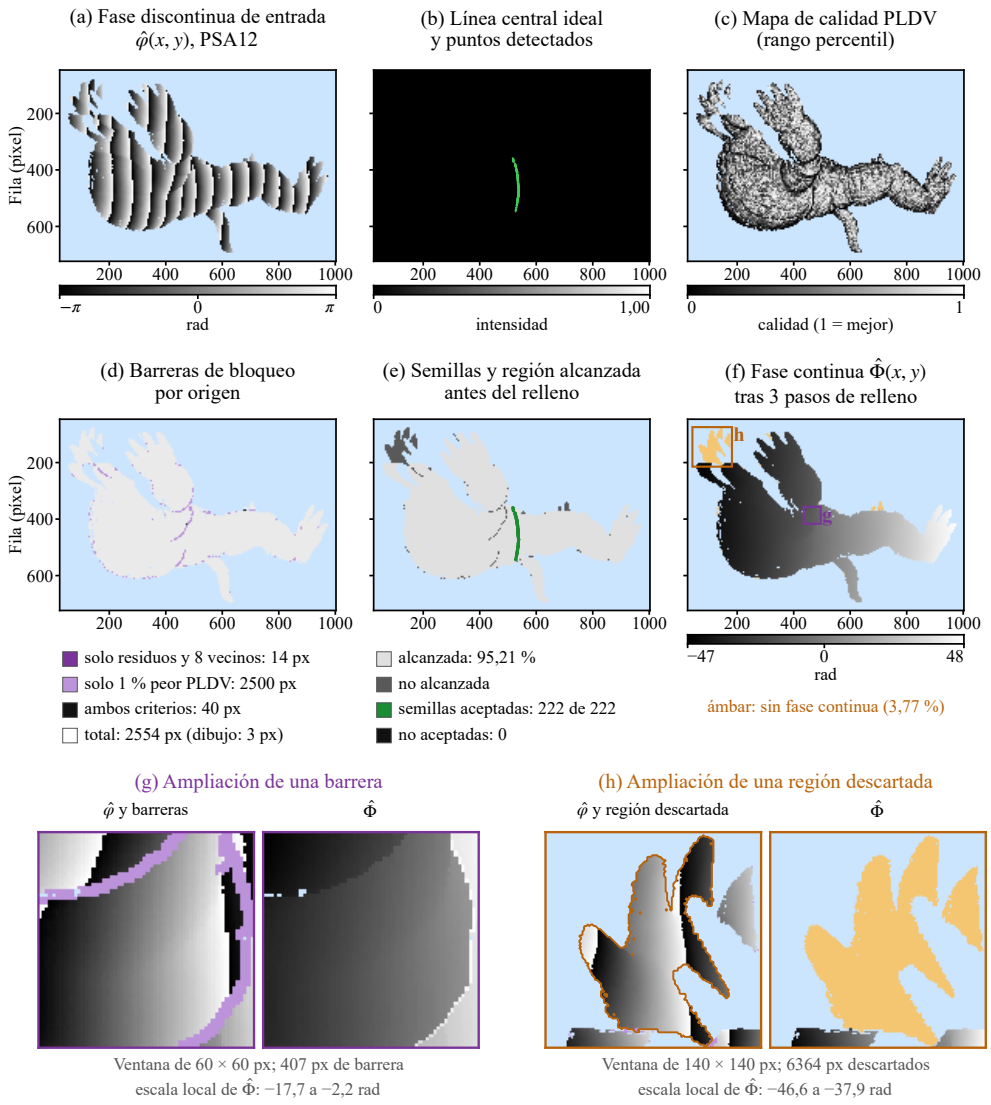
$$\varphi_{\text{ref}}(x,y) = \mathscr{W}\left(\Phi_{\text{geom}}(x,y)\right). \tag{15b}$$

La fase geométrica se conserva como referencia independiente del desenvolvimiento. Por ello, no se reemplaza por una fase continua calculada mediante un algoritmo que pueda cometer errores de orden. Las intensidades ideales y la máscara auxiliar también proceden del simulador; esta disponibilidad de información debe considerarse al interpretar el alcance de la comparación.

La recuperación de fase discontinua conserva las operaciones de los métodos descritas para el Exp1. El cambio principal se encuentra en el desenvolvimiento espacial: para el Exp2 se utilizó una estrategia que construye barreras en regiones con discontinuidades y propaga la fase desde puntos asociados con la línea central del proyector. La figura 18 presenta esta secuencia y distingue los píxeles recuperados de los descartados.

Figura 18

Etapas del desenvolvimiento espacial



Mapas en escala de grises; exterior de la máscara en azul claro; ámbar: dentro de la máscara sin fase continua. Escala de $\hat{\varphi}$: de $-\pi$ a π; de $\hat{\Phi}$ en (f): de −47 a 48 rad; en (g) y (h), escala local. Barreras: residuos con sus 8 vecinos ∪ 1 % de peor PLDV. Propagación con 4 vecinos desde semillas con $|\hat{\varphi}| \leq 0{,}5$ rad; relleno de barrera hasta 3 pasos. Error de orden (diagnóstico con la referencia geométrica, no usada en el método): 0,068 % de los píxeles recuperados.

Nota. Variante con descarte de discontinuidades. Corresponde al Exp2, para la escena de prueba `armadillo_ascii_scene_069`, condición Moderate y fase discontinua $\widehat{\varphi}(x,y)$ estimada con PSA12. (a) Fase de entrada; (b) línea central ideal y puntos detectados; (c) mapa de calidad PLDV; (d) barreras de bloqueo según su origen (residuos con sus vecinos y 1 % de peor calidad); (e) semillas aceptadas y región alcanzada antes del relleno; (f) fase continua estimada $\widehat{\Phi}(x,y)$ tras tres pasos de relleno. En esta escena los 222 puntos detectados fueron aceptados como semillas. (g) Ampliación de una barrera y (h) de una región descartada, con escala local de $\widehat{\Phi}$. La cobertura pasa de 95,21 % a 96,23 % y el error de orden es 0,068 % de los píxeles cubiertos, medido con la referencia geométrica solo como diagnóstico.

La implementación combina residuos de fase dilatados con sus ocho vecinos y el umbral correspondiente al uno por ciento de peor calidad local. Las barreras se conservan con su espesor; el desenvolvimiento utiliza conectividad de cuatro vecinos y no atraviesa estas regiones. Los píxeles válidos de la línea central que admite el procedimiento permiten iniciar la propagación, mientras que las regiones no alcanzadas permanecen sin recuperar. Finalmente se permiten hasta tres pasos de relleno de píxeles de barrera a partir de vecinos cuya fase continua ya fue recuperada. Este procedimiento se documenta como implementación del proyecto, con decisiones propias, y no como reproducción exacta de un algoritmo publicado.

La imagen auxiliar de línea central se utiliza para establecer el origen de fase. En la evaluación de prueba se verifica que su intensidad máxima dentro de la máscara alcance el umbral 50 sobre la escala de 0 a 255. Si no existe una referencia admisible, la cobertura continua se registra como cero y los errores continuos quedan sin definir. La presencia de línea no garantiza que todas las regiones tengan semillas válidas ni que el desenvolvimiento sea correcto. La referencia geométrica no interviene en la propagación; se consulta después para medir el error absoluto, sin aplicar el ajuste global permitido en el Exp1.

Además de la comparación de cinco métodos sobre las 384 escenas de prueba y cuatro condiciones, se conserva una evaluación específica del desenvolvimiento clásico guiado y del desenvolvimiento con descarte sobre las 1920 escenas, únicamente en Moderate. Esta segunda evaluación abarca entrenamiento, validación y prueba y se identifica como análisis completo del conjunto, no como evaluación exclusiva de prueba. La auditoría posterior de la línea central se conserva separada de ambas corridas, con sus criterios y registros por escena. Ninguna escena se excluye únicamente por presentar un error elevado.

Las inspecciones exploratorias previas incluyeron escenas de prueba; por ello, la evaluación final no se describe como una prueba completamente ciega respecto de todas las decisiones de desarrollo. La selección de pesos usa validación, pero esa separación no elimina el carácter exploratorio de las inspecciones del desenvolvimiento. Los diagnósticos que alinean regiones con la fase ideal

se identifican expresamente como análisis con verdad de referencia y no sustituyen las métricas absolutas originales.

8.3 Evaluación cuantitativa y análisis comparativo

La escena constituye la unidad de registro y el bloque de las comparaciones entre métodos. Para una misma escena y condición se conserva la referencia y la máscara de evaluación, mientras que cada procedimiento produce su estimación y su máscara recuperada. Se distingue el error angular de la fase discontinua del error lineal de la fase continua. Esta separación evita que un salto de representación entre $-\pi$ y π se interprete como un error físico de aproximadamente una vuelta.

Sea Ω la región válida de evaluación de una escena y m el método. El error circular y sus dos medidas principales se definen mediante las ecuaciones (16a), (16b) y (16c):

$$e_{c,m}(x,y) = \mathscr{W}\left[\widehat{\varphi}_m(x,y) - \varphi_{\mathrm{ref}}(x,y)\right], \tag{16a}$$

$$\mathrm{RMSE}_{c,m} = \sqrt{\frac{1}{|\Omega|}\sum_{(x,y)\in\Omega} e_{c,m}(x,y)^2}, \tag{16b}$$

$$\mathrm{MAE}_{c,m} = \frac{1}{|\Omega|}\sum_{(x,y)\in\Omega} |e_{c,m}(x,y)|. \tag{16c}$$

Ambas medidas se expresan en radianes. La primera pondera con mayor intensidad las desviaciones grandes; la segunda resume la magnitud media del error. Los indicadores se calculan primero por escena, sin sustituirlos por una única acumulación de todos los píxeles del conjunto.

Para fase continua se utiliza la región recuperada $\Omega_m \subseteq \Omega$. Las ecuaciones (17a), (17b)

y (17c) explicitan la diferencia de protocolo entre experimentos y el cálculo del error:

$$k_m = \begin{cases} \operatorname{round}\left(\dfrac{\operatorname{mediana}_{\Omega_m}(\widehat{\Phi}_m - \Phi_{\text{ref}})}{2\pi}\right), & \text{Exp1}, \\ 0, & \text{Exp2}, \end{cases} \tag{17a}$$

$$e_{\Phi,m} = \widehat{\Phi}_m - 2\pi k_m - \Phi_{\text{ref}}, \tag{17b}$$

$$\text{RMSE}_{\Phi,m} = \sqrt{\frac{1}{|\Omega_m|}\sum_{(x,y)\in\Omega_m} e_{\Phi,m}(x,y)^2}. \tag{17c}$$

El MAE continuo se calcula promediando $|e_{\Phi,m}|$ sobre la misma región. En el Exp2 se toma $\Phi_{\text{ref}} = \Phi_{\text{geom}}$. Si Ω_m está vacía, los errores continuos quedan sin definir; no se les asigna cero. El ajuste global del Exp1 debe acompañar su interpretación y evita presentar sus errores como equivalentes a los errores absolutos sin ajuste del Exp2.

La cobertura cuantifica la parte de la máscara válida del objeto que fue recuperada, no la fracción de toda la imagen. Para el diagnóstico de orden del Exp2 se considera error cuando la desviación absoluta supera π. Las ecuaciones (18a), (18b) y (18c) definen cobertura recuperada, proporción de error entre píxeles recuperados y cobertura correcta respecto a ese criterio:

$$C_m = 100\frac{|\Omega_m|}{|\Omega|}, \tag{18a}$$

$$E_{\text{orden},m} = 100\frac{|\{(x,y)\in\Omega_m : |e_{\Phi,m}| > \pi\}|}{|\Omega_m|}, \tag{18b}$$

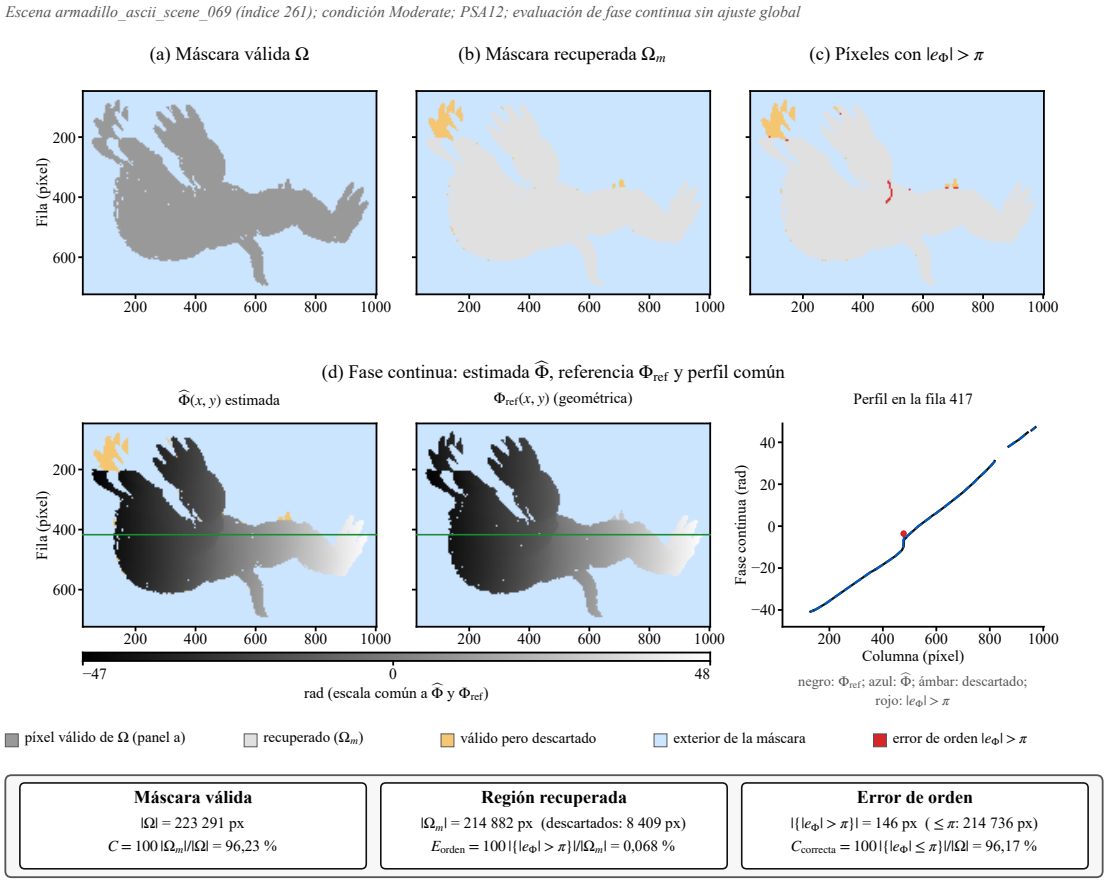
$$C_{\text{correcta},m} = 100\frac{|\{(x,y)\in\Omega_m : |e_{\Phi,m}| \leq \pi\}|}{|\Omega|}. \tag{18c}$$

La expresión «correcta» se refiere aquí al criterio de orden; no garantiza un error residual pequeño ni una exactitud métrica determinada. Una cobertura elevada puede coexistir con numerosos errores, mientras que un error bajo sobre una región pequeña puede ocultar una pérdida importante de información. Por ello, las tres cantidades se interpretan junto con el RMSE y el MAE. Los errores

continuos de los archivos originales se calculan sobre la región recuperada por cada método: si las regiones difieren, también cambia su soporte de evaluación y no debe afirmarse que se compararon exactamente los mismos píxeles. La figura 19 ilustra esta distinción.

Figura 19

Máscara válida, región recuperada y píxeles con error de orden



Nota. Corresponde a la evaluación continua del Exp2, para la escena de prueba `armadillo_ascii_scene_069`, condición Moderate y PSA12 con descarte de discontinuidades. (a) Máscara válida Ω, con $|\Omega| = 223\,291$ píxeles; (b) región recuperada Ω_m, con $|\Omega_m| = 214\,882$ píxeles, y píxeles válidos descartados en ámbar; (c) píxeles recuperados con $|e_\Phi| > \pi$, que son 146 y se dibujan ensanchados para hacerlos visibles; (d) fase continua estimada $\widehat{\Phi}$ y de referencia Φ_{ref} con escala común, y perfil de ambas en la fila 417. Con estos conteos, la cobertura recuperada es $C = 96{,}23\,\%$, el error de orden es $E_{\text{orden}} = 0{,}068\,\%$ y la cobertura correcta es $C_{\text{correcta}} = 96{,}17\,\%$ (ecuaciones (18a) a (18c)).

La similaridad estructural se incorporó como indicador complementario, tomando como antecedente el índice SSIM de Wang et al. (2004). Su aplicación a fase debe declarar la representación

y el soporte espacial. El Exp1 utiliza una adaptación propia: calcula SSIM sobre coseno y seno de la fase y promedia ambos valores. Se emplean ventanas gaussianas de 11×11 píxeles, desviación 1.5, rango de datos 2 y constantes $K_1 = 0.01$ y $K_2 = 0.03$. Solo se promedian ventanas completamente contenidas en la máscara, mediante su erosión de cinco píxeles.

El Exp2 conserva el indicador global implementado sobre la fase discontinua normalizada a $[0, 1]$: utiliza una única ventana formada por los píxeles de la máscara, covarianza poblacional, rango 1 y las mismas constantes K_1 y K_2. Esta definición no elimina la discontinuidad angular en el límite del intervalo. En consecuencia, sus valores no se comparan numéricamente con los del SSIM coseno–seno del Exp1 y se interpretan junto con los errores circulares. Se conserva esta diferencia para documentar los resultados efectivamente calculados.

El costo se estudió separando cantidad de imágenes, tiempo, memoria y energía. La cantidad de imágenes corresponde a las entradas del método; no representa una medición del tiempo físico de adquisición. Los ensayos se realizaron en un equipo con procesador Intel Core i9-9900 y GPU NVIDIA GeForce RTX 4060 Ti de 8 GB, utilizando MATLAB R2026a y Python 3.11.9 con PyTorch 2.6.0 y CUDA 12.4. El procesamiento clásico se ejecutó en CPU y la inferencia neuronal en GPU, por lo que los tiempos describen esas implementaciones y plataformas concretas.

La primera medición temporal registró el algoritmo sin lectura o escritura de archivos ni carga de pesos. Se realizaron diez calentamientos y cinco repeticiones por escena en Moderate, con lote unitario, orden reproducible y sincronización de GPU. La decodificación de U-Net y el cálculo de atan2 de Dual CNN forman parte de la obtención de la fase discontinua. LTPU, por su naturaleza, ya produce una fase continua; esta diferencia debe mantenerse visible al interpretar los tiempos de esta primera medición.

Una medición posterior consideró el procedimiento hasta fase continua con un único cronómetro por pasada. Para las redes incluyó transferencias CPU–GPU, inferencia, retorno a CPU, conversiones de arreglos y llamada al motor de MATLAB. En el Exp1 incluyó el desenvolvimiento clásico y en el Exp2 la detección de línea y el desenvolvimiento con descarte. La lectura

de la imagen auxiliar de línea quedó incluida en el Exp2; se excluyeron la lectura inicial de imágenes y máscaras, carga de pesos, arranque del motor, escritura de salidas e intercambio mediante archivos del flujo de producción. El tiempo total se resume a partir de las pasadas completas, no sumando medianas de etapas.

La memoria se midió en ejecuciones separadas del cronometraje. Para los métodos clásicos se registró el máximo informado por el perfilador de MATLAB para la función objetivo, con las entradas previamente cargadas. Para las redes se registró el pico incremental de memoria GPU asignada por PyTorch sobre una línea base con modelo y entrada residentes. Cada método se midió con diez calentamientos y cinco repeticiones, conservando el máximo. Los indicadores CPU y GPU tienen alcances diferentes y no permiten ordenar conjuntamente los métodos como si fueran una misma medida de memoria total.

La energía del procedimiento completo se midió en 24 escenas de Moderate por experimento, con tres repeticiones independientes y orden de métodos rotado. Se utilizó la potencia del paquete CPU registrada con HWiNFO cada dos segundos y el contador acumulado de energía GPU de NVML, descontando una línea base de reposo. En el Exp1 se repitieron ciclos completos hasta alcanzar bloques de al menos 120 segundos; en el Exp2 se ejecutaron seis ciclos de las 24 escenas por bloque. Para los clásicos del Exp1, la energía GPU se fijó en cero y su actividad se conservó como control. Las mediciones informan energía neta por pasada de CPU y GPU, no consumo eléctrico total en el enchufe ni energía de entrenamiento. Se informa la variación entre las tres repeticiones y la sensibilidad a la elección de la línea base. Los ciclos y subventanas internos no se tratan como nuevas repeticiones independientes.

La precisión numérica se registra para cada medición. En este trabajo, bf16 designa la inferencia con *autocast* de PyTorch en bfloat16, con los pesos almacenados en fp32 y la conversión de la salida a fase realizada en fp32, y fp32 designa la inferencia sin *autocast*. Las fases evaluadas provienen de la U-Net en bf16 y de la Dual CNN en fp32. El costo de cada red se midió en las dos precisiones; por tanto, la Dual CNN en bf16 es una configuración de comparación de costo y no la

que produjo las métricas de precisión, y el costo en fp32 de la U-Net no se atribuye automáticamente a su ejecución en bf16. Las variantes medidas se identifican por separado. Esta distinción permite valorar la relación entre calidad y costo sin combinar una precisión obtenida bajo una configuración con recursos medidos bajo otra.

El costo de las redes en las dos precisiones se informa con dos alcances que no se mezclan. El primero mide solo la inferencia de la red, con el alcance de la primera medición temporal (lote unitario, diez calentamientos y cinco repeticiones), la memoria GPU incremental y la energía de la GPU con el contador de NVML, sin la CPU; su medición en fp32 se contrastó con el costo ya registrado y reproduce su tiempo y su memoria. El segundo es el procedimiento completo hasta fase continua, con la energía neta de CPU y GPU descrita arriba, y solo este es comparable con la energía de los métodos clásicos. El 1 de octubre de 2026 se midieron los métodos clásicos, la U-Net en fp32 y en bf16 y la Dual CNN en fp32. Para completar la comparación, el 3 de octubre se repitió el procedimiento completo únicamente para la Dual CNN en fp32 y en bf16, con el mismo protocolo y con el orden de los dos bloques rotado entre repeticiones, por lo que ese orden no queda balanceado. La energía neta depende del estado del equipo en cada sesión; por ello bf16 y fp32 se comparan solo dentro de una misma sesión. Como control de reproducibilidad entre sesiones se estableció, antes de interpretar las comparaciones, el criterio de que la energía de la Dual CNN en fp32 medida el 3 de octubre debía coincidir con la medida el 1 de octubre: la diferencia entre ambas medias no debía superar dos veces la mayor de las dos desviaciones estándar entre repeticiones. El control se evalúa por separado en cada experimento y su resultado se presenta en el capítulo de resultados. En ambos experimentos, de la medición del 3 de octubre se usa únicamente el cambio relativo entre bf16 y fp32 de esa misma sesión, y frente a los métodos medidos el 1 de octubre se usa la Dual CNN en fp32 de esa fecha.

Para el análisis estadístico, los resultados se conservaron por escena, método y condición. Se emplearon contrastes globales de Friedman y comparaciones pareadas de Wilcoxon con ajuste de Holm, de acuerdo con las familias registradas en los scripts de análisis. Este uso de comparaciones

pareadas y contrastes globales es consistente con la orientación metodológica de Demšar (2006); su aplicación concreta queda limitada por la estructura de este conjunto de escenas. Se acompañan los valores de significación con tamaños de efecto, dirección de las diferencias e intervalos de confianza cuando están disponibles.

Las escenas de prueba proceden de dos geometrías en cada experimento y no equivalen a una muestra de numerosos objetos independientes. Por ello, los resúmenes por figura o modelo complementan los agregados por escena, y la inferencia no se extrapola a toda superficie posible. Los intervalos obtenidos mediante remuestreo de escenas describen la variabilidad condicionada a los objetos incluidos. Asimismo, se distingue la mediana de las diferencias pareadas de la diferencia entre medianas; cada intervalo se interpreta según el estimando registrado por el análisis correspondiente. Las ampliaciones del análisis continuo y las mediciones posteriores del procedimiento completo se identifican como tales, sin presentarlas como decisiones preinscritas.

8.4 Reproducibilidad y alcance de la evaluación

La trazabilidad del trabajo se conserva en el repositorio del proyecto (Mayorga Gamboa et al., 2026), mediante configuraciones, manifiestos, código, historiales, identificadores de versión y registros de evaluación. Los datos completos del simulador permanecen en su ubicación original. Las referencias a sus archivos permiten relacionar cada escena con sus entradas y salidas sin duplicar el conjunto ni modificar las observaciones utilizadas por las corridas congeladas.

Los manifiestos documentan la separación por geometría y las configuraciones identifican las semillas, representaciones, hiperparámetros y pesos seleccionados. La corrida definitiva del Exp1 con semilla 42 se distingue de sus antecedentes. Las semillas y opciones deterministas permiten controlar las fuentes de variación documentadas, pero no implican igualdad numérica universal entre equipos, bibliotecas y versiones diferentes. La partición fija por geometría constituye el procedimiento ejecutado; no se denomina validación cruzada.

La evaluación completa del desenvolvimiento sobre 1920 escenas conserva su código y

protocolo congelados. Los diagnósticos posteriores, incluida la auditoría de presencia de línea central, se registran como análisis separados. Una alineación de regiones mediante la referencia ideal puede ayudar a reconocer desplazamientos enteros de fase, pero no constituye una corrección autónoma ni una salida admisible para reemplazar los errores absolutos. Cualquier exclusión futura requiere un criterio técnico explícito, aplicado por igual a los métodos y documentado sin retirar silenciosamente resultados desfavorables.

El alcance experimental se limita a los datos simulados, a los modelos geométricos y a las condiciones radiométricas descritas. El uso de máscaras y referencias auxiliares ideales, la observación exploratoria de casos de prueba y las diferencias entre protocolos se mantienen como límites de interpretación. Tampoco se infieren efectos causales independientes de las degradaciones a partir de sus combinaciones. La integración de múltiples vistas y el desarrollo de mecanismos autónomos adicionales para identificar errores quedan fuera del trabajo de pregrado.

Con este procedimiento, la evidencia de cumplimiento de los objetivos se apoya en productos verificables: la selección fundamentada de métodos, sus implementaciones y verificaciones, los registros de precisión y costo, y las comparaciones por escena. La valoración final de ese cumplimiento se realizará con los resultados del capítulo 9 y las conclusiones del capítulo 10, respetando los límites de cada experimento.

9. Resultados y discusión

Este capítulo presenta la evidencia obtenida en los dos experimentos descritos en el capítulo 8. Se organiza por experimento y, dentro de cada uno, por tipo de resultado: primero la fase discontinua $\widehat{\varphi}(x,y)$, que permite comparar el error angular de todos los métodos; luego la fase continua $\widehat{\Phi}(x,y)$, que depende además del procedimiento de desenvolvimiento; y finalmente el costo computacional. La sección 9.4 reúne los resultados por criterio y los vincula con los objetivos del trabajo. No se construye una puntuación única, porque la propuesta no asigna pesos relativos al error, a la similaridad estructural, al tiempo, a la memoria, a la energía ni al número de imágenes; por ello cada criterio se discute por separado y no se declara un método ganador absoluto. Las tablas extensas con todas las comparaciones estadísticas se encuentran en el apéndice A.

9.1 Criterios para leer los resultados

La comparación parte de los resultados registrados por escena. Cada valor de error se calcula por escena, método y condición, y se resume sobre las 160 escenas de prueba del Exp1 (80 de `Cono` y 80 de `Prisma`) o sobre las 384 del Exp2 (192 de `armadillo_ascii` y 192 de `lucy`). Las escenas de un mismo objeto no son independientes entre sí y cada experimento solo contiene dos objetos de prueba; por ello los intervalos de confianza del 95 % de las medias, obtenidos remuestreando escenas, y las pruebas estadísticas describen la variabilidad condicionada a esos objetos y no estiman la variabilidad entre objetos. Los resultados por objeto se presentan solo de forma descriptiva.

Para mantener la correspondencia con la metodología, los algoritmos de corrimiento de fase de tres, cuatro y doce imágenes se denominan PSA3, PSA4 y PSA12 en todo el documento; los archivos de resultados del Exp2 los identifican como PS3, PS4 y PS12, y las tablas de este capítulo

ya los muestran con el nombre unificado. La fase discontinua estimada se escribe $\widehat{\varphi}$ y la continua $\widehat{\Phi}$. Las condiciones radiométricas (Ideal, Medium, Moderate y VeryStrong) combinan ruido gaussiano, efecto gamma y saturación con *bloom*, según la Tabla 1; son configuraciones compuestas, de modo que ningún resultado se atribuye a uno solo de esos factores. Las redes se entrenaron únicamente con Moderate: en esa condición se evalúa su desempeño cercano al dominio de entrenamiento y en las demás, su generalización.

Con esta nomenclatura, el RMSE y el MAE circulares de $\widehat{\varphi}$ se calculan con el error circular (ecuaciones (16a), (16b) y (16c)). Para $\widehat{\Phi}$ se utilizan el RMSE y el MAE continuos (ecuaciones (17b) y (17c)) y, en el Exp2, la cobertura C, el error de orden E_{orden} y la cobertura correcta C_{correcta} (ecuaciones (18a), (18b) y (18c)).

El SSIM se calcula con los protocolos propios de cada experimento (SSIM circular sobre coseno y seno en el Exp1 y SSIM global de ventana única en el Exp2). Por esa razón, *el SSIM y las métricas de fase continua no son comparables numéricamente entre experimentos*; entre experimentos solo se comparan órdenes y patrones de comportamiento. En el Exp1 no se definió una métrica de error de orden y la cobertura continua es del 100 % para todos los métodos, de modo que no distingue su calidad.

La comparación de estas métricas se apoya en el análisis estadístico descrito en metodología. Para cada condición y métrica se aplicó una prueba global de Friedman entre métodos, con la escena como bloque, y comparaciones pareadas de Wilcoxon entre cada par de métodos, con ajuste de Holm dentro de cada familia de comparaciones. En el Exp1, las 30 familias globales de $\widehat{\varphi}$ y las 20 de $\widehat{\Phi}$ son significativas tras el ajuste, y 263 de las 288 comparaciones pareadas de $\widehat{\varphi}$ también lo son al nivel 0,05; en el Exp2, lo son las 27 familias globales de $\widehat{\varphi}$ y 208 de las 210 comparaciones pareadas. La elevada proporción de contrastes significativos hace necesario examinar también la magnitud de las diferencias: un valor $p < 0{,}001$ aporta evidencia contra la hipótesis nula del contraste, pero no cuantifica la magnitud de la diferencia.

Por eso la interpretación se apoya en la diferencia entre métodos expresada en la unidad

de la métrica y en el rango biserial r. Aquí Δ es la mediana de las diferencias por escena (primer método menos segundo) y r resume el balance de los rangos de las diferencias con signo; un $|r|$ cercano a 1 indica un predominio de los rangos en una misma dirección, sin informar cuán grande es la diferencia en radianes. Los valores p menores que 0,001 se informan como $< 0{,}001$. Las pruebas de $\widehat{\Phi}$ se tratan como evidencia secundaria, porque se añadieron después de las pruebas sobre $\widehat{\varphi}$. El apéndice A contiene los contrastes globales y las comparaciones pareadas de las condiciones Moderate y VeryStrong.

La precisión se examina junto con el costo computacional, para el cual se distinguen dos alcances. El primero es la *inferencia* sola: tiempo de la red o del algoritmo hasta $\widehat{\varphi}$ (en LTPU incluye ya $\widehat{\Phi}$), memoria (*PeakMem* de MATLAB en los métodos clásicos y pico incremental de VRAM de PyTorch en las redes) y, para las redes, energía de la GPU. El segundo es el *procedimiento completo hasta* $\widehat{\Phi}$, con 24 escenas de Moderate y tres repeticiones independientes, e incluye transferencias, desenvolvimiento y conversiones; solo este alcance pone a todos los métodos en una base común. Su energía es neta, de CPU y GPU con el reposo descontado, y no corresponde al consumo en el enchufe. La energía de la GPU de los métodos clásicos se tomó como cero, con una cota declarada para lo no contabilizado de 1,2 %, 1,0 %, 3,5 % y 2,8 % de la energía neta de PSA3, PSA4, PSA12 y LTPU en el Exp1. Las memorias de CPU y GPU no se comparan entre sí.

Dentro de estos alcances, la precisión numérica de las redes se declara en cada medición: las predicciones evaluadas son las de la U-Net en bf16 y las de la Dual CNN en fp32; la Dual CNN en bf16 es solo una configuración de comparación de costo. Como la energía depende del estado del equipo en cada sesión, bf16 y fp32 se comparan únicamente dentro de una misma sesión (1 y 3 de octubre de 2026).

9.2 Experimento 1: superficies analíticas controladas

Bajo estos criterios, el Exp1 compara PSA3, PSA4, PSA12, LTPU, U-Net y Dual CNN sobre 160 escenas de prueba, con el propósito de examinar su funcionamiento en las superficies

analíticas del entorno controlado. La fase continua se obtiene para las redes y los PSA con un punto semilla en el centroide de la máscara, la máscara y el desenvolvimiento espacial clásico; LTPU entrega directamente la fase continua desde su secuencia temporal.

9.2.1 Fase discontinua

La comparación comienza en Moderate, para relacionar el desempeño con la condición utilizada durante el entrenamiento. La Tabla 3 resume el desempeño en esta condición, y la Figura 20 muestra cómo cambia el error entre condiciones.

Tabla 3
Exp1: error y similaridad de la fase discontinua en Moderate

Método	RMSE circular (rad)	MAE circular (rad)	SSIM circular
PSA3	0,1950 [0,1865; 0,2033]	0,1694 [0,1612; 0,1773]	0,9029
PSA4	0,1058 [0,0967; 0,1148]	0,0865 [0,0782; 0,0948]	0,9681
PSA12	0,0150 [0,0148; 0,0152]	0,0118 [0,0117; 0,0120]	0,9994
LTPU	0,5294 [0,4832; 0,5757]	0,4291 [0,3887; 0,4695]	0,5249
U-Net	0,0225 [0,0195; 0,0262]	0,0111 [0,0106; 0,0116]	0,9976
Dual CNN	0,0322 [0,0272; 0,0377]	0,0124 [0,0117; 0,0132]	0,9968

Nota. Media sobre 160 escenas de prueba; entre corchetes, el intervalo de confianza del 95 %.

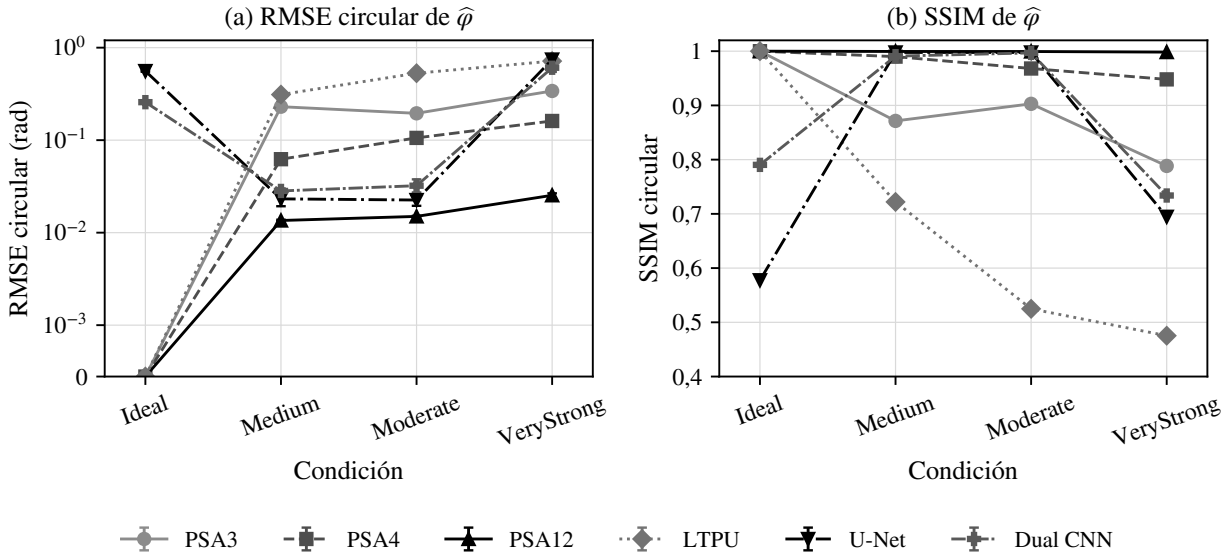
En la condición de entrenamiento, PSA12 presenta el menor RMSE circular y el mayor SSIM, y las dos redes ocupan el segundo y el tercer lugar, por delante de PSA4, PSA3 y LTPU. La comparación entre PSA12 y la U-Net no tiene un ganador único: PSA12 tiene menor RMSE (mediana de las diferencias por escena de $-0{,}001625$ rad, tomando PSA12 menos U-Net, $r = -0{,}43$, $p < 0{,}001$) y mayor SSIM, pero la U-Net tiene menor MAE (mediana de las diferencias por escena de 0,001716 rad, en el mismo orden, $p < 0{,}001$). Las medianas de las diferencias pareadas son de milésimas de radián y el efecto sobre el RMSE es moderado; estos resultados describen la magnitud

observada, sin establecer equivalencia entre los métodos.

Al comparar las dos redes, la U-Net supera a la Dual CNN en las tres métricas ($p < 0{,}001$), mientras que el MAE de la Dual CNN y el de PSA12 no difieren de forma significativa ($p = 0{,}276$). Con menos imágenes que PSA12, PSA3 y PSA4 tienen errores mucho mayores (RMSE de 0,1950 y 0,1058 rad frente a 0,0150 rad) y ambas redes los superan en las tres métricas. LTPU ocupa el último lugar en las tres métricas frente a todos los demás métodos.

Figura 20

Exp1: RMSE circular y SSIM circular de la fase discontinua



Nota. (a) RMSE circular de $\widehat{\varphi}$ (media con intervalo de confianza del 95 %, escala simétrica logarítmica) y (b) SSIM circular, para cada método en las cuatro condiciones, sobre 160 escenas de prueba. Las redes se entrenaron solo con Moderate; Ideal, Medium y VeryStrong evalúan su generalización. La línea que une los puntos de una misma serie facilita la lectura de la tendencia y no representa mediciones intermedias.

Al extender la comparación a las cuatro condiciones de la Figura 20, aparecen diferencias en el comportamiento de los métodos clásicos y las redes. PSA12 es el de menor RMSE y mayor SSIM en Medium, Moderate y VeryStrong, y su deterioro es gradual: su RMSE pasa de 0,0135 rad en Medium a 0,0253 rad en VeryStrong, frente a 0,1612 y 0,3403 rad de PSA4 y PSA3. En Ideal, los cuatro métodos clásicos alcanzan un RMSE de $4{,}6 \times 10^{-8}$ rad, el piso numérico del cálculo, y

sus diferencias no son significativas.

Las redes, en cambio, solo son competitivas cerca de Moderate: su RMSE en Ideal (imagen sin degradar) es de 0,5586 rad para la U-Net y de 0,2575 rad para la Dual CNN, y en VeryStrong sube a 0,7442 y 0,6065 rad. En Ideal no hay ruido de entrada que explique el error, y es coherente con que las redes solo vieron imágenes de Moderate, pero este experimento no varió el dominio de entrenamiento y por eso no permite afirmar esa como la causa. Entre las dos redes el orden depende de la condición: la U-Net es mejor en Medium y Moderate, y la Dual CNN lo es en Ideal y VeryStrong, aunque en estas dos condiciones el error de ambas es grande. En VeryStrong, PSA3 supera en RMSE circular a las dos redes y a LTPU, y la U-Net y LTPU quedan muy próximas (medias de 0,744 y 0,716 rad; efecto pequeño según el rango biserial, $|r| = 0{,}24$). LTPU coincide con los PSA en Ideal y, en las tres condiciones degradadas, queda por debajo de PSA3, PSA4 y PSA12 en las tres métricas; los valores de LTPU en la Figura 20 muestran que su error crece de forma sostenida con la degradación.

9.2.2 Fase continua

La recuperación de la fase continua permite examinar si las diferencias angulares anteriores se conservan después del desenvolvimiento. En el Exp1, la cobertura continua es del 100 % para todos los métodos y condiciones, por lo que no discrimina entre ellos y la comparación recae sobre el error continuo (Tabla 4).

Tabla 4
Exp1: RMSE y MAE continuos en Moderate y VeryStrong

	Moderate		VeryStrong	
Método	RMSE (rad)	MAE (rad)	RMSE (rad)	MAE (rad)
PSA3	0,1950	0,1694	2,174	0,865
PSA4	0,1058	0,0865	0,642	0,207
PSA12	0,0150	0,0118	0,042	0,018
LTPU	0,5294	0,4291	1,609	0,666
U-Net	0,0251	0,0112	13,298	8,280
Dual CNN	0,0381	0,0129	10,592	6,460

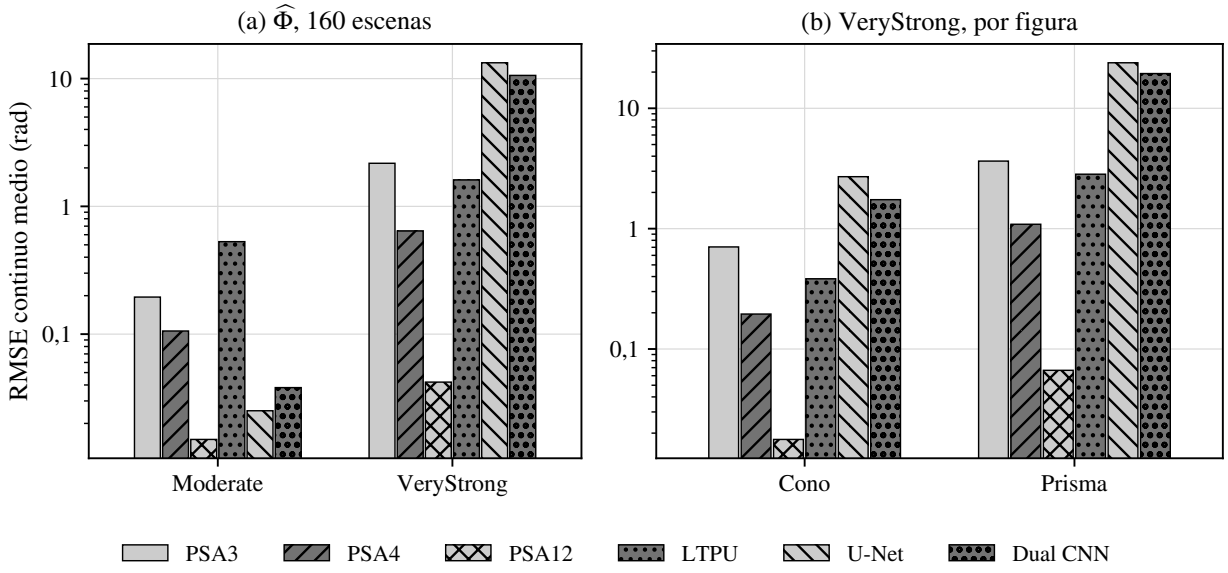
Nota. Media sobre 160 escenas; la cobertura continua es del 100 % en todos los casos. La evaluación admite un único desplazamiento entero global de 2π por escena.

En Moderate, el error continuo de los métodos clásicos coincide con el circular y el de las redes es algo mayor; el orden por RMSE es el mismo que en fase discontinua.

El cambio más marcado aparece en VeryStrong. Allí el RMSE continuo medio de la U-Net y de la Dual CNN, de 13,3 y 10,6 rad, está en el orden de 2π o por encima, valor compatible con errores de desenvolvimiento que se propagan en la región; el RMSE circular, que es insensible a los saltos de 2π, no los muestra. Esto ilustra que estimar la fase y desenvolverla son problemas distintos: un método puede tener un error circular moderado y un error continuo grande. En esa condición el orden por RMSE continuo medio es PSA12, PSA4, LTPU, PSA3, Dual CNN y U-Net; la Dual CNN es mejor que la U-Net ($p < 0{,}001$). LTPU muestra una ventaja específica del desenvolvimiento temporal: su error continuo es menor que el de PSA3 ($p < 0{,}001$, $r = 0{,}52$) y que el de las dos redes, aunque PSA4 y PSA12 mantienen un error menor que el suyo.

La Figura 21 presenta el RMSE continuo en ambas condiciones y permite observar, además, su distribución entre las dos figuras de prueba en VeryStrong.

Figura 21
Exp1: RMSE continuo medio de la fase continua



Nota. Escala logarítmica. (a) Condiciones Moderate y VeryStrong sobre las 160 escenas. (b) Condición VeryStrong separada por figura de prueba (`Cono` y `Prisma`, 80 escenas cada una); la separación es descriptiva y no se contrastó formalmente.

El panel (b) de la Figura 21 muestra que, en VeryStrong, el error de las redes se concentra en `Prisma`: 23,9 y 19,4 rad en la U-Net y la Dual CNN, frente a 2,7 y 1,7 rad en `Cono`. En Moderate, LTPU pasa de 0,233 rad en `Cono` a 0,826 rad en `Prisma`, y PSA3 y PSA4 muestran direcciones opuestas entre sí. Con solo dos figuras de prueba no es posible separar el efecto de la forma del de otras diferencias entre escenas, de modo que estas diferencias se reportan como observación y no como explicación.

9.2.3 Costo computacional

El desempeño anterior debe interpretarse junto con los recursos necesarios para obtener la fase. La Tabla 5 reúne el costo de inferencia y el del procedimiento completo; más adelante, la Figura 24 permite visualizar el tiempo y la energía de cada experimento con sus respectivas escalas.

Tabla 5
Exp1: costo computacional de cada método

			Solo inferencia		Procedimiento completo hasta $\widehat{\Phi}$		
Método	Precisión	Imágenes	Tiempo (ms)	Memoria (MiB)	Sesión	Tiempo (ms)	Energía (J)
PSA3	double	3	13,29	3,67	1 oct	23,1	$0{,}67 \pm 0{,}05$
PSA4	double	4	17,04	3,67	1 oct	26,3	$0{,}74 \pm 0{,}04$
PSA12	double	12	48,69	3,67	1 oct	52,2	$1{,}40 \pm 0{,}09$
LTPU	double	20	55,29	18,35	1 oct	49,5	$2{,}03 \pm 0{,}10$
U-Net	bf16	1	44,57	710,4	1 oct	68,4	$6{,}39 \pm 0{,}12$
U-Net	fp32	1	73,64	919,8	1 oct	97,3	$10{,}16 \pm 0{,}18$
Dual CNN	bf16	1	47,98	293,6	3 oct	70,7	$6{,}38 \pm 0{,}14$
Dual CNN	fp32	1	50,00	378,3	1 oct	73,9	$7{,}37 \pm 0{,}14$
Dual CNN	fp32	1	50,00	378,3	3 oct	72,8	$7{,}27 \pm 0{,}17$

Nota. Tiempos medianos; memoria de CPU en los métodos clásicos y de GPU en las redes; energía neta media $\pm$ desviación entre tres repeticiones.

En el procedimiento completo, los métodos clásicos tardan entre 23,1 y 52,2 ms, LTPU 49,5 ms y las redes entre 68,4 y 97,3 ms en la sesión del 1 de octubre. Que LTPU resulte algo más rápido que PSA12 (49,5 frente a 52,2 ms) es una consecuencia del alcance: el procedimiento de PSA12 incluye el desenvolvimiento espacial, mientras que LTPU entrega la fase continua directamente; en la inferencia sola, que no incluye ese desenvolvimiento, LTPU tarda 55,29 ms y PSA12 48,69 ms, y LTPU sigue siendo más lento que PSA3 y PSA4 en ambos alcances.

La energía separa a las dos familias con más claridad que el tiempo: los métodos clásicos, incluido LTPU, consumen entre 0,67 y 2,03 J por pasada y las redes entre 6,39 y 10,16 J. La U-Net en bf16 consume 3,1 veces la energía de LTPU y 4,6 veces la de PSA12. Por tanto, en este experimento la ventaja de las redes es el número de adquisiciones (una imagen) y no el costo energético.

El cambio de precisión modifica ese costo: bf16 reduce la energía de la U-Net en un 37,1 % y la de la Dual CNN en un 12,3 % dentro de su sesión y en las tres repeticiones (Tabla 9), y reduce el tiempo de inferencia sola de 73,64 a 44,57 ms en la U-Net y de 50,00 a 47,98 ms en la Dual CNN. El efecto de bf16 sobre el tiempo del procedimiento completo es menor en la Dual CNN ($-2{,}8$ %) que en la U-Net ($-29{,}7$ %).

Para interpretar estas reducciones es necesario considerar la estabilidad de las mediciones. El control de reproducibilidad entre sesiones se cumplió en el Exp1 (Tabla 10): la energía de la Dual CNN en fp32 del 3 de octubre (7,27 J) coincide con la del 1 de octubre (7,37 J) dentro de 0,6 desviaciones. Además, la elección de la línea base de reposo (mediana en lugar de promedio) eleva los valores entre 2 y 6 % sin alterar el orden de los siete resultados de energía.

9.3 Experimento 2: geometrías tridimensionales complejas

Los resultados del entorno controlado dan paso al Exp2, que amplía la evaluación a geometrías tridimensionales complejas. En este experimento se comparan PSA3, PSA4, PSA12, U-Net y Dual CNN sobre 384 escenas generadas con `FringePointCloudSimulator`. LTPU no se incluyó, por la delimitación justificada en el capítulo 8.

La fase continua se obtiene con el mismo procedimiento para los cinco métodos: desenvolvimiento con descarte de discontinuidades guiado por la línea central de fase cero (leída de la carpeta de la condición ideal) y por la máscara de la escena, sin desplazamiento global y con cobertura que puede ser menor del 100 %. Es decir, aunque las redes reciben una sola imagen para estimar $\widehat{\varphi}$, la fase absoluta de este experimento emplea además esas dos entradas auxiliares. Los dos experimentos usan geometrías, protocolos de SSIM y de fase continua distintos y tamaños de prueba diferentes, y no se aislaron esos factores; por eso sus diferencias numéricas no se interpretan como el efecto aislado de alguno de esos factores. El SSIM y las métricas de fase continua conservan las restricciones de comparación indicadas al inicio del capítulo.

9.3.1 Fase discontinua

Como en el primer experimento, se examina primero Moderate para evaluar los métodos en la condición de entrenamiento. La Tabla 6 reúne el error y la similaridad de la fase discontinua en Moderate; la Figura 22 amplía la comparación a las cuatro condiciones radiométricas.

Tabla 6
Exp2: error y similaridad de la fase discontinua en Moderate

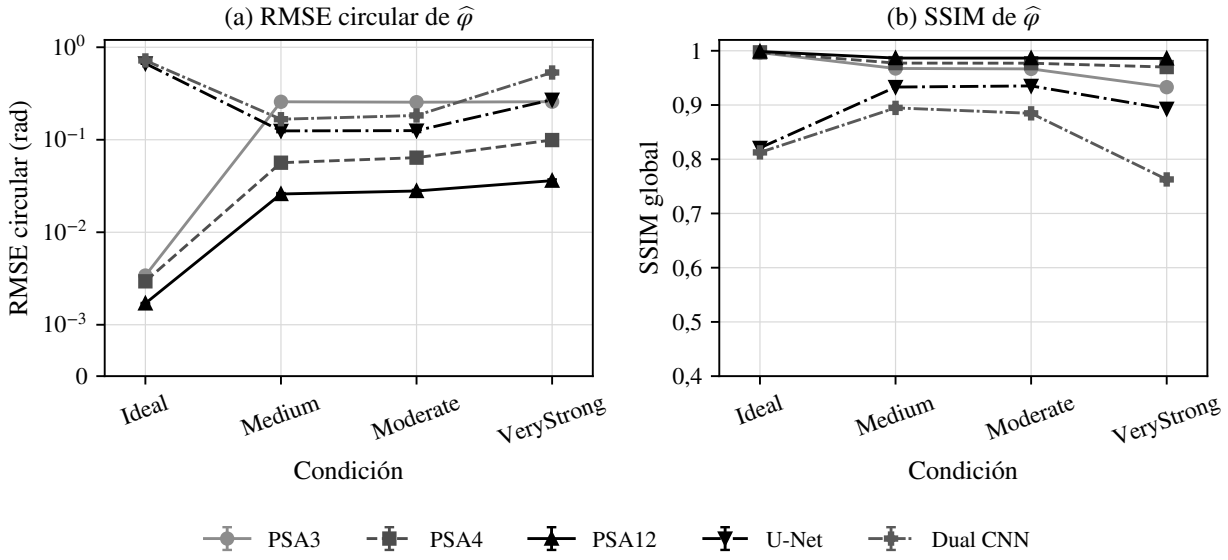
Método	RMSE circular (rad)	MAE circular (rad)	SSIM global
PSA3	0,2548 [0,2538; 0,2557]	0,2236 [0,2230; 0,2242]	0,9666
PSA4	0,0641 [0,0628; 0,0654]	0,0384 [0,0379; 0,0389]	0,9771
PSA12	0,0279 [0,0272; 0,0288]	0,0144 [0,0143; 0,0146]	0,9865
U-Net	0,1255 [0,1212; 0,1300]	0,0635 [0,0615; 0,0655]	0,9352
Dual CNN	0,1830 [0,1783; 0,1877]	0,1105 [0,1078; 0,1132]	0,8846

Nota. Media sobre 384 escenas de prueba; entre corchetes, el intervalo de confianza del 95 %.

En Moderate, el orden por RMSE y por MAE es PSA12, PSA4, U-Net, Dual CNN y PSA3, y las diez comparaciones pareadas de cada métrica son significativas ($p < 0{,}001$) con rangos biseriales de magnitud entre 0,95 y 1,00 en el RMSE (Tabla 6; apéndice, Tabla A5), lo que indica un marcado predominio de las diferencias con signo en la dirección de ese orden. A diferencia del Exp1, aquí ninguna red supera a PSA4 ni a PSA12 en RMSE, MAE o SSIM: el RMSE de la U-Net es 4,49 veces el de PSA12 y el de la Dual CNN es 6,55 veces el de PSA12, y el de la U-Net es 1,96 veces el de PSA4. Frente a PSA3, las dos redes tienen menor RMSE y MAE, pero PSA3 conserva un SSIM mayor que el de ambas. Esta discrepancia entre métricas se observa en los datos; no se analizó su causa.

Figura 22

Exp2: RMSE circular y SSIM global de la fase discontinua



Nota. (a) RMSE circular de $\widehat{\varphi}$ (media con intervalo de confianza del 95 %, escala simétrica logarítmica) y (b) SSIM global, para cada método en las cuatro condiciones, sobre 384 escenas de prueba. Las redes se entrenaron solo con Moderate. La línea entre puntos facilita la lectura de la tendencia y no representa mediciones intermedias.

Entre condiciones (Figura 22), PSA12 es el de mejor RMSE, MAE y SSIM en las cuatro y el de menor deterioro: su RMSE pasa de 0,0017 rad en Ideal a 0,0361 rad en VeryStrong. PSA3 mantiene un error casi constante, entre 0,2548 y 0,2578 rad, en las tres condiciones degradadas, muy superior al de Ideal (0,0034 rad); ese patrón no muestra un aumento gradual del error entre las tres condiciones, aunque no se analizó su causa.

En las redes, el comportamiento vuelve a depender de la condición radiométrica. Tienen un error alto en Ideal y VeryStrong: en Ideal, 0,668 rad en la U-Net y 0,720 rad en la Dual CNN, frente a 0,0017 a 0,0034 rad de los clásicos, con la misma salvedad del Exp1 sobre el dominio de entrenamiento. La U-Net conserva un RMSE medio similar en Medium y Moderate (0,1247 y 0,1255 rad, respectivamente; no se detectó una diferencia significativa, $p = 0{,}254$, $r = 0{,}067$) y empeora en VeryStrong (0,2707 rad); la Dual CNN lo hace más (0,5335 rad), y es menos precisa y

menos robusta que la U-Net en todas las condiciones medidas. El orden de los cinco métodos por RMSE en Moderate es el mismo en `armadillo_ascii` y en `lucy`.

9.3.2 Fase continua

El análisis de la fase continua amplía la comparación anterior al incorporar el desenvolvimiento y la región efectivamente recuperada. La Tabla 7 presenta conjuntamente el error de la fase continua, la cobertura recuperada, el error de orden y la cobertura correcta, para distinguir cuánto se recupera y qué proporción del objeto satisface el criterio de orden.

Tabla 7
Exp2: error continuo y cobertura de la fase continua

Método	RMSE media (rad)	RMSE mediana (rad)	C (%)	E_{orden} media (%)	E_{orden} mediana (%)	C_{correcta} (%)
Moderate						
PSA3	1,189	0,276	94,53	6,76	0,031	88,29
PSA4	0,875	0,121	93,57	4,60	0,027	89,41
PSA12	0,826	0,101	93,15	4,57	0,024	89,03
U-Net	1,034	0,184	93,25	6,12	0,041	87,72
Dual CNN	1,148	0,273	95,00	6,57	0,096	88,91
VeryStrong						
PSA3	1,176	0,286	94,76	6,49	0,049	88,78
PSA4	0,907	0,154	93,94	4,61	0,034	89,75
PSA12	0,833	0,106	93,31	4,50	0,025	89,22
U-Net	1,278	0,569	94,36	7,27	0,685	87,69
Dual CNN	2,015	1,433	95,05	11,97	5,228	83,78

Nota. Condiciones Moderate y VeryStrong, sobre 384 escenas de prueba.

En $\widehat{\Phi}$, las medias del error continuo son mucho mayores que las medianas (Tabla 7), porque una minoría de escenas concentra errores de orden de fase: en Moderate, el E_{orden} medio está entre

4,5 % y 6,8 % mientras que su mediana es de apenas 0,024 a 0,096 %. Por eso se informan ambas, y las medianas describen la escena típica mientras que las medias reflejan el peso de los casos con error de orden. En Moderate, el orden por mediana de RMSE continuo coincide con el de $\widehat{\varphi}$ (PSA12, PSA4, U-Net, Dual CNN, PSA3) y las comparaciones pareadas de RMSE y MAE continuos son significativas en las diez parejas.

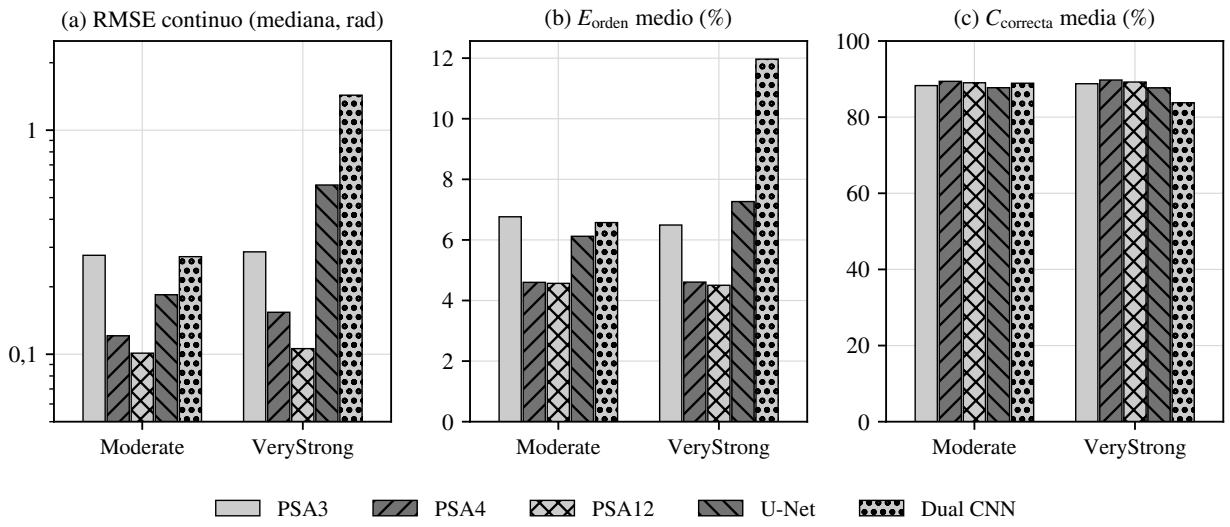
La relación entre ambos tipos de error también se observa en Ideal, donde $\widehat{\varphi}$ es casi perfecta, los tres métodos clásicos tienen un RMSE continuo medio cercano a 0,83 rad: con una fase discontinua casi exacta, el error continuo restante es compatible con errores de orden del procedimiento de desenvolvimiento y con la geometría de los modelos, sin que se haya aislado cada causa.

Además del error, es necesario considerar cuánto se recupera del objeto. La cobertura media en Moderate está entre 93,15 % y 95,00 % y, por sí sola, no demuestra una recuperación correcta: la Dual CNN tiene la mayor cobertura y, a la vez, un error continuo mayor que PSA12, PSA4 y la U-Net. En VeryStrong la Dual CNN es el método con mayor error continuo (RMSE medio de 2,015 rad), mayor error de orden (mediana de 5,23 %) y menor cobertura correcta (83,78 %), mientras que PSA12 conserva 0,833 rad.

La Figura 23 ilustra estas diferencias mediante la mediana del RMSE continuo, el error de orden medio y la cobertura correcta media en Moderate y VeryStrong.

Figura 23

Exp2: RMSE continuo, error de orden y cobertura correcta



Nota. (a) Mediana del RMSE continuo de $\widehat{\Phi}$ (escala logarítmica), (b) error de orden medio E_{orden} y (c) cobertura correcta media C_{correcta} (eje de 0 a 100 %), en Moderate y VeryStrong, sobre 384 escenas de prueba. El eje del panel (c) parte de cero para no exagerar diferencias de pocos puntos porcentuales.

El error continuo también depende del modelo tridimensional, de forma descriptiva: en Moderate, el RMSE continuo medio es mayor en `armadillo_ascii` que en `lucy` para todos los métodos (por ejemplo, 1,2703 frente a 0,3826 rad en PSA12 y 1,5216 frente a 0,5464 rad en la U-Net) sin que cambie el orden entre métodos. Con dos modelos de prueba no se puede aislar la causa de esa diferencia.

9.3.3 Costo computacional

La recuperación de la fase continua también determina una parte importante del costo del Exp2. La Tabla 8 reúne las mediciones de inferencia y del procedimiento completo, y la Figura 24 muestra cómo se distribuyen el tiempo y la energía entre los métodos de cada experimento.

Tabla 8

Exp2: costo computacional de cada método

			Solo inferencia		Procedimiento completo hasta $\widehat{\Phi}$		
Método	Precisión	Imágenes	Tiempo (ms)	Memoria (MiB)	Sesión	Tiempo (ms)	Energía (J)
PSA3	double	3	22,16	6,02	1 oct	1882,0	$43{,}96 \pm 4{,}30$
PSA4	double	4	28,60	6,02	1 oct	1908,3	$47{,}98 \pm 8{,}05$
PSA12	double	12	82,05	6,02	1 oct	1945,6	$44{,}86 \pm 4{,}22$
U-Net	bf16	1	74,04	1164,0	1 oct	2080,2	$83{,}57 \pm 1{,}49$
U-Net	fp32	1	121,38	1504,5	1 oct	2107,4	$99{,}06 \pm 1{,}98$
Dual CNN	bf16	1	78,46	481,7	3 oct	2126,5	$74{,}55 \pm 1{,}24$
Dual CNN	fp32	1	82,34	617,2	1 oct	2178,5	$89{,}43 \pm 3{,}45$
Dual CNN	fp32	1	82,34	617,2	3 oct	2130,3	$78{,}60 \pm 2{,}53$

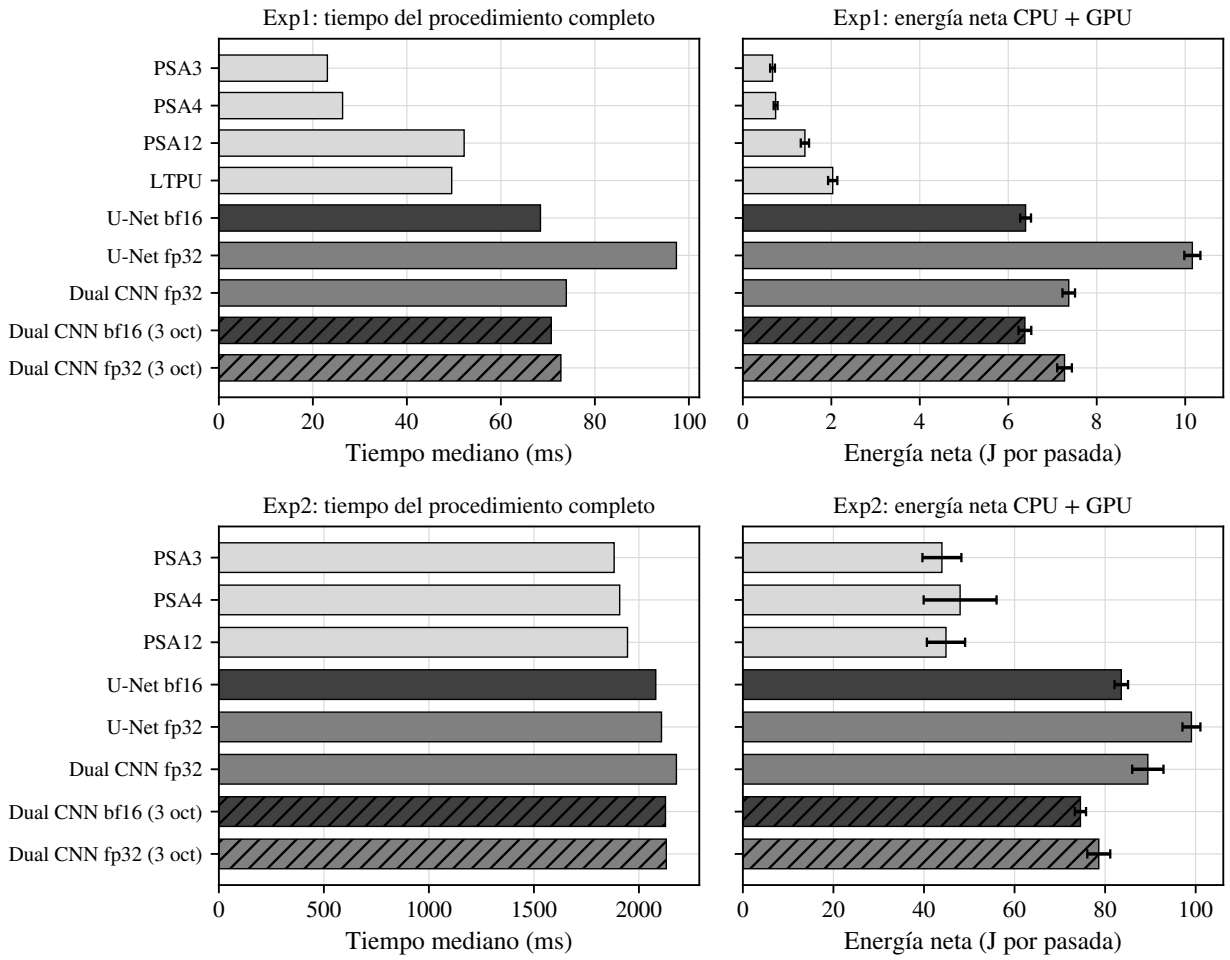
Nota. Se emplean las mismas medidas de la Tabla 5.

En el Exp2, cerca de 1,9 s de cada pasada corresponden al desenvolvimiento en MATLAB, que domina el tiempo del procedimiento completo en todos los métodos. Las medianas van de 1882,0 ms (PSA3) a 2178,5 ms (Dual CNN) y los intervalos de confianza de esas medianas se solapan ampliamente; no se hizo una prueba inferencial de tiempos y estas mediciones no permiten ordenar los métodos por tiempo. Tampoco se afirma que sean iguales.

En cuanto a la energía, las mediciones separan a las familias: los métodos clásicos consumen entre 43,96 y 47,98 J por pasada, sin diferencias distinguibles entre ellos dentro de la incertidumbre entre repeticiones (en PSA4 se registró un episodio de actividad ajena en la primera repetición), y las redes consumen aproximadamente el doble.

Figura 24

Tiempo y energía del procedimiento completo



Nota. Tiempo mediano y energía neta de CPU y GPU del procedimiento completo hasta $\widehat{\Phi}$, para el Exp1 (fila superior) y el Exp2 (fila inferior), en 24 escenas de Moderate con tres repeticiones; las barras de error de la energía son la desviación entre repeticiones. Las barras rayadas corresponden a la sesión del 3 de octubre. Los dos experimentos tienen escalas distintas y no deben compararse entre sí.

Como se resume en la Tabla 9, bf16 reduce la energía de las dos redes en las tres repeticiones de cada sesión: $-15{,}6\,\%$ en la U-Net y $-5{,}2\,\%$ en la Dual CNN, con un efecto casi nulo sobre el tiempo del procedimiento completo ($-1{,}3\,\%$ y $-0{,}2\,\%$), porque el desenvolvimiento domina ese tiempo; en la inferencia sola, bf16 reduce el tiempo de 121,38 a 74,04 ms en la U-Net y de 82,34 a 78,46 ms en la Dual CNN, y la VRAM de 1504,5 a 1164,0 MiB y de 617,2 a 481,7 MiB. Como

cada red se midió en una sesión distinta (U-Net, 1 de octubre; Dual CNN, 3 de octubre, con orden fp32/bf16 no balanceado), los cambios de la Tabla 9 se comparan dentro de cada red y no entre redes.

Tabla 9

Efecto de la precisión bf16 en las redes

Energía del procedimiento completo

Experimento	Red (sesión)	Energía fp32 → bf16 (J)	Cambio (%)	Repeticiones con menor energía
Exp1	U-Net (1 oct)	10,16 → 6,39	−37,1	3 de 3
Exp1	Dual CNN (3 oct)	7,27 → 6,38	−12,3	3 de 3
Exp2	U-Net (1 oct)	99,06 → 83,57	−15,6	3 de 3
Exp2	Dual CNN (3 oct)	78,60 → 74,55	−5,2	3 de 3

Cambios relativos de tiempo

Experimento	Red (sesión)	Procedimiento completo (%)	Solo inferencia (%)
Exp1	U-Net (1 oct)	−29,7	−39,5
Exp1	Dual CNN (3 oct)	−2,8	−4,0
Exp2	U-Net (1 oct)	−1,3	−39,0
Exp2	Dual CNN (3 oct)	−0,2	−4,7

Nota. Cambio porcentual de energía y de tiempo al pasar de fp32 a bf16. El cambio se calcula dentro de cada sesión como $(\text{bf16}/\text{fp32} - 1) \times 100$.

Las reducciones dentro de una sesión deben distinguirse de la reproducibilidad entre sesiones. La Tabla 10 presenta este control para la Dual CNN en fp32 y permite delimitar el uso de las mediciones suplementarias.

Tabla 10
Control de reproducibilidad de la energía de la Dual CNN

Experimento	1 oct (J)	3 oct (J)	Diferencia / mayor desviación	Criterio (≤ 2)
Exp1	7,37 $\pm$ 0,14	7,27 $\pm$ 0,17	0,6	Cumplido
Exp2	89,43 $\pm$ 3,45	78,60 $\pm$ 2,53	3,1	No cumplido

Nota. Medición en fp32. Media $\pm$ desviación entre tres repeticiones, comparada entre sesiones. Criterio: diferencia absoluta de medias no mayor que dos veces la mayor de las dos desviaciones.

El control de reproducibilidad no se cumplió en el Exp2 (Tabla 10): la energía de la Dual CNN en fp32 del 3 de octubre, de 78,60 J, es 10,83 J menor que la del 1 de octubre, de 89,43 J, es decir, unas 3,1 desviaciones, sin causa determinada. Por esta razón, del 3 de octubre solo se usa el cambio relativo entre bf16 y fp32, y la medición de bf16 de esa sesión se muestra como información suplementaria, sin usarla como comparación directa frente a los métodos clásicos. La falta de reproducibilidad se informa como un límite de las cifras absolutas de la Dual CNN en el Exp2; no afecta la conclusión de que las redes consumen más energía que los métodos clásicos, porque ambas mediciones de la Dual CNN superan con amplitud a las de los clásicos, ni la dirección del efecto de bf16, que se observó dentro de cada sesión.

9.4 Comparación integrada y relación con los objetivos

La evidencia de precisión y costo permite reunir las ventajas y limitaciones observadas sin reducirlas a una puntuación única. La Tabla 11 sintetiza, por criterio, lo que muestran los dos experimentos. Se compara el orden y el patrón de los métodos, no los valores de las métricas entre experimentos. La tabla no define un ganador global, porque los criterios no están ponderados.

Tabla 11
Síntesis por criterio de los dos experimentos

Criterio	Exp1 (6 métodos, 160 escenas)	Exp2 (5 métodos, 384 escenas)
Error de $\widehat{\varphi}$ en Moderate	PSA12 menor RMSE y mayor SSIM; U-Net menor MAE (diferencia de milésimas de radián); las redes superan a PSA3 y PSA4	PSA12, PSA4, U-Net, Dual CNN, PSA3 en RMSE y MAE; ninguna red supera a PSA4 ni a PSA12
Robustez fuera de Moderate	PSA12 el más robusto; las redes se degradan fuertemente en Ideal y VeryStrong; PSA3 supera a las redes en VeryStrong	PSA12 el más robusto; las redes tienen error alto en Ideal y en VeryStrong; U-Net conserva un RMSE similar en Medium
Error de $\widehat{\Phi}$	En Moderate, PSA12 tiene menor RMSE, seguido por las redes; en VeryStrong, PSA12 y PSA4 presentan menor RMSE y LTPU supera a PSA3 y las redes	En Moderate, el orden por mediana de RMSE coincide con el de $\widehat{\varphi}$; la cobertura se interpreta junto con el error de orden y la cobertura correcta
Imágenes para estimar $\widehat{\varphi}$	Redes: 1; PSA3: 3; PSA4: 4; PSA12: 12; LTPU: 20	Redes: 1; PSA3: 3; PSA4: 4; PSA12: 12
Tiempo (procedimiento completo)	Clásicos 23 a 52 ms; LTPU 50 ms; redes 68 a 97 ms	Entre 1882 y 2179 ms, sin que las mediciones permitan ordenar los métodos
Memoria	Dos indicadores no comparables (CPU y VRAM); la Dual CNN usa menos VRAM que la U-Net	Igual
Energía (procedimiento completo)	Clásicos (con LTPU) entre 0,67 y 2,03 J; redes entre 6,4 y 10,2 J; bf16 la reduce	Clásicos entre 44,0 y 48,0 J; redes aproximadamente el doble; bf16 la reduce

Nota. «Menor» y «mayor» se refieren a la medida del propio experimento, con los alcances de las secciones 9.2.3 y 9.3.3; ningún criterio se pondera.

La lectura conjunta de los dos experimentos muestra ventajas que dependen del método, la condición y la métrica considerados. PSA12 tiene el menor RMSE circular en la condición de entrenamiento en ambos experimentos y es el más robusto fuera de ella. Ninguna red supera a PSA4

ni a PSA12 de forma general. En Exp1 la U-Net tiene un MAE algo menor que PSA12 en Moderate, aunque con la ventaja de PSA12 en RMSE y SSIM; en Exp2 las redes quedan por detrás de PSA4 y PSA12. En Moderate, las redes alcanzan un error competitivo solo en el Exp1, donde superan a PSA3 y PSA4.

Desde el punto de vista de la adquisición, las redes requieren una sola imagen para estimar la fase discontinua; esa ventaja no implica un menor consumo energético ni una mayor precisión en todas las condiciones. Si se prioriza reducir el número de imágenes con una precisión alta, PSA4 ofrece un compromiso frente a PSA12 en ambos experimentos, aunque en Moderate del Exp1 las redes alcanzan errores menores con una sola imagen; si se prioriza una sola imagen, la U-Net es la mejor opción de las dos redes cerca de su dominio de entrenamiento (Medium y Moderate), y en Moderate del Exp2 con un RMSE circular 1,96 veces el de PSA4. Los métodos clásicos son los de menor energía por pasada.

Por su parte, LTPU, solo en el Exp1, se conserva como referencia de desenvolvimiento temporal y no como método recomendado: PSA4 y PSA12 lo superan en precisión en las condiciones degradadas, con menos imágenes, aunque en VeryStrong su error continuo es menor que el de PSA3 y el de las dos redes.

El error de las redes es mayor en el Exp2 que en el Exp1 (en Moderate, 0,1255 y 0,1830 rad frente a 0,0225 y 0,0322 rad en RMSE circular, y en el Exp2 ya no superan a PSA4). Los dos experimentos difieren a la vez en geometrías, tamaño del conjunto de prueba y protocolos de SSIM y de fase continua, y no se aislaron esos factores; por ello la diferencia se registra como observación y no se atribuye a una sola causa, como la geometría o el cambio de dominio. Entre las dos redes tampoco hay un orden universal: en el Exp1 el orden depende de la condición y en el Exp2 la U-Net es mejor en todas.

Esta comparación permite precisar qué evidencia aporta el trabajo a los objetivos propuestos. El primer objetivo específico requiere revisar y sintetizar el estado del arte, por lo que su cumplimiento debe sustentarse en la revisión bibliográfica y en la selección fundamentada de los

métodos. Las comparaciones experimentales complementan esa selección, pero no sustituyen la evidencia bibliográfica. Para el segundo objetivo, la implementación y verificación se describen en el capítulo 8: se implementaron los seis métodos del Exp1 y los cinco del Exp2, y se entrenaron las redes con semilla 42 en modo determinista. Los resultados aquí presentados complementan esas verificaciones mediante una evaluación con escenas y criterios comunes dentro de cada experimento, respetando las particularidades del desenvolvimiento temporal de LTPU.

Sobre esa base, el tercer objetivo se desarrolla mediante el cálculo del RMSE, MAE y SSIM de $\widehat{\varphi}$, el error de $\widehat{\Phi}$ y, en el Exp2, la cobertura, el error de orden y la cobertura correcta. Estas medidas se acompañan del número de imágenes, el tiempo, la memoria y la energía, con los alcances declarados. La evidencia permite comparar los métodos por criterio en las condiciones y objetos evaluados, aunque no establecer diferencias de tiempo entre métodos en el Exp2, una memoria única comparable entre CPU y GPU, ni consumo de energía en el enchufe.

La aplicación de esos criterios a las mismas escenas y condiciones dentro de cada experimento aporta, a su vez, la evidencia correspondiente al cuarto objetivo. La comparación abarca cuatro configuraciones radiométricas compuestas, primero en superficies analíticas y luego en geometrías tridimensionales complejas. Al combinar ruido, gamma y saturación, los resultados describen el comportamiento bajo esas configuraciones, sin atribuir efectos independientes a cada factor ni extrapolarlos a otras geometrías o a adquisiciones reales.

En conjunto, estos elementos sustentan el objetivo general de comparar cuantitativamente métodos clásicos y redes profundas. La respuesta depende de la condición y de la métrica: las ventajas de las redes en Moderate del Exp1 no se mantienen frente a PSA4 y PSA12 en el Exp2, y la reducción del número de imágenes para estimar la fase discontinua no se traduce en una reducción de la energía del procedimiento completo. Esta lectura conserva las diferencias entre experimentos y debe entenderse dentro de las limitaciones que se presentan a continuación.

9.5 Limitaciones documentadas

El alcance de la comparación está condicionado por los datos y por el diseño experimental. Las condiciones radiométricas combinan ruido, gamma y saturación, por lo que no permiten separar el efecto de cada factor. Además, todos los datos son simulados y cada experimento contiene solo dos objetos de prueba. En consecuencia, las pruebas estadísticas y los intervalos de confianza describen la variabilidad entre las escenas evaluadas, condicionada a esos objetos, y no la variabilidad entre objetos ni el desempeño con adquisiciones reales. La elevada proporción de contrastes significativos refuerza la necesidad de interpretar los valores p junto con la magnitud de las diferencias.

El entrenamiento introduce otro límite al alcance de los resultados. Se utilizó una partición fija por figura o por modelo, con siete objetos de entrenamiento, uno de validación y dos de prueba, una única semilla y un solo entrenamiento de cada red por experimento. Esta elección constituye una desviación respecto de la validación cruzada y no permite estimar la variabilidad por reentrenamiento. Como las redes se entrenaron únicamente con Moderate y no se varió el dominio de entrenamiento, tampoco se puede aislar la causa de su error en las demás condiciones ni de su mayor error en el Exp2.

La interpretación de la fase continua requiere, además, conservar las diferencias entre protocolos. En el Exp1 se emplean un punto semilla y la máscara, y la evaluación admite un único desplazamiento entero global de 2π; en el Exp2 se utilizan la línea central y la máscara de la escena, sin ese ajuste global. Estas entradas auxiliares forman parte del alcance de la recuperación evaluada. Tanto las métricas de fase continua como el SSIM tienen definiciones propias de cada experimento y sus valores no se comparan directamente entre ellos. Asimismo, las inspecciones exploratorias previas del desenvolvimiento incluyeron escenas de prueba del Exp2, de modo que la evaluación final no es completamente ciega respecto de todas las decisiones de desarrollo.

Las mediciones de costo también tienen límites específicos. La energía corresponde al consumo neto de CPU y GPU, no al consumo en el enchufe; la contribución de la GPU de los

métodos clásicos se tomó como cero, con la cota declarada para lo no contabilizado. El control de reproducibilidad de la Dual CNN en fp32 no se cumplió en el Exp2 y el orden fp32/bf16 de la sesión suplementaria no estuvo balanceado. Por otra parte, el desenvolvimiento en MATLAB domina el tiempo del procedimiento completo en ese experimento, y las mediciones disponibles no permiten establecer diferencias temporales entre los métodos. Los indicadores de memoria de CPU y GPU conservan alcances distintos y no constituyen una medida conjunta comparable.

Estas condiciones delimitan la evidencia obtenida y su posible extensión a datos de entrenamiento más variados, otras geometrías, configuraciones de visibilidad o dinámica temporal, que no se evaluó. Con ese alcance, las conclusiones del trabajo se presentan en el capítulo 10.

10. Conclusiones y trabajo futuro

Este capítulo recoge lo que la evidencia del capítulo 9 permite concluir respecto al objetivo general y a los objetivos específicos enunciados en la introducción, y plantea las líneas en las que el estudio puede continuar. Las conclusiones se refieren a los datos simulados, a los objetos de prueba y a las condiciones radiométricas de los dos experimentos descritos en el capítulo 8.

10.1 Conclusiones

La comparación cuantitativa planteada en el objetivo general muestra que ningún método fue preferible en todos los criterios, y que la posición de las redes respecto a los métodos clásicos cambió de un experimento a otro. Cerca de su condición de entrenamiento, las redes mostraron un desempeño favorable en el primer experimento: en Moderate del Exp1, la U-Net y la Dual CNN alcanzaron un RMSE circular de 0,0225 y 0,0322 rad, inferior al de PSA4 (0,1058 rad) y al de PSA3 (0,1950 rad), y la U-Net obtuvo incluso un MAE circular menor que el de PSA12. Esa ventaja frente a PSA4 no se mantuvo sobre las geometrías tridimensionales del Exp2, donde el orden por RMSE fue PSA12, PSA4, U-Net, Dual CNN y PSA3, y el error de la U-Net fue 1,96 veces el de PSA4 y 4,49 veces el de PSA12. Las redes, entrenadas únicamente en Moderate, perdieron además competitividad en las condiciones alejadas de ese dominio; en Ideal, por ejemplo, el RMSE circular de la U-Net fue de 0,5586 rad en el Exp1 y de 0,668 rad en el Exp2, mientras que los métodos clásicos permanecieron en el orden de milésimas de radián o por debajo. Como el dominio de entrenamiento no se varió, este deterioro no puede atribuirse a una causa única.

Entre los métodos clásicos, PSA12 fue el más consistente, con el menor RMSE circular en Moderate de ambos experimentos y el menor deterioro entre condiciones. Esa consistencia no lo hizo superior en todos los criterios. Además del menor MAE de la U-Net en el Exp1, en Moderate

del Exp2 PSA4 presentó una cobertura correcta media ligeramente mayor que la de PSA12 (89,41 % frente a 89,03 %), diferencia que se reporta de forma descriptiva y no como superioridad estadística. Estos casos refuerzan la conclusión central del trabajo: un menor error angular y una mayor cantidad de fase recuperada correctamente son criterios complementarios, y la preferencia por un método depende de cuál de ellos se priorice.

Llegar a estas conclusiones requirió las bases establecidas por los dos primeros objetivos específicos. La revisión del estado del arte, primer objetivo específico, presentada en el capítulo 7, permitió fijar los métodos comparados y explicar parte de sus diferencias, como la sensibilidad de cada PSA a los armónicos que introduce la no linealidad gamma y el contraste entre la estimación directa de la U-Net y la estimación indirecta de la Dual CNN. Las implementaciones del segundo objetivo específico se verificaron antes de comparar: en la condición Ideal del Exp1, los métodos clásicos recuperaron la fase discontinua con un RMSE de $4{,}6 \times 10^{-8}$ rad, el piso numérico del cálculo. La partición fija por geometría, la semilla 42 y el registro de configuraciones y código en el repositorio del proyecto (Mayorga Gamboa et al., 2026), descritos en la sección 8.4, hacen rastreable cada resultado, aunque el uso de una sola semilla no permite estimar la variabilidad por reentrenamiento.

El protocolo de métricas definido en el tercer objetivo específico evaluó la fase discontinua con el RMSE y el MAE circulares, junto con el SSIM circular en el Exp1 y global en el Exp2, y la fase continua con errores lineales, cobertura recuperada, error de orden y cobertura correcta. La evaluación conjunta de estas medidas permitió distinguir la precisión angular de la recuperación efectiva del objeto, algo que el error circular por sí solo no habría mostrado. En VeryStrong del Exp1, el RMSE continuo medio de la U-Net y de la Dual CNN (13,3 y 10,6 rad) superó ampliamente a su RMSE circular (0,7442 y 0,6065 rad), lo que es compatible con errores de orden que la medida circular no registra; y en Moderate del Exp2, la Dual CNN obtuvo la mayor cobertura y, a la vez, un error continuo mayor que el de PSA12, PSA4 y la U-Net. El análisis estadístico mostró también que la significación por sí sola no basta para interpretar las diferencias, pues 263 de las 288

comparaciones pareadas de $\widehat{\varphi}$ en el Exp1 y 208 de las 210 en el Exp2 resultaron significativas; por eso la interpretación se apoyó en la magnitud de las diferencias y no solo en los valores p.

La comparación cuantitativa del cuarto objetivo específico deja ver, por último, el compromiso entre precisión, número de adquisiciones y costo. Reducir la entrada neuronal a una sola imagen no garantizó menor error ni menor consumo energético: en el Exp1, las redes consumieron entre 6,39 y 10,16 J por pasada del procedimiento completo, frente a 0,67 a 2,03 J de los métodos clásicos, y en el Exp2 aproximadamente el doble que estos. La precisión bf16 redujo el consumo de las redes dentro de cada sesión ($-37{,}1\,\%$ en la U-Net del Exp1). Entre los métodos clásicos, el compromiso se expresa en el número de imágenes: PSA12 ofreció la mayor precisión con doce, PSA4 ofreció un compromiso intermedio con cuatro y PSA3, con tres, fue el más sensible a la no linealidad. LTPU, con veinte imágenes, obtuvo en VeryStrong del Exp1 un error continuo menor que el de PSA3 y el de las dos redes, aunque PSA4 y PSA12 lo superaron con menos imágenes. Cuando se requiere una sola imagen, la U-Net fue la mejor de las dos redes cerca de su dominio de entrenamiento. Esa ventaja, sin embargo, corresponde a la estimación de la fase discontinua: la recuperación de la fase continua empleó las entradas auxiliares declaradas en la metodología, como la máscara de evaluación y, en el Exp2, la línea central del proyector, de modo que el procedimiento completo no se basa en una única imagen.

Estas conclusiones deben leerse junto con las limitaciones expuestas en la sección 9.5: todos los datos son simulados, cada experimento contiene dos objetos de prueba, las condiciones radiométricas combinan ruido, gamma y saturación sin separar sus efectos, cada red se entrenó una sola vez y la energía corresponde al consumo neto de CPU y GPU.

10.2 Trabajo futuro

Los resultados obtenidos abren varias líneas para dar continuidad al estudio. La más directa es llevar la comparación a adquisiciones reales de un sistema de proyección de franjas calibrado, lo que permitiría contrastar con mediciones físicas el comportamiento observado en simulación. Otra

consiste en ampliar el dominio de entrenamiento de las redes con varias condiciones radiométricas y examinar si así se reduce la pérdida de competitividad observada fuera de Moderate; en esa misma dirección, un diseño factorial que varíe por separado el ruido, la no linealidad gamma y la saturación ayudaría a identificar qué mecanismo afecta más a cada método.

También resultaría de interés aumentar el número de objetos de prueba y repetir el entrenamiento con varias semillas, para estimar la variabilidad entre objetos y por reentrenamiento, y explorar estrategias de desenvolvimiento que dependan menos de entradas auxiliares, incluido el desenvolvimiento temporal con secuencias multifrecuencia sobre geometrías complejas. Finalmente, la medición de la energía en el enchufe y la integración de múltiples vistas permitirían valorar el costo y la cobertura de los métodos dentro de un sistema de medición completo.

Referencias Bibliográficas

Cui, H., Liao, W., Dai, N., and Cheng, X. (2011). Reliability-guided phase-unwrapping algorithm for the measurement of discontinuous three-dimensional objects. *Optical Engineering*, 50(6):063602.

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7:1–30.

Feng, S., Chen, Q., Gu, G., Tao, T., Zhang, L., Hu, Y., Yin, W., and Zuo, C. (2019). Fringe pattern analysis using deep learning. *Advanced Photonics*, 1(2):025001.

Ghiglia, D. C. and Pritt, M. D. (1998). *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. Wiley, New York.

Gonzalez, A. and Meneses, J. (2019). Accurate calibration method for a fringe projection system by projecting an adaptive fringe pattern. *Applied optics*, 58(17):4610–4615.

González, A. L. G., Fonseca, J. E. M., and Téllez, J. L. (2012). Proyección de franjas en metrología óptica facial. *Inge CuC*, 8(1):191–206.

González, A. L. G. (2020). Reconstrucción tridimensional digital de cráneos humanos para aplicaciones en ciencias forenses. Trabajo de grado de maestría en ingeniería de sistemas e informática, Universidad Industrial de Santander, Bucaramanga.

Hu, W., Miao, H., Yan, K., and Fu, Y. (2021). A fringe phase extraction method based on neural network. *Sensors*, 21(5):1664.

Huntley, J. M. and Saldner, H. (1993). Temporal phase-unwrapping algorithm for automated interferogram analysis. *Applied Optics*, 32(17):3047–3052.

Kando, D., Tomioka, S., Miyamoto, N., and Ueda, R. (2019). Phase extraction from single interferogram including closed-fringe using deep learning. *Applied Sciences*, 9(17):3529.

Li, Y., Qian, J., Feng, S., Chen, Q., and Zuo, C. (2020). Single-shot spatial frequency multiplex fringe pattern for phase unwrapping using deep learning. In *Optics Frontier Online 2020: Optics Imaging and Display*, volume 11571, pages 314–319. SPIE.

Machineni, R. C., Spoorthi, G., Vengala, K. S., Gorthi, S., and Gorthi, R. K. S. S. (2020). End-to-end deep learning-based fringe projection framework for 3d profiling of objects. *Computer vision and image understanding*, 199:103023.

Mayorga Gamboa, Y. F., Silva Solano, J. D., González Gómez, A. L., and Meneses Fonseca, J. E. (2026). recuperacion-fase-dl: recuperación de fase en proyección de franjas mediante métodos clásicos y redes neuronales.

Múnera, N., Lora, G., and García-Sucerquia, J. (2011). Técnicas de proyección de franjas y de escaneo láser para la reconstrucción 3d y generación de archivos cam de piezas dentales. *Revista Colombiana de Física*, 43(3):856.

Rastogi, P. and Hack, E. (2014). *Phase estimation in optical interferometry*. CRC Press.

Rincón, O. J., Nicholls, J. J., and Gaviria, P. (2009). Proyección de franjas con corrimiento de fase: aplicación a las ciencias forenses. *Antistio: revista científica del Instituto Nacional de Medicina Legal y Ciencias Forenses de Colombia*, 1(3):8–13.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.

Saldner, H. O. and Huntley, J. M. (1997). Temporal phase unwrapping: application to surface profiling of discontinuous objects. *Applied Optics*, 36(13):2770–2775.

Surrel, Y. (1997). Additive noise effect in digital phase detection. *Applied Optics*, 36(1):271–276.

Takeda, M., Ina, H., and Kobayashi, S. (1982). Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. *Journal of the Optical Society of America*, 72(1):156–160.

Wang, K., Li, Y., Kemao, Q., Di, J., and Zhao, J. (2019). One-step robust deep learning phase unwrapping. *Optics express*, 27(10):15100–15115.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612.

Yan, K., Yu, Y., Huang, C., Sui, L., Qian, K., and Asundi, A. (2019). Fringe pattern denoising based on deep learning. *Optics Communications*, 437:148–152.

Ye, X., Cheng, H.-B., Wu, H.-Y., Zhou, D.-M., and Tam, H.-Y. (2015). Gamma correction for three-dimensional object measurement by phase measuring profilometry. *Optik*, 126(24):5534–5538.

Zhang, S. (2010). Recent progresses on real-time 3d shape measurement using digital fringe projection techniques. *Optics and Lasers in Engineering*, 48(2):149–158.

Zhang, S. (2013). *Handbook of 3D machine vision: Optical metrology and imaging*. CRC Press.

Zhao, M., Wang, H., and Kemao, Q. (2015). Snake-assisted quality-guided phase unwrapping for discontinuous phase fields. *Applied Optics*, 54(24):7462–7470.

Zuo, C., Feng, S., Huang, L., Tao, T., Yin, W., and Chen, Q. (2018). Phase shifting algorithms for fringe projection profilometry: A review. *Optics and lasers in engineering*, 109:23–59.

Apéndice A. Tablas complementarias de resultados

Este apéndice reúne las tablas extensas de inferencia estadística que sustentan el capítulo 9. Las tablas de contrastes globales (Tablas A1 y A2) listan todas las familias; las de comparaciones pareadas se limitan a las condiciones Moderate y VeryStrong (Tablas A3 a A6). La definición e interpretación de Δ, p y r se presentan en la sección 9.1. Los métodos del Exp2 se identifican con la nomenclatura unificada (PSA3, PSA4 y PSA12).

Tabla A1
Contrastes globales de Friedman del Exp1

Contraste	Métrica	Condición o método	Estadístico de Friedman	p ajustado (Holm)
Entre métodos	RMSE circular	Ideal	573,3	$< 0{,}001$
Entre métodos	RMSE circular	Moderate	655,2	$< 0{,}001$
Entre métodos	RMSE circular	Medium	722,0	$< 0{,}001$
Entre métodos	RMSE circular	VeryStrong	699,3	$< 0{,}001$
Entre condiciones	RMSE circular	PSA3	476,4	$< 0{,}001$
Entre condiciones	RMSE circular	PSA4	476,4	$< 0{,}001$
Entre condiciones	RMSE circular	PSA12	476,4	$< 0{,}001$
Entre condiciones	RMSE circular	LTPU	476,4	$< 0{,}001$
Entre condiciones	RMSE circular	U-Net	380,9	$< 0{,}001$
Entre condiciones	RMSE circular	Dual CNN	399,0	$< 0{,}001$
Entre métodos	MAE circular	Ideal	572,8	$< 0{,}001$
Entre métodos	MAE circular	Moderate	654,0	$< 0{,}001$
Entre métodos	MAE circular	Medium	722,3	$< 0{,}001$

Contraste	Métrica	Condición o método	Estadístico de Friedman	p ajustado (Holm)
Entre métodos	MAE circular	VeryStrong	619,4	$< 0{,}001$
Entre condiciones	MAE circular	PSA3	362,4	$< 0{,}001$
Entre condiciones	MAE circular	PSA4	476,4	$< 0{,}001$
Entre condiciones	MAE circular	PSA12	476,4	$< 0{,}001$
Entre condiciones	MAE circular	LTPU	476,4	$< 0{,}001$
Entre condiciones	MAE circular	U-Net	380,1	$< 0{,}001$
Entre condiciones	MAE circular	Dual CNN	415,2	$< 0{,}001$
Entre métodos	SSIM	Ideal	717,7	$< 0{,}001$
Entre métodos	SSIM	Moderate	723,5	$< 0{,}001$
Entre métodos	SSIM	Medium	747,7	$< 0{,}001$
Entre métodos	SSIM	VeryStrong	706,3	$< 0{,}001$
Entre condiciones	SSIM	PSA3	426,4	$< 0{,}001$
Entre condiciones	SSIM	PSA4	476,4	$< 0{,}001$
Entre condiciones	SSIM	PSA12	476,4	$< 0{,}001$
Entre condiciones	SSIM	LTPU	441,7	$< 0{,}001$
Entre condiciones	SSIM	U-Net	379,9	$< 0{,}001$
Entre condiciones	SSIM	Dual CNN	419,4	$< 0{,}001$
Entre métodos	RMSE continuo	Ideal	592,7	$< 0{,}001$
Entre métodos	RMSE continuo	Medium	718,8	$< 0{,}001$
Entre métodos	RMSE continuo	Moderate	643,3	$< 0{,}001$
Entre métodos	RMSE continuo	VeryStrong	734,8	$< 0{,}001$
Entre condiciones	RMSE continuo	PSA3	476,4	$< 0{,}001$
Entre condiciones	RMSE continuo	PSA4	476,4	$< 0{,}001$

Contraste	Métrica	Condición o método	Estadístico de Friedman	p ajustado (Holm)
Entre condiciones	RMSE continuo	PSA12	476,4	$< 0{,}001$
Entre condiciones	RMSE continuo	LTPU	476,4	$< 0{,}001$
Entre condiciones	RMSE continuo	U-Net	382,3	$< 0{,}001$
Entre condiciones	RMSE continuo	Dual CNN	435,5	$< 0{,}001$
Entre métodos	MAE continuo	Ideal	590,0	$< 0{,}001$
Entre métodos	MAE continuo	Medium	722,3	$< 0{,}001$
Entre métodos	MAE continuo	Moderate	653,4	$< 0{,}001$
Entre métodos	MAE continuo	VeryStrong	748,8	$< 0{,}001$
Entre condiciones	MAE continuo	PSA3	460,4	$< 0{,}001$
Entre condiciones	MAE continuo	PSA4	476,4	$< 0{,}001$
Entre condiciones	MAE continuo	PSA12	476,4	$< 0{,}001$
Entre condiciones	MAE continuo	LTPU	476,4	$< 0{,}001$
Entre condiciones	MAE continuo	U-Net	382,3	$< 0{,}001$
Entre condiciones	MAE continuo	Dual CNN	456,2	$< 0{,}001$

Nota. 160 escenas como bloques, con p ajustado por Holm sobre todas las familias del Exp1. «Entre métodos» compara los seis métodos dentro de una condición; «entre condiciones» compara las cuatro condiciones dentro de un método. Las filas de fase continua corresponden al análisis secundario (RMSE y MAE continuos).

Tabla A2
Contrastes globales de Friedman del Exp2

Contraste	Métrica	Condición o método	Estadístico de Friedman	*p* ajustado (Holm)
Entre métodos	RMSE circular	Ideal	1531,3	$< 0{,}001$
Entre métodos	RMSE circular	Medium	1508,6	$< 0{,}001$
Entre métodos	RMSE circular	Moderate	1492,3	$< 0{,}001$
Entre métodos	RMSE circular	VeryStrong	1454,6	$< 0{,}001$
Entre condiciones	RMSE circular	PSA3	801,9	$< 0{,}001$
Entre condiciones	RMSE circular	PSA4	1141,5	$< 0{,}001$
Entre condiciones	RMSE circular	PSA12	1005,2	$< 0{,}001$
Entre condiciones	RMSE circular	U-Net	1033,9	$< 0{,}001$
Entre condiciones	RMSE circular	Dual CNN	1066,1	$< 0{,}001$
Entre métodos	MAE circular	Ideal	1528,2	$< 0{,}001$
Entre métodos	MAE circular	Medium	1531,2	$< 0{,}001$
Entre métodos	MAE circular	Moderate	1527,3	$< 0{,}001$
Entre métodos	MAE circular	VeryStrong	1409,6	$< 0{,}001$
Entre condiciones	MAE circular	PSA3	1046,5	$< 0{,}001$
Entre condiciones	MAE circular	PSA4	1152,0	$< 0{,}001$
Entre condiciones	MAE circular	PSA12	1132,4	$< 0{,}001$
Entre condiciones	MAE circular	U-Net	1089,6	$< 0{,}001$
Entre condiciones	MAE circular	Dual CNN	1135,5	$< 0{,}001$
Entre métodos	SSIM	Ideal	1449,5	$< 0{,}001$
Entre métodos	SSIM	Medium	1529,6	$< 0{,}001$

Contraste	Métrica	Condición o método	Estadístico de Friedman	p ajustado (Holm)
Entre métodos	SSIM	Moderate	1528,8	$< 0{,}001$
Entre métodos	SSIM	VeryStrong	1523,7	$< 0{,}001$
Entre condiciones	SSIM	PSA3	1056,8	$< 0{,}001$
Entre condiciones	SSIM	PSA4	992,8	$< 0{,}001$
Entre condiciones	SSIM	PSA12	794,9	$< 0{,}001$
Entre condiciones	SSIM	U-Net	1032,9	$< 0{,}001$
Entre condiciones	SSIM	Dual CNN	991,6	$< 0{,}001$
Entre métodos	RMSE continuo	Ideal	1131,3	$< 0{,}001$
Entre métodos	RMSE continuo	Medium	981,1	$< 0{,}001$
Entre métodos	RMSE continuo	Moderate	1011,0	$< 0{,}001$
Entre métodos	RMSE continuo	VeryStrong	1100,8	$< 0{,}001$
Entre condiciones	RMSE continuo	PSA3	560,9	$< 0{,}001$
Entre condiciones	RMSE continuo	PSA4	729,0	$< 0{,}001$
Entre condiciones	RMSE continuo	PSA12	452,2	$< 0{,}001$
Entre condiciones	RMSE continuo	U-Net	767,5	$< 0{,}001$
Entre condiciones	RMSE continuo	Dual CNN	708,0	$< 0{,}001$
Entre métodos	MAE continuo	Ideal	1388,3	$< 0{,}001$
Entre métodos	MAE continuo	Medium	1284,2	$< 0{,}001$
Entre métodos	MAE continuo	Moderate	1284,5	$< 0{,}001$
Entre métodos	MAE continuo	VeryStrong	1222,7	$< 0{,}001$
Entre condiciones	MAE continuo	PSA3	855,8	$< 0{,}001$
Entre condiciones	MAE continuo	PSA4	1006,2	$< 0{,}001$
Entre condiciones	MAE continuo	PSA12	927,6	$< 0{,}001$

Contraste	Métrica	Condición o método	Estadístico de Friedman	p ajustado (Holm)
Entre condiciones	MAE continuo	U-Net	942,7	$< 0{,}001$
Entre condiciones	MAE continuo	Dual CNN	866,7	$< 0{,}001$
Entre métodos	Error de orden E_{orden}	Ideal	937,0	$< 0{,}001$
Entre métodos	Error de orden E_{orden}	Medium	178,6	$< 0{,}001$
Entre métodos	Error de orden E_{orden}	Moderate	237,7	$< 0{,}001$
Entre métodos	Error de orden E_{orden}	VeryStrong	877,8	$< 0{,}001$
Entre condiciones	Error de orden E_{orden}	PSA3	238,0	$< 0{,}001$
Entre condiciones	Error de orden E_{orden}	PSA4	111,4	$< 0{,}001$
Entre condiciones	Error de orden E_{orden}	PSA12	25,8	$< 0{,}001$
Entre condiciones	Error de orden E_{orden}	U-Net	521,2	$< 0{,}001$
Entre condiciones	Error de orden E_{orden}	Dual CNN	563,3	$< 0{,}001$
Entre métodos	Cobertura C	Ideal	226,4	$< 0{,}001$
Entre métodos	Cobertura C	Medium	129,4	$< 0{,}001$
Entre métodos	Cobertura C	Moderate	104,6	$< 0{,}001$

Contraste	Métrica	Condición o método	Estadístico de Friedman	p ajustado (Holm)
Entre métodos	Cobertura C	VeryStrong	77,7	$< 0{,}001$
Entre condiciones	Cobertura C	PSA3	91,7	$< 0{,}001$
Entre condiciones	Cobertura C	PSA4	25,6	$< 0{,}001$
Entre condiciones	Cobertura C	PSA12	9,8	0,060
Entre condiciones	Cobertura C	U-Net	116,0	$< 0{,}001$
Entre condiciones	Cobertura C	Dual CNN	13,3	0,020
Entre métodos	Cobertura correcta C_{correcta}	Ideal	30,4	$< 0{,}001$
Entre métodos	Cobertura correcta C_{correcta}	Medium	31,9	$< 0{,}001$
Entre métodos	Cobertura correcta C_{correcta}	Moderate	48,4	$< 0{,}001$
Entre métodos	Cobertura correcta C_{correcta}	VeryStrong	308,4	$< 0{,}001$
Entre condiciones	Cobertura correcta C_{correcta}	PSA3	3,5	0,642
Entre condiciones	Cobertura correcta C_{correcta}	PSA4	11,1	0,044
Entre condiciones	Cobertura correcta C_{correcta}	PSA12	2,2	0,642
Entre condiciones	Cobertura correcta C_{correcta}	U-Net	38,2	$< 0{,}001$

Contraste	Métrica	Condición o método	Estadístico de Friedman	p ajustado (Holm)
Entre condiciones	Cobertura correcta C_{correcta}	Dual CNN	353,5	$< 0{,}001$

Nota. 384 escenas como bloques, con p ajustado por Holm sobre todas las familias del Exp2. «Entre métodos» compara los cinco métodos dentro de una condición; «entre condiciones» compara las cuatro condiciones dentro de un método. Las filas de fase continua incluyen RMSE, MAE, error de orden y coberturas.

Tabla A3

Comparaciones pareadas de Wilcoxon del Exp1 para la fase discontinua

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r
RMSE circular, condición Moderate				
PSA3 frente a PSA4	PSA4	0,1029	$< 0{,}001$	0,54
PSA3 frente a PSA12	PSA12	0,1897	$< 0{,}001$	1,00
PSA3 frente a LTPU	PSA3	−0,3063	$< 0{,}001$	−0,57
PSA3 frente a U-Net	U-Net	0,1759	$< 0{,}001$	1,00
PSA3 frente a Dual CNN	Dual CNN	0,1737	$< 0{,}001$	1,00
PSA4 frente a PSA12	PSA12	0,0868	$< 0{,}001$	1,00
PSA4 frente a LTPU	PSA4	−0,4080	$< 0{,}001$	−1,00
PSA4 frente a U-Net	U-Net	0,0449	$< 0{,}001$	1,00
PSA4 frente a Dual CNN	Dual CNN	0,0396	$< 0{,}001$	0,99
PSA12 frente a LTPU	PSA12	−0,4960	$< 0{,}001$	−1,00
PSA12 frente a U-Net	PSA12	−0,0016	$< 0{,}001$	−0,43
PSA12 frente a Dual CNN	PSA12	−0,0043	$< 0{,}001$	−0,77

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r
LTPU frente a U-Net	U-Net	0,4828	$< 0{,}001$	1,00
LTPU frente a Dual CNN	Dual CNN	0,4373	$< 0{,}001$	1,00
U-Net frente a Dual CNN	U-Net	$-0{,}0033$	$< 0{,}001$	$-0{,}55$
MAE circular, condición Moderate				
PSA3 frente a PSA4	PSA4	0,0996	$< 0{,}001$	0,55
PSA3 frente a PSA12	PSA12	0,1649	$< 0{,}001$	1,00
PSA3 frente a LTPU	PSA3	$-0{,}2137$	$< 0{,}001$	$-0{,}50$
PSA3 frente a U-Net	U-Net	0,1637	$< 0{,}001$	1,00
PSA3 frente a Dual CNN	Dual CNN	0,1627	$< 0{,}001$	1,00
PSA4 frente a PSA12	PSA12	0,0634	$< 0{,}001$	1,00
PSA4 frente a LTPU	PSA4	$-0{,}3071$	$< 0{,}001$	$-1{,}00$
PSA4 frente a U-Net	U-Net	0,0641	$< 0{,}001$	1,00
PSA4 frente a Dual CNN	Dual CNN	0,0625	$< 0{,}001$	1,00
PSA12 frente a LTPU	PSA12	$-0{,}3786$	$< 0{,}001$	$-1{,}00$
PSA12 frente a U-Net	U-Net	0,0017	$< 0{,}001$	0,48
PSA12 frente a Dual CNN	Dual CNN	0,0006	0,276	0,10
LTPU frente a U-Net	U-Net	0,3789	$< 0{,}001$	1,00
LTPU frente a Dual CNN	Dual CNN	0,3789	$< 0{,}001$	1,00
U-Net frente a Dual CNN	U-Net	$-0{,}0007$	$< 0{,}001$	$-0{,}38$
SSIM, condición Moderate				
PSA3 frente a PSA4	PSA4	$-0{,}0812$	$< 0{,}001$	$-0{,}62$
PSA3 frente a PSA12	PSA12	$-0{,}1103$	$< 0{,}001$	$-1{,}00$
PSA3 frente a LTPU	PSA3	0,3311	$< 0{,}001$	0,98

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r
PSA3 frente a U-Net	U-Net	−0,1070	< 0,001	−1,00
PSA3 frente a Dual CNN	Dual CNN	−0,1077	< 0,001	−1,00
PSA4 frente a PSA12	PSA12	−0,0255	< 0,001	−1,00
PSA4 frente a LTPU	PSA4	0,4048	< 0,001	1,00
PSA4 frente a U-Net	U-Net	−0,0249	< 0,001	−0,97
PSA4 frente a Dual CNN	Dual CNN	−0,0238	< 0,001	−0,93
PSA12 frente a LTPU	PSA12	0,4303	< 0,001	1,00
PSA12 frente a U-Net	PSA12	0,0008	< 0,001	1,00
PSA12 frente a Dual CNN	PSA12	0,0017	< 0,001	0,99
LTPU frente a U-Net	U-Net	−0,4297	< 0,001	−1,00
LTPU frente a Dual CNN	Dual CNN	−0,4287	< 0,001	−1,00
U-Net frente a Dual CNN	U-Net	0,0009	< 0,001	0,56
RMSE circular, condición VeryStrong				
PSA3 frente a PSA4	PSA4	0,1789	< 0,001	1,00
PSA3 frente a PSA12	PSA12	0,2897	< 0,001	1,00
PSA3 frente a LTPU	PSA3	−0,4064	< 0,001	−1,00
PSA3 frente a U-Net	PSA3	−0,4117	< 0,001	−0,83
PSA3 frente a Dual CNN	PSA3	−0,2725	< 0,001	−0,59
PSA4 frente a PSA12	PSA12	0,1364	< 0,001	1,00
PSA4 frente a LTPU	PSA4	−0,5812	< 0,001	−1,00
PSA4 frente a U-Net	PSA4	−0,5670	< 0,001	−1,00
PSA4 frente a Dual CNN	PSA4	−0,4277	< 0,001	−1,00
PSA12 frente a LTPU	PSA12	−0,7200	< 0,001	−1,00

Par comparado	Método favorecido por la mediana	Δ mediana	*p* ajustado (Holm)	*r*
PSA12 frente a U-Net	PSA12	−0,7100	< 0,001	−1,00
PSA12 frente a Dual CNN	PSA12	−0,5631	< 0,001	−1,00
LTPU frente a U-Net	U-Net	0,0215	0,008	−0,24
LTPU frente a Dual CNN	Dual CNN	0,1362	< 0,001	0,97
U-Net frente a Dual CNN	Dual CNN	0,1086	< 0,001	1,00
MAE circular, condición VeryStrong				
PSA3 frente a PSA4	PSA4	0,1598	< 0,001	1,00
PSA3 frente a PSA12	PSA12	0,2205	< 0,001	1,00
PSA3 frente a LTPU	PSA3	−0,2932	< 0,001	−0,84
PSA3 frente a U-Net	PSA3	−0,0987	< 0,001	−0,50
PSA3 frente a Dual CNN	PSA3	−0,0682	< 0,001	−0,50
PSA4 frente a PSA12	PSA12	0,0875	< 0,001	1,00
PSA4 frente a LTPU	PSA4	−0,4317	< 0,001	−1,00
PSA4 frente a U-Net	PSA4	−0,2163	< 0,001	−0,84
PSA4 frente a Dual CNN	PSA4	−0,1824	< 0,001	−0,98
PSA12 frente a LTPU	PSA12	−0,5356	< 0,001	−1,00
PSA12 frente a U-Net	PSA12	−0,3226	< 0,001	−1,00
PSA12 frente a Dual CNN	PSA12	−0,2863	< 0,001	−1,00
LTPU frente a U-Net	U-Net	0,1707	< 0,001	0,94
LTPU frente a Dual CNN	Dual CNN	0,1720	< 0,001	1,00
U-Net frente a Dual CNN	Dual CNN	0,0095	< 0,001	0,50
SSIM, condición VeryStrong				
PSA3 frente a PSA4	PSA4	−0,1410	< 0,001	−1,00

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r
PSA3 frente a PSA12	PSA12	−0,1773	< 0,001	−1,00
PSA3 frente a LTPU	PSA3	0,2976	< 0,001	1,00
PSA3 frente a U-Net	PSA3	0,0744	< 0,001	0,50
PSA3 frente a Dual CNN	PSA3	0,0461	< 0,001	0,50
PSA4 frente a PSA12	PSA12	−0,0517	< 0,001	−1,00
PSA4 frente a LTPU	PSA4	0,4842	< 0,001	1,00
PSA4 frente a U-Net	PSA4	0,2069	< 0,001	1,00
PSA4 frente a Dual CNN	PSA4	0,1768	< 0,001	1,00
PSA12 frente a LTPU	PSA12	0,5450	< 0,001	1,00
PSA12 frente a U-Net	PSA12	0,2586	< 0,001	1,00
PSA12 frente a Dual CNN	PSA12	0,2285	< 0,001	1,00
LTPU frente a U-Net	U-Net	−0,2072	< 0,001	−1,00
LTPU frente a Dual CNN	Dual CNN	−0,2238	< 0,001	−1,00
U-Net frente a Dual CNN	Dual CNN	−0,0168	< 0,001	−0,50

Nota. Condiciones Moderate y VeryStrong, sobre 160 escenas. Δ es la mediana de las diferencias por escena (primer método menos segundo, en rad para RMSE y MAE; sin unidad para SSIM); p está ajustado por Holm dentro de cada familia; r es el rango biserial, con el signo de Δ.

Tabla A4

Comparaciones pareadas de Wilcoxon del Exp1 para la fase continua

Par comparado	Método favorecido por la mediana	Δ mediana	*p* ajustado (Holm)	*r*	IC 95 % de Δ mediana
RMSE continuo, condición Moderate					
PSA3 frente a PSA4	PSA4	0,1029	$< 0{,}001$	0,54	[0,0039; 0,1951]
PSA3 frente a PSA12	PSA12	0,1897	$< 0{,}001$	1,00	[0,1407; 0,2323]
PSA3 frente a LTPU	PSA3	−0,3063	$< 0{,}001$	−0,57	[−0,6191; −0,0038]
PSA3 frente a U-Net	U-Net	0,1759	$< 0{,}001$	0,99	[0,1383; 0,2303]
PSA3 frente a Dual CNN	Dual CNN	0,1737	$< 0{,}001$	0,94	[0,1361; 0,2245]
PSA4 frente a PSA12	PSA12	0,0868	$< 0{,}001$	1,00	[0,0377; 0,1375]
PSA4 frente a LTPU	PSA4	−0,4080	$< 0{,}001$	−1,00	[−0,6207; −0,2018]
PSA4 frente a U-Net	U-Net	0,0443	$< 0{,}001$	0,98	[0,0402; 0,1187]
PSA4 frente a Dual CNN	Dual CNN	0,0390	$< 0{,}001$	0,93	[0,0349; 0,0665]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA12 frente a LTPU	PSA12	−0,4960	$< 0{,}001$	−1,00	[−0,7594; −0,2399]
PSA12 frente a U-Net	PSA12	−0,0016	$< 0{,}001$	−0,43	[−0,0034; 0,0002]
PSA12 frente a Dual CNN	PSA12	−0,0043	$< 0{,}001$	−0,77	[−0,0070; −0,0026]
LTPU frente a U-Net	U-Net	0,4628	$< 0{,}001$	1,00	[0,2425; 0,7412]
LTPU frente a Dual CNN	Dual CNN	0,2995	$< 0{,}001$	1,00	[0,2392; 0,7242]
U-Net frente a Dual CNN	U-Net	−0,0033	$< 0{,}001$	−0,51	[−0,0040; −0,0026]
MAE continuo, condición Moderate					
PSA3 frente a PSA4	PSA4	0,0996	$< 0{,}001$	0,55	[0,0046; 0,1844]
PSA3 frente a PSA12	PSA12	0,1649	$< 0{,}001$	1,00	[0,1170; 0,2085]
PSA3 frente a LTPU	PSA3	−0,2137	$< 0{,}001$	−0,50	[−0,5037; 0,0442]
PSA3 frente a U-Net	U-Net	0,1637	$< 0{,}001$	1,00	[0,1168; 0,2098]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA3 frente a Dual CNN	Dual CNN	0,1627	$< 0{,}001$	1,00	[0,1188; 0,2073]
PSA4 frente a PSA12	PSA12	0,0634	$< 0{,}001$	1,00	[0,0243; 0,1092]
PSA4 frente a LTPU	PSA4	$-0{,}3071$	$< 0{,}001$	$-1{,}00$	[$-0{,}5023$; $-0{,}1409$]
PSA4 frente a U-Net	U-Net	0,0641	$< 0{,}001$	1,00	[0,0265; 0,1089]
PSA4 frente a Dual CNN	Dual CNN	0,0582	$< 0{,}001$	1,00	[0,0254; 0,1089]
PSA12 frente a LTPU	PSA12	$-0{,}3786$	$< 0{,}001$	$-1{,}00$	[$-0{,}6203$; $-0{,}1645$]
PSA12 frente a U-Net	U-Net	0,0017	$< 0{,}001$	0,43	[0,0010; 0,0020]
PSA12 frente a Dual CNN	Dual CNN	0,0006	0,420	0,07	[0,0000; 0,0013]
LTPU frente a U-Net	U-Net	0,3789	$< 0{,}001$	1,00	[0,1650; 0,6193]
LTPU frente a Dual CNN	Dual CNN	0,3789	$< 0{,}001$	1,00	[0,1635; 0,6198]
U-Net frente a Dual CNN	U-Net	$-0{,}0007$	$< 0{,}001$	$-0{,}36$	[$-0{,}0013$; $-0{,}0001$]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
RMSE continuo, condición VeryStrong					
PSA3 frente a PSA4	PSA4	1,0348	$< 0{,}001$	1,00	[0,7815; 1,7876]
PSA3 frente a PSA12	PSA12	1,4422	$< 0{,}001$	1,00	[0,8739; 2,4846]
PSA3 frente a LTPU	LTPU	0,3879	$< 0{,}001$	0,52	[0,3364; 0,4240]
PSA3 frente a U-Net	PSA3	$-6{,}6645$	$< 0{,}001$	$-1{,}00$	[$-13{,}2395$; $-2{,}7481$]
PSA3 frente a Dual CNN	PSA3	$-5{,}4420$	$< 0{,}001$	$-1{,}00$	[$-12{,}6160$; $-1{,}3637$]
PSA4 frente a PSA12	PSA12	0,4720	$< 0{,}001$	1,00	[0,3937; 0,6494]
PSA4 frente a LTPU	PSA4	$-0{,}3322$	$< 0{,}001$	$-0{,}89$	[$-0{,}3880$; $-0{,}3014$]
PSA4 frente a U-Net	PSA4	$-8{,}4162$	$< 0{,}001$	$-1{,}00$	[$-16{,}1145$; $-3{,}1281$]
PSA4 frente a Dual CNN	PSA4	$-7{,}3807$	$< 0{,}001$	$-1{,}00$	[$-15{,}9683$; $-1{,}9839$]
PSA12 frente a LTPU	PSA12	$-0{,}6766$	$< 0{,}001$	$-1{,}00$	[$-1{,}5566$; $-0{,}4413$]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA12 frente a U-Net	PSA12	−9,1796	$< 0{,}001$	−1,00	[−17,5412; −3,5143]
PSA12 frente a Dual CNN	PSA12	−8,3669	$< 0{,}001$	−1,00	[−17,1660; −2,1625]
LTPU frente a U-Net	LTPU	−7,8702	$< 0{,}001$	−1,00	[−14,8635; −3,0881]
LTPU frente a Dual CNN	LTPU	−7,3795	$< 0{,}001$	−1,00	[−13,8466; −1,7394]
U-Net frente a Dual CNN	Dual CNN	1,0700	$< 0{,}001$	0,65	[0,8680; 1,2677]
MAE continuo, condición VeryStrong					
PSA3 frente a PSA4	PSA4	0,3637	$< 0{,}001$	1,00	[0,2752; 0,6055]
PSA3 frente a PSA12	PSA12	0,5017	$< 0{,}001$	1,00	[0,3228; 0,8250]
PSA3 frente a LTPU	LTPU	0,0694	$< 0{,}001$	0,51	[0,0535; 0,0819]
PSA3 frente a U-Net	PSA3	−4,8967	$< 0{,}001$	−0,99	[−10,2302; −0,5838]
PSA3 frente a Dual CNN	PSA3	−4,7927	$< 0{,}001$	−0,98	[−9,9156; −0,2931]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA4 frente a PSA12	PSA12	0,1277	$< 0{,}001$	1,00	[0,0651; 0,2190]
PSA4 frente a LTPU	PSA4	$-0{,}2368$	$< 0{,}001$	$-1{,}00$	[$-0{,}4872$; $-0{,}2084$]
PSA4 frente a U-Net	PSA4	$-5{,}4546$	$< 0{,}001$	$-1{,}00$	[$-11{,}4117$; $-0{,}8252$]
PSA4 frente a Dual CNN	PSA4	$-5{,}1815$	$< 0{,}001$	$-1{,}00$	[$-10{,}9808$; $-0{,}5638$]
PSA12 frente a LTPU	PSA12	$-0{,}5633$	$< 0{,}001$	$-1{,}00$	[$-0{,}8867$; $-0{,}2620$]
PSA12 frente a U-Net	PSA12	$-5{,}6018$	$< 0{,}001$	$-1{,}00$	[$-11{,}6188$; $-0{,}9003$]
PSA12 frente a Dual CNN	PSA12	$-5{,}3123$	$< 0{,}001$	$-1{,}00$	[$-11{,}2647$; $-0{,}6148$]
LTPU frente a U-Net	LTPU	$-4{,}8566$	$< 0{,}001$	$-1{,}00$	[$-10{,}4892$; $-0{,}6220$]
LTPU frente a Dual CNN	LTPU	$-4{,}5255$	$< 0{,}001$	$-0{,}99$	[$-10{,}1539$; $-0{,}3491$]
U-Net frente a Dual CNN	Dual CNN	0,2485	$< 0{,}001$	0,70	[0,1953; 0,3437]

Nota. Condiciones Moderate y VeryStrong, sobre 160 escenas, como análisis secundario. Δ es la mediana de las diferencias por escena (primer método menos segundo, en rad), con su intervalo de confianza del 95 % por *bootstrap*; *p* está ajustado por Holm dentro de cada familia; *r* es el rango biserial.

Tabla A5

Comparaciones pareadas de Wilcoxon del Exp2 para la fase discontinua

Par comparado	Método favorecido por la mediana	Δ mediana	*p* ajustado (Holm)	*r*
RMSE circular, condición Moderate				
PSA3 frente a PSA4	PSA4	0,1927	$< 0{,}001$	1,00
PSA3 frente a PSA12	PSA12	0,2279	$< 0{,}001$	1,00
PSA3 frente a U-Net	U-Net	0,1415	$< 0{,}001$	1,00
PSA3 frente a Dual CNN	Dual CNN	0,0822	$< 0{,}001$	0,95
PSA4 frente a PSA12	PSA12	0,0350	$< 0{,}001$	1,00
PSA4 frente a U-Net	PSA4	−0,0512	$< 0{,}001$	−0,97
PSA4 frente a Dual CNN	PSA4	−0,1069	$< 0{,}001$	−1,00
PSA12 frente a U-Net	PSA12	−0,0864	$< 0{,}001$	−1,00
PSA12 frente a Dual CNN	PSA12	−0,1438	$< 0{,}001$	−1,00
U-Net frente a Dual CNN	U-Net	−0,0557	$< 0{,}001$	−0,99
MAE circular, condición Moderate				
PSA3 frente a PSA4	PSA4	0,1869	$< 0{,}001$	1,00
PSA3 frente a PSA12	PSA12	0,2099	$< 0{,}001$	1,00
PSA3 frente a U-Net	U-Net	0,1666	$< 0{,}001$	1,00
PSA3 frente a Dual CNN	Dual CNN	0,1217	$< 0{,}001$	1,00
PSA4 frente a PSA12	PSA12	0,0231	$< 0{,}001$	1,00
PSA4 frente a U-Net	PSA4	−0,0196	$< 0{,}001$	−0,96
PSA4 frente a Dual CNN	PSA4	−0,0642	$< 0{,}001$	−1,00

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r
PSA12 frente a U-Net	PSA12	−0,0431	< 0,001	−1,00
PSA12 frente a Dual CNN	PSA12	−0,0877	< 0,001	−1,00
U-Net frente a Dual CNN	U-Net	−0,0443	< 0,001	−1,00
SSIM, condición Moderate				
PSA3 frente a PSA4	PSA4	−0,0103	< 0,001	−1,00
PSA3 frente a PSA12	PSA12	−0,0197	< 0,001	−1,00
PSA3 frente a U-Net	PSA3	0,0254	< 0,001	0,99
PSA3 frente a Dual CNN	PSA3	0,0719	< 0,001	1,00
PSA4 frente a PSA12	PSA12	−0,0091	< 0,001	−1,00
PSA4 frente a U-Net	PSA4	0,0351	< 0,001	1,00
PSA4 frente a Dual CNN	PSA4	0,0823	< 0,001	1,00
PSA12 frente a U-Net	PSA12	0,0459	< 0,001	1,00
PSA12 frente a Dual CNN	PSA12	0,0924	< 0,001	1,00
U-Net frente a Dual CNN	U-Net	0,0456	< 0,001	1,00
RMSE circular, condición VeryStrong				
PSA3 frente a PSA4	PSA4	0,1609	< 0,001	1,00
PSA3 frente a PSA12	PSA12	0,2209	< 0,001	1,00
PSA3 frente a U-Net	U-Net	0,0493	0,248	0,07
PSA3 frente a Dual CNN	PSA3	−0,2247	< 0,001	−1,00
PSA4 frente a PSA12	PSA12	0,0605	< 0,001	1,00
PSA4 frente a U-Net	PSA4	−0,1069	< 0,001	−1,00
PSA4 frente a Dual CNN	PSA4	−0,3815	< 0,001	−1,00
PSA12 frente a U-Net	PSA12	−0,1698	< 0,001	−1,00

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r
PSA12 frente a Dual CNN	PSA12	−0,4450	$< 0{,}001$	−1,00
U-Net frente a Dual CNN	U-Net	−0,2464	$< 0{,}001$	−1,00
MAE circular, condición VeryStrong				
PSA3 frente a PSA4	PSA4	0,1637	$< 0{,}001$	1,00
PSA3 frente a PSA12	PSA12	0,2023	$< 0{,}001$	1,00
PSA3 frente a U-Net	U-Net	0,1106	$< 0{,}001$	0,84
PSA3 frente a Dual CNN	PSA3	−0,0362	$< 0{,}001$	−0,50
PSA4 frente a PSA12	PSA12	0,0379	$< 0{,}001$	1,00
PSA4 frente a U-Net	PSA4	−0,0514	$< 0{,}001$	−1,00
PSA4 frente a Dual CNN	PSA4	−0,1964	$< 0{,}001$	−1,00
PSA12 frente a U-Net	PSA12	−0,0921	$< 0{,}001$	−1,00
PSA12 frente a Dual CNN	PSA12	−0,2357	$< 0{,}001$	−1,00
U-Net frente a Dual CNN	U-Net	−0,1378	$< 0{,}001$	−1,00
SSIM, condición VeryStrong				
PSA3 frente a PSA4	PSA4	−0,0321	$< 0{,}001$	−1,00
PSA3 frente a PSA12	PSA12	−0,0465	$< 0{,}001$	−1,00
PSA3 frente a U-Net	PSA3	0,0297	$< 0{,}001$	0,94
PSA3 frente a Dual CNN	PSA3	0,1581	$< 0{,}001$	1,00
PSA4 frente a PSA12	PSA12	−0,0141	$< 0{,}001$	−1,00
PSA4 frente a U-Net	PSA4	0,0621	$< 0{,}001$	1,00
PSA4 frente a Dual CNN	PSA4	0,1923	$< 0{,}001$	1,00
PSA12 frente a U-Net	PSA12	0,0781	$< 0{,}001$	1,00
PSA12 frente a Dual CNN	PSA12	0,2090	$< 0{,}001$	1,00

Par comparado	Método favorecido por la mediana	Δ mediana	*p* ajustado (Holm)	*r*
U-Net frente a Dual CNN	U-Net	0,1223	$< 0{,}001$	1,00

Nota. Condiciones Moderate y VeryStrong, sobre 384 escenas. Δ es la mediana de las diferencias por escena (primer método menos segundo, en rad para RMSE y MAE; sin unidad para SSIM); *p* está ajustado por Holm dentro de cada familia; *r* es el rango biserial, con el signo de Δ.

Tabla A6

Comparaciones pareadas de Wilcoxon del Exp2 para la fase continua

Par comparado	Método favorecido por la mediana	Δ mediana	*p* ajustado (Holm)	*r*	IC 95 % de Δ mediana
RMSE continuo, condición Moderate					
PSA3 frente a PSA4	PSA4	0,1816	$< 0{,}001$	0,91	[0,1743; 0,1867]
PSA3 frente a PSA12	PSA12	0,2209	$< 0{,}001$	0,89	[0,2078; 0,2238]
PSA3 frente a U-Net	U-Net	0,1118	$< 0{,}001$	0,71	[0,1027; 0,1211]
PSA3 frente a Dual CNN	Dual CNN	0,0439	$< 0{,}001$	0,44	[0,0360; 0,0536]
PSA4 frente a PSA12	PSA12	0,0219	$< 0{,}001$	0,82	[0,0176; 0,0265]
PSA4 frente a U-Net	PSA4	$-0{,}0459$	$< 0{,}001$	$-0{,}68$	[$-0{,}0533$; $-0{,}0383$]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA4 frente a Dual CNN	PSA4	−0,1144	$< 0{,}001$	−0,86	[−0,1271; −0,1000]
PSA12 frente a U-Net	PSA12	−0,0780	$< 0{,}001$	−0,83	[−0,0859; −0,0654]
PSA12 frente a Dual CNN	PSA12	−0,1487	$< 0{,}001$	−0,87	[−0,1603; −0,1373]
U-Net frente a Dual CNN	U-Net	−0,0539	$< 0{,}001$	−0,67	[−0,0611; −0,0486]
MAE continuo, condición Moderate					
PSA3 frente a PSA4	PSA4	0,1861	$< 0{,}001$	0,96	[0,1853; 0,1867]
PSA3 frente a PSA12	PSA12	0,2094	$< 0{,}001$	0,94	[0,2089; 0,2097]
PSA3 frente a U-Net	U-Net	0,1557	$< 0{,}001$	0,81	[0,1529; 0,1610]
PSA3 frente a Dual CNN	Dual CNN	0,1096	$< 0{,}001$	0,71	[0,1020; 0,1155]
PSA4 frente a PSA12	PSA12	0,0228	$< 0{,}001$	0,88	[0,0226; 0,0231]
PSA4 frente a U-Net	PSA4	−0,0275	$< 0{,}001$	−0,75	[−0,0316; −0,0238]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA4 frente a Dual CNN	PSA4	−0,0805	$< 0{,}001$	−0,89	[−0,0866; −0,0747]
PSA12 frente a U-Net	PSA12	−0,0525	$< 0{,}001$	−0,90	[−0,0562; −0,0490]
PSA12 frente a Dual CNN	PSA12	−0,1082	$< 0{,}001$	−0,90	[−0,1140; −0,1002]
U-Net frente a Dual CNN	U-Net	−0,0464	$< 0{,}001$	−0,78	[−0,0483; −0,0444]
Error de orden E_{orden}, condición Moderate					
PSA3 frente a PSA4	PSA4	0,0020	$< 0{,}001$	0,52	[0,0014; 0,0029]
PSA3 frente a PSA12	PSA12	0,0023	$< 0{,}001$	0,54	[0,0017; 0,0038]
PSA3 frente a U-Net	sin diferencia	0,0000	0,743	−0,02	[0,0000; 0,0000]
PSA3 frente a Dual CNN	sin diferencia	0,0000	0,001	−0,22	[−0,0019; 0,0000]
PSA4 frente a PSA12	sin diferencia	0,0000	$< 0{,}001$	0,28	[0,0000; 0,0000]
PSA4 frente a U-Net	PSA4	−0,0019	$< 0{,}001$	−0,46	[−0,0031; −0,0003]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA4 frente a Dual CNN	PSA4	−0,0061	$< 0{,}001$	−0,64	[−0,0138; −0,0032]
PSA12 frente a U-Net	PSA12	−0,0036	$< 0{,}001$	−0,63	[−0,0063; −0,0018]
PSA12 frente a Dual CNN	PSA12	−0,0077	$< 0{,}001$	−0,71	[−0,0177; −0,0040]
U-Net frente a Dual CNN	sin diferencia	0,0000	$< 0{,}001$	−0,25	[−0,0018; 0,0000]
Cobertura correcta C_{correcta}, condición Moderate					
PSA3 frente a PSA4	PSA3	0,0047	0,053	0,15	[0,0000; 0,0089]
PSA3 frente a PSA12	PSA3	0,0035	0,019	0,17	[−0,0008; 0,0116]
PSA3 frente a U-Net	PSA3	0,0426	$< 0{,}001$	0,39	[0,0257; 0,0759]
PSA3 frente a Dual CNN	PSA3	0,0190	0,206	0,11	[0,0091; 0,0295]
PSA4 frente a PSA12	PSA12	−0,0056	0,904	−0,02	[−0,0086; −0,0018]
PSA4 frente a U-Net	PSA4	0,0254	$< 0{,}001$	0,31	[0,0099; 0,0385]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA4 frente a Dual CNN	PSA4	0,0086	0,904	−0,04	[−0,0022; 0,0217]
PSA12 frente a U-Net	PSA12	0,0188	$< 0{,}001$	0,26	[0,0096; 0,0257]
PSA12 frente a Dual CNN	PSA12	0,0074	0,127	−0,13	[−0,0106; 0,0190]
U-Net frente a Dual CNN	Dual CNN	−0,0293	$< 0{,}001$	−0,33	[−0,0788; −0,0079]
RMSE continuo, condición VeryStrong					
PSA3 frente a PSA4	PSA4	0,1425	$< 0{,}001$	0,90	[0,1377; 0,1487]
PSA3 frente a PSA12	PSA12	0,2130	$< 0{,}001$	0,87	[0,2028; 0,2170]
PSA3 frente a U-Net	PSA3	−0,0077	$< 0{,}001$	−0,23	[−0,0309; 0,0071]
PSA3 frente a Dual CNN	PSA3	−0,3894	$< 0{,}001$	−0,80	[−0,4875; −0,2951]
PSA4 frente a PSA12	PSA12	0,0462	$< 0{,}001$	0,86	[0,0400; 0,0502]
PSA4 frente a U-Net	PSA4	−0,1560	$< 0{,}001$	−0,87	[−0,1850; −0,1224]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA4 frente a Dual CNN	PSA4	−0,5748	$< 0{,}001$	−0,93	[−0,6999; −0,4895]
PSA12 frente a U-Net	PSA12	−0,2179	$< 0{,}001$	−0,91	[−0,2604; −0,1806]
PSA12 frente a Dual CNN	PSA12	−0,6496	$< 0{,}001$	−0,93	[−0,8104; −0,5483]
U-Net frente a Dual CNN	U-Net	−0,3417	$< 0{,}001$	−0,78	[−0,3867; −0,2808]
MAE continuo, condición VeryStrong					
PSA3 frente a PSA4	PSA4	0,1594	$< 0{,}001$	0,94	[0,1570; 0,1629]
PSA3 frente a PSA12	PSA12	0,2011	$< 0{,}001$	0,92	[0,1998; 0,2027]
PSA3 frente a U-Net	U-Net	0,0892	$< 0{,}001$	0,42	[0,0743; 0,0986]
PSA3 frente a Dual CNN	PSA3	−0,1083	$< 0{,}001$	−0,61	[−0,1650; −0,0712]
PSA4 frente a PSA12	PSA12	0,0394	$< 0{,}001$	0,89	[0,0383; 0,0413]
PSA4 frente a U-Net	PSA4	−0,0840	$< 0{,}001$	−0,92	[−0,0961; −0,0683]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA4 frente a Dual CNN	PSA4	−0,2856	$< 0{,}001$	−0,95	[−0,3557; −0,2500]
PSA12 frente a U-Net	PSA12	−0,1283	$< 0{,}001$	−0,94	[−0,1424; −0,1137]
PSA12 frente a Dual CNN	PSA12	−0,3387	$< 0{,}001$	−0,95	[−0,4139; −0,2939]
U-Net frente a Dual CNN	U-Net	−0,1806	$< 0{,}001$	−0,80	[−0,2068; −0,1661]
Error de orden E_{orden}, condición VeryStrong					
PSA3 frente a PSA4	PSA4	0,0056	$< 0{,}001$	0,61	[0,0038; 0,0076]
PSA3 frente a PSA12	PSA12	0,0079	$< 0{,}001$	0,67	[0,0061; 0,0120]
PSA3 frente a U-Net	PSA3	−0,0600	$< 0{,}001$	−0,51	[−0,0992; −0,0367]
PSA3 frente a Dual CNN	PSA3	−1,0601	$< 0{,}001$	−0,74	[−1,7385; −0,6407]
PSA4 frente a PSA12	PSA12	0,0011	$< 0{,}001$	0,44	[0,0002; 0,0017]
PSA4 frente a U-Net	PSA4	−0,1173	$< 0{,}001$	−0,82	[−0,1662; −0,0828]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA4 frente a Dual CNN	PSA4	−1,6844	$< 0{,}001$	−0,92	[−2,8112; −1,2484]
PSA12 frente a U-Net	PSA12	−0,1530	$< 0{,}001$	−0,87	[−0,2515; −0,1059]
PSA12 frente a Dual CNN	PSA12	−1,9822	$< 0{,}001$	−0,91	[−3,0559; −1,3200]
U-Net frente a Dual CNN	U-Net	−0,8527	$< 0{,}001$	−0,69	[−1,3200; −0,5887]
Cobertura correcta C_{correcta}, condición VeryStrong					
PSA3 frente a PSA4	PSA3	0,0022	0,673	0,06	[−0,0034; 0,0068]
PSA3 frente a PSA12	PSA3	0,0020	0,011	0,17	[−0,0056; 0,0095]
PSA3 frente a U-Net	PSA3	0,1676	$< 0{,}001$	0,49	[0,1105; 0,2330]
PSA3 frente a Dual CNN	PSA3	0,7870	$< 0{,}001$	0,61	[0,5152; 1,5272]
PSA4 frente a PSA12	PSA12	−0,0050	0,673	0,03	[−0,0082; −0,0000]
PSA4 frente a U-Net	PSA4	0,1265	$< 0{,}001$	0,45	[0,0775; 0,2194]

Par comparado	Método favorecido por la mediana	Δ mediana	p ajustado (Holm)	r	IC 95 % de Δ mediana
PSA4 frente a Dual CNN	PSA4	0,7601	$< 0{,}001$	0,65	[0,4221; 1,5316]
PSA12 frente a U-Net	PSA12	0,0592	$< 0{,}001$	0,28	[0,0457; 0,1146]
PSA12 frente a Dual CNN	PSA12	0,6140	$< 0{,}001$	0,53	[0,4040; 1,1668]
U-Net frente a Dual CNN	U-Net	0,4203	$< 0{,}001$	0,45	[0,2868; 0,7151]

Nota. Condiciones Moderate y VeryStrong, sobre 384 escenas. Δ es la mediana de las diferencias por escena (primer método menos segundo; en rad para RMSE y MAE, en puntos porcentuales para E_{orden} y C_{correcta}), con su intervalo de confianza del 95 % por *bootstrap*; p está ajustado por Holm dentro de cada familia; r es el rango biserial.