

## GENERACIÓN DE IMÁGENES HOLOGRÁFICAS SINTÉTICAS CON DESPLAZAMIENTO DE FASE MEDIANTE AUTOENCODERS Y PREPROCESAMIENTO DE RUIDO CON MATCHING PURSUIT

María del Rocío Hernández Flores<sup>1</sup>, Bárbara Emma Sánchez Rinza<sup>2</sup>, Mario Rossainz Lopez<sup>1</sup>

Benemérita Universidad Autónoma de Puebla

Faculta de Ciencias de la Computación<sup>1</sup>, Faculta de Ciencias Físico Matemáticas<sup>2</sup>

Ciudad Universitaria

hf223470495@alm.buap.mx, barbara.sanchez@correo.buap.mx, mario.rossainz@correo.buap.mx

### RESUMEN

La holografía digital permite capturar información de ondas luminosas, pero las cámaras registran únicamente la intensidad, perdiendo la información de fase. Para mitigar esta limitación, se emplea Phase Shifting Holography, que requiere múltiples imágenes con desplazamientos de fase, aumentando la complejidad experimental. Este trabajo propone un enfoque híbrido que combina preprocesamiento de imágenes mediante el algoritmo Matching Pursuit para reducción de ruido estructurado, optimizado mediante algoritmos genéticos, con un modelo generativo tipo autoencoder capaz de generar imágenes holográficas sintéticas con desplazamiento de fase a partir de una sola imagen de referencia. Se describe la formación de la base de datos de entrenamiento, el diseño de la arquitectura del autoencoder y el proceso de entrenamiento, incluyendo métricas de desempeño. Se presentan resultados que muestran la eficacia del preprocesamiento en la reducción de ruido y las limitaciones actuales en la generación de imágenes sintéticas, señalando la necesidad de un conjunto de entrenamiento más robusto para obtener reconstrucciones confiables de fase.

**Palabras Clave:** aprendizaje profundo, matching pursuit, modelos generativos

### ABSTRACT

Digital holography enables capturing information about light waves, although cameras record only intensity, losing phase information. To mitigate this limitation, Phase Shifting Holography is employed, which requires multiple images with phase shifts, increasing experimental complexity. This work proposes a hybrid approach that combines Matching Pursuit (MP) for structured noise reduction, optimized using a genetic algorithm, with a multi-output convolutional autoencoder capable of generating synthetic holographic images with phase shifts from a single reference image. The formation of the training dataset, the autoencoder architecture, the training process, and performance metrics are described. Results show the effectiveness of the preprocessing in noise reduction and highlight current limitations in generating synthetic images, emphasizing the need for a robust training dataset to achieve reliable phase reconstruction.

**Keywords:** Deep learning, matching pursuit, generative models

### 1. INTRODUCCIÓN

La reconstrucción precisa de imágenes holográficas a partir de datos interferométricos representa un desafío fundamental en óptica digital, especialmente en condiciones donde la

adquisición de múltiples interferogramas resulta inviable o cuando el ruido degrada significativamente la calidad de las imágenes. En años recientes, diversos enfoques basados en aprendizaje profundo han sido desarrollados para abordar estos problemas, destacando su capacidad para sintetizar interferogramas adicionales o suprimir ruido sin necesidad de modificar el hardware experimental.

Zhao et al. [1] introdujeron MSPS-Net, una red neuronal convolucional de múltiples etapas basada en el esquema pix2pix, capaz de generar interferogramas con desplazamientos de fase a partir de una sola imagen. Este enfoque permite implementar esquemas interferométricos de recuperación de fase tipo one-shot, eliminando la dependencia de adquisiciones múltiples y facilitando la captura de fenómenos dinámicos. De manera similar, Huang et al. [2] propusieron una arquitectura tipo Y-Net para el caso de la holografía de correlación incoherente de Fresnel (FINCH), generando interferogramas faltantes con desplazamientos de fase  $2/3 \pi$  y  $4/3 \pi$ . Al integrar esta red con un algoritmo tradicional de retro propagación y demodulación de tres pasos, lograron reconstrucciones precisas a partir de una única captura.

Por otro lado, la presencia de ruido tipo speckle, característico en la reconstrucción de hologramas digitales, limita la calidad visual y cuantitativa de las imágenes obtenidas. Para abordar este problema, Jeon et al. [3] propusieron una red neuronal convolucional multiescala, entrenada para realizar la supresión de speckle directamente sobre la imagen reconstruida. Este método, que opera sobre una sola toma, supera en rendimiento y velocidad a técnicas tradicionales de filtrado espacial, demostrando alta eficacia en la reducción de ruido sin comprometer detalles estructurales.

En este contexto, el presente trabajo propone un enfoque híbrido que combina la eficiencia del algoritmo Matching Pursuit (MP) como método de reducción de ruido estructurado en imágenes interferométricas, con la capacidad generativa de un autoencoder entrenado para sintetizar interferogramas con desplazamiento de fase a partir de una imagen de referencia. El algoritmo MP permite descomponer la señal en una base dispersa adaptada, filtrando componentes dominados por ruido sin alterar la información significativa. Posteriormente, el autoencoder se encarga de reconstruir los interferogramas con fases artificiales

simuladas, permitiendo la recuperación de fase bajo esquemas similares a los métodos de múltiples exposiciones, pero a partir de una única entrada.

Este enfoque apunta a ofrecer una solución integral que reduzca la carga experimental, mejore la calidad de las imágenes reconstruidas y habilite nuevas aplicaciones de holografía digital en entornos dinámicos o ruidosos, manteniendo compatibilidad con técnicas clásicas de recuperación de fase por desplazamiento.

Este trabajo propone un enfoque híbrido que integra Matching Pursuit (MP) para preprocesamiento y reducción de ruido con un autoencoder convolucional para generar interferogramas sintéticos con desplazamientos de fase a partir de una imagen de referencia. El objetivo es reducir la carga experimental, mejorar la calidad de los hologramas y habilitar aplicaciones en entornos dinámicos o ruidosos, manteniendo compatibilidad con técnicas clásicas de recuperación de fase por desplazamiento.

## 2. FORMACIÓN DE LA BASE DE DATOS

El conjunto de datos se formó a partir de hologramas experimentales adquiridos en laboratorio con distintos desplazamientos de fase. Cada holograma se almacenó en formato .tif de 16 bits.

Para entrenar el autoencoder, se generaron pares entrada–target:

- La entrada corresponde al holograma preprocesado con MP.
- Los targets son los hologramas experimentales correspondientes a los distintos desplazamientos de fase.

En total, el dataset consta de 20 hologramas, con 4 imágenes por desplazamiento. Esta información permite reproducir los entrenamientos y evaluar la capacidad generativa del modelo.

## 3. ESTRUCTURA DEL PROYECTO

La Figura 1 muestra el diagrama general del flujo de trabajo propuesto. En la primera etapa, los hologramas obtenidos experimentalmente en el laboratorio son sometidos a un proceso de preprocesamiento mediante el algoritmo de matching pursuit, con el objetivo de mejorar la representación estructural y reducir el ruido presente en las imágenes originales. Posteriormente, los hologramas resultantes de este preprocesamiento son utilizados como conjunto de datos para el entrenamiento de un modelo autoencoder, cuya arquitectura está diseñada para aprender a generar hologramas a partir de una imagen de referencia.

Una vez completado el entrenamiento del autoencoder, se procede a la evaluación de la calidad de los hologramas generados. Para ello, se aplica el método de recuperación de fase de cuatro pasos (four-step phase-shifting) tanto a los hologramas originales como a los generados. Esta técnica permite calcular la fase de las ondas luminosas asociadas a cada conjunto de imágenes, lo cual resulta fundamental para evaluar la fidelidad de los hologramas sintéticos. Finalmente, se realiza una comparación cuantitativa entre la fase obtenida a partir de los hologramas originales y aquella derivada de los hologramas

generados, con el propósito de validar la eficacia del modelo propuesto en la reconstrucción precisa de la información de fase.

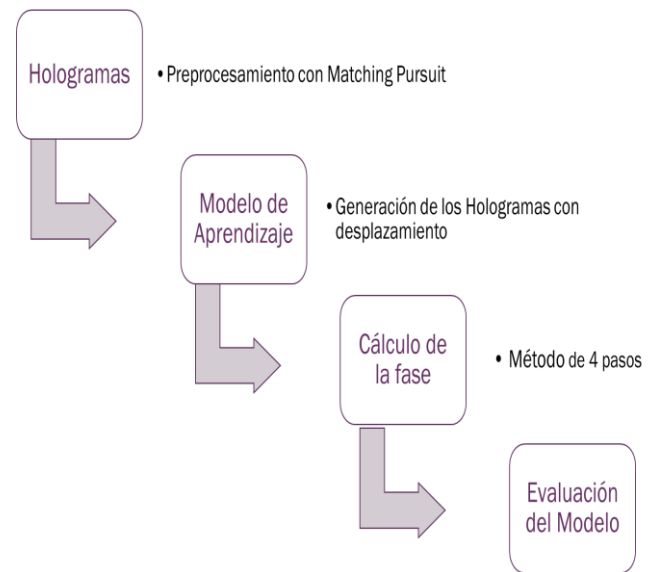


Figura 1. Diagrama del flujo de trabajo propuesto.

Los hologramas adquiridos experimentalmente son preprocesados mediante matching pursuit para mejorar su calidad. Estos hologramas procesados se emplean para entrenar un autoencoder, cuya salida se evalúa mediante el cálculo de fase utilizando el método de cuatro pasos. La fase resultante se compara con la obtenida a partir de los hologramas originales, permitiendo validar la fidelidad de los hologramas generados.

## 4. MATCHING PURSUIT

### 4.1. Matching Pursuit como método de preprocesamiento

Matching Pursuit es un algoritmo de aproximación utilizado en el procesamiento de señales e imágenes, que permite descomponer una señal compleja en una combinación lineal de funciones o "átomos" de una base redundante, también conocida como diccionario [4]. La idea detrás de Matching Pursuit es encontrar de manera iterativa las mejores aproximaciones de la señal original usando una base de señales sencillas (generalmente senoidales, funciones de onda, o funciones con características específicas). Este proceso resulta útil en aplicaciones donde la señal puede ser compleja y es deseable representar solo sus componentes más significativos, facilitando la compresión o el análisis [5].

### 4.2. Entrenamiento de los diccionarios

En este trabajo, se busca entrenar los diccionarios como parte del preprocesamiento de imágenes holográficas utilizando la técnica de Dictionary Learning. Para ello, se recopilieron conjuntos de imágenes correspondientes a diferentes desplazamientos de fase,

las cuales fueron redimensionadas y segmentadas en parches de tamaño fijo con solapamiento. Cada parche es vectorizado y empleado como muestra de entrada para el entrenamiento. Posteriormente, para cada conjunto de imágenes asociadas a un mismo desplazamiento, previamente se determinan parámetros óptimos para ajustar un modelo de aprendizaje de diccionario. El entrenamiento se lleva a cabo mediante el algoritmo orthogonal matching pursuit (OMP), con el objetivo de obtener un conjunto de átomos que permitieran representar de forma dispersa la información contenida en los parches. Finalmente, los diccionarios obtenidos para cada conjunto son almacenados para su uso posterior en tareas de reconstrucción y análisis de imágenes holográficas.

#### 4.2.1 Optimización de los parámetros

En este estudio, se plantea el uso de algoritmos genéticos como estrategia de optimización para determinar los parámetros más adecuados en el entrenamiento de diccionarios aplicados a imágenes holográficas. Concretamente, se hace uso del marco DEAP para implementar un algoritmo evolutivo capaz de explorar el espacio de parámetros que definen el modelo de Dictionary Learning, tales como el número de componentes ( $n\_components$ ), el número máximo de iteraciones ( $n\_iter$ ), el parámetro de regularización ( $\alpha$ ), la tolerancia de convergencia ( $tol$ ) y el algoritmo de transformación ( $transform\_algorithm$ ). La función objetivo del algoritmo es minimizar el error cuadrático medio entre la imagen original y su reconstrucción a partir de parches codificados dispersamente. Para evaluar cada configuración candidata, se aplica validación cruzada sobre conjuntos de imágenes obtenidas con diferentes desplazamientos de fase, utilizando ventanas de Hann y segmentación por parches con solapamiento. La reconstrucción de cada imagen se realiza empleando la transformada de Fourier y la técnica de matching pursuit sobre cada parche, y los errores se promediaron para estimar la aptitud de cada individuo. El proceso evolutivo considera operadores de cruce, mutación y selección por torneo, y paralelizado mediante múltiples núcleos de procesamiento para mejorar la eficiencia computacional. Finalmente, los mejores individuos —es decir, los conjuntos de parámetros con menor error promedio— son seleccionados para cada grupo de imágenes y para distintos tamaños de resolución. Estos parámetros óptimos son almacenados para su posterior uso en el entrenamiento de diccionarios específicos por tamaño e índice de imagen, asegurando así una representación dispersa adaptada a las características estadísticas de los hologramas.

## 5. MODELO GENERATIVO

### 5.1. Autoencoders para la generación de imágenes

La generación de imágenes mediante inteligencia artificial ha experimentado un crecimiento notable en los últimos años, en parte gracias al desarrollo y la popularización de modelos de aprendizaje profundo. Entre estos, los autoencoders (AE) se destacan como una de las arquitecturas más efectivas para aprender representaciones compactas y útiles de datos. Se

utilizan en diversas aplicaciones como la compresión de imágenes, la reducción de ruido y la generación de imágenes. Un autoencoder es una red neuronal diseñada para aprender una representación de los datos de entrada en un espacio de menor dimensión y luego reconstruir esos datos a partir de esa representación.

La eficiencia de los autoencoders se basa en la capacidad de capturar las características más relevantes de los datos mientras descarta la información menos significativa, lo que los hace especialmente útiles para tareas de compresión y generación.

### 5.2. Modelo generativo propuesto

El modelo propuesto consiste en un autoencoder convolucional con una arquitectura ramificada, diseñado para generar tres salidas distintas a partir de una única imagen de entrada. Las imágenes empleadas se almacenan en formato .tif y son procesadas mediante una función personalizada que redimensiona y normaliza los datos a una resolución reducida ( $288 \times 528$ ), preservando la información de profundidad de 16 bits. La entrada al modelo se organiza mediante un generador que agrupa los datos en lotes, donde cada conjunto consta de una imagen original y tres imágenes correspondientes a distintas transformaciones o desplazamientos, usadas como salidas esperadas. La estructura del autoencoder incluye un codificador común que extrae una representación latente compartida, la cual alimenta tres decodificadores independientes, permitiendo la reconstrucción simultánea de las tres imágenes de salida. Cada decodificador emplea capas densas y convoluciones transpuestas para aumentar progresivamente la resolución. El modelo se entrena utilizando el optimizador Adam y una función de pérdida MSE independiente para cada salida, incorporando además mecanismos de control como checkpointing y early stopping para guardar la mejor versión del modelo y prevenir el sobreajuste. Esta implementación resulta especialmente útil para aprender representaciones profundas multisalida en contextos donde la imagen de entrada se asocia a múltiples reconstrucciones esperadas.

## 6. RESULTADOS

En esta sección se presentan los resultados obtenidos tras aplicar el esquema de preprocesamiento basado en matching pursuit sobre un conjunto de cuatro hologramas experimentales. Para esta etapa, se utilizó un enfoque de optimización evolutiva mediante un algoritmo genético, cuyo objetivo fue ajustar únicamente tres parámetros: el número de átomos del diccionario, el número de iteraciones del algoritmo y el valor de penalización  $\alpha$ .

En la Figura 2 se muestran los resultados visuales de este proceso para cada una de las imágenes utilizadas. En ellas se puede notar una pérdida notable de detalle en las zonas de alto contraste y una distorsión general en los patrones de interferencia, lo cual sugiere que la configuración de parámetros optimizada no es suficiente para preservar la información de fase. Esto pone de manifiesto la necesidad de incluir criterios adicionales de calidad en la función objetivo del algoritmo genético, o bien, incorporar

otras variables que permitan una representación más fiel de los hologramas durante la etapa de preprocesamiento.

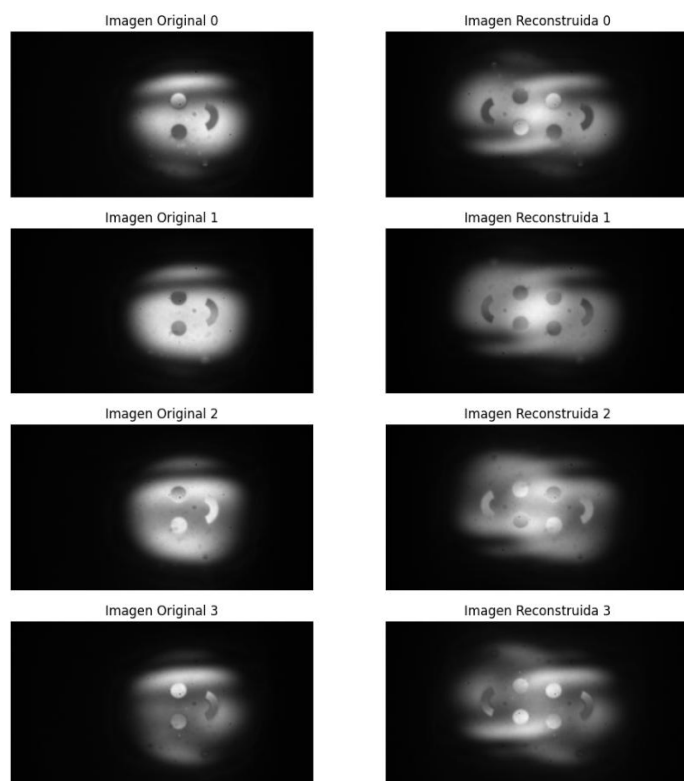


Figura 2. Comparación entre los hologramas originales y las reconstrucciones obtenidas mediante matching pursuit para cada uno de los cuatro desplazamientos de fase (índices 0 a 3).

En cada par de imágenes, la izquierda corresponde al holograma original y la derecha a su reconstrucción. Las reconstrucciones presentan pérdida de detalle y distorsiones en los patrones de interferencia, lo que afecta la precisión en el cálculo de la fase.

Si bien el algoritmo logró una convergencia adecuada en términos computacionales, las imágenes reconstruidas no conservaron de forma precisa las estructuras interferométricas presentes en los hologramas originales. Esta deficiencia en la reconstrucción impacta directamente en el cálculo de la fase. Al aplicar el método de cuatro pasos para la recuperación de fase sobre los hologramas reconstruidos, se observaron discrepancias significativas respecto a la fase obtenida a partir de los hologramas originales como se muestra en la Figura 3.

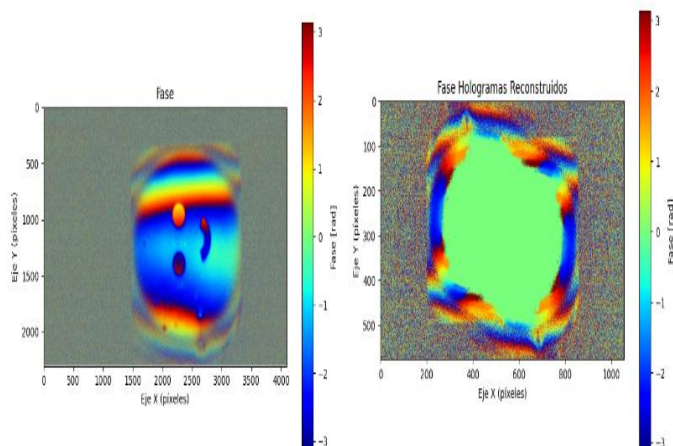


Figura 3. Comparación de la fase calculada mediante el método de cuatro pasos utilizando los hologramas originales (izquierda) y los hologramas reconstruidos mediante matching pursuit (derecha).

La fase obtenida a partir de las reconstrucciones presenta errores significativos y pérdida de continuidad, lo que evidencia que la calidad de los hologramas generados no es suficiente para una recuperación de fase precisa.

## 7. CONCLUSIONES

En este trabajo se presentó un esquema de generación de hologramas basado en el preprocesamiento de datos mediante matching pursuit y el entrenamiento de un autoencoder, con el objetivo de validar la calidad de los hologramas generados a través del cálculo de fase utilizando el método de cuatro pasos. Los resultados obtenidos mostraron que el uso de un algoritmo genético para optimizar únicamente el número de átomos del diccionario, el número de iteraciones y el parámetro  $\alpha$  no fue suficiente para lograr una reconstrucción precisa de los hologramas. Las imágenes reconstruidas presentaron pérdida de detalle y distorsiones en los patrones de interferencia, lo que se tradujo en errores significativos en la recuperación de fase.

Esto evidencia que, aunque el enfoque evolutivo ofrece ventajas en la exploración del espacio de parámetros, es necesario incorporar criterios de calidad más específicos durante el proceso de optimización. Además, se sugiere considerar la inclusión de variables adicionales en la búsqueda de hiperparámetros, tales como el tipo de ventana o la selección del subconjunto de entrenamiento, así como posibles restricciones físicas que refuercen la conservación de la información interferométrica.

Por otro lado, se identificaron limitaciones importantes en el entrenamiento del autoencoder. A pesar de contar con una arquitectura funcional, el proceso de entrenamiento no logró converger ni en una computadora personal con recursos estándar ni utilizando plataformas en la nube como Google Colab. Esto indica que el modelo requiere una mayor capacidad de cómputo,

lo que representa un desafío a superar para lograr una implementación efectiva y eficiente del sistema propuesto.

Como trabajo futuro, se plantea refinar el proceso de preprocesamiento con objetivos más alineados a la preservación de fase, mejorar la infraestructura computacional para el entrenamiento del modelo, y evaluar el impacto de estas mejoras en el desempeño general del sistema de generación de hologramas.

## REFERENCIAS

- [1] Zhao, Y., Hu, K. & Liu, F. W. One-shot phase retrieval method for interferometry using a multi-stage phase-shifting network. *IEEE Photonics Technol. Lett.* 35,577–580 (2023).
- [2] Huang, T. et al. Single-shot Fresnel incoherent correlation holography via deep learning based phase-shifting technology. *Opt. Express* 31, 12349–12356 (2023)
- [3] Jeon, W. et al. Speckle noise reduction for digital holographic images using multi-scale convolutional neural networks. *Opt. Lett.* 43,4240–4243 (2018)
- [4] Mallat, S., & Zhang, Z. (1993). Matching Pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12), 3397-3415.
- [5] Tropp, J. A. (2004). Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10), 2231-2242.