

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
PLAN DE TRABAJO DE GRADO

FECHA DE PRESENTACIÓN: Bucaramanga, 31 de julio de 2021

TITULO:	Algoritmo de clasificación de objetos en imágenes difractivas basado en medidas cuadráticas codificadas usando un enfoque de aprendizaje profundo
MODALIDAD:	Trabajo de investigación
AUTOR:	David Santiago Morales Norato, código: 2170102
DIRECTOR:	PhD.(c). Andrés Felipe Jerez Ariza
CODIRECTOR:	Ph.D. Henry Arguello Fuentes
ENTIDADES INTERESADAS:	Universidad Industrial de Santander, Escuela de Ingeniería de Sistemas e Informática, y Grupo de Investigación en Diseño de Algoritmos y Procesamiento de Datos Multidimensionales (HDSP)

COMITÉ DE TRABAJOS DE GRADO:

EVALUADOR ASIGNADO:

CONCEPTO DEL EVALUADOR:

APROBACIÓN DEL COMITÉ:

FECHA: _____

ACTA NO. _____

CONTENIDO

LISTA DE FIGURAS	2
LISTA DE TABLAS	3
1 INTRODUCCIÓN	4
2 PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA	6
3 OBJETIVOS	7
3.1 OBJETIVO GENERAL	7
3.2 OBJETIVOS ESPECÍFICOS	7
4 MARCO DE REFERENCIA.	8
4.1 ADQUISICIÓN DE LA INFORMACIÓN DE FASE.	8
4.1.1 SISTEMA ÓPTICO DE DIFRACCIÓN	8
4.1.2 PROBLEMA DE RECUPERACIÓN DE FASE	9
4.1.3 MEDIDAS CUADRÁTICAS CODIFICADAS	10
4.2 ALGORITMOS DE RECUPERACIÓN DE FASE.	10
4.2.1 FORMULACIONES CONVEXAS	10
4.2.2 FORMULACIONES NO CONVEXAS	11
4.3 SISTEMAS DE CLASIFICACIÓN	12
4.3.1 MÁQUINAS DE SOPORTE VECTORIAL	12
4.3.2 K VECINOS MÁS CERCANOS	13
4.3.3 REDES NEURONALES	14
4.3.4 CLASIFICACIÓN USANDO MEDIDAS CUADRÁTICAS	15
5 METODOLOGÍA	16
6 CRONOGRAMA	17
7 PRESUPUESTO	18
8 Bibliografía	19

LISTA DE FIGURAS

Figura 1 Sistema óptico codificado de difracción	8
Figura 2 Experimento intercambiando las fases de la transformada de Fourier de 2 imágenes	9
Figura 3 Representación de una SVM en \mathbb{R}^2 . La recta $\mathbf{w}\mathbf{x} - b = 0$ en rojo representa el plano óptimo que soluciona el problema de optimización (11).	13
Figura 4 Representación de KNN en \mathbb{R}^2 con $K = 5$	14
Figura 5 Arquitectura de una arquitectura de red neuronal $f : \mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$	15

LISTA DE TABLAS

Tabla 1 Cronograma de actividades	17
Tabla 2 Recursos humanos.	18
Tabla 3 Equipos e infraestructura.	18
Tabla 4 Recursos y software.	18
Tabla 5 Presupuesto general.	18

1 INTRODUCCIÓN

Los algoritmos computacionales basados en aprendizaje profundo han sido ampliamente estudiados en la literatura, especialmente, la clasificación de objetos en imágenes ha sido una de las tareas computacionales más abordadas en ese tópico [1, 2, 3]. Los enfoques de aprendizaje profundo utilizan arquitecturas de redes neuronales, que consisten en la concatenación de múltiples capas compuestas de unidades mínimas llamadas neuronas. Cada neurona realiza una combinación lineal entre las entradas, para posteriormente usar una función no lineal en la salida. Las salidas de cada neurona en una capa funcionan como entrada de las neuronas ubicadas en la siguiente capa, creando así, una arquitectura de red neuronal profunda [4].

En general, la clasificación de objetos se realiza sobre imágenes en escala de grises [5], RGB [6], o más recientemente, imágenes espectrales [1]. Sin embargo, los enfoques de clasificación de objetos basados en aprendizaje profundo incorporan como entrada de la arquitectura de red neuronal, la información de la intensidad de la luz incidente sobre el sensor, omitiendo la información de fase, que resulta fundamental en aplicaciones como cristalografía de rayos-x [7], astronomía [8], holografía [9], entre otras. Esta limitación se atañe a los sistemas ópticos que dependen de la conversión de fotones a electrones, puesto que no permiten una adquisición directa de la información de fase [10]. Por lo tanto, la obtención de esta información de fase requiere la implementación de algoritmos computacionales que logren recuperar los datos perdidos.

Los algoritmos de recuperación de fase permiten reconstruir la información de un campo óptico inicial con base en la adquisición de medidas de intensidad que siguen un modelo cuadrático de propagación según diferentes campos de difracción, tales como campo cercano, medio y lejano [11]. Dentro de estos sistemas ópticos de difracción se han incorporado máscaras de fase para la modulación del campo óptico inicial, puesto que, la literatura ha demostrado que la inclusión de este tipo de elementos ópticos durante el proceso de adquisición, genera redundancia en las medidas captadas, garantizando la recuperación de fase en hasta una constante unimodular [12]. Estas medidas adquiridas a través de sistemas ópticos de difracción que incluyen máscaras de fase, se denominan medidas cuadráticas codificadas. Las imágenes recuperadas usando medidas cuadráticas codificadas se conocen como imágenes ópticas difractivas. Diversos algoritmos iterativos han sido propuestos para resolver la reconstrucción de imágenes difractivas a partir del problema de recuperación de fase. Tradicionalmente, estos algoritmos se construyen bajo formulaciones convexas, tales como el *PhaseLift* [13] y *PhaseMax* [14]; o formulaciones no convexas, tales como, *Wirtinger Flow* [15], *Truncated Wirtinger Flow* [16], *Truncated Amplitude Flow* [17] y *Reweighted Amplitude Flow* [18].

Actualmente, se han involucrado arquitecturas de redes neuronales profundas en el campo de formación de imágenes ópticas, puesto que, permite incorporar el modelo de adquisición como una capa de la misma arquitectura de red neuronal para promover una mejor reconstrucción de las medidas adquiridas. La inclusión del modelo de adquisición ha permitido el diseño de sistemas ópticos de adquisición a través de su implementación en configuraciones de red neuronal donde los pesos entrenables representan las variables optimizables del modelo de propagación [19].

Por otra parte, algunos trabajos han estudiado la clasificación de objeto usando únicamente medidas de intensidad captadas mediante sistemas ópticos definidos matemáticamente a través de operadores lineales, por ejemplo, sistemas de tomografía computarizada [20], imágenes de un solo píxel [21], entre otros. De manera que, los sistemas de clasificación que omiten el proceso de reconstrucción de las medidas obtenidas disminuyen el tiempo de inferencia, puesto que, se elimina la etapa de reconstrucción de la imagen, que generalmente, implica un alto costo computacional en los sistemas de clasificación. A pesar de que, se han desarrollado arquitecturas de redes neuronales para la clasificación usando medidas de intensidad bajo un modelo cuadrático [22, 23], en la literatura de imágenes difractivas que resultan del problema de recuperación de fase no se han abordado modelos de redes neuronales para la detección de objetos basados en medidas cuadráticas codificadas. Así que, surge el interés de estudiar sistemas de clasificación que permitan la discriminación de objetos con base en medidas cuadráticas codificadas.

Por lo tanto, este trabajo de investigación propone el diseño de un algoritmo de clasificación de objetos en imágenes difractivas sobre medidas cuadráticas codificadas mediante el uso de aprendizaje profundo. La arquitectura de red neuronal de clasificación propuesta incluirá el modelo matemático que describe la adquisición de medidas cuadráticas para la clasificación de objetos en imágenes difractivas. Además, la evaluación de la arquitectura de red neuronal propuesta se realizará a través de bases de datos de la literatura y medidas simuladas. Asimismo, el algoritmo computacional de clasificación propuesto se comparará con técnicas del estado del arte.

Este documento se encuentra organizado de la siguiente manera: En sección 2, se describe el planteamiento del problema, incluyendo la justificación de este trabajo de investigación. La sección 3 presenta los objetivos tanto general como específicos de este trabajo. La sección 4 corresponde a la descripción de los conceptos generales y teóricos en relación con los sistemas ópticos de difracción, los algoritmos de recuperación de fase y sistemas de clasificación. La sección 5 describe la metodología planteada para alcanzar los objetivos propuestos. La sección 6 ilustra el cronograma de actividades que se realizarán durante el desarrollo del proyecto. Finalmente, la sección 7 describe el presupuesto del proyecto.

2 PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA

En diferentes campos de la ciencia e ingeniería [7, 8, 9], la información de fase es imprescindible. Sin embargo, debido a las altas frecuencias de la luz, los sensores únicamente captan la intensidad de la luz incidente en el sensor, perdiendo la información de fase del campo óptico subyacente. En particular, el modelo de propagación que siguen los sistemas ópticos de difracción está dado por $y_k = |\mathbf{A}_k \mathbf{z}|^2$ con $k = 1, \dots, K$ proyecciones, donde $\mathbf{z} \in \mathbb{C}^n$ representa el objeto de interés, $\mathbf{y}_k \in \mathbb{R}^m$ son las medidas adquiridas y $\mathbf{A}_k \in \mathbb{C}^{n \times n}$ describe la propagación del frente de onda de la luz hasta el sensor en cada proyección k . Cabe destacar que, este sistema de ecuaciones presenta múltiples soluciones, debido a diferentes ambigüedades triviales que resultan de las medidas adquiridas y para diferentes transformaciones de \mathbf{z} [24, 10], las cuales se describen a continuación:

1. Desplazamiento de la fase global $\mathbf{z}[n] \leftarrow \mathbf{z}[n] \cdot e^{j\Phi_0}$,
2. Reflexión sobre un eje $\mathbf{z}[n] \leftarrow \bar{\mathbf{z}}[-n]$,
3. Desplazamiento espacial $\mathbf{z}[n] \leftarrow \mathbf{z}[n + n_0]$,

donde $\bar{(\cdot)}$ corresponde al operador conjugado y $j = \sqrt{-1}$. Para recuperar la información de fase hasta una constante de fase global en imágenes difractivas formadas a partir de sistemas ópticos basados en medidas cuadráticas, la literatura ha incluido un elemento de modulación denominado máscaras de fase, que generan redundancia en las medidas captadas, las cuales se conocen como medidas cuadráticas codificadas. No obstante, el proceso de reconstrucción implica un alto costo computacional para tareas computacionales de inferencia sobre imágenes difractivas. Por consiguiente, algunos trabajos han estudiado la clasificación de objetos usando exclusivamente las medidas adquiridas [21, 20, 22, 23], evitando el proceso de reconstrucción de las imágenes. Este tipo de enfoques permiten reducir el tiempo de inferencia de los algoritmos. Sin embargo, estos esquemas de clasificación no se han abordado previamente sobre sistemas ópticos que producen medidas cuadráticas codificadas. Este trabajo propone el diseño de un algoritmo de clasificación de objetos en imágenes difractivas sobre medidas cuadráticas codificadas mediante el uso de aprendizaje profundo. Finalmente, este proyecto se desarrollará con el apoyo del grupo de investigación en diseño de algoritmo y procesamiento de datos multidimensionales (HDSP). El grupo HDSP es experto en el área de procesamiento de imágenes y señales de alta dimensionalidad. Este grupo de investigación se encuentra clasificado actualmente en categoría A1 por Colciencias, adscrito a la Escuela de Ingeniería de Sistemas e Informática de la Universidad Industrial de Santander.

Pregunta de investigación: ¿Cómo incorporar enfoques de aprendizaje profundo para la clasificación de objetos en sistemas ópticos difractivos basados en medidas cuadráticas codificadas?

3 OBJETIVOS

3.1 OBJETIVO GENERAL

Desarrollar un algoritmo de clasificación de objetos en imágenes difractivas basado en medidas cuadráticas codificadas usando un enfoque de aprendizaje profundo.

3.2 OBJETIVOS ESPECÍFICOS

1. Modelar matemáticamente el proceso de adquisición de medidas cuadráticas codificadas utilizando máscaras de fase.
2. Diseñar e implementar un algoritmo de clasificación de objetos en medidas cuadráticas codificadas a partir de aprendizaje profundo que incorpore el modelo de adquisición.
3. Simular una configuración óptica difractiva para la adquisición de medidas cuadráticas codificadas usando máscaras de fase.
4. Evaluar el algoritmo de clasificación de medidas cuadráticas codificadas basado en aprendizaje profundo frente a otras técnicas del estado del arte usando medidas adquiridas con la configuración óptica simulada.

4 MARCO DE REFERENCIA

Esta sección presenta una descripción de los conceptos generales y teóricos en relación con la adquisición de la información de fase, los algoritmos de recuperación de fase y sistemas de clasificación.

4.1 ADQUISICIÓN DE LA INFORMACIÓN DE FASE

Los sistemas ópticos tradicionales captan únicamente la información de la intensidad de la luz incidente sobre el sensor, perdiendo con esto la información de fase, que resulta fundamental en aplicaciones como cristalografía de rayos-x [7], astronomía [8], holografía [9], entre otras. Por lo tanto, la obtención de la información de fase requiere la implementación de sistemas ópticos que logren codificar la fase del campo óptico, preservándola implícita en las medidas de intensidad adquiridas.

4.1.1. SISTEMA ÓPTICO DE DIFRACCIÓN

En múltiples áreas de la ciencia e ingeniería se presenta la adquisición de medidas de intensidad haciendo uso de sistemas ópticos de difracción [8, 7, 9]. La Figura 1 muestra un esquema común de los sistemas ópticos de difracción, los cuales se componen por un objeto representado con el vector z iluminado por luz coherente, modulado en fase por una máscara de fase D , produciendo las medidas cuadráticas codificadas. Dependiendo de la distancia del sensor L , la longitud de onda λ y el radio de apertura d , los modelos de propagación se denominan campo cercano, medio y lejano.

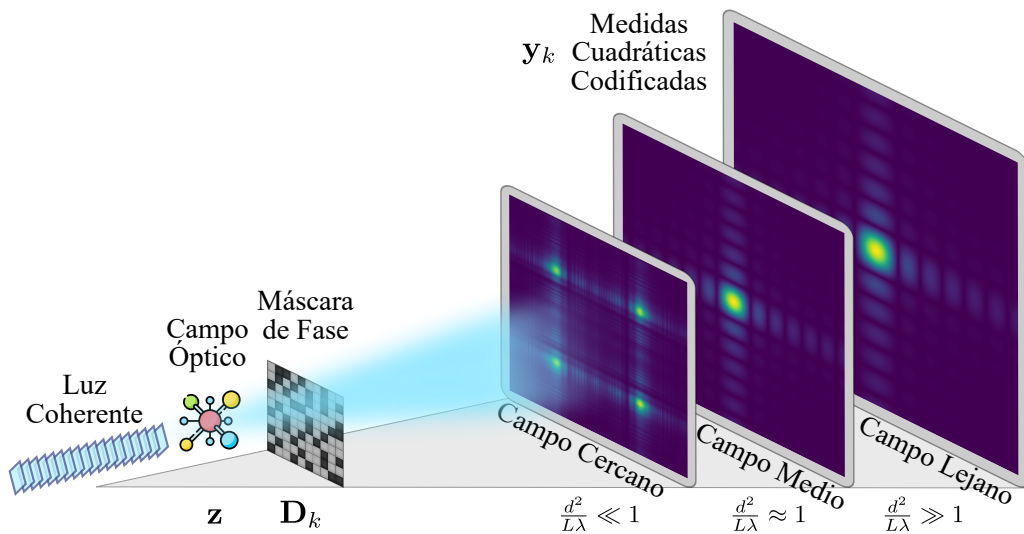


Figura 1: Sistema óptico codificado de difracción

4.1.2. PROBLEMA DE RECUPERACIÓN DE FASE

Debido a que los sensores únicamente captan la información de la intensidad de la luz, el modelo de adquisición en los sistemas ópticos de difracción se describe como

$$\mathbf{y}_k = |\mathbf{A}_k \mathbf{z}|^2, \quad (1)$$

donde $\mathbf{z} \in \mathbb{C}^n$ es una vectorización del objeto de interés, $\mathbf{y}_k \in \mathbb{R}^n$ son las medidas adquiridas y $\mathbf{A}_k \in \mathbb{C}^{n \times n}$ son las matrices que describe la propagación del frente de onda de la luz hasta el sensor, aquí, $k = 1, \dots, K$, siendo K el número total de proyecciones. De acuerdo con la teoría de difracción [25], la propagación del frente de onda del objeto se modela entre los campos, cercano, medio y lejano dependiendo del número de Fresnel $F = \frac{d^2}{L\lambda}$, que determina el modelado matemático acorde a la zona de difracción en que se captan las medidas.

$$\mathbf{y}_k = \begin{cases} |\mathbf{F} \mathbf{T} \mathbf{F}^H \mathbf{D}_k \mathbf{z}|^2, & (\text{Campo Cercano, } F \ll 1), \\ |\mathbf{F}^H \mathbf{Q} \mathbf{D}_k \mathbf{z}|^2, & (\text{Campo Medio, } F \approx 1), \\ |\mathbf{F} \mathbf{D}_k \mathbf{z}|^2, & (\text{Campo Lejano, } F \gg 1), \end{cases}$$

donde $|\cdot|$ representa el operador de magnitud, \mathbf{F} corresponde a la transformada discreta de Fourier, la matriz diagonal $\mathbf{D}_k \in \mathbb{C}^{n \times n}$ representa las máscaras de fase, finalmente $\mathbf{T} \in \mathbb{C}^{n \times n}$ y $\mathbf{Q} \in \mathbb{C}^{n \times n}$ son matrices ortogonales que modelan la función de transferencia espacial del campo medio y cercano, respectivamente [25, 11].

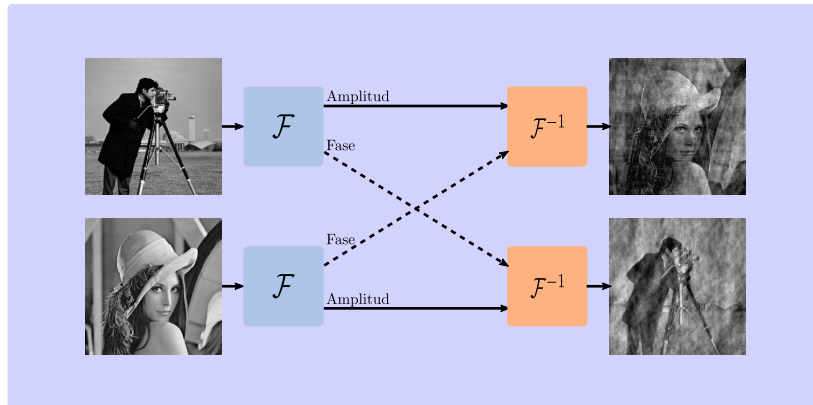


Figura 2: Experimento intercambiando las fases de la transformada de Fourier de 2 imágenes .

Cabe señalar la importancia de la información fase en Fourier a partir de la Figura 2. En esta Figura se muestra un experimento desarrollado en [10] donde se intercambia la información de fase de la transformada de Fourier de dos imágenes, posteriormente se aplica la transformada de Fourier inversa a este intercambio. En este ejemplo, se observa que en la imagen resultante predomina la información de las fases contrarias.

4.1.3. MEDIDAS CUADRÁTICAS CODIFICADAS

Dado la naturaleza cuadrática de la formulación en la Ec. (1), este problema se cataloga como NP-difícil. Una forma de mitigar la complejidad de este problema, es incluyendo elementos de modulación de fase aleatorios en la formulación, puesto que, estas máscaras aleatorias permiten generar redundancia en las medidas, que conlleva a la obtención una solución exacta con alta probabilidad si se tienen las suficientes muestras [12].

Las entradas aleatorias de la matriz \mathbf{D}_k son *i.i.d* copias de una variable aleatoria $d \in \mathbb{C}$ que satisfacen las siguientes condiciones

$$|d| \leq M, \quad \mathbb{E}[d] = 0, \quad \mathbb{E}[d^2] = 0, \quad \mathbb{E}[|d|^4] = 2\mathbb{E}[|d|^2]^2, \quad (2)$$

donde $\mathbb{E}[\cdot]$ corresponde al valor esperado. Aquí, se busca que $M = 1$ para que la codificación no incremente la potencia de las medidas cuadráticas.

4.2 ALGORITMOS DE RECUPERACIÓN DE FASE

Para recuperar el campo óptico inicial a partir de medidas cuadráticas codificadas, la literatura ha planteado diferentes algoritmos basados en formulaciones convexas y no convexas.

4.2.1. FORMULACIONES CONVEXAS

Estos tipo de algoritmos relajan el problema recuperación de fase a un problema convexo equivalente.

- **PhaseLift [13]**: Este algoritmo plantea el problema de recuperación de la fase como una minimización de la traza de la siguiente forma

$$\begin{aligned} & \underset{\mathbf{z} \in \mathbb{C}^n}{\text{minimize}} && \text{Tr}(\mathbf{z}\mathbf{z}^H), \\ & \text{subject to} && \mathcal{A}(\mathbf{z}\mathbf{z}^H) = \mathbf{b}, \\ & && \mathbf{z}\mathbf{z}^H \succeq 0, \end{aligned} \quad (3)$$

donde $\mathcal{A}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es un operador lineal y $\text{Tr}(\cdot)$ representa la traza de una matriz.

- **PhaseMax [14]**: Sea $\hat{\mathbf{z}} \in \mathbb{C}^n$ un vector aproximación de la señal original \mathbf{z} , de modo que, la señal reconstruida se obtiene solucionando el siguiente problema convexo

$$\begin{aligned} & \underset{\mathbf{z} \in \mathbb{C}^n}{\text{maximize}} && \langle \mathbf{z}, \hat{\mathbf{z}} \rangle_{\mathbb{R}}, \\ & \text{subject to} && |\langle \mathbf{a}_i, \mathbf{z} \rangle| \leq \phi_i, \end{aligned} \quad (4)$$

donde $\langle \cdot, \cdot \rangle_{\mathbb{R}}$ denota la parte real del producto interno y $\phi_i = \sqrt{y_i}$

4.2.2. FORMULACIONES NO CONVEXAS

Las formulaciones no convexas calculan el gradiente siguiendo la diferenciación de Wirtinger como los mostrados a continuación.

- **TRUNCATED WIRTINGER FLOW (TWF)**: El algoritmo TWF propuesto en [16], basa el modelo de muestreo según un variables aleatorias que siguen una distribución de Poisson de la forma:

$$y_i \sim \text{Poisson}(|\langle \mathbf{a}_i, \mathbf{z} \rangle|^2), \quad i = 1, \dots, m. \quad (5)$$

TWF busca minimizar la máxima estimación de probabilidad

$$\underset{\mathbf{z} \in \mathbb{C}^n}{\text{minimize}} - \sum_{i=1}^m \ell(\mathbf{z}; y_i), \quad (6)$$

donde $\ell(\mathbf{z}; y_i) = y_i \log(|\mathbf{a}_i^H \mathbf{z}|^2) - |\mathbf{a}_i^H \mathbf{z}|^2$ con $(\cdot)^H$ el operador de conjugada transpuesta

- **TRUNCATED AMPLITUDE FLOW (TAF)**:

El algoritmo TAF [17] adopta un criterio de mínimos cuadrados para recuperar \mathbf{z} basado en las medidas sin fase \mathbf{y}

$$\underset{\mathbf{z} \in \mathbb{C}^n}{\text{minimize}} \frac{1}{m} \sum_{i=1}^m (|\langle \mathbf{a}_i, \mathbf{z} \rangle| - \phi_i)^2, \quad (7)$$

Este algoritmo asume que las medidas y_i provienen de un sistema gaussiano de la forma $y_i \sim \mathcal{N}(|\langle \mathbf{a}_i, \mathbf{z} \rangle|^2, 1)$

■ REWEIGHTED AMPLITUDE FLOW (RAF):

El algoritmo RAF formulado en [18] sigue el criterio de maximizar la estimación de probabilidad de la forma

$$\underset{z \in \mathbb{C}^n}{\text{minimize}} - \sum_{i=1}^m \ell(\mathbf{z}; \phi_i / y_i), \quad (8)$$

donde en el caso de un muestreo con ruido gaussiano basado en la amplitud $\ell(\mathbf{z}; y_i) = (|\langle \mathbf{a}_i, \mathbf{z} \rangle| - \phi_i)^2$ o basado en la intensidad $\ell(\mathbf{z}; y_i) = (|\langle \mathbf{a}_i, \mathbf{z} \rangle|^2 - y_i)^2$. Por otra parte, basado en un muestreo con distribución de Poisson, $\ell(\mathbf{z}; y_i) = y_i \log(|\mathbf{a}_i^H \mathbf{z}|^2) - |\mathbf{a}_i^H \mathbf{z}|^2$

4.3 SISTEMAS DE CLASIFICACIÓN

La clasificación ha sido una de las tareas computacionales más abordadas en el estado del arte. Específicamente, los algoritmos de clasificación se pueden separar en los más tradicionales como las SVM y KNN [26] y los algoritmos basados en redes neuronales que recientemente han dominado diferentes campos [1, 2, 3]

4.3.1. MÁQUINAS DE SOPORTE VECTORIAL

Las máquinas de soporte vectorial (SVM, por su sigla en inglés) [27] son un método de clasificación binaria, donde cada punto n dimensional \mathbf{x}_i le corresponde una etiqueta de clase $c_i \in \{1, -1\}$.

Suponiendo que los datos de ambas clases son separables linealmente, este método propone separar los datos usando el hiper plano $\mathbf{w}\mathbf{x}_i + b = 0$. En la Figura 3, se muestra una representación de una SVM en \mathbb{R}^2

$$\begin{aligned} \mathbf{w}\mathbf{x}_i + b &\geq 1 \quad \text{si} \quad c_i = 1, \\ \mathbf{w}\mathbf{x}_i + b &\leq -1 \quad \text{si} \quad c_i = -1. \end{aligned} \quad (9)$$

Cabe resaltar que para todos los elementos del conjunto de datos se cumple que:

$$c_i(\mathbf{w}\mathbf{x}_i + b) \geq 1, \quad i = 1, \dots, m \quad (10)$$

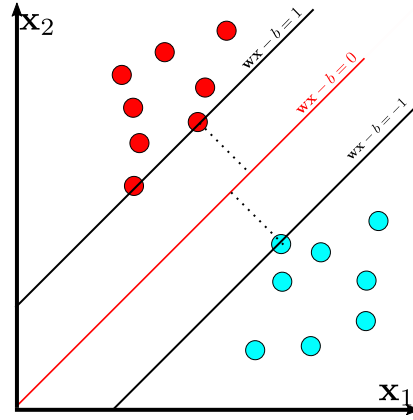


Figura 3: Representación de una SVM en \mathbb{R}^2 . La recta $\mathbf{w}\mathbf{x} - b = 0$ en rojo representa el plano óptimo que soluciona el problema de optimización (11).

El problema de optimización se plantea de la siguiente forma:

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && \|\mathbf{w}\|, \\ & \text{subject to} && c_i(\mathbf{w}\mathbf{x}_i + b) \geq 1. \end{aligned} \quad (11)$$

Usualmente no es posible separar los datos linealmente, por esta razón, se puede incluir una función no lineal ϕ que transforme los datos a un conjunto de características donde las clases sean separables linealmente. El problema de optimización para una SVM usando un kernel ϕ se formula como

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && \|\mathbf{w}\|, \\ & \text{subject to} && c_i(\mathbf{w}\phi(\mathbf{x}_i) + b) \geq 1. \end{aligned} \quad (12)$$

4.3.2. K VECINOS MÁS CERCANOS

K vecinos más cercanos (KNN, por sus siglas en inglés) Propone que un conjunto $D = \{(\mathbf{x}_i, c_i)\}_1^n$, siendo \mathbf{x}_i el vector de características e c_i la clase correspondiente. Para un nuevo vector a clasificar $\hat{\mathbf{x}}$, el algoritmo KNN encuentra los K puntos más cercanos del dataset. La Figura 4 muestra una representación de la clasificación de dos nuevas muestras en \mathbb{R}^2 con $K = 5$. Usualmente la función de distancia usada corresponde a la distancia euclidiana

$$d(\mathbf{x}_p, \mathbf{x}_q) = \|\mathbf{x}_p - \mathbf{x}_q\|_2 \quad (13)$$

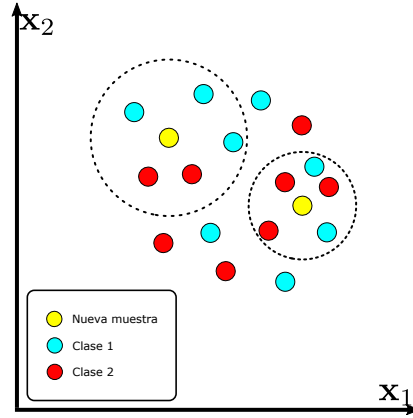


Figura 4: Representación de KNN en \mathbb{R}^2 con $K = 5$.

Posteriormente, haciendo uso de las clases de los K puntos encontrados, de tal forma que $C \subset D$ y $|C| = K$, se cuantifica la cantidad de veces que aparece cada clase y se clasifica la nueva muestra \hat{x} con la clase que más veces aparezca de la forma

$$\text{clase}(\hat{x}) = \arg \max_{\hat{c}} \left\{ \sum_{\hat{c} \in C} \delta(C, \hat{c}) \right\}, \quad (14)$$

donde $\delta(\cdot, \cdot)$ corresponde a la función delta de Kronecker, dada por

$$\delta(a, b) = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases} \quad (15)$$

4.3.3. REDES NEURONALES

Los enfoques de aprendizaje profundo han generado un gran progreso en problemas muy complejos en los últimos años [28, 1, 2, 3, 29]. El aprendizaje profundo busca encontrar una función $f : \mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$. La función f se suele llamar arquitectura de red neuronal profunda, puesto que, consiste en la concatenación de múltiples capas compuestas de unidades mínimas llamadas neuronas. Cada neurona realiza una combinación lineal entre las entradas, para posteriormente usar una función no lineal en la salida. Las salidas de cada neurona en una capa funcionan como entrada de las neuronas ubicadas en la siguiente capa, creando así, una arquitectura de red neuronal profunda [4]. En la Figura 5, se muestra una arquitectura de una arquitectura de red neuronal $f : \mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$

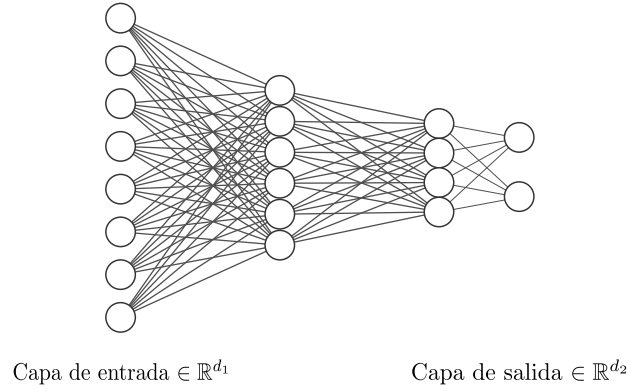


Figura 5: Arquitectura de una arquitectura de red neuronal $f : \mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$.

En la Ec. (16), se muestra el modelado matemático de una red neuronal sencilla.

$$\{f_{\theta}(\mathbf{x}) = \sigma_L(\mathbf{W}_L \sigma_{L-1}(\mathbf{W}_{L-1}(\dots \sigma_2(\mathbf{W}_2 \sigma_1(\mathbf{W}_1 \mathbf{x})))) \mid \theta = \{\mathbf{W}_1 \dots \mathbf{W}_L\}\}, \quad (16)$$

donde para cada capa $1 \leq \ell \leq L$, σ_{ℓ} corresponde a una función no lineal en dicha capa y \mathbf{W}_{ℓ} La matriz de pesos. Para entrenar los pesos θ bajo un enfoque de aprendizaje supervisado, este método hace uso de un conjunto de entrenamiento $\{(\mathbf{x}_i, c_i)\}_{i=1}^N$ y una función de costo $\mathcal{L}(c_i, f_{\theta}(\mathbf{x}_i))$ para plantear el problema de optimización

$$\underset{\theta}{\text{minimize}} \frac{1}{N} \sum_{i=1}^N \mathcal{L}(c_i, f_{\theta}(\mathbf{x}_i)) \quad (17)$$

4.3.4. CLASIFICACIÓN USANDO MEDIDAS CUADRÁTICAS

Recientemente, en el campo de tomografía computarizada [20] e imágenes de un solo píxel [21], se han propuesto sistemas de clasificación que usan únicamente medidas obtenidas a través del sistema lineal de adquisición, esto debido al enfoque basado en el aprendizaje profundo en el que la arquitectura incluye una capa con el sistema óptico de adquisición. Este enfoque ha sido estudiado igualmente en imágenes difractivas como en holografía [22] y cristalografía de rayos-x [23], donde las medidas obtenidas son difícilmente reconocidas por el ojo humano.

Sin embargo, estos enfoques propuestos no han sido abordados en imágenes de medidas cuadráticas codificadas, de modo que, este trabajo propone el desarrollo de un algoritmo de clasificación usando medidas cuadráticas codificadas basado en aprendizaje profundo.

5 METODOLOGÍA

Con el fin de alcanzar los objetivos planteados, este trabajo de investigación se llevará a cabo mediante las siguientes etapas:

- **Revisión del estado del arte:** La revisión de la literatura es un paso fundamental que se llevará a cabo durante todo el desarrollo del proyecto, con el propósito de estar al tanto de los últimos avances en la literatura. Específicamente, los temas de interés de este proyecto incluyen los conceptos, técnicas y trabajos realizados por otros investigadores, de manera que, se garantice la originalidad del trabajo realizado.
- **Modelado matemático del sistema de adquisición:** En esta etapa, se desarrollará el modelo matemático para el proceso de adquisición de imágenes difractivas usando medidas cuadráticas codificadas. En particular, el modelo de propagación obtenido incluirá las máscaras de fase que permitan la modulación del campo óptico.
- **Implementación del método de clasificación:** En este paso se diseñará e implementará un algoritmo de clasificación a partir de medidas cuadráticas codificadas.
- **Simulaciones:** En esta etapa, se realizarán las simulaciones del proceso de adquisición de las medidas cuadráticas codificadas, adicionalmente, se entrenará el algoritmo de clasificación propuesto usando las imágenes simuladas. Para evaluar cuantitativamente el rendimiento la tarea de clasificación bajo el algoritmo propuesto, se hará uso de métricas tales como la exactitud, la matriz de confusión y el puntaje F1
- **Análisis de resultados** Los resultados de la clasificación sobre diferentes conjuntos de datos serán analizados para incluir los posibles cambios al algoritmo propuesto con el objetivo de mejorar la clasificación.
- **Presentación y documentación de resultados:** Finalmente, se realizará la documentación de los resultados para su respectiva presentación en el documento final.

6 CRONOGRAMA

En la tabla 1 se presenta el cronograma de actividades con las tareas que se ejecutarán durante el desarrollo del trabajo de grado.

Actividad	Semana															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Revisión del estado del arte																
Modelado matemático del sistema de adquisición																
Implementación del método de clasificación																
Simulaciones																
Análisis de resultados																
Presentación y documentación de resultados																

Tabla 1: Cronograma de actividades

7 PRESUPUESTO

En las Tablas 2 - 5 se presentan el presupuesto asociado al desarrollo de este proyecto de investigación. Los costos relacionados con el personal involucrado se enuncian en la Tabla 2. El presupuesto de equipos e infraestructura se exhibe en la Tabla 3. El cálculo de los costos de recursos y licencias de software se presenta en la Tabla 4. Finalmente, el presupuesto general se resume en la Tabla 5.

NOMBRE	FUNCIÓN	DEDICACIÓN (HORAS)	COSTO POR HORA	TOTAL	PATROCINADOR
PhD.(c). Andrés Felipe Jerez Ariza	Director	64	\$ 30.000	\$ 1.920.000	UIS
Ph.D. Henry Arguello Fuentes	Codirector	32	\$ 305.000	\$ 9.760.000	UIS
David Santiago Morales Norato	Autor	128	\$ 10.000	\$ 1.280.000	UIS
Subtotal:				\$ 12.960.000	

Tabla 2: Recursos humanos.

ELEMENTO	COSTO COMERCIAL	COSTO POR USO	PATROCINADOR
Intel Core i5 8250u 1.6Ghz + 12 GB RAM	\$ 1.350.000	\$ 270.000	HDSP
Subtotal:		\$ 270.000	

Tabla 3: Equipos e infraestructura.

ELEMENTO	COSTO COMERCIAL	PATROCINADOR
Acceso a bases de datos	\$ 2.000.000	UIS
Licencia Matlab	\$ 1.029.228	UIS
Subtotal:	\$ 3.029.228	

Tabla 4: Recursos y software.

ELEMENTO	EFFECTIVO	ESPECIE	PATROCINADOR
Recursos humanas	\$ 0	\$ 12.960.000	UIS
Equipos	\$ 0	\$ 270.000	HDSP
Recursos y software	\$ 0	\$ 3.029.228	UIS
Subtotal:	\$ 0	\$ 16.259.228	
Costos inesperados (3 %)	\$ 0	\$ 487.776	
Total por tipo	\$ 0	\$ 16.747.004	
TOTAL	\$ 16.747.004		

Tabla 5: Presupuesto general.

8 Bibliografía

- [1] Shutao Li, Weiwei Song, Leyuan Fang, Yushi Chen, Pedram Ghamisi, and Jón Atli Benediktsson. Deep learning for hyperspectral image classification: An overview. IEEE Transactions on Geoscience and Remote Sensing, 57(9):6690–6709, 2019.
- [2] Ying Li, Haokui Zhang, Xizhe Xue, Yenan Jiang, and Qiang Shen. Deep learning for remote sensing image classification: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(6):e1264, 2018.
- [3] Wei Wang, Yujing Yang, Xin Wang, Weizheng Wang, and Ji Li. Development of convolutional neural network and its application in image classification: a survey. Optical Engineering, 58(4):040901, 2019.
- [4] Jianqing Fan, Cong Ma, and Yiqiao Zhong. A selective overview of deep learning. Statistical Science, 36(2):264–290, 2021.
- [5] Hieu Minh Bui, Margaret Lech, Eva Cheng, Katrina Neville, and Ian S Burnett. Using grayscale images for object recognition with convolutional-recursive neural network. In 2016 IEEE Sixth International Conference on Communications and Electronics (ICCE), pages 321–325. IEEE, 2016.
- [6] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6):84–90, 2017.
- [7] Samuel Pinilla, Juan Poveda, and Henry Arguello. Coded diffraction system in x-ray crystallography using a boolean phase coded aperture approximation. Optics Communications, 410:707–716, 2018.
- [8] C Fienup and J Dainty. Phase retrieval and image reconstruction for astronomy. Image recovery: theory and application, 231:275, 1987.
- [9] Yair Rivenson, Yibo Zhang, Harun Günaydın, Da Teng, and Aydogan Ozcan. Phase recovery and holographic image reconstruction using deep learning in neural networks. Light: Science & Applications, 7(2):17141–17141, 2018.
- [10] Yoav Shechtman, Yonina C Eldar, Oren Cohen, Henry Nicholas Chapman, Jianwei Miao, and Mordechai Segev. Phase retrieval with application to optical imaging: a contemporary overview. IEEE signal processing magazine, 32(3):87–109, 2015.
- [11] Joseph W Goodman. Introduction to Fourier optics. Roberts and Company Publishers, 2005.

- [12] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval from coded diffraction patterns. Applied and Computational Harmonic Analysis, 39(2):277–299, 2015.
- [13] Emmanuel J Candes, Thomas Strohmer, and Vladislav Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. Communications on Pure and Applied Mathematics, 66(8):1241–1274, 2013.
- [14] Tom Goldstein and Christoph Studer. Phasemax: Convex phase retrieval via basis pursuit. IEEE Transactions on Information Theory, 64(4):2675–2689, 2018.
- [15] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval via wirtinger flow: Theory and algorithms. IEEE Transactions on Information Theory, 61(4):1985–2007, 2015.
- [16] Yuxin Chen and Emmanuel J Candès. Solving random quadratic systems of equations is nearly as easy as solving linear systems. Communications on pure and applied mathematics, 70(5):822–883, 2017.
- [17] Gang Wang, Georgios B Giannakis, and Yonina C Eldar. Solving systems of random quadratic equations via truncated amplitude flow. IEEE Transactions on Information Theory, 64(2):773–794, 2017.
- [18] Gang Wang, Georgios B Giannakis, Yousef Saad, and Jie Chen. Phase retrieval via reweighted amplitude flow. IEEE Transactions on Signal Processing, 66(11):2818–2833, 2018.
- [19] Jorge Bacca, Tatiana Gelvez, and Henry Arguello. Transmittance regularizer for binary coded aperture design in a computational imaging end-to-end approach. In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 1470–1474. IEEE, 2021.
- [20] Clément Douarre, Carlos F Crispim-Junior, Anthony Gelibert, Laure Tougne, and David Rousseau. On the value of ctis imagery for neural-network-based classification: a simulation perspective. Applied optics, 59(28):8697–8710, 2020.
- [21] Jorge Bacca, Laura Galvis, and Henry Arguello. Coupled deep learning coded aperture design for compressive image classification. Optics express, 28(6):8528–8540, 2020.
- [22] Sung-Jin Kim, Chuangqi Wang, Bing Zhao, Hyungsoon Im, Jouha Min, Hee June Choi, Joseph Tadros, Nu Ri Choi, Cesar M Castro, Ralph Weisleder, et al. Deep transfer learning-based hologram classification for molecular diagnostics. Scientific reports, 8(1):1–12, 2018.

- [23] Angelo Ziletti, Devinder Kumar, Matthias Scheffler, and Luca M Ghiringhelli. Insightful classification of crystal structures using deep learning. Nature communications, 9(1):1–10, 2018.
- [24] Robert Beinert and Gerlind Plonka. Ambiguities in one-dimensional discrete phase retrieval from fourier magnitudes. Journal of Fourier Analysis and Applications, 21(6):1169–1198, 2015.
- [25] Ting-Chung Poon and Jung-Ping Liu. Introduction to modern digital holography: with MATLAB. Cambridge University Press, 2014.
- [26] JINHO Kim¹, BS Kim, and Silvio Savarese. Comparing image classification methods: K-nearest-neighbor and support-vector-machines. In Proceedings of the 6th WSEAS international conference on Computer Engineering and Applications, and Proceedings of the 2012 American conference on Applied Mathematics, volume 1001, pages 48109–2122, 2012.
- [27] Shan Suthaharan. Support vector machine. In Machine learning models and algorithms for big data classification, pages 207–235. Springer, 2016.
- [28] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.
- [29] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In Advances in neural information processing systems, pages 5998–6008, 2017.