

TESIS TESIS TESIS TESIS TESIS



Centro de Ciencias Básicas
Departamento de Matemáticas y Física

TESIS

Evaluación de la composición interna de obras de arte usando
algoritmos de aprendizaje supervisado y espectroscopía terahertz

PRESENTA

Claudia Collazo Guma

PARA OBTENER EL GRADO DE:

Maestría en Ciencias con Opciones a la Computación, Matemáticas Aplicadas

TUTORES

Dra. Mariana Alfaro Gómez
Dra. Luz Judith Rodríguez Esparza

INTEGRANTE DEL COMITÉ TUTORAL

Dr. Enrique Castro Camus

Aguascalientes, Ags., Julio de 2026

TESIS TESIS TESIS TESIS TESIS

Autorizaciones

CARTA DE VOTO APROBATORIO

Mtro. Guillermo Domínguez Aguilar

DECANO DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como **DIRECTORA** designado de la estudiante **CLAUDIA COLLAZO GUMA** con ID 557721 quien realizó la tesis titulada: **EVALUACIÓN DE LA COMPOSICIÓN INTERNA DE OBRAS DE ARTE USANDO ALGORITMOS DE APRENDIZAJE SUPERVISADO Y ESPECTROSCOPIA TERAHERTZ**, un trabajo propio, innovador, relevante e inédito y con fundamento en la fracción IX del Artículo 43 del Reglamento General de Posgrados, doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que ella pueda continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE

"Se Lumen Proferre"

Aguascalientes, Ags., a 30 de junio de 2026.



Dra. Mariana Alfaro Gómez

Directora de tesis

c.c.p.- Interesado
c.c.p.- Coordinación del Programa de Posgrado

CARTA DE VOTO APROBATORIO

Mtro. Guillermo Domínguez Aguilar
DECANO DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como **CODIRECTOR** designado del estudiante **CLAUDIA COLLAZO GUMA** con ID 557721 quien realizó la tesis titulado/a: **EVALUACIÓN DE LA COMPOSICIÓN INTERNA DE OBRAS DE ARTE USANDO ALGORITMOS DE APRENDIZAJE SUPERVISADO Y ESPECTROSCOPÍA TERAHERTZ**, un trabajo propio, innovador, relevante e inédito y con fundamento en la fracción IX del Artículo 43 del Reglamento General de Posgrados, doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que ella pueda continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"

Aguascalientes, Ags., a 29 de junio de 2026.

Luz Judith R.E.

Dra. Luz Judith Rodríguez Esparza
Co director de tesis

c.c.p.- Interesado
c.c.p.- Coordinación del Programa de Posgrado

CARTA DE VOTO APROBATORIO

Mtro. Guillermo Domínguez Aguilar
DECANO DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como **ASESOR** designado del estudiante **CLAUDIA COLLAZO GUMA** con ID 557721 quien realizó la tesis titulado/a: **EVALUACIÓN DE LA COMPOSICIÓN INTERNA DE OBRAS DE ARTE USANDO ALGORITMOS DE APRENDIZAJE SUPERVISADO Y ESPECTROSCOPÍA TERAHERTZ**, un trabajo propio, innovador, relevante e inédito y con fundamento en la fracción IX del Artículo 43 del Reglamento General de Posgrados, doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que *el* pueda continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"

Aguascalientes, Ags., a 30 de junio de 2026.


Dr. Enrique Castro Camus
Asesor de tesis

c.c.p.- Interesado
c.c.p.- Coordinación del Programa de Posgrado

Dictamen de liberación académica

**DICTAMEN DE LIBERACIÓN ACADÉMICA
PARA INICIAR LOS TRÁMITES DEL EXAMEN DE GRADO**

Fecha de dictaminación (dd/mm/aaaa): 22/06/2026

NOMBRE: Claudia Collazo Guma **ID:** 557721

PROGRAMA: Maestría en Ciencias con Opciones a la Computación, Matemáticas Aplicadas **LGAC (del posgrado):** Computación - Inteligencia Artificial

MODALIDAD DEL PROYECTO DE GRADO: Tesis Tradicional (x) *Tesis por artículos científicos () **Tesis por Patente () Trabajo Práctico ()

TÍTULO: Evaluación de la composición interna de obras de arte usando algoritmos de aprendizaje supervisado y espectroscopia Terahertz.

IMPACTO SOCIAL (señalar el impacto logrado): Este trabajo explora el uso de la espectroscopia junto con aprendizaje automático para detectar la composición interna de obras de arte, como el número de capas, con su correspondiente grosor e índice de refracción, de una manera no invasiva.

INDICAR SEGÚN CORRESPONDA:

Elementos para la revisión académica del trabajo de tesis o trabajo práctico:	
SI	El trabajo es congruente con las LGAC del programa de posgrado
SI	La problemática fue abordada desde un enfoque multidisciplinario
SI	Existe coherencia, continuidad y orden lógico del tema central con cada apartado
SI	Los resultados del trabajo dan respuesta a las preguntas de investigación o a la problemática que aborda
SI	Los resultados presentados en el trabajo son de gran relevancia científica, tecnológica o profesional según el área
SI	El trabajo demuestra más de una aportación original al conocimiento de su área
SI	Las aportaciones responden a los problemas prioritarios del país
NO	Generó transferencia del conocimiento o tecnológica
SI	Cumple con la ética para la investigación (reporte de la herramienta antiplagio)
El egresado cumple con lo siguiente:	
SI	Cumple con lo señalado por el Reglamento General de Posgrados
SI	Cumple con los requisitos señalados en el plan de estudios (créditos curriculares, optativos, actividades complementarias, estancia, predoctoral, etc.)
SI	Cuenta con los votos aprobatorios del comité tutorial
SI	Cuenta con la carta de satisfacción del Usuario (En caso de que corresponda)
SI	Coincide con el título y objetivo registrado
SI	Tiene congruencia con cuerpos académicos
SI	Tiene el CVU de la SECINT actualizado
NA	Tiene el o los artículos aceptados o publicados y cumple con los requisitos institucionales (en caso de que proceda)
*En caso de Tesis por artículos científicos publicados (completar solo si la tesis fue por artículos)	
NA	Aceptación o Publicación de los artículos en revistas indexadas de alto impacto según el nivel del programa
NA	El (la) estudiante es el primer autor(a)
NA	El (la) autor(a) de correspondencia es el Director (a) del Núcleo Académico
NA	En los artículos se ven reflejados los objetivos de la tesis, ya que son producto de este trabajo de investigación.
NA	Los artículos integran los capítulos de la tesis y se presentan en el idioma en que fueron publicados
**En caso de Tesis por Patente	
NA	Cuenta con la evidencia de solicitud de patente en el Departamento de Investigación (anexarla al presente formato)

Con base en estos criterios, se autoriza continuar con los trámites de titulación y programación del examen de grado:

FIRMAS

Elaboró:

*NOMBRE Y FIRMA DEL(LA) CONSEJERO(A) SEGÚN LA LGAC DE ADSCRIPCIÓN: DR. ROGELIO SALINAS GUTIÉRREZ

* En caso de conflicto de intereses, firmará un revisor miembro del NA de la LGAC correspondiente distinta al director o miembro del comité tutorial, asignado por el Decano.

NOMBRE Y FIRMA DEL COORDINADOR DE POSGRADO: DRA. MARIANA ALFARO GÓMEZ

Revisó:

NOMBRE Y FIRMA DEL SECRETARIO DE INVESTIGACIÓN Y POSGRADO: DRA. ILIANA ERNESTINA MEDINA RAMÍREZ

Autorizó:

NOMBRE Y FIRMA DEL DECANO: MTRO. GUILLERMO DOMÍNGUEZ AGUILAR

Nota: procede el trámite para el Depto. de Apoyo al Posgrado

En cumplimiento con el Art. 24 fracción V del Reglamento General de Posgrado, que a la letra señala entre las funciones del Consejo Académico: Proponer criterios y mecanismos de selección, permanencia, egreso y titulación de estudiantes para asegurar la eficiencia terminal y la titulación y el Art. 28 fracción IV, atender, asesorar y dar el seguimiento del estudiantado desde su ingreso hasta su titulación.

Elaborado por: D. Apoyo al Posg.
Revisado por: D. Control Escolar/D. Gestión de Calidad.
Aprobado por: D. Control Escolar/ D. Apoyo al Posg.

Código: DO-SEE-FQ-15
Actualización: 02
Emisión: 11/08/25

Agradecimientos

Agradezco al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT), hoy Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI), el respaldo económico recibido a través de la beca de posgrado (CVU 2057087), sin el cual no habría podido cursar la maestría ni dedicarme de lleno a esta investigación.

A la Universidad Autónoma de Aguascalientes le debo la oportunidad de incorporarme a su programa de posgrado, junto con la formación, la infraestructura y los recursos que estuvieron a mi disposición durante todo el proceso.

Mi reconocimiento al comité tutorial que acompañó este trabajo de principio a fin. A mi tutora, la Dra. Mariana Alfaro Gómez, por su dirección constante; a mi co-tutora, la Dra. Luz Judith Rodríguez Esparza, por su mirada atenta; y a mi asesor, el Dr. Enrique Castro Camus, por su consejo experto. Sus observaciones fueron decisivas para dar forma y rigor a esta tesis.

Índice general

Índice de tablas	3
Índice de figuras	4
Acrónimos	5
Resumen en Español	6
Abstract	7
1 Introducción	8
2 Marco teórico	10
2.1 Estado del arte	10
2.2 Fundamentos teóricos	11
2.2.1 Principios de la espectroscopia THz en el dominio del tiempo	11
2.2.2 Análisis estratigráfico y procesamiento de la señal en reflexión	13
2.2.3 Aprendizaje automático para el análisis espectroscópico	14
3 Metodología	17
3.1 Modelo físico directo de reflexión multicapa	19
3.1.1 Pulso de referencia y dominio de la frecuencia	19
3.1.2 Recursión de Fresnel–Rouard con propagación de fase	20
3.1.3 Descripción de la capa: índice real y espesor	22
3.1.4 Generación del pulso reflejado simulado	22
3.1.5 Verificación numérica del generador	24
3.2 Generación del conjunto de datos sintético	26
3.2.1 Parámetros físicos y rangos de muestreo	26
3.2.2 Muestreo balanceado por número de capas	26
3.2.3 Capas degeneradas y casi degeneradas	28
3.2.4 Modelado de la clase sin pintura ($L = 0$)	28
3.2.5 Soporte plástico y ruido de medición	28
3.2.6 Volumen, partición por clases y solapamiento entre conjuntos	30
3.3 Modelo: CNN multitarea	31
3.3.1 Representación de entrada	31
3.3.2 Arquitectura de la red	31
3.3.3 Cabezas de salida: conteo y parámetros	32
3.4 Entrenamiento	33
3.4.1 Función de pérdida multitarea (enmascarada y normalizada)	33
3.4.2 Aumento de datos con ruido	33
3.4.3 Optimización, <i>scheduler</i> y parada temprana	34
3.5 Línea base tradicional: deconvolución de Wiener y detección de picos	35
3.6 Protocolo de evaluación	36
3.6.1 Conjunto de prueba y condiciones de inferencia	36

3.6.2	Métricas reportadas y esquema oracle	36
3.7	Reproducibilidad	38
4	Resultados	39
4.1	Desempeño del sistema principal sobre el conjunto sintético	39
4.1.1	Conteo de capas	39
4.1.2	Regresión de parámetros en modo <i>oracle</i>	40
4.1.3	Dónde ubica la red los ecos	40
4.2	Validación con mediciones reales de pintura automotriz	40
4.2.1	Estratigrafía verdadera por microscopía de corte	40
4.2.2	Predicciones sobre las muestras reales	43
4.2.3	Contraste con el método tradicional de deconvolución	45
5	Discusión de resultados	47
5.1	Respuesta a la pregunta de investigación	47
5.2	Comparación con trabajos relacionados	47
6	Conclusiones	49
	Referencias	50

Índice de tablas

3.1	Verificación numérica del generador de pulsos	24
3.2	Rangos de muestreo del simulador	26
3.3	Configuración de la rama convolucional de la CNN	32
3.4	Hiperparámetros de entrenamiento de la CNN	35
4.1	Exactitud de conteo del sistema principal (sintético)	39
4.2	Regresión de parámetros en modo oracle	40
4.3	Estratigrafía verdadera de las cinco muestras (microscopía)	40
4.4	Predicciones del sistema principal sobre mediciones reales	45
4.5	Salidas de la línea base tradicional (Wiener)	45
4.6	Comparación contra la estratigrafía verdadera (microscopía)	46
5.1	Comparación con trabajos relacionados de medición de estructura por THz	47

Índice de figuras

3.1	Visión general del método propuesto	18
3.2	Pulso de referencia del sistema THz-TDS	19
3.3	Modelo físico directo de reflexión multicapa	21
3.4	El pulso reflejado como superposición de ecos	23
3.5	Verificación del generador sobre dos simulaciones	25
3.6	Distribución del número de capas en el conjunto sintético	27
3.7	Efecto del ruido de medición sobre un pulso	29
3.8	Composición de los subconjuntos por número de capas	30
3.9	Arquitectura de la CNN multitarea	31
3.10	Evolución de la tasa de aprendizaje (scheduler coseno)	34
3.11	Evaluación en modo oracle de la cabeza de regresión	37
4.1	Matriz de confusión del conteo de capas	39
4.2	Ubicación de los ecos sobre pulsos sintéticos	41
4.3	Micrografías de los cortes transversales	42
4.4	Entradas de la validación con mediciones reales	43
4.5	Ubicación de los ecos predichos sobre mediciones reales	44

Acrónimos

ANN	Red neuronal artificial (<i>artificial neural network</i>).
A-scan	Registro temporal completo de la señal THz en un punto de la superficie; unidad básica de los datos.
B-scan	Imagen de corte transversal construida acumulando A-scans a lo largo de una línea de barrido.
BIC	Criterio de información bayesiano (<i>Bayesian information criterion</i>).
BN	Normalización por lotes (<i>batch normalization</i>).
C-scan	Imagen en planta de la superficie, asignando a cada píxel un parámetro extraído de su forma de onda.
CNN	Red neuronal convolucional (<i>convolutional neural network</i>).
CPU	Unidad central de procesamiento (<i>central processing unit</i>).
CUDA	Plataforma de cómputo paralelo de NVIDIA (denominación comercial).
GPU	Unidad de procesamiento gráfico (<i>graphics processing unit</i>).
JSON	Formato de intercambio de datos basado en texto (<i>JavaScript Object Notation</i>).
MAE	Error absoluto medio (<i>mean absolute error</i>); los subíndices d y n indican el parámetro evaluado.
MSE	Error cuadrático medio (<i>mean squared error</i>).
ReLU	Unidad lineal rectificada (<i>rectified linear unit</i>).
RMS	Valor cuadrático medio (<i>root mean square</i>).
RMSE	Raíz del error cuadrático medio (<i>root mean squared error</i>).
SNR	Relación señal-ruido (<i>signal-to-noise ratio</i>).
SVM	Máquina de vectores de soporte (<i>support vector machine</i>).
THz	Terahercio (10^{12} Hz); por extensión, la banda espectral de 0.1 a 10 THz y su radiación.
THz-TDS	Espectroscopía de terahercios en el dominio del tiempo (<i>terahertz time-domain spectroscopy</i>).
ToF	Tiempo de vuelo (<i>time of flight</i>).
XPU	Acelerador de cómputo de Intel (denominación del fabricante para sus dispositivos aceleradores).

Resumen

El diagnóstico no invasivo de pinturas exige conocer su estructura interna —el número de capas y el espesor de cada una—. Los métodos tradicionales recuperan esa información a costa de muestreo destructivo, o recurren a imagen por rayos X, que identifica materiales pero no resuelve el espesor capa por capa. La espectroscopía de terahercios en el dominio del tiempo (THz-TDS) penetra las capas internas sin contacto ni daño, pero recuperar la disposición de las capas a partir de su señal constituye un problema inverso: las capas delgadas o de bajo contraste producen señales casi indistinguibles. Este trabajo aplica aprendizaje supervisado a datos THz-TDS para caracterizar pinturas multicapa, estimando el número de capas, el espesor y el índice de refracción de cada una.

Se desarrolló una red neuronal convolucional multitarea, entrenada íntegramente con pulsos sintéticos generados por un modelo físico de Fresnel-Rouard con referencia instrumental real y ruido medido sobre el equipo. Ante la ausencia de obras reales con estructura conocida, la validación se efectuó sobre pintura automotriz multicapa sobre soporte plástico, un sistema físico sustituto de espesores comparables.

Sobre el conjunto de prueba sintético, el conteo de capas alcanza una exactitud global del 77.4%, con el 92.6% de los errores concentrado entre clases vecinas. Contrastado contra microscopía de corte transversal, el sistema recupera el número de capas en dos de las cinco pinturas reales medidas, frente a ninguna de la línea base tradicional. Las capas finas se identifican como un límite de información, no del modelo. Dentro del dominio declarado, el aprendizaje supervisado resulta una técnica eficaz para caracterizar las capas de las pinturas por THz.

Palabras clave: espectroscopía de terahercios (THz-TDS), aprendizaje profundo, CNN 1D, problema inverso, diagnóstico de pinturas.

Abstract

The non-invasive diagnosis of paintings requires knowledge of their internal structure, namely the number of layers and the thickness of each, information that traditional methods recover only through destructive sampling, or by X-ray imaging, which identifies materials but does not resolve layer-by-layer thickness. Terahertz time-domain spectroscopy (THz-TDS) penetrates internal layers without contact or damage, but recovering the layer arrangement from its signal is an inverse problem: thin or low-contrast layers produce nearly indistinguishable signals. This work applies supervised learning to THz-TDS data to characterize multilayer paint, estimating the number of layers and the thickness and refractive index of each.

A multitask convolutional neural network was developed, trained entirely on synthetic pulses generated by a Fresnel-Rouard physical model with a real instrumental reference and noise measured on the instrument. In the absence of real artworks with known structure, validation was performed on multilayer automotive paint on a plastic substrate, a surrogate system of comparable thicknesses.

On the synthetic test set, layer counting reaches a global accuracy of 77.4%, with 92.6% of the errors concentrated between adjacent classes. Validated against cross-section microscopy, the system recovers the layer count in two of the five real paintings measured, versus none for the traditional baseline. Thin layers are identified as an information limit, not a model limit. Within the declared domain, supervised learning proves an effective technique for characterizing paint layers by THz.

Keywords: terahertz spectroscopy (THz-TDS), deep learning, 1D CNN, inverse problem, painting diagnosis.

Capítulo 1 Introducción

En la restauración y conservación de pinturas artísticas, conocer la composición interna de la obra, cuántas capas tiene y qué tan gruesas son es necesario para tomar decisiones de intervención. Determinar con precisión el número de capas y su espesor exige un corte transversal, esto es, extraer una muestra física de la obra (Artesani, Abate et al., 2023; Reyes-Reyes, 2020). Las técnicas de imagen no invasivas más habituales ofrecen información distinta: la radiografía X revela estructuras internas y detecta materiales inorgánicos (Fukunaga et al., 2007), y la reflectografía infrarroja revela el dibujo subyacente y penetra solo hasta la imprimatura (Artesani, Abate et al., 2023; Guillet et al., 2017). En este contexto, la técnica de espectroscopia de terahercios en el dominio del tiempo (THz-TDS) alcanza las capas de preparación más profundas y permite visualizar capas internas que otros métodos no alcanzan (Guillet et al., 2017).

La radiación de terahercios (THz) se encuentra en la región del espectro electromagnético situada entre las microondas y el infrarrojo lejano. Por su baja energía de fotón, las ondas THz son no ionizantes y no degradan los materiales que atraviesan, por lo que su uso no implica riesgo ni para la obra ni para quien maneja el equipo; además, al propagarse a través de medios ópticamente opacos, dan acceso a la estructura interna de la pieza (Jäckel et al., 2025; Reyes-Reyes et al., 2024). Las ondas THz pueden excitar modos vibracionales de la materia y, en algunos materiales como ciertos pigmentos inorgánicos, producen picos de absorción característicos que ayudan a identificarlos, aunque no todos los materiales presentan una firma espectral específica en esta banda (Fukunaga et al., 2007; Kubiczek & Balzer, 2022). La medición de capas, en cambio, no depende de esas firmas, sino de que la pintura es relativamente transparente al THz: el pulso la atraviesa y se refleja parcialmente en cada interfaz entre capas (Jäckel et al., 2025; Krimi et al., 2016). En la geometría de reflexión que emplea este trabajo, la amplitud de cada eco está dada por el coeficiente de reflexión de Fresnel de la interfaz, que depende del contraste de índice de refracción entre las capas contiguas, mientras que el retardo entre ecos consecutivos es proporcional al espesor óptico de la capa intermedia (Krimi et al., 2016).

Extraer el número de capas y sus espesores a partir de la señal THz es un problema inverso, en el que a partir de una señal medida se reconstruyen los parámetros del sistema que la generó. Cuando las capas son muy delgadas o el contraste entre materiales adyacentes es pequeño, los ecos reflejados en las interfaces se solapan y configuraciones distintas producen señales casi indistinguibles (Reyes-Reyes et al., 2024). El procesamiento tradicional resuelve esta inversión con métodos iterativos de búsqueda de raíces, como el de Newton-Raphson, que son computacionalmente lentos y pueden no converger, mientras que las aproximaciones analíticas que los evitan sacrifican exactitud (Klokkou et al., 2022). La dificultad crece en obras con superficies irregulares como los frescos, donde la distorsión geométrica altera la fase y debe corregirse antes de la extracción en un paso que no siempre resulta directo (Artesani, Abate et al., 2023). El ajuste de los parámetros de procesamiento exige además un análisis detallado de los datos caso por caso, lo que dificulta su aplicación sistemática (Klokkou et al., 2022; Lambert et al., 2022).

Para superar estas limitaciones, este trabajo plantea el siguiente objetivo general: Aplicar algoritmos de aprendizaje supervisado en el análisis de imágenes y datos espectroscópicos en la banda de terahertz (THz), con el fin de identificar estructuras internas, materiales y capas en obras de arte mediante una técnica no invasiva. El aprendizaje automático es una rama de la inteligencia artificial que se basa en entrenar sistemas computacionales para que identifiquen patrones complejos en grandes volúmenes de datos, aprendiendo de ellos sin necesidad de ser programados con reglas matemáticas explícitas (Park & Son, 2021). Dentro de este paradigma, las redes neuronales artificiales (ANN) destacan como una herramienta especialmente adecuada para el procesamiento de señales espectroscópicas, pues pueden aprender directamente la compleja relación entre los datos THz y los parámetros físicos de los materiales, eliminando la necesidad de múltiples pasos de preprocesamiento manual y ofreciendo una vía natural para abordar el problema inverso descrito anteriormente (Klokkou et al., 2022). Un modelo entrenado bajo este enfoque mejora la precisión de los métodos analíticos basados en aproximaciones, resulta más rápido que los métodos iterativos y se mantiene más estable que los métodos clásicos en el rango de baja frecuencia, donde la relación señal-ruido decae (Klokkou et al., 2022).

Este objetivo general se concreta en los siguientes objetivos específicos:

1. Determinar la relación entre los datos obtenidos de la espectroscopía y las propiedades de los materiales de las muestras de pintura multicapa usando algoritmos de aprendizaje supervisado.
2. Determinar el número de capas de una muestra de pintura multicapa y su espesor a partir de los datos de espectroscopía usando algoritmos de aprendizaje supervisado.
3. Evaluar el nivel de desempeño del uso de algoritmos de aprendizaje supervisado para el análisis de pintura multicapa con espectroscopía terahertz.

Validar un método de estimación multicapa requiere muestras cuya estructura interna se conozca, y esa información solo se obtiene por análisis destructivos como el corte transversal, imposibles de aplicar sobre obras de arte reales y, justamente, el tipo de intervención que la técnica busca evitar. Ante la carencia de muestras pictóricas reales con estructura conocida, la validación experimental de este trabajo se realiza sobre sistemas de pintura multicapa automotriz depositada sobre soporte plástico, que operan como sistema físico sustituto. Estas muestras comparten con las capas pictóricas su naturaleza estratificada de capas poliméricas pigmentadas y espesores del mismo orden, propiedades por las que la pintura automotriz multicapa es un sistema modelo habitual en la literatura de medición de espesores multicapa por THz (Krimi et al., 2016; van Mechelen et al., 2014).

Con base en lo anterior, este trabajo busca responder a la siguiente pregunta de investigación: ¿puede el aprendizaje supervisado, aplicado a datos espectroscópicos en la banda de THz, identificar de forma no invasiva la estructura interna de obras de arte, es decir, el número de capas y, para cada una, su espesor e índice de refracción? Para abordar esta cuestión, el presente documento se estructura de la siguiente manera: en la sección de Marco Teórico se revisa la literatura existente sobre el análisis de arte con THz; la sección de Metodología detalla el enfoque cuantitativo y experimental seguido; en la sección de Resultados se presentan los hallazgos y, finalmente, las Conclusiones resumen las implicaciones del estudio.

Capítulo 2 Marco teórico

El análisis no invasivo de obras de arte mediante terahercios se apoya en los trabajos previos que han aplicado la técnica al patrimonio y en los fundamentos físicos y de aprendizaje automático sobre los que se construye este trabajo.

2.1. Estado del arte

El análisis no invasivo de obras de arte requiere técnicas que sean no destructivas, sin contacto y capaces de revelar información de la estructura interna. Cada técnica disponible presenta limitaciones que otras compensan: la radiografía y el análisis por rayos X revelan estructuras internas ocultas y detectan pigmentos inorgánicos densos como los de plomo, omnipresentes en la pintura histórica, aunque no penetran fácilmente a través de ellos y no resuelven las capas situadas por debajo (Abraham et al., 2010); la espectroscopía infrarroja media identifica materiales orgánicos comparándolos con bases de datos de referencia, aunque sus espectros son difíciles de interpretar y, en mezclas de pigmentos, las bandas de absorción de los distintos componentes acaban solapándose (Fukunaga et al., 2007); la espectroscopía Raman es eficaz para detectar vibraciones simétricas de compuestos inorgánicos, pero presenta dificultades para acceder a modos de baja frecuencia y el aumento de potencia del láser puede dañar superficies pictóricas (Artesani, Lamuraglia et al., 2023); y las microondas, pese a resolver bien en profundidad, no alcanzan la resolución lateral que exige el detalle pictórico (Abraham et al., 2010). Estas técnicas se diferencian además en su profundidad de penetración: el ultravioleta no pasa del barniz, la luz visible queda en las capas pictóricas y el infrarrojo llega hasta la imprimatura, mientras que el THz es la única que alcanza las capas de preparación más profundas y describe la estratigrafía completa del objeto, de la superficie al soporte (Guillet et al., 2017; Lambert et al., 2022). A esta capacidad se suma la portabilidad de los sistemas THz, que permite llevar el equipo directamente a los museos o lugares de conservación sin necesidad de trasladar las obras (Lambert et al., 2022).

Frente a estas alternativas, la espectroscopía THz se ha aplicado al estudio de obras de arte en tareas que van desde la identificación de materiales hasta la automatización del análisis estratigráfico.

Fukunaga et al. (2007) sentaron las bases de la aplicación de THz al patrimonio cultural al construir la primera base de datos de espectros THz de materiales artísticos (pigmentos, resinas y polímeros sintéticos), con la que demostraron el potencial de la técnica para identificar y distinguir los materiales de una obra.

A partir de este fundamento espectral, investigaciones posteriores extendieron la técnica hacia el análisis estratigráfico de obras con geometrías más complejas. Artesani, Abate et al. (2023) abordaron el problema de la curvatura superficial en frescos, que distorsiona la fase de la señal THz e impide determinar con precisión el coeficiente de absorción. Para solucionarlo, propusieron un método autorreferenciado, un paso de posprocesamiento computacional que corrige la distorsión en la fase mediante la interpolación de la curvatura con una función polinómica, cuya calidad de ajuste debe ser verificada por el investigador. Una vez corregida la señal, la transformaron al dominio de la frecuencia mediante la transformada de Fourier, calcularon el coeficiente de absorción por píxel para crear mapas de distribución de materiales, e identificaron pigmentos a partir de sus picos de absorción. Para estudiar las capas internas usaron deconvolución temporal, una técnica que separa las señales reflejadas por las distintas interfaces del fresco. Aun así, los autores admiten que el método pierde precisión cuando la superficie tiene una curvatura muy compleja.

En esta misma dirección, Reyes-Reyes et al. (2024) aplicaron THz-TDS en geometría de reflexión para analizar un ícono sobre madera. Escanearon la superficie píxel por píxel, asociando a cada uno una forma de onda THz completa y, con esos datos, hicieron dos análisis para caracterizarla: una imagen de amplitud pico a pico para localizar elementos metálicos y un análisis de tiempo de vuelo (ToF) para estudiar las deformaciones (Reyes-Reyes et al., 2024). Para la estructura interna, obtuvieron escaneos transversales (B-scans) y los procesaron con un algoritmo de deconvolución dispersa, que reconstruye los ecos asumiendo que solo unas pocas interfaces reflejan señal, para mejorar la resolución y separar los ecos de capas delgadas. No obstante, varias etapas exigieron intervención humana, como inspeccionar los espectros a simple vista para confirmar si había un pigmento (Reyes-Reyes et al., 2024).

Aunque la obtención de datos ha mejorado, todavía hay que identificar las capas y calcular su grosor con métodos

manuales o cálculos iterativos. Tradicionalmente, la estructura de una obra se determina a través del tiempo de vuelo (ToF), midiendo cuánto tardan en regresar los ecos reflejados en cada interfaz, como la de aire-barniz o la de barniz-pigmento (Guillet et al., 2017). En este método clásico, el investigador debe reconocer visualmente los picos de reflexión en la forma de onda o aplicar deconvolución estándar para intentar distinguir las señales que se solapan cuando las capas son muy delgadas. Posteriormente, para conocer el espesor físico real, hace falta estimar o saber de antemano el índice de refracción de cada capa (Guillet et al., 2017). Este proceso tradicional se vuelve poco escalable en escaneos de gran formato y pierde precisión ante señales ruidosas o capas extremadamente delgadas. (Abraham et al., 2010; Lambert et al., 2022).

Como primer paso para reducir esta dependencia manual, Lambert et al. (2022) presentaron un caso práctico que une el análisis de THz con un algoritmo capaz de automatizar el procesamiento de los datos. Al evaluar una pintura sobre tabla del siglo XV, buscaban localizar las separaciones entre las capas de pintura y comprobar si una consolidación previa había funcionado. Lo nuevo de su trabajo fue un algoritmo que analiza automáticamente las formas de onda de THz para identificar áreas con defectos mediante una regla fija, que marcaba como defecto todo punto donde la segunda reflexión conservaba una amplitud de al menos el 80 % de la primera. Con este método generaron mapas de delaminación que mostraban visualmente los daños, descubriendo incluso defectos que los conservadores no habían detectado. Si bien este enfoque no utiliza aprendizaje automático, demostró que la automatización mediante reglas fijas ya supone un avance frente a los métodos manuales.

Sin embargo, la rigidez de las reglas fijas dificulta adaptarse a la enorme variabilidad de las obras de arte reales. Para resolver este problema, el análisis de datos espectroscópicos ha avanzado hacia el aprendizaje automático que, en lugar de programar reglas matemáticas explícitas, estas técnicas aprenden a reconocer patrones complejos por sí solas a partir de datos de entrenamiento (Park & Son, 2021). Dentro de este campo, las redes neuronales artificiales (ANN) son uno de los modelos de aprendizaje supervisado aplicados a este fin: ajustan iterativamente sus pesos y sesgos para minimizar el error de predicción, aprendiendo de los datos (Park & Son, 2021).

Una de las primeras aplicaciones de este enfoque en THz fue en el trabajo de Busboom et al. (2020), que aborda un problema muy concreto pero importante, determinar cuántas capas de recubrimiento tiene una muestra cuando ese dato se desconoce, un paso previo necesario para muchos algoritmos que miden el espesor. Su propuesta consiste en usar una red neuronal profunda para clasificar la señal de THz y estimar si la muestra tiene una, dos o tres capas. Una clave de su metodología fue entrenar el modelo con un gran conjunto de mediciones simuladas, ante la dificultad de obtener datos reales. La red alcanzó una exactitud del 86.7 % sobre mediciones reales cuando estas se alinean temporalmente (sin ese alineamiento baja al 70.7 %), lo que muestra que un modelo entrenado con datos sintéticos puede generalizar a casos prácticos. El propio estudio reporta, no obstante, dos limitaciones ante muestras reales irregulares. El desempeño se degrada al aumentar el número de capas y, cuando el modelo se equivoca, tiende a subestimar ese número, prediciendo menos capas de las presentes (Busboom et al., 2020).

Mientras Busboom et al. (2020) se centraron en clasificar el número de capas, Klokou et al. (2022) abordaron la lentitud y la complejidad del análisis mediante redes neuronales para la extracción de parámetros a partir de datos THz-TDS. Su método entrena la red con datos simulados de interacciones luz-materia para que aprenda a relacionar la señal THz con las propiedades del material. El resultado es un modelo más preciso que los métodos analíticos basados en aproximaciones y más rápido que los iterativos de búsqueda de raíces, capaz de aliviar pasos costosos del análisis convencional como el desenrollado de fase o los tiempos de cómputo prolongados.

Li et al. (2020) abordan un problema distinto: la predicción del espectro de una estructura y su diseño inverso a partir de un espectro deseado. Los autores proponen usar redes neuronales para predecir el espectro de reflexión y diseñar una microestructura con una respuesta óptica específica, ya que los métodos tradicionales son lentos. Generaron 15,000 espectros de reflexión simulados y dividieron el problema en dos tareas: predicción directa y diseño inverso. En ambas obtuvieron un buen desempeño, demostrando que las redes neuronales superan a los métodos de simulación convencionales en velocidad y mantienen una alta precisión.

2.2. Fundamentos teóricos

2.2.1. Principios de la espectroscopía THz en el dominio del tiempo

La espectroscopía es una técnica analítica que estudia cómo la radiación electromagnética es absorbida o transmitida por la materia en función de su frecuencia, lo que permite identificar materiales y caracterizar sus propiedades (Fukunaga et al., 2007). En el contexto de la conservación del patrimonio cultural, esta capacidad de identificar de forma no invasiva los pigmentos, aglutinantes y barnices de una obra resulta esencial, pues estos materiales son característicos de cada época y de cada intervención: determinarlos permite estimar el período en que se creó la obra, reconocer cómo se ha conservado y detectar los materiales empleados en restauraciones previas (Fukunaga et al., 2007). Entre las distintas regiones del espectro

electromagnético, la banda de terahercios abarca de 0.1 a 10 THz, entre las microondas y el infrarrojo lejano (Abraham et al., 2010). Su utilidad para identificar materiales depende de qué se mueve al absorber la radiación: mientras en el infrarrojo responden sobre todo las vibraciones de enlaces y grupos moleculares, en el rango THz lo hace la molécula entera, y esa respuesta colectiva resulta más característica de cada material (Fukunaga et al., 2007). Por ser no ionizante, la radiación THz no daña los materiales que examina (Lambert et al., 2022); además atraviesa numerosos materiales dieléctricos con poca dispersión y cubre un ancho de banda amplio (Abraham et al., 2010).

En los sistemas THz-TDS, la generación del pulso THz se basa en la fotoexcitación de portadores libres en una antena fotoconductora polarizada por un láser de femtosegundo. Los portadores acelerados por el campo de polarización generan un pulso THz de banda ancha que se dirige hacia la muestra (Kubiczek & Balzer, 2022). En el dominio del tiempo, este pulso es un transitorio de muy corta duración, de un solo ciclo del campo eléctrico (Reyes-Reyes et al., 2024); por la relación inversa entre duración y contenido espectral, esa brevedad es la que le da el amplio ancho de banda que cubre la región THz. Para la detección, una fracción del haz óptico activa una segunda antena fotoconductora que actúa como receptor: la fotocorriente inducida en esta antena es proporcional a la amplitud del campo eléctrico THz en ese instante (Kubiczek & Balzer, 2022). Una línea de retardo óptica en el brazo del receptor permite muestrear la forma de onda del pulso THz punto a punto en el tiempo, reconstruyendo así su evolución temporal (Kubiczek & Balzer, 2022).

Lo que distingue a esta técnica de las convencionales es la información que guarda. Una medición de absorción o transmisión solo dice cuánta radiación pasa o se absorbe en cada frecuencia; en cambio, THz-TDS conserva toda la evolución del campo eléctrico en el tiempo, en cada punto de la superficie (Caballero Mendoza, 2010). A ese registro tomado en un punto se le llama A-scan, y es la unidad básica de los datos THz (Lambert et al., 2022). Tener la onda completa es lo que permite estudiar materiales formados por capas: cada límite entre ellas devuelve un eco en un instante distinto del A-scan, y esa secuencia de ecos revela cómo se ordenan las capas de la muestra. (Guillet et al., 2017).

La señal guarda también las propiedades ópticas del material, y estas se extraen directamente de un único registro THz, sin necesidad de aplicar las relaciones de dispersión de Kramers-Kronig (Artesani, Lamuraglia et al., 2023). Al aplicar la transformada de Fourier sobre la forma de onda registrada,

$$\tilde{E}(\omega) = \mathcal{F}\{E(t)\} = \int_{-\infty}^{\infty} E(t) e^{-i\omega t} dt = |\tilde{E}(\omega)| e^{i\varphi(\omega)}, \quad (2.1)$$

se recuperan a la vez la amplitud espectral $|\tilde{E}(\omega)|$ y la fase $\varphi(\omega)$ del campo eléctrico (Artesani, Lamuraglia et al., 2023).

La técnica puede operar tanto en geometría de transmisión como de reflexión. En transmisión, el emisor y el detector se sitúan a ambos lados de la muestra y el pulso la atraviesa, por lo que se requieren objetos delgados y accesibles por las dos caras. En reflexión, en cambio, emisor y detector se ubican del mismo lado y se registra el pulso que la muestra devuelve; esta es la geometría que se emplea para el análisis estratigráfico por tiempo de vuelo, pues permite examinar la sucesión de capas desde la cara pictórica de la obra. Resulta especialmente conveniente cuando el soporte impide la transmisión (como una tabla o un muro), aunque en lienzos delgados también es posible obtener imágenes en transmisión (Guillet et al., 2017). De cada píxel pueden derivarse descriptores distintos. La amplitud (máximo, mínimo o pico a pico) informa de la reflectividad local, el instante del pulso principal codifica la profundidad y el espectro obtenido tras la transformada de Fourier da acceso a la composición (Abraham et al., 2010).

El espectro THz permite acceder a la composición de la obra porque numerosos pigmentos inorgánicos presentan picos de absorción característicos en esta banda, los llamados espectros de huella dactilar, que distinguen incluso materiales de color visualmente idéntico, como el ultramar natural frente al artificial o el blanco de titanio frente al de plomo (Fukunaga et al., 2007). Este carácter distintivo no es universal, muchos polímeros y aglutinantes carecen de picos distintivos en la banda THz (Kubiczek & Balzer, 2022), y la lectura de la huella depende además de las condiciones de medición. Cuando varios materiales aparecen mezclados, sus espectros se superponen, de modo que los picos de los pigmentos base reaparecen en el espectro del conjunto (Fukunaga et al., 2007). A partir de la amplitud y la fase de la señal THz en el dominio de la frecuencia se obtiene el índice de refracción complejo del material, definido como:

$$\tilde{n}(\omega) = n(\omega) + i\kappa(\omega) \quad (2.2)$$

donde $n(\omega)$ es el índice de refracción real y $\kappa(\omega)$ el coeficiente de extinción. En geometría de reflexión, ambos se obtienen a partir del coeficiente de reflexión de Fresnel, que para una interfaz a incidencia normal entre el aire (de índice $\tilde{n}_0 = 1$) y la muestra se expresa como

$$r(\omega) = \frac{\tilde{n}_0(\omega) - \tilde{n}(\omega)}{\tilde{n}_0(\omega) + \tilde{n}(\omega)}. \quad (2.3)$$

Como $r(\omega)$ se recupera de la función de transferencia entre la señal de la muestra y la de una referencia altamente reflectante, esta expresión puede invertirse para despejar el índice de refracción complejo del material (Artesani, Lamuraglia et al., 2023):

$$\tilde{n}(\omega) = \tilde{n}_0 \frac{1 - r(\omega)}{1 + r(\omega)}. \quad (2.4)$$

Este principio de aditividad espectral y la extracción directa de los parámetros ópticos sostienen el análisis de obras de arte, donde pigmentos, aglutinantes y barnices se superponen en capas o aparecen mezclados. Aun así, la señal no depende solo de la composición, el tamaño de partícula del pigmento modifica el espectro de absorción, porque con granos gruesos la onda THz se dispersa y los picos salen más anchos y difusos (Fukunaga et al., 2007). El espesor de la capa, en cambio, no introduce nuevos picos de absorción; sino que cambia la señal de otra forma. En transmisión, un espesor elevado en un pigmento muy absorbente atenúa la señal hasta impedir la observación de la huella espectral (Fukunaga et al., 2007); en geometría de reflexión, el espesor se manifiesta en el dominio temporal como el retardo entre los ecos de las dos interfaces de la capa, dado por $\tau = 2nd/c$, donde n es el índice de refracción y d el espesor físico. Por ello es importante distinguir entre el espesor óptico, que la señal THz permite determinar directamente a partir del tiempo de vuelo del pulso, y el espesor físico real de la capa, cuya obtención requiere conocer el índice de refracción del material correspondiente (Artesani, Abate et al., 2023).

2.2.2. Análisis estratigráfico y procesamiento de la señal en reflexión

En geometría de reflexión, el pulso THz incide sobre la superficie de la muestra y cada interfaz entre materiales de distinto índice de refracción genera un eco reflejado. La posición temporal de cada eco indica directamente la presencia de una interfaz a lo largo de la dirección de propagación del haz, y el tiempo transcurrido entre ecos consecutivos es proporcional al espesor óptico de la capa comprendida entre esas dos interfaces (Guillet et al., 2017). Este principio permite reconstruir la secuencia de capas a partir del retardo entre ecos, sin requerir una medición de referencia. El análisis por tiempo de vuelo (ToF) se apoya en tres supuestos. Las capas casi no dispersan ni difractan el haz, las interfaces reflejan poco (de modo que las reflexiones múltiples pueden despreciarse) y el índice de refracción de cada capa es conocido y homogéneo (Guillet et al., 2017). Este último supuesto es una simplificación, ya que el índice no suele conocerse de antemano pero puede estimarse a partir de la amplitud de cada eco mediante el coeficiente de reflexión de Fresnel (Ec. 2.4). Sin embargo, por debajo de cierto espesor los ecos de las dos interfaces de la capa se solapan en el tiempo, de modo que el sistema no los resuelve como picos separados y queda fijado el límite inferior de resolución (Abraham et al., 2010).

Para superar este límite se recurre a la deconvolución, que busca recuperar la respuesta al impulso del sistema: una sucesión de picos estrechos, uno por cada interfaz, sin el ensanchamiento temporal que introduce la forma del pulso incidente (Dong et al., 2016). El punto de partida es modelar la señal reflejada medida $E_s(t)$ como la convolución entre el pulso incidente $e_{\text{ref}}(t)$ y la respuesta al impulso $h(t)$, que codifica la estructura estratificada de la muestra:

$$E_s(t) = (e_{\text{ref}} * h)(t). \quad (2.5)$$

Como la convolución en el tiempo equivale a un producto en frecuencia, la respuesta al impulso se recupera dividiendo el espectro de la señal medida $E_s(\omega)$ entre el del pulso de referencia $E_{\text{ref}}(\omega)$ y aplicando la transformada inversa:

$$h(t) = \mathcal{F}^{-1} \left\{ \frac{E_s(\omega)}{E_{\text{ref}}(\omega)} \right\}, \quad (2.6)$$

donde el cociente $E_s(\omega)/E_{\text{ref}}(\omega)$ es la función de transferencia del sistema (Dong et al., 2016),

$$H(\omega) = \frac{E_s(\omega)}{E_{\text{ref}}(\omega)} = A(\omega)e^{i\varphi(\omega)}, \quad (2.7)$$

cuya amplitud $A(\omega)$ y fase $\varphi(\omega)$ codifican la respuesta óptica del material (Artesani, Lamuraglia et al., 2023). Esta inversión directa, sin embargo, es inestable. En las frecuencias donde el pulso de referencia tiene poca energía, la división por valores muy pequeños amplifica el ruido y genera artefactos de alta frecuencia, que en el dominio del tiempo aparecen como oscilaciones espurias (*ringing*) (Dong et al., 2016).

Para estabilizar el proceso se introduce un filtro $F(\omega)$ que atenúa las frecuencias poco fiables antes de antitransformar:

$$h_f(t) = \mathcal{F}^{-1} \left\{ F(\omega) \frac{E_s(\omega)}{E_{\text{ref}}(\omega)} \right\}. \quad (2.8)$$

Esto equivale a reconstruir la señal con un pulso de entrada más corto y suave que el incidente, lo que estrecha los ecos y separa reflexiones que en la traza medida aparecían solapadas, y puede mejorar la resolución por debajo del ancho del pulso incidente cuando la relación señal-ruido lo permite. Como filtro suelen emplearse uno doble gaussiano, el de Wiener o una ventana de Hanning con una frecuencia de corte f_c ajustable, una elección que equilibra resolución temporal y supresión de ruido (Chen et al., 2010; Dong et al., 2016).

Una vez recuperada la respuesta al impulso, la estructura de capas se obtiene mediante detección de picos sobre $h_f(t)$: cada pico cuya amplitud supera un umbral fijado se interpreta como una interfaz, de modo que el número de picos determina el número de capas (Dong et al., 2016). El retardo entre picos consecutivos Δt proporciona el espesor óptico de la capa intermedia, y el espesor físico se obtiene como $\Delta x = c \Delta t / (2n)$ una vez conocido el índice de refracción n del material (Dong et al., 2017; Reyes-Reyes et al., 2024). Cuando las capas son tan delgadas que el filtrado no basta para separarlas, se recurre a variantes que incorporan información a priori sobre la estructura, como la deconvolución dispersa, que asume que solo unas pocas interfaces reflejan señal y favorece soluciones con un número reducido de ecos (Dong et al., 2017; Reyes-Reyes et al., 2024).

En reflexión, el procesamiento de la señal THz sigue una secuencia de pasos hasta obtener información espectroscópica y estructural fiable. El primero normaliza la señal reflejada por la muestra con la de una superficie de referencia altamente reflectante, lo que da la función de transferencia $H(\omega)$ (Ec. 2.7), cuyo espectro coincide con el de la respuesta al impulso empleada en la deconvolución (Ec. 2.6). El segundo paso aísla el pulso principal acotando la traza temporal con una ventana gaussiana centrada en él, lo que atenúa las reflexiones secundarias de las interfaces internas (Artesani, Lamuraglia et al., 2023). Ya en el dominio de la frecuencia (Ec. 2.1), el tercer paso obtiene el coeficiente de absorción del material mediante las ecuaciones de Fresnel para el índice de refracción complejo (Artesani, Lamuraglia et al., 2023).

En THz-TDS de reflexión, lo que se mide en un único punto de la muestra es la evolución temporal del pulso reflejado; esa traza recibe el nombre de A-scan, y sus ecos sucesivos marcan las interfaces del interior, cada una a su profundidad (Lambert et al., 2022). Para el análisis de la estructura interna, los A-scans individuales se acumulan a lo largo de una línea de escaneo para construir imágenes de corte transversal denominadas B-scans, a las que se asigna un mecanismo de contraste de color en función de la amplitud de los picos del campo eléctrico (Lambert et al., 2022). En estas imágenes, una bolsa de aire aparece como una franja de amplitud casi nula flanqueada por interfaces de reflexión bien marcada (Lambert et al., 2022). Si el escaneo cubre toda un área superficial y a cada píxel se le asigna un parámetro extraído de su forma de onda (por ejemplo, la amplitud pico a pico, que informa sobre la reflectancia local y permite localizar elementos metálicos), se obtiene una imagen en planta de la superficie analizada (Reyes-Reyes et al., 2024); este tipo de imagen de área se denomina habitualmente C-scan. Cuando la superficie presenta curvatura, los datos temporales pueden corregirse numéricamente alineando las posiciones máximas de cada forma de onda a un tiempo de referencia común (Guillet et al., 2017).

Finalmente, deben considerarse dos fuentes adicionales de distorsión: el rango dinámico del instrumento, que establece una frecuencia máxima por encima de la cual el coeficiente de absorción no puede determinarse de forma confiable (Artesani, Lamuraglia et al., 2023), y la absorción del vapor de agua atmosférico, que degrada la identificación espectroscópica y por ello suele eliminarse purgando el sistema con nitrógeno (Artesani, Abate et al., 2023).

2.2.3. Aprendizaje automático para el análisis espectroscópico

El aprendizaje automático aborda el análisis de señales espectroscópicas desde un enfoque basado en datos: en lugar de programar reglas explícitas, un modelo aprende, a partir de un conjunto de ejemplos, la relación entre la señal y la propiedad de interés (Gezimati & Singh, 2024; Park & Son, 2021). En el aprendizaje supervisado, esta relación adopta dos formas según el tipo de salida. Si la salida es una etiqueta discreta, como la categoría o el tipo de material de una muestra, la tarea es de clasificación; si es una magnitud continua, como el espesor o el índice de refracción, es de regresión (Busboom et al., 2020; Klokhou et al., 2022). Entre los modelos supervisados, las redes neuronales destacan en el procesamiento de señales espectroscópicas. Frente a clasificadores clásicos como las máquinas de vectores de soporte (SVM), rinden mejor cuando los materiales tienen propiedades dieléctricas parecidas y aprenden por sí solas las características relevantes de la señal, sin una etapa previa de extracción manual (Kubiczek & Balzer, 2022). Son además muy rápidas al inferir, ya que una vez entrenado el modelo resuelve en segundos lo que a los métodos convencionales les llevaría mucho más cómputo (Li et al., 2020).

El modelo base sobre el que se construyen arquitecturas más especializadas es la red neuronal feedforward, llamada

así porque la información fluye en una sola dirección, de la entrada a la salida y sin ciclos. Sus unidades de cómputo se organizan en una capa de entrada, una o más capas ocultas y una capa de salida. Cada neurona recibe las señales de la capa anterior, las combina en una suma ponderada y aplica al resultado una función de activación no lineal antes de transmitir su salida (Busboom et al., 2020; Ige & Sibiya, 2024). La activación más habitual es la unidad lineal rectificadora (ReLU), que anula los valores negativos y conserva los positivos:

$$\text{ReLU}(x) = \max(0, x). \quad (2.9)$$

Ha desplazado a otras como la sigmoide porque no se satura y atenúa el desvanecimiento del gradiente, lo que facilita entrenar redes de muchas capas (Ige & Sibiya, 2024). Los pesos de estas conexiones son los parámetros que el modelo aprende.

Cuando la entrada es una señal con estructura secuencial, como una serie temporal, se emplea una arquitectura especializada, la red neuronal convolucional (CNN), que es una red feedforward donde algunas capas realizan convoluciones (Ige & Sibiya, 2024). En una CNN unidimensional, un filtro o *kernel* pequeño se desliza a lo largo del único eje de la señal y en cada posición combina, mediante una suma ponderada, las muestras consecutivas que cubre. Al recorrer toda la señal se obtiene un mapa de características:

$$(x * w)(t) = \sum_{i=0}^{k-1} x(t+i) w(i), \quad (2.10)$$

donde x es la señal de entrada, w el filtro de tamaño k y $(x * w)(t)$ el valor del mapa de características en la posición t (Ige & Sibiya, 2024). El número de muestras que avanza el filtro en cada paso se denomina paso (*stride*), y el segmento de la señal que influye en una salida dada se conoce como campo receptivo. Como todas las posiciones se procesan con el mismo filtro, una capa convolucional aprende muchos menos parámetros que una capa totalmente conectada (compartición de parámetros) y reconoce un patrón con independencia de la posición en que aparezca; cada capa suele aplicar varios filtros distintos, de modo que cada uno se especializa en detectar un patrón local diferente (Gezimati & Singh, 2024; Wu, 2017).

Tras la convolución se aplica una función de activación no lineal y, con frecuencia, una capa de submuestreo (*pooling*), que reduce la dimensión del mapa de características agregando valores vecinos (por ejemplo, conservando el máximo de cada ventana); esto disminuye el costo de cómputo y aporta cierta tolerancia a pequeños desplazamientos de la señal (Gezimati & Singh, 2024; Wu, 2017). Al apilar varios bloques de este tipo, la red aprende de forma jerárquica. Las primeras capas detectan patrones locales simples y las profundas los combinan en representaciones más complejas, a la vez que el campo receptivo se amplía y abarca regiones cada vez mayores de la señal (Gezimati & Singh, 2024; Wu, 2017). Finalmente, una o varias capas totalmente conectadas integran las características extraídas para producir la salida, que será una etiqueta en una tarea de clasificación o un valor continuo en una de regresión (Ige & Sibiya, 2024).

Tanto la red feedforward como la CNN se entrenan de la misma manera. El aprendizaje lo guía la función de pérdida, que mide la discrepancia entre la predicción del modelo y el valor correcto, y entrenar consiste en ajustar los pesos para minimizarla (Goodfellow et al., 2016). Para ello se recurre a la retropropagación y al descenso de gradiente, que obtienen el gradiente de la pérdida respecto a cada peso y lo actualizan en la dirección que reduce el error,

$$w \leftarrow w - \eta \frac{\partial \mathcal{L}}{\partial w}, \quad (2.11)$$

donde η es la tasa de aprendizaje y $\mathcal{L}$ la función de pérdida. En la práctica, esa actualización se delega en un optimizador como Adam (Kingma & Ba, 2015; Kubiczek & Balzer, 2022; Wu, 2017). El proceso depende de hiperparámetros que no se aprenden de los datos, sino que se fijan de antemano, como la tasa de aprendizaje, el tamaño del lote y el número de épocas (Kubiczek & Balzer, 2022). La capacidad del modelo (su profundidad y su número de unidades) es otra decisión de diseño. Una red más grande capta relaciones más complejas y suele mejorar la precisión, pero suma parámetros y alarga el entrenamiento, de modo que la arquitectura equilibra exactitud y costo computacional (Goodfellow et al., 2016).

Aumentar la capacidad eleva el riesgo de sobreajuste, esto es, que el modelo memorice los datos de entrenamiento en lugar de aprender patrones que generalicen. Para contenerlo se usan técnicas de regularización como el *dropout*, que durante el entrenamiento desactiva al azar una fracción de las neuronas (Ige & Sibiya, 2024; Kubiczek & Balzer, 2022). El control del sobreajuste se apoya además en repartir los datos en tres subconjuntos (entrenamiento, validación y prueba) con funciones distintas (Busboom et al., 2020). El de entrenamiento ajusta los pesos, mientras que el de validación, que no interviene en ese ajuste, vigila el desempeño a lo largo del entrenamiento y permite detenerlo de forma anticipada (parada temprana) cuando la pérdida de validación deja de mejorar. El de prueba, del todo independiente, se reserva para evaluar el modelo

final una sola vez, de modo que la evaluación refleje su capacidad de generalizar a datos no vistos (Busboom et al., 2020; Kubiczek & Balzer, 2022).

La función de pérdida dirige la optimización, pero el desempeño del modelo se juzga con métricas de evaluación, que cuantifican la calidad de las predicciones y dependen del tipo de tarea. En las expresiones que siguen, n es el número de muestras, y_i el valor real de la i -ésima muestra, $\hat{y}_i$ el valor predicho por el modelo y $\bar{y}$ la media de los valores reales. En clasificación, la métrica más directa es la exactitud (*accuracy*), la proporción de muestras bien clasificadas sobre el total (Busboom et al., 2020):

$$\text{Exactitud} = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(\hat{y}_i = y_i), \quad (2.12)$$

donde $\mathbb{I}(\cdot)$ vale 1 cuando la predicción coincide con la etiqueta verdadera y 0 en caso contrario. La pérdida que suele optimizarse en clasificación es la entropía cruzada, que mide la distancia entre la distribución de probabilidad predicha y la verdadera sobre las C clases (Ige & Sibiyá, 2024):

$$\mathcal{L}_{\text{CE}} = - \sum_{c=1}^C y_c \log(\hat{y}_c). \quad (2.13)$$

En regresión, la pérdida más común es el error cuadrático medio (MSE), que al elevar al cuadrado penaliza con más fuerza los errores grandes (Ige & Sibiyá, 2024):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (2.14)$$

Para reportar el desempeño se añaden varias métricas derivadas, empezando por la raíz cuadrada del MSE, el RMSE, que expresa el error en las mismas unidades que la variable y resulta más interpretable (Park & Son, 2021):

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (2.15)$$

El error absoluto medio (MAE) promedia los valores absolutos de las diferencias y es menos sensible a los errores aislados que el MSE (Ige & Sibiyá, 2024):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (2.16)$$

Por último, el coeficiente de determinación (R^2) indica qué proporción de la variabilidad de los valores reales explica el modelo:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (2.17)$$

Su valor máximo es 1, que corresponde a un ajuste perfecto. Un valor de 0 indica que el modelo no predice mejor que la media de los datos, y los valores negativos, que predice peor que ella (Gezimati & Singh, 2024).

La preparación de los datos de entrada también influye directamente en la precisión alcanzable, y la normalización de las señales es un paso especialmente relevante. Estandarizar las entradas a media cero y varianza unitaria (la normalización *z-score*) lleva todas las variables a una escala común y mejora de forma medible la exactitud del modelo (Goodfellow et al., 2016). A esto se suma la elección del dominio de representación, temporal o frecuencial, que condiciona la capacidad de generalización (Kubiczek & Balzer, 2022). Como conseguir suficientes mediciones reales etiquetadas resulta costoso, una estrategia consolidada en THz consiste en generar los datos de entrenamiento por simulación física, modelando la interacción de la radiación con sistemas multicapa de parámetros controlados y agregando ruido para aproximar las condiciones de una medición real (Busboom et al., 2020; Li et al., 2020). El conjunto así generado se reparte en los subconjuntos de entrenamiento, validación y prueba ya descritos (Busboom et al., 2020).

Capítulo 3 Metodología

Este capítulo describe el proceso desarrollado para estimar el número de capas de una pintura multicapa y los parámetros ópticos de cada una, su espesor d_j e índice de refracción n_j , a partir del pulso reflejado $E_{\text{refl}}(t)$ registrado por THz-TDS en configuración de reflexión. En su forma directa, la física de la reflexión multicapa predice el pulso a partir de los parámetros del conjunto de capas; aquí se busca lo contrario, recuperar esos parámetros del pulso medido. Se trata, por tanto, de un problema inverso, ya que invierte la dirección habitual del modelo físico.

Para resolver ese problema inverso, un modelo físico directo basado en los coeficientes de Fresnel genera un conjunto de entrenamiento sintético. Con ese conjunto se entrena el modelo propuesto, una red neuronal convolucional unidimensional (CNN 1D) multitarea con dos cabezas de salida. Una cabeza determina cuántas capas hay, tarea de clasificación en la que se elige entre un número discreto de clases, y la otra estima los valores continuos de cada capa (su espesor e índice de refracción), lo que se conoce como regresión. La red resuelve ambas tareas de forma conjunta y en un solo paso (*end-to-end*), en vez de encadenar por separado un clasificador con varios regresores.

El uso de datos sintéticos responde a la ausencia de un conjunto real etiquetado a la escala requerida. Etiquetar una medición real exige conocer la estructura del punto medido, información que solo se obtiene por métodos destructivos como el corte transversal, imposibles de aplicar a las decenas de miles de ejemplos que exige entrenar una red. Solo un simulador con la física directa del problema permite generar ese volumen con etiquetas exactas.

Aunque la motivación de esta tesis es el diagnóstico de las capas internas de las obras de arte (Capítulo 1), su validación experimental se realiza sobre pintura multicapa automotriz depositada sobre soporte plástico. La razón es que validar un estimador de estructura interna exige muestras cuya composición de capas se conozca de antemano, y este trabajo no dispone de ellas para obras de arte reales, pero sí para el sistema automotriz. Este último es, además, el banco de pruebas establecido de la medición multicapa por THz-TDS (Krimi et al., 2016; van Mechelen et al., 2014), con capas poliméricas y pigmentadas como las pictóricas, espesores (5–60 μm , con barnices que superan los 90 μm (Krimi et al., 2016)) del mismo orden que los de interés, y muestras que pueden prepararse y medirse sin restricciones. Esta elección de validación condiciona dos aspectos del modelo. El sustrato se modela como plástico (Sección 3.2.5) y los rangos de espesor e índice del muestreo se fijan con datos de pintura automotriz (Sección 3.2.1). Como ambos quedan atados a ese sistema, los resultados valen para el caso automotriz, y su extensión al diagnóstico de obras de arte se plantea como límite del alcance y línea de trabajo futuro.

La exposición sigue el orden lógico del flujo de trabajo: el modelo físico directo de reflexión multicapa (Sección 3.1); la generación del conjunto de datos sintético (Sección 3.2); el modelo de CNN multitarea, incluida su representación de entrada (Sección 3.3); el protocolo de entrenamiento (Sección 3.4); la línea base tradicional de deconvolución contra la que se contrasta el método propuesto (Sección 3.5); el esquema de evaluación (Sección 3.6); y, por último, las condiciones de reproducibilidad (Sección 3.7).

La Figura 3.1 sirve como mapa de ruta de las fases centrales de ese flujo, desde el modelo físico directo hasta la evaluación del método propuesto.

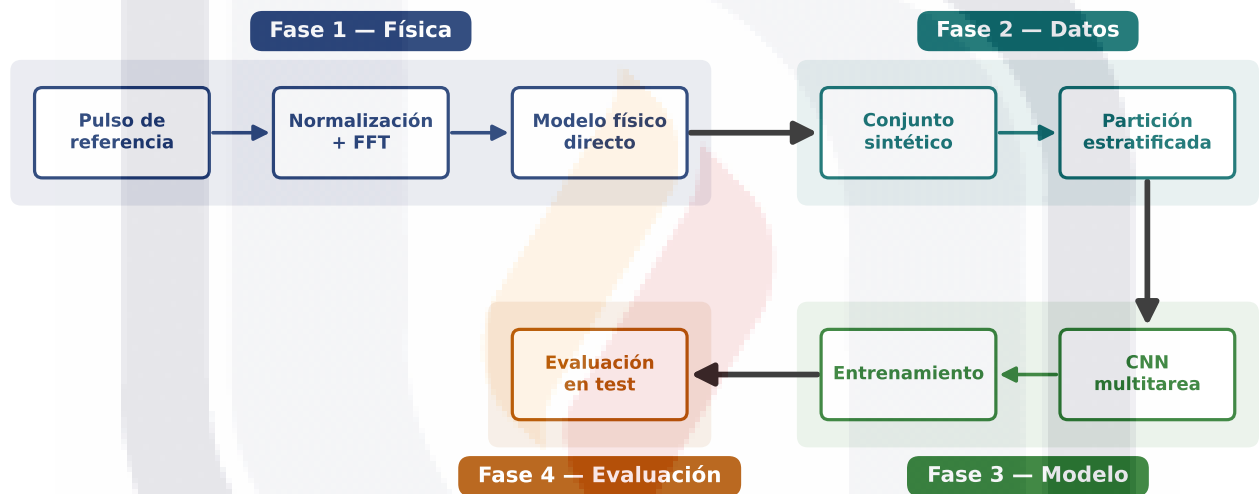


Figura 3.1. Visión general del método propuesto, organizado en cuatro fases. La fase física (azul) sintetiza el pulso reflejado de una pila multicapa mediante la recursión de Fresnel–Rouard (Sección 3.1). La fase de datos (verde azulado) construye el conjunto sintético balanceado por número de capas L y lo particiona en entrenamiento, validación y prueba (Sección 3.2). La fase de aprendizaje (verde) entrena la CNN multitarea para predicción conjunta de L y los parámetros (d, n) de cada capa (Secciones 3.3 y 3.4). La fase de evaluación (naranja) cuantifica la exactitud de conteo y el MAE / R^2 de regresión en modo *oracle* (Sección 3.6).

3.1. Modelo físico directo de reflexión multicapa

El modelo físico directo funciona como simulador y, a partir del índice de refracción y el espesor de cada capa de la muestra, calcula el pulso reflejado que registraría el equipo THz-TDS. Para ello parte de un pulso de referencia real medido en aire y propaga el campo a través de la pila de capas mediante la recursión de Fresnel-Rouard, que compone las reflexiones de todas las interfaces con la propagación dentro de cada capa. Las subsecciones siguientes detallan la construcción de la referencia y el paso al dominio de la frecuencia, la recursión de Fresnel-Rouard, la descripción de cada capa por su índice y espesor, y la síntesis final del pulso reflejado simulado.

3.1.1. Pulso de referencia y dominio de la frecuencia

El procedimiento parte de una señal medida, no simulada: el pulso de referencia $E_{\text{ref}}(t)$. Usar una medición real como punto de partida traslada a los datos sintéticos el ancho de banda y la forma del equipo.

Esta señal se obtiene de tres barridos en aire adquiridos con el sistema de imágenes terahertz, cada uno una imagen de 50×50 píxeles con una forma de onda por píxel. El pulso de referencia se construye promediando las 7,500 formas de onda de los tres barridos. Como las réplicas son casi idénticas (su correlación supera 0.999), el promedio cancela el ruido aleatorio de cada píxel sin deformar la señal. El eje temporal asociado se construye como una sucesión uniforme de T muestras con paso constante Δt , es decir $t_k = k \Delta t$ para $k = 0, 1, \dots, T - 1$, donde Δt se lee directamente de los metadatos binarios del instrumento.

Al promedio se le aplica además una inversión global de signo, $E_{\text{ref}} = -\bar{E}_{\text{aire}}$, que corrige la polaridad de la referencia. El barrido «en aire» del instrumento no registra el pulso incidente, sino su reflexión sobre un soporte metálico, como confirma su correlación de +0.97 con una medición directa de metal desnudo. Esa reflexión equivale a multiplicar por un coeficiente de Fresnel $r \approx -1$, de modo que invertir el signo del promedio recupera el pulso incidente, algo que se verifica sobre el caso sin pintura ($L = 0$). Sin la corrección, el pulso simulado queda invertido respecto de la medición real de metal (-0.97), mientras que con ella ambos coinciden (+0.968 tras el alineamiento temporal), valor que se conserva como prueba de regresión del simulador. La Figura 3.2 muestra el pulso de referencia resultante tras el promediado y la inversión de signo.

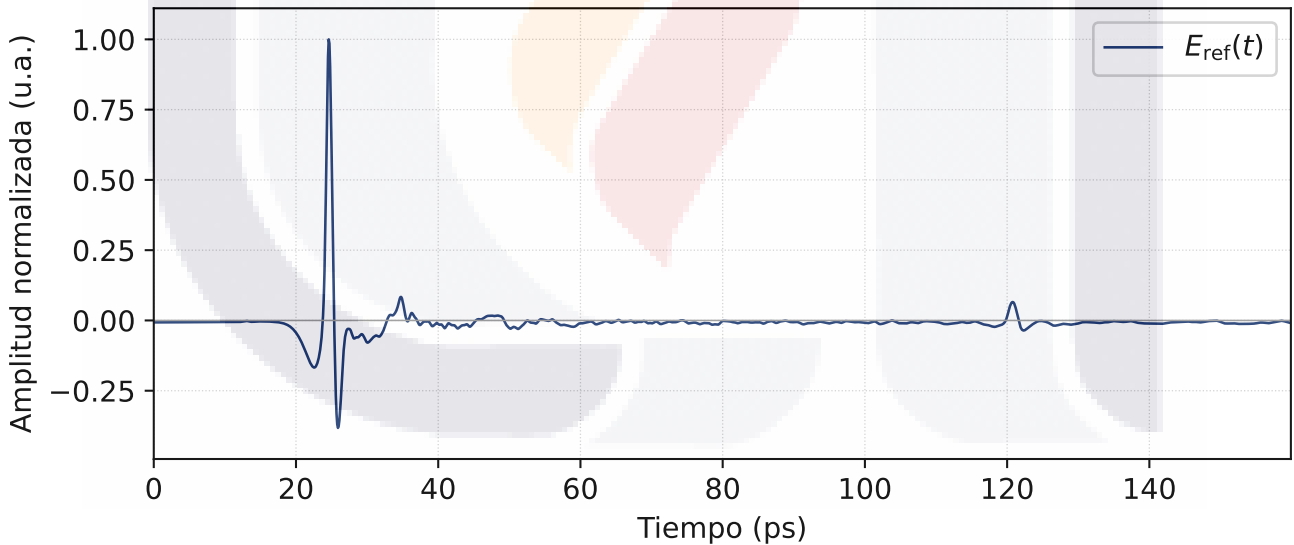


Figura 3.2. Pulso de referencia $E_{\text{ref}}(t)$ del sistema THz-TDS, construido como el promedio de las 7,500 formas de onda de tres barridos en aire ($3 \times 50 \times 50$ píxeles) con inversión global de signo (archivos `referencia_aire_{1,2,3}.picoimage`). El pulso mostrado es la referencia incidente $E_{\text{ref}} = -\bar{E}_{\text{aire}}$, con pico dominante positivo; se obtiene invirtiendo el signo del promedio crudo del aire, cuyo pico es negativo antes de la corrección de polaridad. La forma de onda se adquiere con $T = 1,600$ muestras y paso temporal $\Delta t = 0.1$ ps, lo que define una ventana de 159.9 ps; la amplitud se muestra normalizada al intervalo $[-1, 1]$ según la Ec. (3.1). En la ventana de adquisición completa, la referencia está dominada por un único pulso centrado en ≈ 24.6 ps.

La normalización del pulso al intervalo $[-1, 1]$ se realiza dividiendo cada muestra por la amplitud máxima absoluta de la señal,

$$\hat{e}_{\text{ref}}[t] = \frac{e_{\text{ref}}[t]}{\max_t |e_{\text{ref}}[t]|}, \quad (3.1)$$

operación que preserva la forma de onda y la fase relativa, y solo reescala la amplitud. Esta normalización se justifica porque la amplitud absoluta que entrega el instrumento está en unidades de cuantización arbitrarias (enteros del conversor analógico-digital), sin calibración a campo eléctrico. Lo que importa físicamente no es esa escala absoluta, sino la forma del pulso y el tamaño relativo de sus ecos, que vienen dados por los coeficientes de reflexión. Por eso normalizar elimina solo una escala arbitraria y conserva toda la información útil. Los pulsos simulados se expresan luego como fracción de esa referencia, de modo que comparten una escala común sin perder su amplitud relativa; respecto de ella se define la desviación del ruido σ .

El pulso se transforma al dominio de la frecuencia mediante la transformada rápida de Fourier para señales reales (`rfft`), que produce $\lfloor T/2 \rfloor + 1$ *bins* de frecuencia. Como la referencia y el pulso simulado comparten esos mismos *bins*, sus espectros pueden multiplicarse para sintetizar el pulso reflejado y compararse muestra a muestra. Usar un pulso de referencia y trabajar en el dominio de la frecuencia es el procedimiento estándar de extracción de parámetros en espectroscopía de THz-TDS (Duvillaret et al., 1996):

$$E_{\text{ref}}(f) = \mathcal{F}\{\hat{e}_{\text{ref}}(t)\}, \quad \omega = 2\pi f, \quad (3.2)$$

donde la transformada devuelve el espectro en función de la frecuencia natural f (en THz, por expresar el tiempo en picosegundos), y la frecuencia angular $\omega = 2\pi f$ (en rad ps^{-1}) es la que emplea la fase de propagación dentro del simulador multicapa. El tiempo se expresa en picosegundos y las distancias en micrómetros, un sistema de unidades consistente que se mantiene en todo el simulador.

3.1.2. Recursión de Fresnel-Rouard con propagación de fase

El simulador modela la respuesta de un sistema multicapa bajo incidencia normal. Esta hipótesis de incidencia se corresponde con la geometría de adquisición del equipo, que mide en configuración de reflexión con el haz perpendicular a la muestra. Con incidencia normal, además, la reflexión es la misma para cualquier polarización de la luz, así que el modelo puede ignorar esa variable (Born & Wolf, 1999). La meta es saber cuánto refleja la estructura completa, y se llega en dos pasos. Primero se calcula cuánto refleja cada interfaz por separado y después cómo se combinan esas reflexiones a lo largo de todas las capas. La estructura se ordena de la cara de entrada hacia el fondo,

$$[\text{Aire}, \text{Capa}_1, \dots, \text{Capa}_L, \text{Sustrato}],$$

descrita por el vector de índices $\{n_0, n_1, \dots, n_L, n_{\text{sus}}\}$ con $n_0 = 1$ (aire) y por los espesores $\{d_1, \dots, d_L\}$ de las L capas internas. Con $M = L + 1$ interfaces en total, cada interfaz i separa dos medios de índices n_i y n_{i+1} y refleja según el coeficiente de Fresnel (Born & Wolf, 1999):

$$r_i(\omega) = \frac{n_i - n_{i+1}}{n_i + n_{i+1}}. \quad (3.3)$$

La Figura 3.3 resume esta geometría, con la estructura multicapa, sus $M = L + 1$ interfaces, los índices n_j , los espesores d_j y el eco parcial que cada interfaz aporta al pulso reflejado.

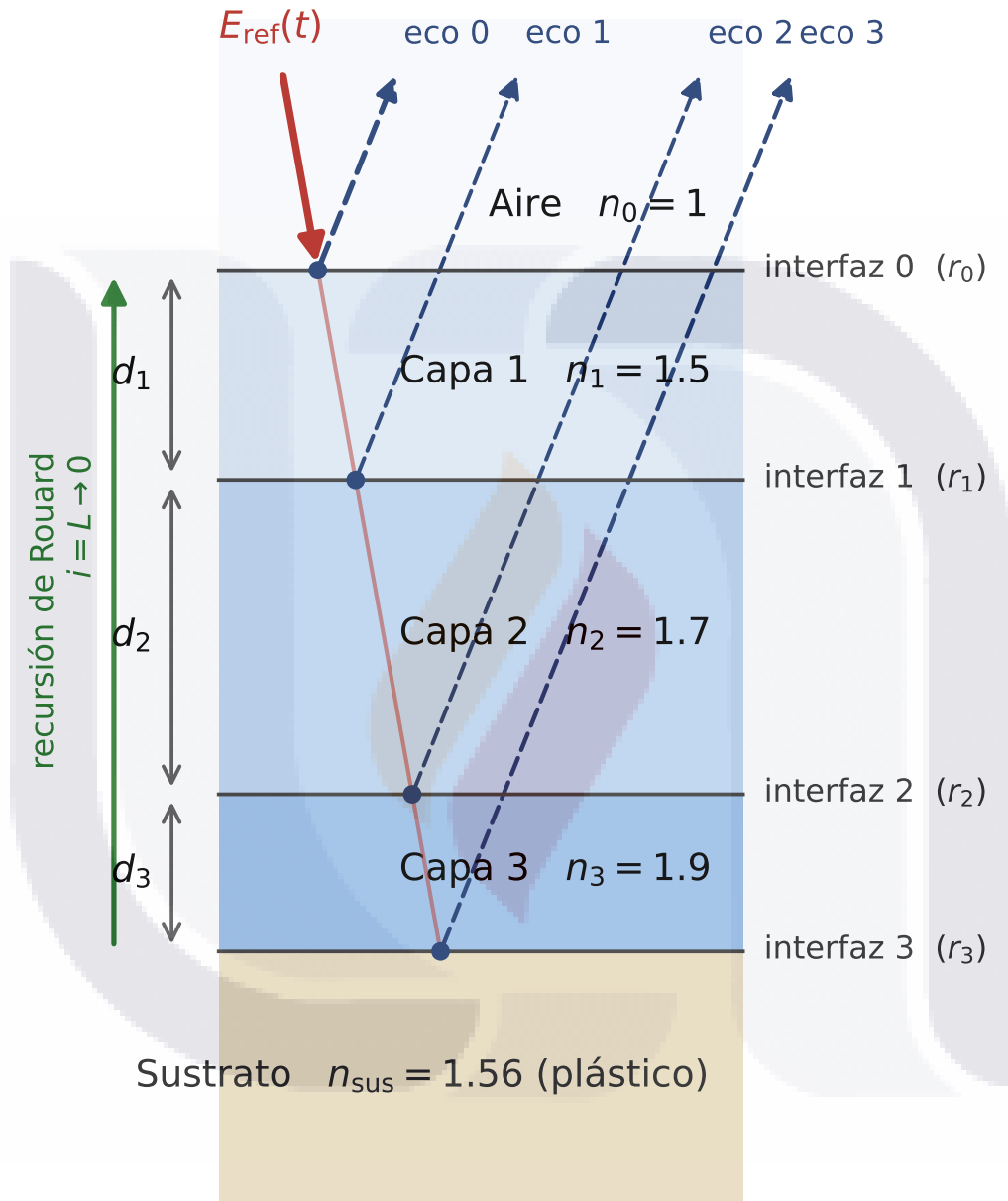


Figura 3.3. Esquema del modelo físico directo de reflexión multicapa, para un ejemplo de $L = 3$ capas sobre sustrato plástico. La estructura se ordena de la cara de entrada (aire, $n_0 = 1$) hacia el fondo (sustrato, $n_{\text{sus}} = 1.56$), con $M = L + 1$ interfaces numeradas $i = 0, \dots, L$. Cada interfaz i separa los medios de índices n_i y n_{i+1} y tiene coeficiente de Fresnel r_i (Ec. (3.3)). La altura de cada capa es proporcional a su espesor d_j . El pulso incidente $E_{\text{ref}}(t)$ genera en cada interfaz un eco parcial (líneas azules) de grosor proporcional a $|r_i|$, calculado con los índices del ejemplo. Por eso el eco frontal (aire–capa 1, $r_0 \approx -0.20$) es el de mayor amplitud, mientras que los ecos internos y el del sustrato plástico ($r_3 \approx +0.10$) son débiles. La flecha verde marca el sentido de la recursión de Rouard, que compone los coeficientes desde la interfaz del fondo ($i = L$) hacia la superior ($i = 0$). La incidencia es normal, y los rayos se dibujan inclinados solo para distinguirlos.

La reflectividad total del sistema, $r_{\text{tot}}(\omega)$, se construye con la recursión de Rouard, un método recursivo de composición de coeficientes de Fresnel para medios multicapa (Krimi et al., 2016), equivalente al método de matriz de transferencia empleado habitualmente para medios estratificados (Born & Wolf, 1999). La recursión itera desde la última interfaz (capa L -sustrato) hacia la primera (aire-capas 1). La recursión se inicializa en la interfaz del fondo, donde la reflectividad total coincide con el coeficiente de Fresnel puro de esa interfaz, sin término de propagación adicional:

$$r_{\text{tot}}^{(L)}(\omega) = r_L(\omega). \quad (3.4)$$

Para cada interfaz superior $i = L - 1, L - 2, \dots, 0$ la reflectividad total se actualiza mediante

$$r_{\text{tot}}^{(i)}(\omega) = \frac{r_i(\omega) + r_{\text{tot}}^{(i+1)}(\omega) e^{-j\varphi_{i+1}}}{1 + r_i(\omega) r_{\text{tot}}^{(i+1)}(\omega) e^{-j\varphi_{i+1}}}, \quad \varphi_{i+1}(\omega) = \frac{2\omega n_{i+1} d_{i+1}}{c}, \quad (3.5)$$

donde $\varphi_{i+1}(\omega)$ es el desfase de ida y vuelta que el campo acumula al atravesar la capa de índice n_{i+1} y espesor d_{i+1} , esto es, la capa material interpuesta entre las interfaces i e $i + 1$. El factor 2 recoge el doble trayecto (ida y vuelta) dentro de la capa. La velocidad de la luz se emplea en la forma $c = 299.792458 \mu\text{m ps}^{-1}$, que es numéricamente idéntica al valor convencional $299,792,458 \text{ m s}^{-1}$ pero expresada en las unidades del modelo; esta elección es la que vuelve dimensionalmente consistente la fase φ_{i+1} , con ω en rad ps^{-1} y d_{i+1} en μm . La reflectividad total del sistema es $r_{\text{tot}}(\omega) \equiv r_{\text{tot}}^{(0)}(\omega)$.

En la implementación, los espesores de las L capas materiales se almacenan en un arreglo que empieza a numerar desde cero, mientras que en este capítulo las capas se numeran desde uno; por eso el espesor d_{i+1} de la Ec. (3.5) corresponde al elemento i de ese arreglo. Como el término de fase solo se evalúa para las interfaces $i = 0, \dots, L - 1$ (la interfaz del fondo se inicializa sin propagación, Ec. (3.4)), la capa asociada es siempre una capa material: nunca el aire ni el sustrato, que no tienen espesor asignado en el modelo.

3.1.3. Descripción de la capa: índice real y espesor

Cada capa interna k ($k = 1, \dots, L$) queda descrita por su índice de refracción n_k y su espesor d_k . Ese par basta para fijar la señal, porque de él dependen las dos cantidades que hacen distinguible cada capa en el pulso. Los retardos de los ecos de interfaz vienen dados por el camino óptico de ida y vuelta de cada capa,

$$\tau_k = \frac{2 n_k d_k}{c}, \quad (3.6)$$

mientras que sus amplitudes relativas dependen de los contrastes de índice en las interfaces (Ec. (3.3)).

El índice de refracción se toma como real ($\kappa = 0$), de modo que el modelo no representa la absorción del material. El generador sortea el índice y el espesor de cada capa de forma independiente y uniforme, dentro de rangos físicamente plausibles.

3.1.4. Generación del pulso reflejado simulado

Con la reflectividad total $r_{\text{tot}}(\omega)$ ya disponible, el pulso reflejado se sintetiza modulando con ella el espectro de referencia y volviendo al dominio temporal con la transformada inversa de Fourier para señales reales (`irfft`), con el número de muestras fijado en T :

$$y_{\text{sim}}(t) = \mathcal{F}^{-1}\{E_{\text{ref}}(\omega) \cdot r_{\text{tot}}(\omega)\}, \quad \text{con } T \text{ muestras.} \quad (3.7)$$

Fijar T es imprescindible. Sin ese dato, `irfft` deduce la longitud de salida del número de *bins* de frecuencia y devuelve un pulso de longitud distinta de la referencia, con un desfase que pasaría inadvertido. Por eso el simulador exige T y aborta si no se le da, lo que mantiene el pulso alineado con la referencia.

Lo que se almacena como muestra es la salida de la Ec. (3.7) con su amplitud cruda, sin normalizar por pulso. Como la referencia ya está normalizada a pico unitario (Ec. (3.1)), la amplitud del pulso simulado queda expresada como fracción del pulso incidente y refleja directamente la intensidad de reflexión de cada eco, dada por los coeficientes de Fresnel. Conservar esa amplitud es esencial, porque normalizar cada pulso por separado borraría la información de reflexión que distingue las estructuras de capas y, en particular, la reflexión de fondo débil del soporte plástico.

La Figura 3.4 muestra dos pulsos así sintetizados sobre el pulso de referencia real y un sustrato plástico ($n_{\text{sus}} = 1.56$), y deja ver cómo la disposición de las capas fija la estructura temporal de ecos. Cada interfaz produce un eco, retardado respecto del frontal por la suma de los caminos ópticos τ_k (Ec. (3.6)) de las capas que el pulso ya cruzó.

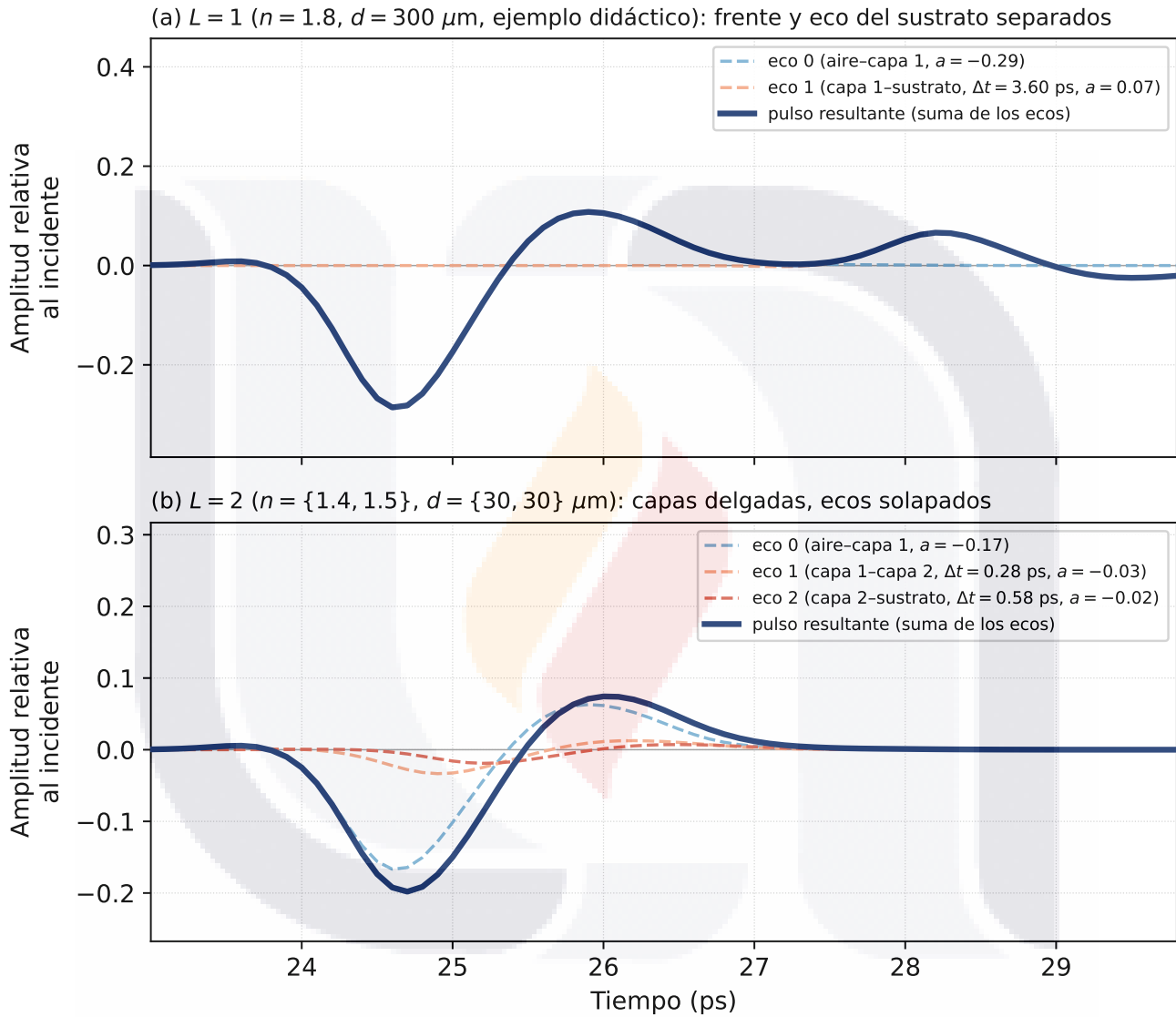


Figura 3.4. El pulso reflejado como superposición de ecos, en amplitud relativa al pulso incidente. Cada curva a trazos es un eco por interfaz, escalado por su amplitud de reflexión a_i (indicada en la leyenda junto con su retardo); la curva azul sólida es su suma, el pulso reflejado resultante. (a) Con una sola capa gruesa ($n = 1.8$, $d = 300 \mu\text{m}$), tomada como ejemplo didáctico, el eco frontal y el del sustrato ($\Delta t = 3.60 \text{ ps}$) se separan con claridad. (b) Con dos capas delgadas ($n = \{1.4, 1.5\}$, $d = \{30, 30\} \mu\text{m}$), los retardos ($\sim 0.3\text{--}0.6 \text{ ps}$) son muy inferiores al ancho de una reflexión ($\sim 0.9 \text{ ps}$ a media altura), de modo que los ecos se solapan y el pulso queda dominado por el frente. Solo se dibujan los ecos de primer orden; el modelo completo (Ec. (3.7)) añade reflexiones internas más débiles. El ejemplo (a) está por encima del rango de muestreo, solo para separar los ecos; el (b) cae dentro ($n \in [1.4, 2.0]$, $d \in [20, 100] \mu\text{m}$).

La señal $y_{\text{sim}}(t)$ así obtenida, tras el preprocesamiento, constituye la entrada del modelo inverso.

3.1.5. Verificación numérica del generador

Antes de generar los datos, se comprobó que el simulador reproduce la física que debe. Cada comprobación corre su propio código sobre un caso simple y compara la salida con la fórmula cerrada que le corresponde. La Tabla 3.1 resume las cuatro.

Las dos primeras se ven en la Figura 3.5, que muestra dos simulaciones sobre un pulso incidente. En cada una, el pulso reflejado tiene dos ecos: el del frente, cuya amplitud es el coeficiente de Fresnel $(1 - n)/(1 + n)$ de la interfaz aire-capas, y el del fondo, que llega retrasado el tiempo de ida y vuelta $\tau = 2nd/c$ de la capa (Ec. (3.6)). Los dos valores medidos coinciden con las fórmulas: la amplitud con un error del orden de 10^{-14} (la precisión de la máquina) y el retardo dentro del paso de muestreo.

La tercera es la suma de los rebotes internos. Dentro de una capa el pulso rebota muchas veces entre sus dos caras, y la recursión de Fresnel-Rouard (Ec. (3.5)) reúne todos esos rebotes en una sola expresión. Su salida coincide con la de un cálculo independiente que suma los ecos uno a uno.

La cuarta es la conservación de energía. Un sistema sin pérdidas no devuelve más señal de la que recibe, así que su reflectividad cumple $|r_{\text{tot}}(\omega)| \leq 1$. La comprobación da como máximo 0.9998 para un reflector metálico, 0.22 para el plástico y 0.38 para el aire, siempre por debajo de uno.

Tabla 3.1. Verificación numérica del generador de pulsos. Cada comprobación contrasta la salida del simulador, obtenida ejecutando su propio código, con la expresión cerrada que debe reproducir.

Comprobación	Referencia teórica	Resultado
Amplitud del frente	$ (1 - n)/(1 + n) $	error $\sim 10^{-14}$
Retardo del eco	$\tau = 2nd/c$	error 0.07 %
Recursión vs suma de ecos	suma de rebotes internos	RMS 0 % ($K = 30$)
Pasividad	$ r_{\text{tot}}(\omega) \leq 1$	máx. 0.9998 (metal)

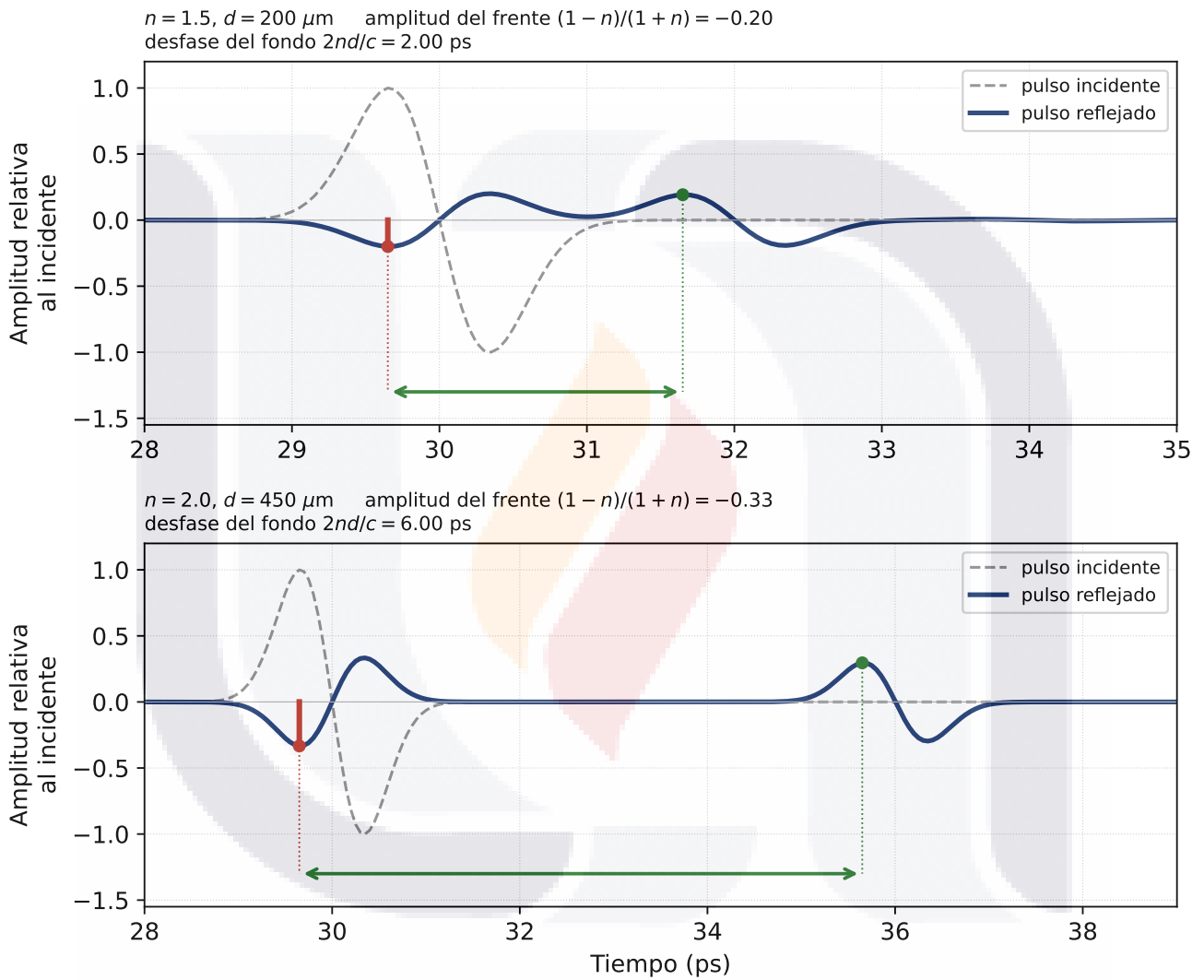


Figura 3.5. Verificación del generador sobre dos simulaciones. En cada panel se superponen el pulso incidente (gris a trazos) y el pulso reflejado que devuelve el simulador (azul) para una capa distinta. La amplitud del eco frontal es el coeficiente de Fresnel $(1 - n)/(1 + n)$ (marca roja) y el eco del fondo llega retrasado el tiempo de ida y vuelta $2nd/c$ (flecha verde); los valores, anotados sobre cada panel, coinciden con las fórmulas. Los dos casos ($n = 1.5, d = 200 \mu\text{m}$ y $n = 2.0, d = 450 \mu\text{m}$) difieren tanto en amplitud como en retardo. El pulso incidente es un monociclo sintético ($\text{FWHM} \approx 0.8 \text{ ps}$).

3.2. Generación del conjunto de datos sintético

El mismo simulador físico genera el conjunto de datos sintético, y se construye con dos generadores. El generador base (**GeneradorBase**) se encarga de sortear los parámetros de cada muestra y de balancear las clases por número de capas, mientras que el generador de producción (**GeneradorRealista**) le añade el muestreo continuo de los parámetros y los elementos que dan realismo físico a la señal.

3.2.1. Parámetros físicos y rangos de muestreo

La Tabla 3.2 resume los rangos admisibles del número de capas L , el espesor por capa d_k y el índice real por capa n'_k , los tres parámetros que describen cada estructura simulada.

Tabla 3.2. Rangos de muestreo de los parámetros del simulador en la configuración realista.

Parámetro	Símbolo	Rango
Espesor por capa	d_k	$[20, 100] \mu\text{m}$
Índice real por capa	n'_k	$[1.4, 2.0]$
Número de capas	L	$\{0, 1, 2, 3, 4\}$

Cada parámetro se sortea dentro de su rango, con los límites que fija la Tabla 3.2, y el número de capas llega como máximo a $L_{\text{máx}} = 4$. Este tope no es arbitrario, ya que los sistemas de pintura automotriz e industrial se estructuran típicamente en hasta cuatro capas funcionales (Krimi et al., 2016; van Mechelen et al., 2014), de modo que admitir $L \in \{0, \dots, 4\}$ cubre desde el soporte desnudo hasta el sistema completo.

El rango de índice $[1.4, 2.0]$ no proviene de medir la pintura real, sino de valores típicos de resinas poliméricas pigmentadas en la banda THz, cercanos a 1.5 y hasta 2.0 con pigmento. Estos se toman de composites similares (Cai et al., 2017), de modo que el intervalo es físicamente plausible pero no medido.

3.2.2. Muestreo balanceado por número de capas

El conjunto se construye con la misma cantidad de muestras para cada número de capas, de modo que el modelo no se incline hacia los valores de L más frecuentes. Esa cantidad es $\lfloor N/(L_{\text{máx}} + 1) \rfloor$ muestras sobre un total de N , y las pocas que sobran cuando la división no es exacta se reparten de a una entre las primeras clases, así que ninguna clase termina con más de una muestra de diferencia respecto de las demás.

Para alcanzarlo, el generador sortea un valor de L y acumula muestras hasta completar lo asignado a cada clase. El recuento se hace sobre el número de capas final de cada muestra, el que queda tras fundir las capas de igual índice, no sobre el que se sorteó. Por eso una fusión solo obliga a sortear otra muestra, sin romper el balance del conjunto. La Figura 3.6 lo confirma sobre el conjunto generado.

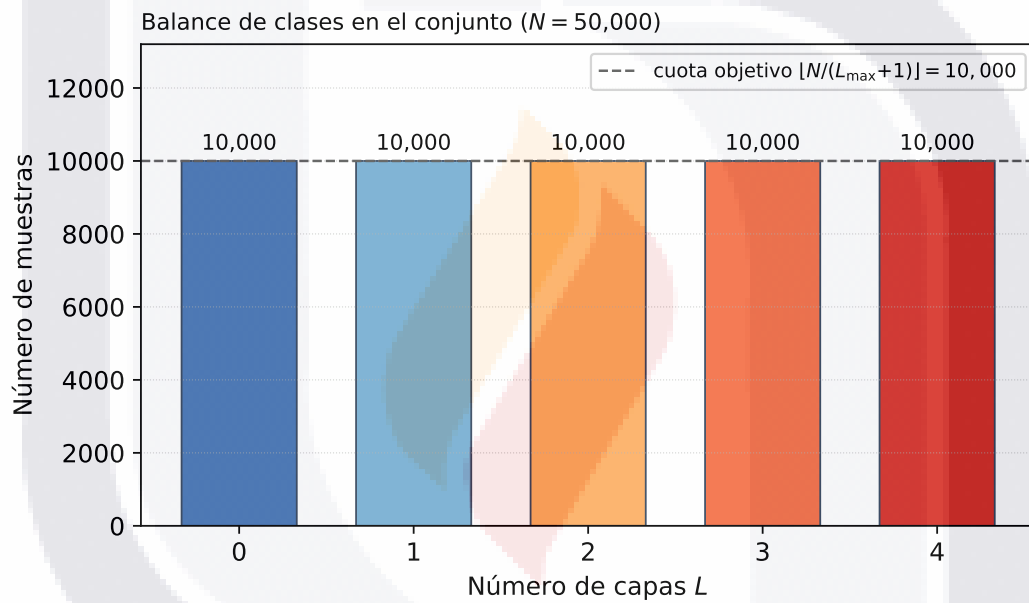


Figura 3.6. Distribución del número de capas L en el conjunto sintético generado ($N = 50,000$, semilla 42, configuración realista). El conteo se realiza sobre el L efectivo (posterior a la fusión) de cada muestra almacenada. Las cinco clases contienen exactamente 10,000 muestras, coincidiendo con la cantidad objetivo $\lfloor N/(L_{\text{máx}} + 1) \rfloor = 10,000$ (línea discontinua): el muestreo balanceado no deja ninguna clase sobre- ni sub-representada.

3.2.3. Capas degeneradas y casi degeneradas

Dos capas consecutivas con el mismo índice de refracción no forman una interfaz que refleje y son ópticamente una sola capa. Para no pedirle al modelo que separe capas físicamente indistinguibles, el generador incluye una regla que fusiona todo par consecutivo de índice idéntico en una capa equivalente, de espesor la suma de ambos, y reetiqueta la muestra con su número de capas efectivo.

Con el muestreo continuo el índice se sortea como número real, así que dos capas nunca reciben un valor exactamente igual y esa fusión no llega a actuar. En su lugar aparece la casi-degeneración: pares de capas con un contraste de índice tan pequeño que el eco de su interfaz no se separa del ruido. Sobre el conjunto generado (60,000 pares de capas adyacentes), el 11.10% tiene un coeficiente de reflexión $|r| < 2\sigma$ (esto es, $|r| < 0.010$, el doble de la desviación del ruido $\sigma = 0.005$), de modo que su interfaz no alcanza el nivel de una detección fiable aunque la etiqueta las cuente como capas separadas. Es un límite de información, no del modelo: ninguna técnica podría contar con fiabilidad una capa cuyo eco no se separa del ruido.

Esta fusión es propia del muestreo discreto del índice. Con el muestreo continuo de la clase base, donde dos índices nunca coinciden exactamente, no ocurre.

3.2.4. Modelado de la clase sin pintura ($L = 0$)

La clase $L = 0$ corresponde a la zona sin pintura, donde el soporte plástico queda expuesto. Aun sin capas, ese soporte refleja el pulso THz, de modo que una medición real de $L = 0$ no es una señal nula sino un eco débil característico (la interfaz aire-plástico, con $r \approx -0.22$). Por eso el generador no la representa con ruido ni con ceros, sino con la reflexión del soporte plástico desnudo, el pulso del sistema [Aire, Sustrato] sin capas intermedias, que se obtiene simulando con la lista de capas vacía. Modelarla así reproduce lo que de verdad se mediría sobre el soporte expuesto y evita que el clasificador aprenda un atajo falso, el de asociar «ausencia de señal» con $L = 0$.

Como esa reflexión es determinística, todas las muestras de $L = 0$ comparten el mismo pulso limpio, y su única variabilidad viene del ruido que se les inyecta después como aumento de datos.

3.2.5. Soporte plástico y ruido de medición

El sustrato se modela como un plástico semi-infinito de índice $n_{\text{sus}} = 1.56$, el del soporte real sobre el que están depositadas las pinturas de validación. El metal interviene en la medición únicamente como espejo de referencia, para recuperar el pulso incidente (Sección 3.1.1), no como fondo de las muestras. Con $n_{\text{sus}} = 1.56$ el soporte refleja débilmente: la interfaz aire-plástico del sustrato desnudo tiene $r \approx -0.22$ por la Ec. (3.3), y las interfaces entre las capas y el sustrato son todavía más débiles ($|r| \lesssim 0.1$), lejos del $r \approx -1$ de un metal; el modelo reproduce así la reflexión de fondo débil que devuelven las pinturas reales. El plástico físico es finito, pero su eco posterior cae muy lejos de la región temporal de las capas y no interviene en el conteo ni en la estimación del espesor, por lo que se representa como semi-infinito.

El ruido de medición es gaussiano y aditivo, con desviación $\sigma = 0.005$ relativa a la amplitud del pulso incidente. Este valor es medido por el propio instrumento, a partir del residuo entre los tres barridos en aire, de 0.0037 a 0.0057, y de la ventana sin señal de la medición de metal desnudo, de 0.0012 a 0.0038, siempre relativos a la amplitud máxima. El $\sigma = 0.005$ adoptado es la cota superior de ese residuo. Un valor anterior de 0.011, sin respaldo experimental, sobreestimaba el ruido por un factor de 2 a 3.

Este ruido no se guarda en el conjunto de datos, que contiene solo los pulsos limpios, sino que se agrega durante el entrenamiento y se vuelve a generar en cada época.

La Figura 3.7 muestra el efecto de este ruido sobre un pulso simulado.

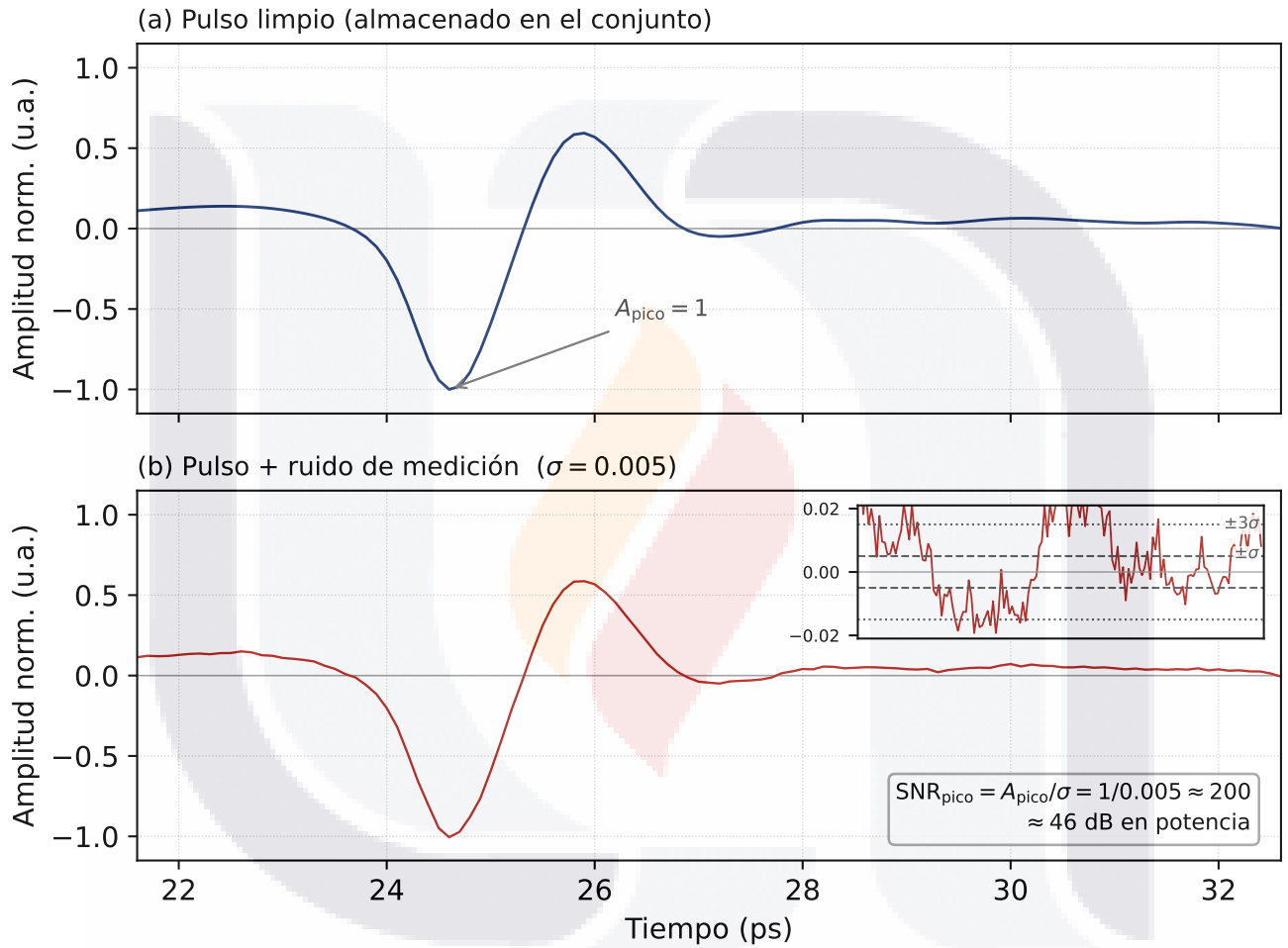


Figura 3.7. Efecto del ruido de medición sobre un pulso simulado ($L = 1$; $n_1 = 1.8$, $d_1 = 100 \mu\text{m}$, sustrato plástico). (a) Pulso limpio, tal como se almacena en el conjunto, normalizado de modo que su amplitud de pico es $A_{\text{pico}} = 1$. (b) Una realización del mismo pulso con ruido gaussiano aditivo de desviación $\sigma = 0.005$, el mismo aumento de datos que se aplica en entrenamiento e inferencia (Ec. (3.15)). A la escala del pico, el ruido es casi imperceptible; el recuadro amplía una ventana de la línea de base (sin señal) para revelar su magnitud real frente a las bandas $\pm\sigma$ y $\pm 3\sigma$ (relación señal-ruido de pico $1/\sigma = 200$; véase la Sección 3.4.2). El ruido mostrado corresponde a una realización con semilla fija para reproducibilidad de la figura; el código de producción consume el generador aleatorio global.

3.2.6. Volumen, partición por clases y solapamiento entre conjuntos

El conjunto tiene $N = 50,000$ muestras. Ese número se fija en cada corrida y queda registrado en su archivo de resultados, así que siempre se sabe cuántas muestras se usaron.

La partición se hace en dos etapas, ambas estratificadas, que conservan la proporción de cada clase de L . Se aplica una sola vez y sobre los índices del conjunto, no sobre los pulsos, lo que fija un conjunto de prueba estable y reproducible. Las fracciones son las habituales en aprendizaje supervisado, y estos son los tamaños absolutos que producen:

- Prueba: el 20 %, 10,000 muestras.
- Validación: el 15 % del resto, 6,000 muestras (1,200 por clase), las justas para estabilizar la pérdida de validación.
- Entrenamiento: las 34,000 restantes, donde se ajustan los ~ 1.2 millones de parámetros del modelo.

Una sola semilla (por defecto, 42) controla la generación de las muestras y las dos etapas de la partición, así que todo el conjunto es reproducible. La Figura 3.8 muestra que la proporción de cada clase se conserva en los tres subconjuntos.

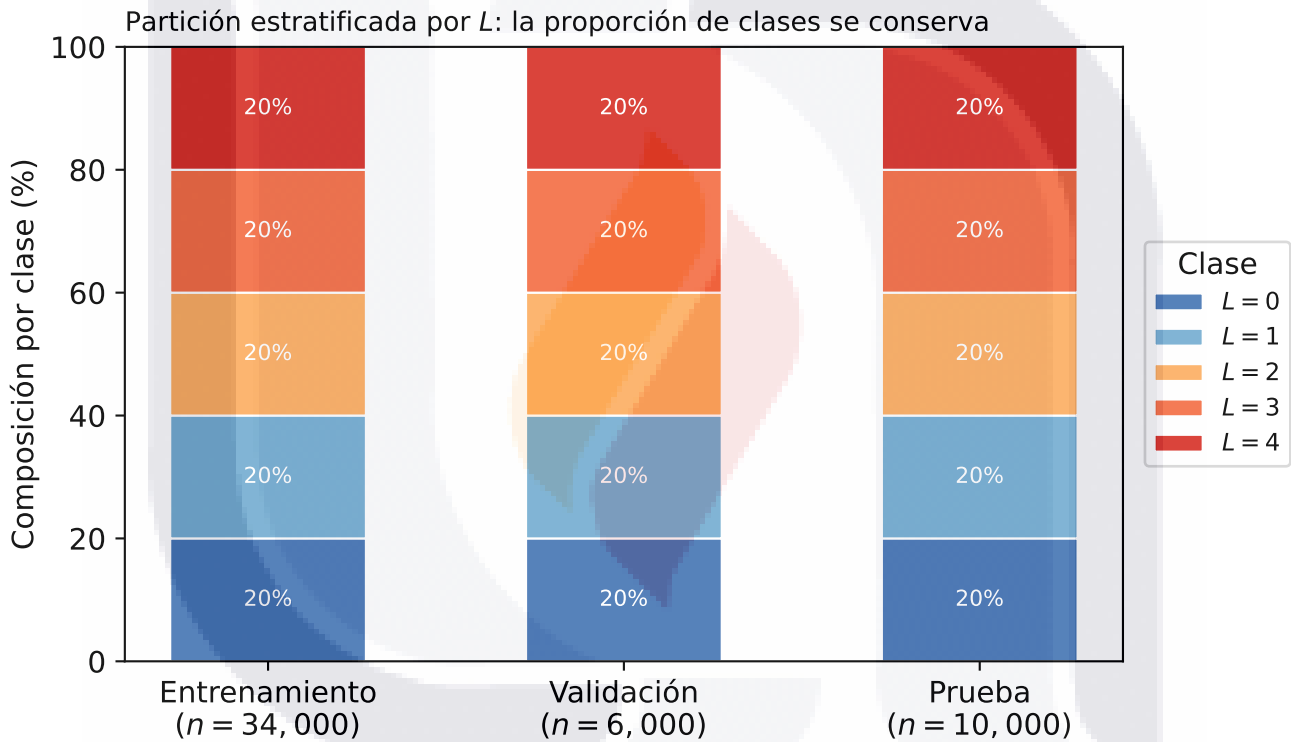


Figura 3.8. Composición por número de capas L de los tres subconjuntos. Cada barra es un subconjunto normalizado al 100 %, y la altura de cada segmento es la fracción de muestras de cada clase L . La proporción del 20 % por clase se conserva en entrenamiento ($n = 34,000$), validación ($n = 6,000$) y prueba ($n = 10,000$).

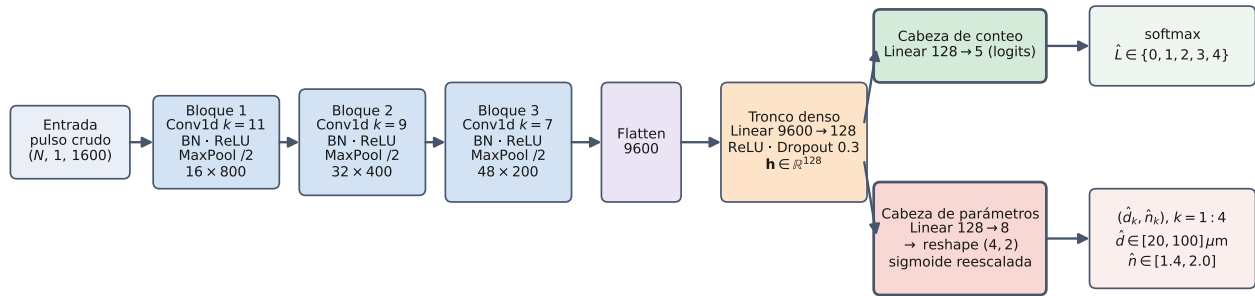

Aprendizaje multitarea: el tronco comparte $\mathbf{h}$ entre ambas cabezas

Figura 3.9. Arquitectura de la CNN multitarea. El pulso crudo de entrada recorre una rama convolucional de tres bloques que extrae características y acorta la longitud temporal. Tras el aplanamiento, un tronco denso genera la representación compartida $\mathbf{h}$, de la que parten las dos cabezas de salida, una para el conteo de capas y otra para la regresión de sus parámetros. Las dimensiones de cada etapa figuran en el diagrama y en la Tabla 3.3.

El solapamiento entre conjuntos (*data leakage*) puede darse de dos maneras. La del aumento de datos queda descartada por construcción, ya que el ruido se añade después de la partición, así que ninguna versión ruidosa de una muestra cae en dos subconjuntos. La otra viene del muestreo discreto. Como cada parámetro toma un número finito de valores (Sección 3.2.1), dos muestras distintas pueden compartir la misma configuración física $(L, \{n'_k\}, \{d_k\})$ y repartirse entre entrenamiento y prueba, algo que partir por índices no evita. El sistema principal elimina esta fuente al muestrear los parámetros de forma continua (Sección 3.2.1), con lo que la probabilidad de configuraciones repetidas es nula.

3.3. Modelo: CNN multitarea

El modelo propuesto es una CNN unidimensional multitarea que resuelve, en un solo modelo, el conteo de capas y la regresión de sus parámetros.

El pulso crudo entra por una rama convolucional, cuya salida pasa a un tronco denso compartido que se divide en dos cabezas de salida, una por tarea. Compartir el tronco entre las dos tareas, en lugar de entrenar modelos separados por número de capas, es lo propio del aprendizaje multitarea (Caruana, 1997), y se justifica por dos razones. Por un lado, ambas tareas dependen de la misma información, la estructura de ecos del pulso, ya que contar capas es ver cuántos ecos hay y estimar el espesor y el índice es medir sus retardos y amplitudes; así, un tronco compartido extrae esa estructura una sola vez y cada tarea regulariza a la otra. Por otro, evita mantener y sincronizar tantos modelos como clases de capas habría que cubrir.

3.3.1. Representación de entrada

La entrada es el pulso temporal, en un solo canal, que se presenta a la red como un tensor de forma $(N, 1, T)$, con T la longitud temporal del pulso, y conserva la amplitud cruda con que sale del simulador.

La longitud temporal T , el número de muestras de cada pulso, no la elige el diseño de la red sino el instrumento de medición. Por eso cualquier valor de longitud que aparezca en esta sección corresponde al equipo empleado.

3.3.2. Arquitectura de la red

La Figura 3.9 ofrece una vista completa del modelo, desde el pulso de entrada hasta las dos cabezas de salida.

La rama convolucional recorre el pulso con filtros cortos que aprenden a detectar rasgos locales de la señal, como los ecos y sus flancos, y capa tras capa los resume en tramos temporales cada vez más largos. Se implementa como un extractor de características formado por tres bloques sucesivos $\text{Conv1D} \rightarrow \text{BatchNorm} \rightarrow \text{ReLU} \rightarrow \text{MaxPool1D}$ (la normalización por lotes (Ioffe & Szegedy, 2015) estabiliza y acelera el entrenamiento) que reciben la señal en un solo canal, con la configuración de la Tabla 3.3. Esos valores siguen el patrón convencional de las CNN, con canales que crecen y *kernels* que se reducen con la profundidad, de modo que los bloques tempranos detectan motivos amplios y los tardíos los componen (Goodfellow et al., 2016; LeCun et al., 2015). También encajan con la escala temporal del problema, ya que el *kernel* inicial de 11 muestras abarca 1.1 ps, del orden de medio ciclo del pulso de referencia. No salen de una búsqueda de hiperparámetros, sino del diseño base del trabajo. El único componente que se fijó con un criterio físico explícito fue el *pooling*, sin afirmar que la arquitectura resultante sea la óptima.

Tabla 3.3. Configuración de la rama convolucional de la CNN. La columna de longitud temporal se expresa en función de T (longitud determinada por el instrumento) y, entre paréntesis, su valor para el equipo empleado, con $T = 1,600$.

Bloque	Canales salida	Kernel	Pooling	Longitud temporal
Entrada	1	—	—	T (1,600)
1	16	11	2	$T/2$ (800)
2	32	9	2	$T/4$ (400)
3	48	7	2	$T/8$ (200)

Cada Conv1D usa relleno (*padding*) de $\lfloor k/2 \rfloor$, de modo que la longitud temporal solo se acorta en las capas MaxPool1D . Para el $T = 1,600$ del instrumento, el último bloque entrega un mapa de 48 canales y 200 muestras, es decir 9,600 activaciones.

Ese *pooling* se fijó en (2, 2, 2), con factor de submuestreo total 8, por un criterio de resolución temporal. Con paso de adquisición $\Delta t = 0.1$ ps, cada punto del mapa final cubre 0.8 ps. Como los retardos entre ecos que distinguen una estructura de capas de otra valen entre ≈ 0.19 ps (una capa de $20 \mu\text{m}$ con $n = 1.4$) y ≈ 1.3 ps en los rangos del problema (Tabla 3.2), un submuestreo más agresivo fundiría esa estructura fina en una sola activación, mientras que el paso de 0.8 ps la conserva en parte. Como contrapartida, el mapa final mantiene 9,600 activaciones, con lo que el tronco denso concentra la mayor parte de los 1,246,429 parámetros entrenables del modelo en su capa lineal.

El tronco denso proyecta esas activaciones, con ReLU y *dropout* (Srivastava et al., 2014) de probabilidad $p = 0.3$, en la representación compartida:

$$\mathbf{h} = \text{Dropout}_{0.3}(\text{ReLU}(W \text{vec}(\mathbf{z}) + \mathbf{b})) \in \mathbb{R}^{128}, \quad (3.8)$$

donde $\mathbf{z}$ es el mapa de características del último bloque convolucional de la rama y $\text{vec}(\cdot)$ denota su aplanamiento.

3.3.3. Cabezas de salida: conteo y parámetros

A partir de la representación compartida $\mathbf{h}$ se desprenden dos cabezas lineales.

La cabeza de conteo produce un vector de *logits* $\mathbf{u} = W_c \mathbf{h} + \mathbf{b}_c \in \mathbb{R}^{L_{\text{máx}}+1} = \mathbb{R}^5$, interpretado como una clasificación categórica sobre $L \in \{0, 1, 2, 3, 4\}$. La cabeza entrega los *logits* sin normalizar; la *softmax* que los convierte en probabilidades no se aplica dentro del modelo, sino en la función de pérdida (*cross-entropy*). Conceptualmente, la distribución y la predicción de conteo quedan dadas por

$$p_\theta(L = \ell | \mathbf{y}) = \frac{e^{u_\ell}}{\sum_{m=0}^{L_{\text{máx}}} e^{u_m}}, \quad \hat{L} = \arg \max_{\ell} p_\theta(L = \ell | \mathbf{y}). \quad (3.9)$$

Al plantear el conteo como una sola clasificación multiclase, y no como una presencia independiente por *slot*, la salida es por construcción un único valor de L , sin combinaciones imposibles como «hay capa 1 y capa 3, pero no capa 2».

La cabeza de parámetros aplica otra proyección lineal sobre $\mathbf{h}$ y produce una salida cruda $\mathbf{r} = W_p \mathbf{h} + \mathbf{b}_p \in \mathbb{R}^{L_{\text{máx}} \times 2} = \mathbb{R}^{4 \times 2}$, dos valores (espesor e índice) por *slot*. La cabeza produce siempre $L_{\text{máx}}$ *slots*, con independencia del número verdadero de capas de la muestra; los *slots* asociados a índices mayores que el conteo de capas presente no corresponden a ninguna capa física.

Cada par crudo $(\tilde{d}_k, \tilde{n}_k)$ se acota a su rango físico mediante una sigmoide reescalada:

$$\hat{d}_k = \sigma(\tilde{d}_k) (d_{\text{máx}} - d_{\text{mín}}) + d_{\text{mín}}, \quad d_{\text{mín}} = 20, \quad d_{\text{máx}} = 100 \text{ (}\mu\text{m)}; \quad (3.10)$$

$$\hat{n}'_k = \sigma(\tilde{n}_k) (n_{\text{máx}} - n_{\text{mín}}) + n_{\text{mín}}, \quad n_{\text{mín}} = 1.4, \quad n_{\text{máx}} = 2.0, \quad (3.11)$$

con $\sigma(x) = 1/(1 + e^{-x})$. Estos rangos físicos son los mismos que usa el generador de datos (Tabla 3.2), de modo que toda predicción cae siempre dentro del rango declarado y no hace falta penalizar las violaciones en la función de pérdida. Por último, la cabeza apila ambos valores acotados en la forma $(L_{\text{máx}}, 2)$.

3.4. Entrenamiento

El entrenamiento de la CNN avanza por épocas y se detiene cuando la pérdida de validación deja de mejorar. Lo gobiernan tres piezas, la pérdida que se minimiza, el ruido que se añade a la entrada en cada paso y las reglas que ajustan los pesos y deciden cuándo parar.

3.4.1. Función de pérdida multitarea (enmascarada y normalizada)

Se optimiza una pérdida compuesta por un término de conteo y un término de regresión sobre los parámetros, balanceados por un hiperparámetro λ :

$$\mathcal{L} = \mathcal{L}_{\text{conteo}} + \lambda \mathcal{L}_{\text{params}}, \quad \lambda = 1. \quad (3.12)$$

El término de conteo es la entropía cruzada categórica (Goodfellow et al., 2016) (Ec. (2.13)) sobre las $L_{\text{máx}} + 1$ clases de capas, evaluada en las probabilidades $p_{\theta}(L | \mathbf{y}_b)$ de la cabeza de conteo (Ec. (3.9)):

$$\mathcal{L}_{\text{conteo}} = -\frac{1}{B} \sum_{b=1}^B \log p_{\theta}(L = L_b | \mathbf{y}_b), \quad (3.13)$$

donde B es el tamaño del lote y L_b el número real de capas de la muestra b .

El término de regresión es el error cuadrático medio (MSE) enmascarado y normalizado por el rango físico de cada parámetro, sobre los pares $(\hat{d}_{b,k}, \hat{n}_{b,k})$ que predice la cabeza de parámetros. La máscara de presencia $m_{b,k} \in \{0, 1\}$ vale 1 si la capa k existe en la muestra b (es decir, si $k \leq L_b$) y 0 en caso contrario; con ella y los rangos físicos

$$\Delta d = d_{\text{máx}} - d_{\text{mín}} = 100 - 20 = 80 \text{ } \mu\text{m}, \quad \Delta n = n_{\text{máx}} - n_{\text{mín}} = 2.0 - 1.4 = 0.6,$$

el término de regresión se escribe:

$$\mathcal{L}_{\text{params}} = \frac{1}{\text{máx}(\sum_{b,k} m_{b,k}, 1)} \sum_{b=1}^B \sum_{k=1}^{L_{\text{máx}}} m_{b,k} \left[\left(\frac{\hat{d}_{b,k} - d_{b,k}}{\Delta d} \right)^2 + \left(\frac{\hat{n}_{b,k} - n_{b,k}}{\Delta n} \right)^2 \right]. \quad (3.14)$$

La máscara $m_{b,k}$ se construye a partir de la presencia real (*ground truth*) de cada muestra, no de la presencia predicha por la red, de modo que los *slots* inexistentes no contribuyen al gradiente de regresión. El denominador de la normalización se acota inferiormente a 1 con el operador $\text{máx}(\cdot, 1)$ de la ecuación (3.14), para evitar la división por cero en lotes donde ninguna muestra presenta capas, es decir cuando $\sum_{b,k} m_{b,k} = 0$. La normalización por rango pone a d (escala de decenas de μm) y a n (escala ~ 1) en magnitudes comparables, evitando que el espesor domine la optimización.

Gracias a la normalización por rango, ambos términos de la pérdida quedan en escalas comparables, lo que justifica el peso neutro $\lambda = 1$. La entropía cruzada de un clasificador de 5 clases y un error cuadrático de variables normalizadas a $[0, 1]$ son ambos de orden unidad al inicio del entrenamiento, de modo que ningún término domina al otro. Normalizar por rango es justamente lo que evita compensar escalas dispares ajustando λ . De todos modos, $\lambda = 1$ es el valor por defecto del diseño y no se ajustó ni se probaron otros valores.

3.4.2. Aumento de datos con ruido

Durante el entrenamiento, la red recibe el pulso limpio y, antes de extraer las representaciones que alimentan a la CNN, se le añade una realización fresca de ruido gaussiano. En este experimento la única representación activa es el pulso crudo, por lo que el ruido se aplica directamente sobre la señal de entrada $\mathbf{y}_b$:

$$\mathbf{y}_b^{\text{aug}} = \mathbf{y}_b + \boldsymbol{\xi}_b, \quad \boldsymbol{\xi}_b \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}). \quad (3.15)$$

La desviación del ruido es $\sigma = 0.005$, la medida sobre el instrumento (Sección 3.2.5). Corresponde a una relación señal-ruido (SNR) de pico $1/\sigma = 200$ sobre el pulso normalizado, es decir el cociente entre la amplitud pico de la señal y la desviación del ruido, equivalente a ≈ 46 dB en potencia (no una SNR de potencia de 200 dB). La Figura 3.7 muestra el efecto de este ruido y su escala relativa al pico de la señal.

El ruido se vuelve a sortear en cada lote y cada época, de modo que la red nunca ve dos veces la misma realización y aprende a ser robusta frente al ruido de medición.

El conjunto de validación también se evalúa con ruido, pero de forma distinta. Mientras el de entrenamiento cambia en cada lote y época, el de validación es una única realización fija, sorteada de un generador propio con semilla fija y construida una sola vez antes de entrenar. Esto se hace por dos razones. Por un lado, evaluar la validación con ruido es coherente con el uso real, ya que la pérdida que decide la parada temprana debe medir el desempeño sobre señales ruidosas, no sobre un pulso limpio que el modelo nunca enfrentará. Por otro, fijar esa realización hace la selección de modelo justa y reproducible, porque todas las épocas se comparan contra la misma perturbación, de modo que las diferencias de pérdida reflejan al modelo y no la suerte del sorteo.

3.4.3. Optimización, scheduler y parada temprana

Los hiperparámetros de optimización se resumen en la Tabla 3.4. Se usa el optimizador AdamW, es decir, Adam (Kingma & Ba, 2015) con *weight decay* desacoplado (Loshchilov & Hutter, 2019). La tasa de aprendizaje inicial (10^{-3} , el valor por defecto de Adam), el *weight decay* (10^{-5}) y el tamaño de lote (128) son valores convencionales para AdamW en redes de este tamaño, heredados del diseño base sin búsqueda específica, y se mantienen fijos durante todo el entrenamiento. La tasa de aprendizaje η se modula con un *scheduler* de recocido por coseno (*cosine annealing*) (Loshchilov & Hutter, 2017), que la reduce desde su valor inicial hasta $\eta_{\min} = 0$ a lo largo del entrenamiento:

$$\eta_e = \frac{\eta}{2} \left(1 + \cos \frac{\pi e}{E} \right), \quad e = 0, 1, \dots, E, \quad (3.16)$$

con E el número máximo de épocas. El horizonte del recocido se fija en $T_{\max} = E$, es decir el número de épocas efectivas del experimento ($E = 100$), y la actualización se realiza una vez por época, no por lote, de modo que al cabo de E épocas la tasa alcanza exactamente 0. La Figura 3.10 representa esta evolución.

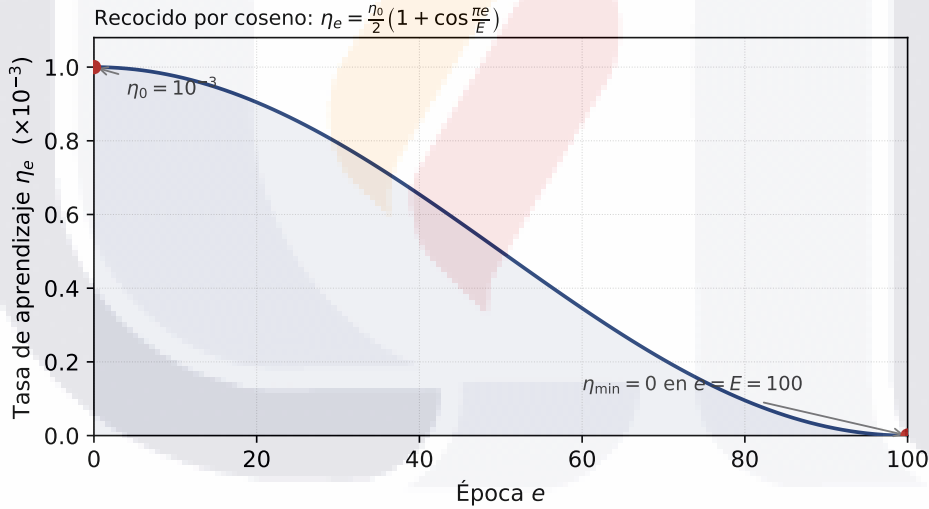


Figura 3.10. Evolución de la tasa de aprendizaje η_e bajo el *scheduler* de recocido por coseno (Ec. (3.16)) con $\eta_0 = 10^{-3}$, $\eta_{\min} = 0$ y $E = 100$ épocas. La tasa parte de su valor máximo, decae de forma suave (con derivada nula en los extremos) y alcanza exactamente 0 en la última época. La actualización ocurre una vez por época.

Tras cada época se evalúa la pérdida de validación, y si no mejora en más de $\text{mín } \Delta = 10^{-4}$ durante 8 épocas consecutivas, el entrenamiento se detiene (*early stopping*). La parada temprana cumple aquí su doble función habitual. Por un lado regulariza, al detener el ajuste antes de que el modelo memorice particularidades del entrenamiento que no generalizan (Goodfellow et al., 2016); por otro, evita gastar cómputo en épocas que ya no mejoran. Sus dos umbrales son convencionales. El $\text{mín } \Delta = 10^{-4}$ define qué cuenta como mejora real frente a la fluctuación de la pérdida, y la paciencia de 8 épocas da margen para que el recocido de la tasa (Ec. (3.16)) recupere mesetas pasajeras antes de declarar el estancamiento. Ninguno se ajustó por búsqueda.

Al finalizar, se restaura el modelo de la mejor época, la de menor pérdida de validación. El estado también se guarda en disco al final de cada época, lo que permite reanudar o inspeccionar cualquiera de ellas.

Tabla 3.4. Hiperparámetros de entrenamiento de la CNN.

Hiperparámetro	Valor
Optimizador	AdamW
Tasa de aprendizaje η	10^{-3}
<i>Weight decay</i>	10^{-5}
<i>Scheduler</i>	<i>Cosine annealing</i> ($T_{\text{máx}} = E$, $\eta_{\text{mín}} = 0$)
Tamaño de lote B	128
Épocas máximas E	100
Peso de regresión λ	1
Fracción de validación	0.15
<i>Early stopping</i>	paciencia 8, $\text{mín } \Delta = 10^{-4}$
Selección de modelo	mejor época por pérdida de validación

3.5. Línea base tradicional: deconvolución de Wiener y detección de picos

La línea base tradicional reproduce el procedimiento clásico de análisis estratigráfico, la deconvolución filtrada con detección de picos descrita en el marco teórico (Sección 2.2.2, Ec. (2.8)). No requiere entrenamiento ni datos de referencia (*ground truth*), y opera únicamente sobre el pulso de muestra y el pulso de referencia de las mediciones reales de pintura, y consta de seis pasos.

El primer paso transforma al dominio de la frecuencia el pulso de muestra E_s y el de referencia E_{ref} mediante la transformada rápida de Fourier para señales reales, obteniendo $E_s(\omega)$ y $E_{\text{ref}}(\omega)$ sobre el mismo eje de frecuencias.

El segundo paso realiza el filtro $F(\omega)$ de la Ec. (2.8) como un pasabanda suave construido con dos funciones de error (filtro *doble-erf*), que atenúa de forma gradual las frecuencias fuera de la banda útil:

$$F(f) = \frac{1}{2} \left[1 + \text{erf} \left(\frac{f-f_1}{\sigma_1} \right) \right] \cdot \frac{1}{2} \left[1 + \text{erf} \left(\frac{f_2-f}{\sigma_2} \right) \right], \quad (3.17)$$

con $f_1 = 0.1$ THz, $\sigma_1 = 0.05$ THz (flanco inferior) y $f_2 = 1.5$ THz, $\sigma_2 = 0.3$ THz (flanco superior). Estos valores se fijan de antemano para acotar la banda de trabajo y no amplificar ruido de alta frecuencia.

El tercer paso es la deconvolución de Wiener, porque la inversión directa $1/E_{\text{ref}}$ de la Ec. (2.6) es inestable donde la referencia tiene poca energía. Para regularizarla se emplea un filtro de Wiener, que sustituye esa división por una forma estabilizada y la combina con el pasabanda $F(\omega)$:

$$H(\omega) = \frac{E_s(\omega) E_{\text{ref}}^*(\omega)}{|E_{\text{ref}}(\omega)|^2 + Q^2} F(\omega), \quad Q^2 = \varepsilon \max_{\omega} |E_{\text{ref}}(\omega)|^2, \quad (3.18)$$

donde E_{ref}^* es el conjugado complejo y Q^2 es el término de regularización de tipo Wiener (Chen et al., 2010; Dong et al., 2016), fijado en este trabajo como una fracción $\varepsilon = 10^{-3}$ de la energía espectral máxima de la referencia. La respuesta al impulso se recupera por transformada inversa, $h(t) = \mathcal{F}^{-1}\{H(\omega)\}$.

El cuarto paso centra la traza, ya que la deconvolución sitúa el pico de la primera interfaz en $t \approx 0$, donde queda partido por el borde del arreglo. Se aplica un desplazamiento circular de $N/2$ muestras para centrar la traza, y se trabaja con su valor absoluto $|h(t)|$.

El quinto paso detecta y filtra los picos, para lo cual sobre $|h(t)|$ se buscan los máximos locales que superan un umbral de prominencia

$$p_{\text{mín}} = \max(0.05 \max_t |h(t)|, 4\sigma_n), \quad (3.19)$$

donde σ_n estima el ruido de fondo, exigiendo además una separación mínima de 9 muestras entre picos consecutivos. A continuación se descartan los pre-ecos y los ecos de Fabry-Perot. Se conserva como primera interfaz el pico más cercano al centro y como última la interfaz de salida, el último pico cuya amplitud supera el 25 % del máximo posterior a la primera, y se elimina todo pico fuera de ese intervalo.

El sexto paso obtiene el conteo y los espesores, ya que cada par de picos consecutivos delimita una capa, y el número de picos retenidos determina el número de capas; el retardo Δt_k entre los picos k y $k+1$ se convierte en espesor físico

$$d_k = \frac{c \Delta t_k}{2 n_k}, \quad (3.20)$$

donde c es la velocidad de la luz y n_k el índice de refracción de la capa, asumido como dato de entrada (escalar común o un valor por capa). La Ec. (3.20) es la misma relación tiempo de vuelo-espesor empleada en el análisis clásico (Sección 2.2.2) (Reyes-Reyes et al., 2024).

3.6. Protocolo de evaluación

Esta sección describe cómo se construye el conjunto de prueba, bajo qué condiciones de inferencia se aplica el ruido de medición y qué métricas se computan.

3.6.1. Conjunto de prueba y condiciones de inferencia

La evaluación se realiza sobre el conjunto de prueba definido en la Sección 3.2.6: las 10,000 muestras (20 % de N), que mantienen la proporción de clases del conjunto completo y nunca se vieron durante el entrenamiento.

Los pulsos se almacenan en su forma limpia, sin ruido. El ruido de medición se introduce exclusivamente en tiempo de inferencia, como una perturbación aditiva gaussiana de desviación estándar $\sigma = 0.005$ aplicada a cada muestra del conjunto de prueba en una única realización, con el propósito de emular las condiciones reales de adquisición; la desviación es la medida sobre el instrumento (Sección 3.2.5). Esa realización se sortea de un generador aleatorio dedicado, sembrado con la semilla del experimento desplazada en dos, de modo que el ruido de prueba es reproducible bit a bit e independiente tanto del estado del generador global como de los sorteos consumidos durante el entrenamiento. La perturbación se aplica una sola vez, antes de la extracción de características, que no añade ruido propio: no existe doble perturbación. Usar una única realización por muestra, en lugar de promediar varias, reproduce la situación operativa real, en la que cada punto medido se observa una vez; la incertidumbre que esa realización única induce sobre las métricas agregadas queda acotada por el tamaño del conjunto de prueba (10,000 muestras, Sección 3.2.6).

3.6.2. Métricas reportadas y esquema oracle

El desempeño se evalúa sobre la partición de prueba con dos familias de métricas, las de conteo de capas y las de regresión de parámetros.

Para el conteo de capas se reporta la exactitud (*accuracy*) global,

$$\text{Acc} = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \mathbb{1}[\hat{L}_i = L_i], \quad (3.21)$$

calculada como la fracción de muestras de prueba cuyo conteo predicho $\hat{L}_i$ coincide con el conteo verdadero L_i . Adicionalmente se computa la exactitud condicional por clase,

$$\text{Acc}(L) = P(\hat{L} = L \mid L), \quad (3.22)$$

para cada L presente en el conjunto de prueba, y la matriz de confusión de tamaño 5×5 sobre las clases $\{0, 1, 2, 3, 4\}$, correspondiente a $L_{\text{máx}} = 4$.

La cabeza de regresión se evalúa en modo *oracle*, es decir, condicionada al L verdadero. Para cada $L \in \{1, \dots, L_{\text{máx}}\}$ se selecciona el subconjunto $\{i : L_i = L\}$ y se leen los primeros L slots de la cabeza de parámetros de la red, que se comparan posicionalmente (el k -ésimo slot predicho contra la k -ésima capa real) contra los L parámetros verdaderos de cada muestra. El caso $L = 0$ se excluye de esta métrica, pues no posee parámetros que estimar. La Figura 3.11 esquematiza este pareo posicional.

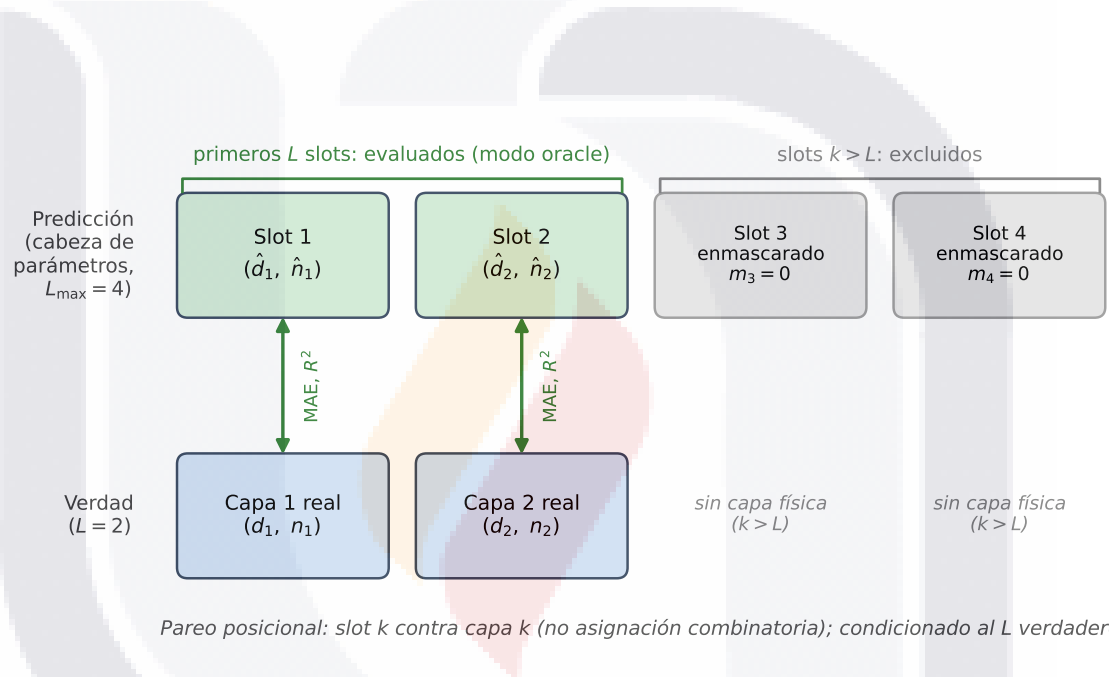


Figura 3.11. Esquema de la evaluación en modo *oracle* de la cabeza de parámetros, para un ejemplo con $L = 2$ capas reales y $L_{\max} = 4$. La red emite siempre $L_{\max}$ slots; condicionada al L verdadero, la métrica lee los primeros L slots y los compara posicionalmente (slot k contra capa k) para computar el MAE y el R^2 . Los slots con $k > L$ no corresponden a ninguna capa física (máscara $m_k = 0$) y quedan excluidos. El pareo es posicional, no una asignación combinatoria óptima: la métrica supone, además del conteo correcto, que el orden de los slots predichos coincide con el de las capas reales.

Sobre cada subconjunto se reportan el error absoluto medio (MAE) y el coeficiente de determinación R^2 , por separado para el espesor d y para el índice de refracción n' :

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |\hat{\theta}_i - \theta_i|, \quad R^2 = 1 - \frac{\sum_i (\hat{\theta}_i - \theta_i)^2}{\sum_i (\theta_i - \bar{\theta})^2}, \quad (3.23)$$

donde θ denota genéricamente el parámetro evaluado (d o n'), $\hat{\theta}_i$ su valor predicho, θ_i su valor real y $\bar{\theta}$ la media de los valores reales del subconjunto. La elección de este par de métricas es complementaria. El MAE se expresa en las unidades físicas del parámetro (micrómetros para d , adimensional para n'), lo que permite confrontarlo directamente con los espesores reales del problema y con la precisión que una aplicación exigiría; y el R^2 , al ser adimensional y relativo a la varianza del subconjunto, permite comparar la calidad de regresión entre parámetros de escalas distintas. Se prefiere el MAE sobre la raíz del error cuadrático medio como métrica principal porque pondera todos los errores linealmente y es menos sensible a valores atípicos, lo que lo hace más representativo del error típico por capa. Ambas métricas se calculan sobre el vector aplanado que reúne todos los *slots* del subconjunto: para un subconjunto con L capas, cada muestra aporta L pares $(\hat{\theta}, \theta)$, y el conjunto de todos ellos constituye las m observaciones del cálculo. Como todas las capas se sortean del mismo rango (Tabla 3.2), las posiciones comparten distribución y el R^2 agregado no queda sesgado por diferencias sistemáticas de escala entre *slots*. El espesor d se reporta en micrómetros (μm), unidad en la que se almacenan los valores de referencia.

El carácter *oracle* de esta evaluación significa que la métrica aísla la regresión de los aciertos del conteo. Al condicionar al L verdadero, los errores no se inflan por fallos en la predicción del número de capas, y el esquema mide el error de regresión puro bajo conteo perfecto. El pareo posicional es válido porque el orden de las capas no es arbitrario, ya que el simulador las apila en orden de profundidad y el *ground truth* las almacena en ese mismo orden, de modo que el *slot* k corresponde siempre a la k -ésima capa desde la superficie. Además, la pérdida de entrenamiento (Ec. (3.14)) usa ese mismo pareo, así que la red aprende a emitir sus salidas en esa convención. En un escenario real con L desconocido, el error podría ser mayor por dos fuentes adicionales, los errores de conteo y los de asignación de *slots*.

3.7. Reproducibilidad

Para que los resultados reportados en este trabajo puedan ser auditados y replicados, se publican en el repositorio de Github <https://github.com/claudiacollazoguma/thz-cnn-multilayer>.

Capítulo 4 Resultados

Este capítulo presenta los resultados obtenidos a lo largo de la investigación. Se empieza por el sistema principal sobre datos sintéticos, donde se conoce la estructura verdadera de cada muestra y, por lo tanto, el error puede medirse de forma exacta (Sección 4.1). Por último, el sistema se aplica a mediciones reales de pintura (Sección 4.2).

4.1. Desempeño del sistema principal sobre el conjunto sintético

El sistema principal (Sección 3.3.2) se evaluó sobre el conjunto de prueba de la Sección 3.6.1: 10,000 muestras nunca vistas, con ruido de medición $\sigma = 0.005$ aplicado una sola vez con semilla fija, de modo que la evaluación es repetible.

4.1.1. Conteo de capas

La exactitud global de conteo es del 77.4%. La Tabla 4.1 la desglosa por clase y la Figura 4.1 presenta la matriz de confusión.

Tabla 4.1. Exactitud de conteo del sistema principal (semilla 42) sobre el conjunto de prueba sintético (2,000 muestras por clase).

	$L = 0$	$L = 1$	$L = 2$	$L = 3$	$L = 4$	Global
Exactitud (%)	98.7	91.8	74.8	55.0	66.9	77.4

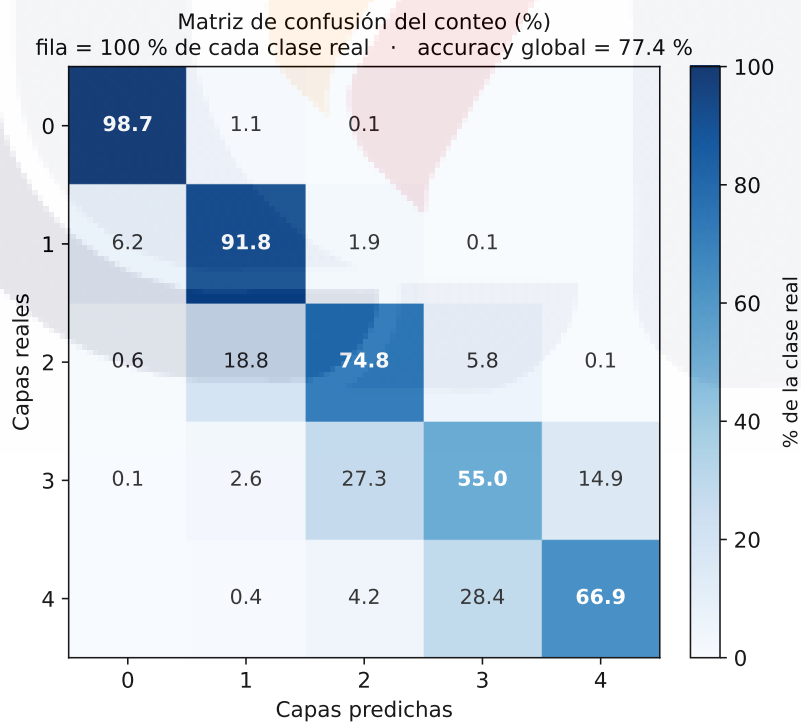


Figura 4.1. Matriz de confusión del conteo de capas (sistema principal, semilla 42), normalizada por fila (cada fila suma el 100 % de su clase real). El 92.6 % de los errores confunde clases adyacentes ($\hat{L} = L \pm 1$).

Las clases $L = 0$ y $L = 1$ se resuelven de forma casi perfecta: el caso sin pintura es determinista salvo por el ruido,

y una capa única produce un par de ecos bien separados. La dificultad crece con el número de capas y los errores no son aleatorios: el 92.6 % de las confusiones ocurre entre clases adyacentes, el comportamiento esperable cuando el eco de una capa fina o de bajo contraste queda enmascarado por el solapamiento temporal. Que $L = 3$ sea la clase más difícil, y no $L = 4$ (el caso con más capas), se explica porque recibe errores desde ambos lados (se lee como 2 o como 4), mientras que $L = 4$ solo puede confundirse hacia abajo.

4.1.2. Regresión de parámetros en modo oracle

La Tabla 4.2 reporta la cabeza de regresión condicionada al conteo verdadero (protocolo de la Sección 3.6.2).

Tabla 4.2. Regresión de parámetros del sistema principal (semilla 42) en modo *oracle*, por número de capas verdadero. MAE en las unidades de cada parámetro; 2,000 muestras por clase.

L	MAE_d (μm)	R_d^2	MAE_n	R_n^2
1	9.4	0.72	0.021	0.98
2	15.0	0.36	0.048	0.85
3	16.9	0.24	0.068	0.72
4	17.3	0.19	0.079	0.64

El error crece con la profundidad del problema, de 9.4 a $\sim 17 \mu\text{m}$ en espesor y de 0.021 a 0.079 en índice, con R_d^2 que cae hasta 0.19 en $L = 4$: con muchas capas, la CNN captura la estructura gruesa pero no resuelve los parámetros finos.

4.1.3. Dónde ubica la red los ecos

La red no busca los picos uno por uno, sino que predice los parámetros directamente del pulso. Para ver cómo lo hace, la Figura 4.2 muestra dos cosas sobre ejemplos del test sintético. Por un lado, dónde caen los ecos según la estratigrafía que predice la red, calculados con sus $\hat{d}_k$ y $\hat{n}_k$, frente a dónde caen los ecos reales. Por otro, un mapa de saliencia, que resalta las partes del pulso que más pesan en la decisión de cuántas capas hay.

La saliencia se concentra en el grupo de ecos, no en el ruido ni en la línea de base, lo que muestra que la red se fija en la parte físicamente relevante del pulso. Además, la diferencia entre los ecos predichos y los reales es pequeña con una capa y crece a medida que hay más, en línea con el error de regresión de la Tabla 4.2, ya que con más capas hay más solapamiento y los ecos profundos se ubican peor.

4.2. Validación con mediciones reales de pintura automotriz

Esta sección confronta el sistema entrenado con mediciones reales del dominio de validación declarado en la metodología. Son cinco muestras de pintura automotriz, denominadas por su color (amarillo, azul, blanco, mandarina y naranja), preparadas con el proceso típico de pintado sobre soporte plástico, más una medición del metal desnudo que opera como espejo de referencia. Para estas cinco muestras se midió la estratigrafía verdadera por microscopía de corte transversal (Tabla 4.3), de modo que la comparación puede hacerse contra ella, es decir, contra el número de capas y el espesor por capa medidos en las propias muestras.

4.2.1. Estratigrafía verdadera por microscopía de corte

La estratigrafía de cada muestra se determinó por microscopía óptica del corte transversal, con la escala calibrada contra un patrón de 0.050 mm. La Tabla 4.3 reporta el número de capas y el espesor de cada una, y la Figura 4.3 muestra las micrografías de los cortes de las que se obtuvieron. Esta es la referencia contra la cual se contrastan las predicciones del modelo.

Tabla 4.3. Estratigrafía verdadera de las cinco muestras de pintura, por microscopía de corte transversal (escala calibrada con patrón de 0.050 mm). d_k : espesores por capa en el orden medido; L : número de capas; Σd : espesor total. Espesores en μm .

Muestra	L	d_k (μm)	Σd (μm)
Azul	4	12.2, 14.2, 14.2, 9.5	50.1
Naranja	4	28.2, 14.4, 43.1, 4.73	90.4
Blanco	3	38.6, 50.0, 11.8	100.4
Amarillo	4	32.7, 7.5, 61.0, 3.45	104.7
Mandarina	3	22.2, 25.5, 7.1	54.8

Cada medición se llevó a la misma representación que las señales de entrenamiento. Se le removió el nivel de continua estimado en la ventana previa al pulso y se alineó en el tiempo contra el eje del simulador. El corrimiento se estimó

Dónde ubica la CNN los ecos — ejemplos representativos por L (test sintético plástico; eco predicho vs real; sombreado = saliencia)

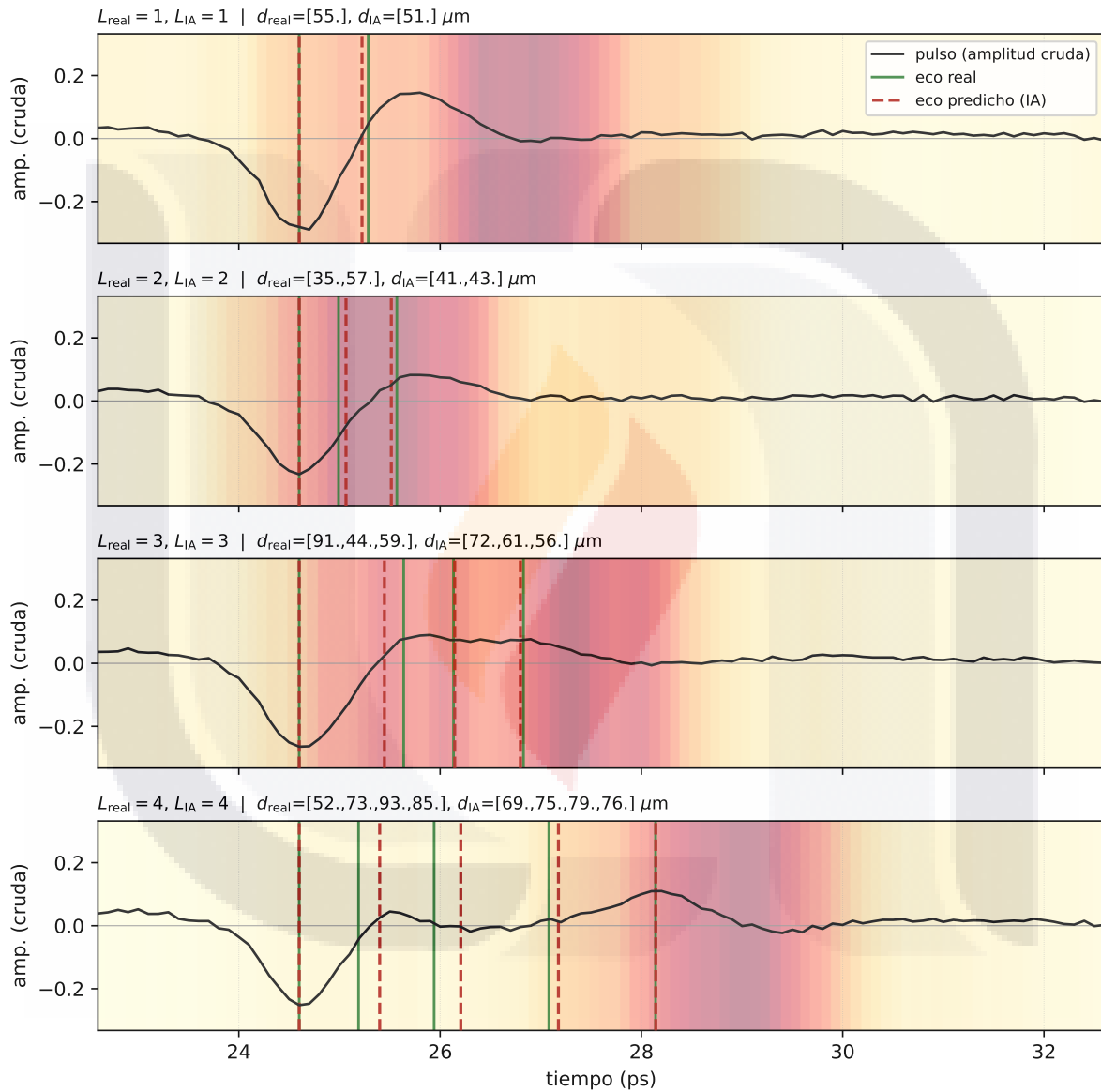
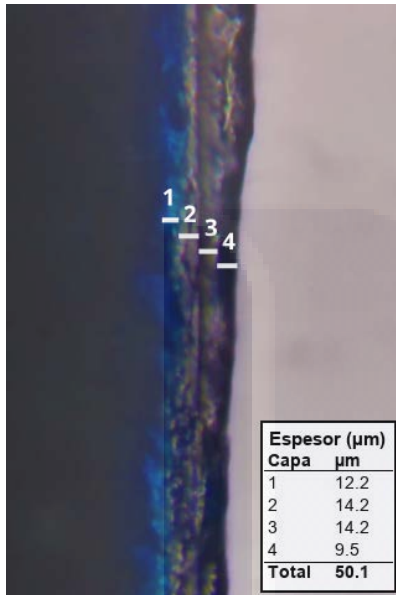
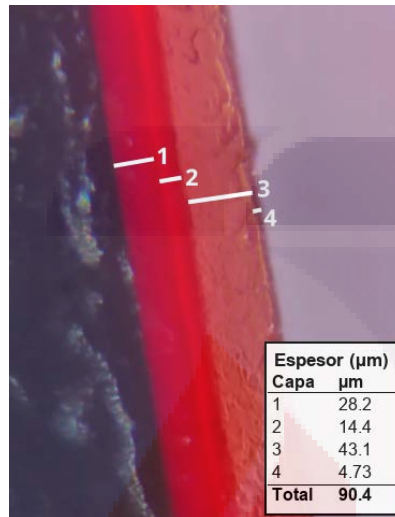


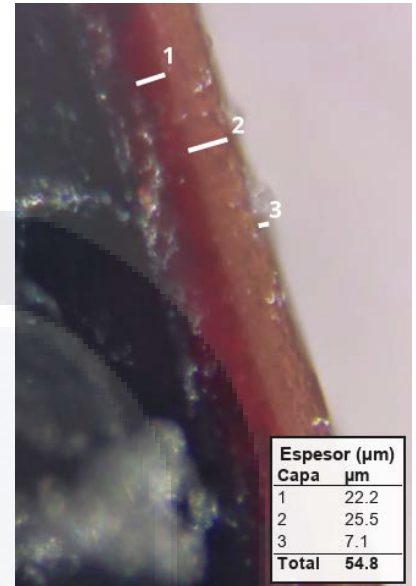
Figura 4.2. Ubicación de los ecos sobre pulsos sintéticos representativos (uno por clase, correctamente contados; pulso en amplitud cruda, tal como lo recibe la red). Líneas verdes continuas: ecos de la estratigrafía verdadera. Líneas rojas discontinuas: ecos de la estratigrafía predicha por la CNN. Sombreado: saliencia $|\partial \text{logit}_{\hat{L}} / \partial \text{pulso}|$ normalizada. La saliencia se concentra en el paquete de ecos principal, y el desajuste predicho-verdadero crece con L , en línea con la Tabla 4.2.



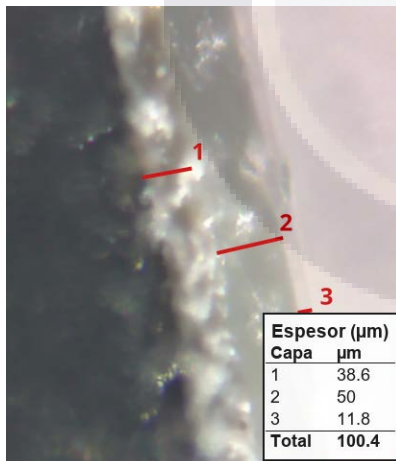
(a) Azul ($L = 4$)



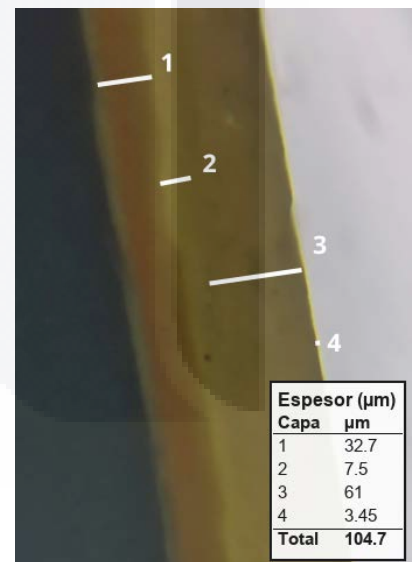
(b) Naranja ($L = 4$)



(c) Mandarina ($L = 3$)



(d) Blanco ($L = 3$)



(e) Amarillo ($L = 4$)

Figura 4.3. Micrografías ópticas de los cortes transversales de las cinco muestras de pintura sobre sustrato plástico, base de la estratigrafía de la Tabla 4.3: (a) azul, (b) naranja, (c) mandarina, (d) blanco, (e) amarillo. En cada corte, el recubrimiento es la banda de color contigua al borde del sustrato; la numeración sobre cada micrografía señala las capas medidas y el recuadro reporta su espesor en μm .

correlacionando la medición de metal desnudo con el pulso de referencia sintético, resultó en -7.6 ps y se aplicó idéntico a todas las muestras. Por último se aplicó la misma representación que en el entrenamiento. Tras el alineamiento, la correlación de Pearson entre el metal medido y la referencia simulada es $r = 0.968$, lo que valida de extremo a extremo la cadena de referencia, polaridad y preprocesamiento. El metal desnudo no funciona como control de la clase $L = 0$, porque el modelo se entrenó sobre sustrato plástico y una medición de metal queda fuera de su dominio: en efecto la clasifica como $L = 4$. La clase $L = 0$ corresponde al soporte plástico desnudo, que no se midió por separado. La Figura 4.4 muestra el acuerdo metal medido-simulado y las cinco señales de pintura tal como las recibe el modelo.

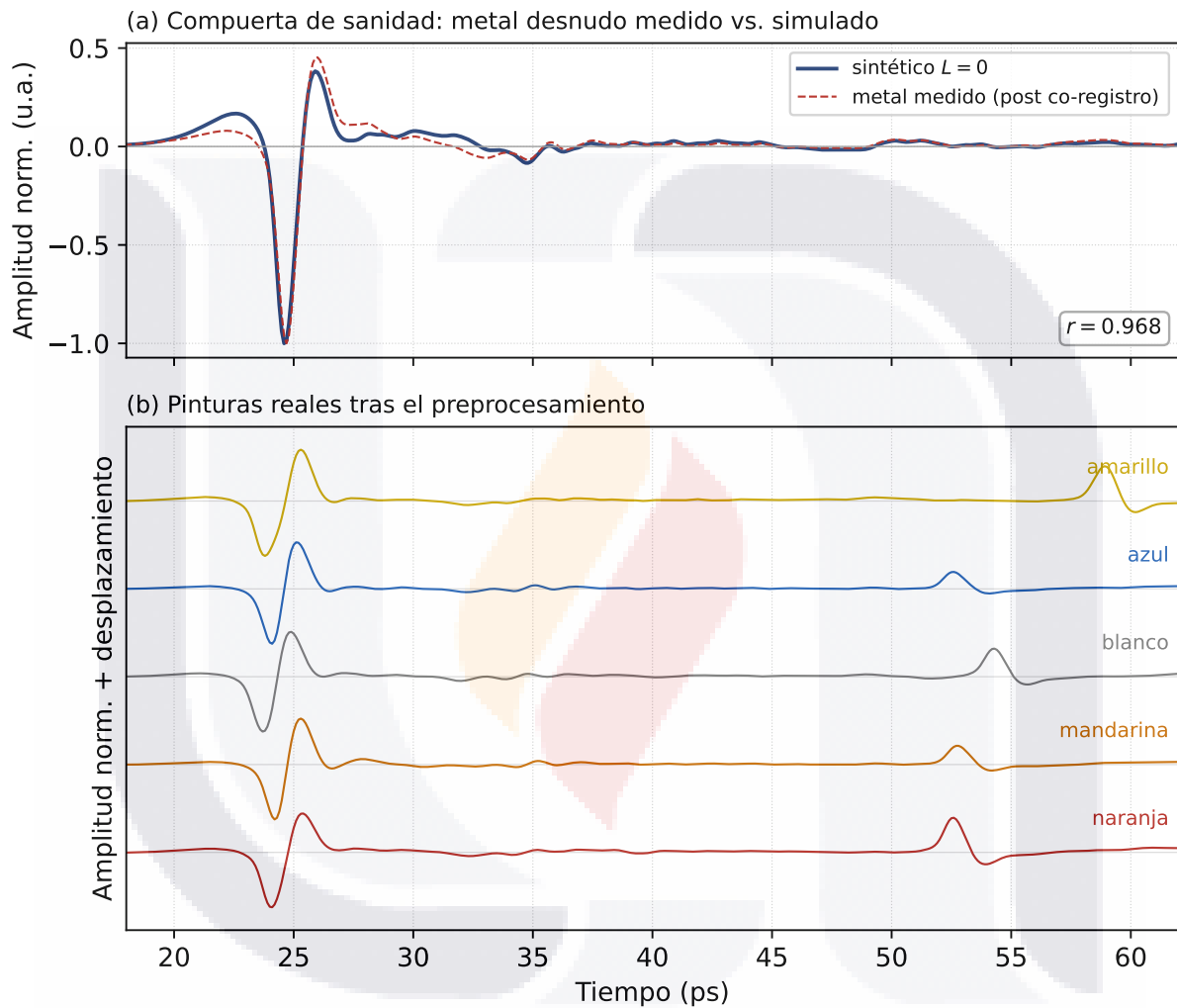


Figura 4.4. Entradas de la validación con mediciones reales. (a) Comprobación básica: medición real de metal desnudo tras el preprocesamiento (línea discontinua) superpuesta al pulso de referencia sintético (línea continua); la correlación de Pearson tras el alineamiento es $r = 0.968$. (b) Las cinco mediciones de pintura tras la misma cadena de preprocesamiento, desplazadas verticalmente. En todas se aprecia el paquete de ecos principal (≈ 24 – 28 ps) y un eco posterior (≈ 52 – 56 ps); la medición del amarillo muestra la anomalía de polaridad discutida en el texto.

4.2.2. Predicciones sobre las muestras reales

La Tabla 4.4 presenta las predicciones del sistema principal sobre las cinco pinturas, con la clase de conteo, su probabilidad *softmax* y los espesores por *slot*. La Figura 4.5 las hace visibles sobre el propio pulso, mostrando dónde sitúa la red los ecos predichos y dónde concentra su saliencia.

El conteo de capas es lo mejor validado. Frente a la estratigrafía verdadera (Tabla 4.3), el sistema principal acierta el número de capas en dos de las cinco muestras (amarilla y mandarina), con un error medio de 0.6 capas. En las tres restantes subcuenta por una sola capa (azul $\hat{L} = 3$ frente a 4, naranja $\hat{L} = 3$ frente a 4 y blanco $\hat{L} = 2$ frente a 3). El subconteo es el error físicamente esperable (Tabla 4.3): cuando una capa cae por debajo del piso de $20\mu\text{m}$, sus ecos son los más débiles y

Dónde la CNN ubica los ecos — mediciones reales de pintura (eco predicho; sombreado = saliencia; sin verdad de d,n)

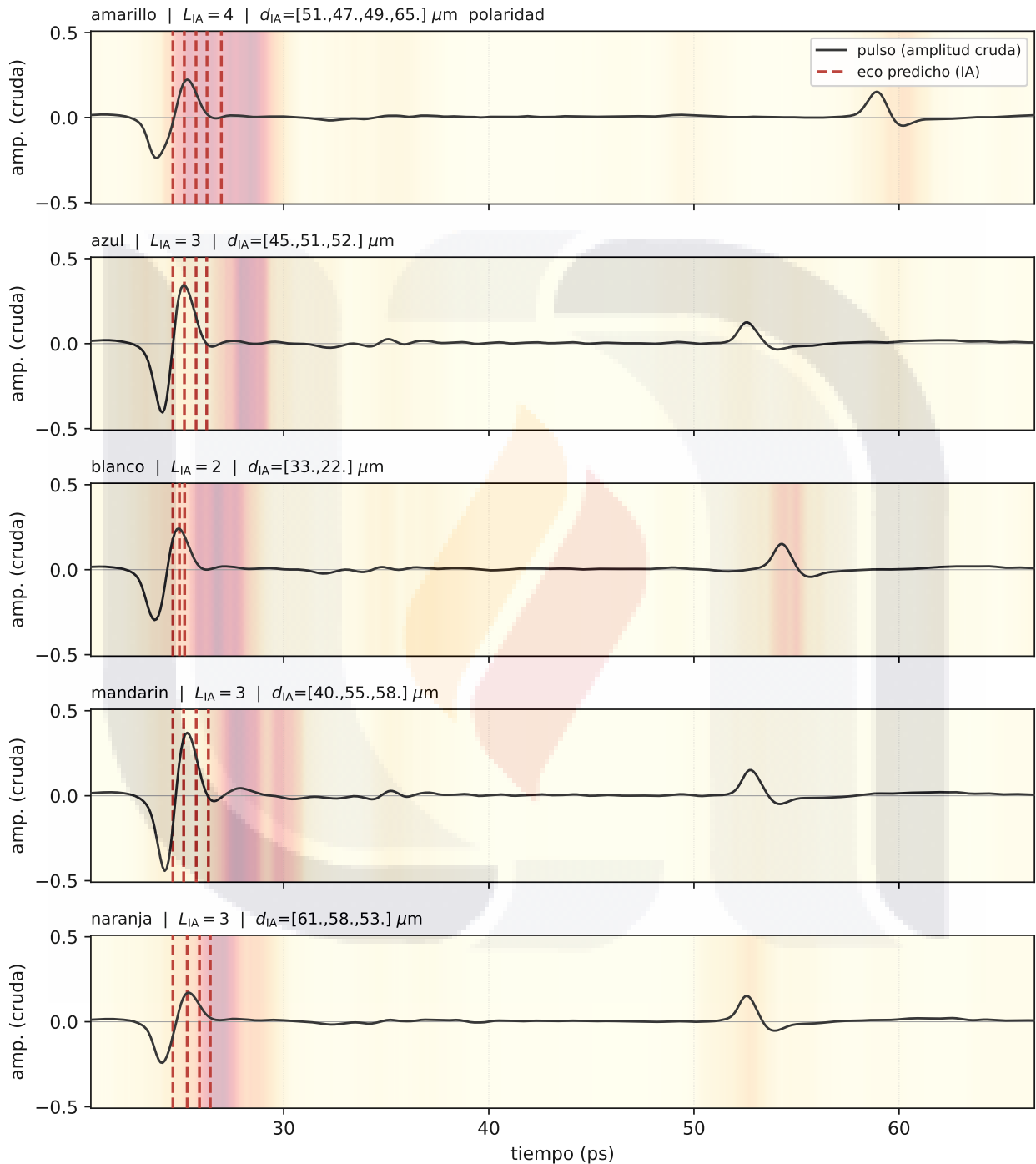


Figura 4.5. Ubicación de los ecos predichos sobre las mediciones reales (pulso en amplitud cruda; líneas rojas discontinuas: ecos de la estratigrafía predicha; sombreado: saliencia del logit de conteo). La microscopía aporta el espesor por capa pero no el índice n —necesario para situar los ecos verdaderos en el tiempo—, por lo que solo se grafican los ecos predichos. La saliencia se concentra en el paquete principal ($\approx 24-28$ ps); el eco posterior ($\approx 52-56$ ps) queda fuera de la región atendida por la red.

Tabla 4.4. Predicciones del sistema principal sobre las mediciones reales. P es la probabilidad *softmax* de la clase predicha. Los espesores $\hat{d}_k$ se listan del *slot* 1 al $\hat{L}$ (rango de entrenamiento $d \in [20, 100]$ μm). $\Sigma \hat{d}$ es el espesor total predicho.

Muestra	$\hat{L}$	P	$\hat{d}_k$ (μm)	$\Sigma \hat{d}$ (μm)
Amarillo ^a	4	0.934	51.4, 46.6, 49.0, 65.5	212.5
Azul	3	0.779	45.1, 51.5, 51.6	148.2
Blanco	2	0.490	32.5, 21.9	54.4
Mandarina	3	0.512	39.6, 54.8, 57.9	152.2
Naranja	3	0.329	61.5, 58.1, 53.4	173.0

^a La medición del amarillo presenta una anomalía de polaridad (pico principal positivo) respecto del resto de las muestras y del dominio de entrenamiento; su predicción debe leerse con reserva adicional (véase el texto).

próximos, y dos capas vecinas se confunden en una sola. Además, varias predicciones son poco confiadas, con probabilidad *softmax* apenas por encima del azar (naranja 0.33, blanco 0.49, mandarina 0.51). La Tabla 4.6 resume esta comparación, junto con la de la línea base tradicional.

En los espesores, el modelo sobreestima de forma sistemática. Los espesores predichos se concentran entre 22 y 66 μm , sin bajar del piso de entrenamiento de 20 μm . Pero diez de las dieciocho capas reales miden menos de 20 μm (hasta 3.45 μm , Tabla 4.3), de modo que esas capas finas se predicen engrosadas hasta el rango aprendido. Como consecuencia, los totales predichos (54–213 μm) no siguen a los reales (50–105 μm , Tabla 4.3): salvo en blanco, donde el subconteo los deja cortos, los sobreestiman por un factor de 1.9 a 3.0.

Los índices de refracción resultan poco informativos en este caso. Como no existe medición directa del índice THz de estas pinturas contra la cual comparar (la cota [1.4, 2.0] proviene de materiales similares, Sección 3.2.1), no se extraen conclusiones de los $\hat{n}_k$ sobre muestras reales.

4.2.3. Contraste con el método tradicional de deconvolución

Para situar el sistema principal frente al procedimiento clásico, se aplicó la línea base de deconvolución de Wiener con detección de picos (Sección 3.5) a las mismas cinco pinturas. El método divide en el dominio de la frecuencia el espectro de cada medición por el del pulso de metal desnudo, con la regularización de Wiener de la Ec. (3.18), y asume un índice de refracción $n = 1.60$, ya que requiere n a priori para convertir retardos en espesores. Ese valor cae dentro del rango de índices que la literatura de estratigrafía THz reporta para capas de pintura y barniz (≈ 1.5 –1.85) (Cai et al., 2017; Dong et al., 2017). A diferencia del sistema principal, este método usa una referencia, y su propio desplazamiento circular alinea la traza, por lo que no interviene el corrimiento temporal estimado para la red. La Tabla 4.5 reúne sus salidas, con el número de picos detectados, las capas plausibles tras descartar los huecos fuera de [10, 500] μm y los espesores resultantes.

Tabla 4.5. Salidas de la línea base tradicional (deconvolución de Wiener con detección de picos, Sección 3.5) sobre las mediciones reales, con $n = 1.60$ asumido y el metal desnudo como referencia. «#picos» es el número de máximos detectados; L , las capas plausibles tras filtrar los huecos fuera de [10, 500] μm ; en negrita, los espesores fijados en 84.3 μm , que es el valor mínimo que el método puede reportar —no una medición real, sino el límite inferior del procedimiento (véase el texto). Σd es el espesor total.

Muestra	#picos	L	d_k (μm)	Σd (μm)
Amarillo	5	3	131.2, 93.7, 84.3	309.2
Azul	3	1	84.3	84.3
Blanco	4	2	121.8, 84.3	206.1
Mandarina	3	1	84.3	84.3
Naranja	3	1	84.3	84.3

En el conteo, la línea base subcuenta de forma sistemática. Sobre las cinco pinturas, de tres o cuatro capas según la microscopía (Tabla 4.3), la línea base detecta entre una y tres capas, por debajo tanto de la estratigrafía verdadera como de los $\hat{L} \in \{2, 3, 4\}$ del sistema principal (Tabla 4.4), y no acierta el conteo en ninguna muestra. En tres de las cinco (azul, mandarina y naranja) colapsa toda la estratigrafía a un único par de interfaces.

En espesor, el método queda clavado en su valor mínimo. Obtiene el grosor de cada capa a partir de la distancia entre dos picos consecutivos, de modo que a mayor distancia corresponde mayor grosor. Pero su detector no admite dos picos separados por menos de nueve muestras, así que el grosor más pequeño que puede reportar es $c(9\Delta t)/(2n) = 84.3$ μm (con $\Delta t = 0.1$ ps y $n = 1.60$), y por debajo de ese valor le resulta imposible. Las cinco muestras presentan al menos un espesor fijado exactamente en ese mínimo de 84.3 μm , y en azul, mandarina y naranja es el único valor, lo que indica que el método

no midió la capa, sino que devolvió su límite inferior porque no logró separar los ecos. Los pocos espesores por encima de ese mínimo (93.7–131.2 μm) carecen de contraparte verdadera con la cual contrastarlos, y el único total comparable, el del amarillo (309 μm), casi triplica el total real de esa muestra (104.7 μm , Tabla 4.3).

En conjunto, contra la microscopía (Tabla 4.6), el procedimiento clásico de detección de picos no acierta el conteo en ninguna de las cinco muestras, con un error medio de 2.0 capas frente a 0.6 del sistema principal, y no resuelve la estratigrafía de capas finas de estas pinturas. Esto concuerda con el límite de resolución por solapamiento de ecos descrito en el marco teórico (Sección 2.2.2). Este contraste motiva el enfoque aprendido, que infiere la estructura desde el pulso crudo sin depender de la separabilidad de los picos en la traza deconvolucionada. En espesor absoluto, en cambio, ninguno de los dos métodos es fiable sobre estas mediciones reales. La línea base queda clavada en su mínimo de 84.3 μm y el sistema principal sobreestima los totales al predecir espesores en el rango aprendido para capas que en realidad son más finas. Reducir esa distancia entre el dominio sintético y las mediciones reales queda como trabajo futuro.

Tabla 4.6. Comparación contra la estratigrafía verdadera por microscopía (Tabla 4.3): conteo de capas L y espesor por capa d_k del sistema principal (CNN) y de la línea base tradicional (deconvolución de Wiener). Como el número de capas difiere entre métodos, los d_k se listan en su propio orden, sin implicar una correspondencia capa a capa. En negrita, los conteos que coinciden con la microscopía. El sistema principal acierta el conteo en 2 de 5 muestras (error medio 0.6 capas); el tradicional, en ninguna (error medio 2.0, subconteo sistemático). Ningún método estima bien el espesor absoluto sobre muestras reales (véase el texto). Espesores en μm .

Muestra	Microscopía (verdad)		CNN (principal)		Tradicional	
	L	d_k	$\hat{L}$	$\hat{d}_k$	L	d_k
Azul	4	12.2, 14.2, 14.2, 9.5	3	45.1, 51.5, 51.6	1	84.3
Naranja	4	28.2, 14.4, 43.1, 4.73	3	61.5, 58.1, 53.4	1	84.3
Blanco	3	38.6, 50.0, 11.8	2	32.5, 21.9	2	121.8, 84.3
Amarillo	4	32.7, 7.5, 61.0, 3.45	4	51.4, 46.6, 49.0, 65.5	3	131.2, 93.7, 84.3
Mandarina	3	22.2, 25.5, 7.1	3	39.6, 54.8, 57.9	1	84.3
Aciertos de conteo	—		2/5		0/5	

Capítulo 5 Discusión de resultados

5.1. Respuesta a la pregunta de investigación

¿Puede el aprendizaje supervisado, sobre datos THz, recuperar la estratigrafía (número de capas, espesores e índices) de una pintura multicapa? Sí, con condiciones. Sobre datos sintéticos, el sistema cuenta y estima los parámetros (Sección 4.1); sobre las cinco pinturas reales, contrastadas contra microscopía de corte, recupera el número de capas en dos, frente a ninguna del método tradicional, pero no el espesor absoluto. Esas condiciones, en particular el rango de espesor, delimitan el trabajo pendiente.

5.2. Comparación con trabajos relacionados

Los trabajos que recuperan estructura por THz se reparten en dos familias y este se ubica entre ambas, la que ajusta un modelo físico a la señal medida y la que entrena una red para invertirla. La Tabla 5.1 las resume y deja claro lo principal, que las cifras no son directamente comparables, porque cada trabajo resuelve una tarea distinta, sobre un tipo de muestra distinto y bajo supuestos distintos.

Tabla 5.1. Comparación de este trabajo con trabajos relacionados de medición de estructura por THz. Cada resultado se reporta tal como lo declara su fuente. Las cifras no son comparables entre sí: varían la tarea, el tipo de muestra y la información que cada método supone conocida (véase el texto).

Trabajo	Tarea	Enfoque y entrada	Información a priori	Resultado reportado
Este trabajo	Nº de capas (0–4), espesor e índice por capa	CNN 1D sobre el pulso crudo	Ninguna	77.4 % de conteo (sintético); 2 de 5 en pinturas reales
Busboom et al. (2020)	Nº de capas (1–3)	Red densa; pulso deconvolucionado y alineado	Deconvolución y alineamiento temporal	96.5 % (sintético); 86.7 % (real)
Klokkou et al. (2022)	Índice n, k (una capa)	Red; espectro de amplitud y fase	Referencia y espesor conocidos	Sin error numérico; $\sim 2\times$ más rápida que Newton–Raphson
Krimi et al. (2016)	Espesor por capa (≤ 4)	Modelo de Rouard con optimización	Nº de capas e índices calibrados	Error por capa $< 1 \mu\text{m}$
van Mechelen et al. (2014)	Espesor por capa (1–3)	Modelo disperso estratificado	Número de capas; índices no calibrados	Total THz $96.7 \mu\text{m}$ (referencia $98 \mu\text{m}$)
Dong et al. (2016)	Estratigrafía (≤ 2 capas)	Deconvolución frecuencia–wavelet	Referencia metálica; índice constante	Separa una capa oculta (cualitativo)
Jäckel et al. (2025)	Espesor por capa (3)	Matriz de transferencia; THz continuo de 6 frecuencias	Nº de capas, índices y espesores nominales	$\sigma < 5 \mu\text{m}$ por capa en $< 1 \text{ ms}$

En la primera, Krimi et al. (2016) mide el espesor de pinturas automotrices de hasta cuatro capas sobre metal con el método de Rouard y optimización evolutiva, y reporta una exactitud por capa habitualmente menor a $1 \mu\text{m}$. van Mechelen et al. (2014) recupera los espesores de un sistema automotriz de tres capas con un modelo disperso estratificado, con un total THz de $96.7 \mu\text{m}$ frente a los $98 \mu\text{m}$ de la referencia mecánica. Jäckel et al. (2025) determina los tres espesores de una muestra recubierta a partir de solo seis frecuencias de THz de onda continua, con una desviación estándar por capa menor a $5 \mu\text{m}$ en menos de un milisegundo. Dong et al. (2016) reconstruye la estratigrafía de una pintura artística real mediante

deconvolución en el dominio frecuencia-wavelet, igual que otros trabajos de la misma línea (Chen et al., 2010; Dong et al., 2017). Estos métodos alcanzan una precisión de espesor que este trabajo no iguala, pero a cambio de un supuesto que aquí se evita, ya que todos parten de conocer la estructura del sistema, el número de capas y, según el caso, los índices calibrados o los espesores nominales que la Tabla 5.1 detalla por trabajo. Ninguno infiere cuántas capas hay.

En la segunda, Busboom et al. (2020) cuenta capas, de una a tres, sobre sustrato metálico con una red densa, y alcanza el 96.5% de exactitud sobre datos sintéticos y el 86.7% sobre mediciones reales. Es el trabajo más cercano al conteo de este sistema, pero las cifras no se comparan de forma directa. Su red distingue tres clases y la de aquí cinco, de cero a cuatro capas; sus muestras reales tienen capas de 100 a 500 μm , mucho más gruesas que las decenas de micrómetros de la pintura, de modo que sus ecos se separan en el tiempo sin el solapamiento que limita este problema; y su red recibe la señal deconvolucionada contra un pulso idealizado y alineada en el tiempo, mientras que la de aquí recibe el pulso crudo. Coincide, eso sí, en el tipo de error, porque su red también tiende a subcontar. Klokko et al. (2022) extrae el índice de refracción complejo de un material de una sola capa con una red neuronal, a partir del espectro de amplitud y fase y con el espesor como dato conocido. No informa un error numérico de la red, solo que resulta unas dos veces más rápida que el ajuste iterativo de Newton-Raphson. Su problema es más acotado que el de aquí, con una capa, sin conteo y con referencia y espesor conocidos, y los propios autores plantean como trabajo futuro operar directamente sobre el dominio del tiempo sin referencia ni espesor, que es justamente el régimen en el que este sistema trabaja.

Frente a este panorama, el aporte de este trabajo no está en la precisión de espesor sino en el alcance de la tarea con menos información a priori. Es el único de los comparados que estima de forma conjunta el número de capas, el espesor y el índice de cada una a partir del pulso THz crudo, sin referencia, sin índices calibrados y sin fijar el número de capas. Ese alcance tiene un costo, la menor exactitud absoluta sobre las pinturas reales (Sección 4.2).

Capítulo 6 Conclusiones

Este trabajo usó aprendizaje supervisado sobre datos de espectroscopía THz en el dominio del tiempo para recuperar, sin contacto, la estratigrafía de una pintura multicapa, es decir, cuántas capas tiene y el espesor e índice de cada una. Como no se cuenta con muestras de obras reales cuya estructura se conozca, el sistema se desarrolló y validó sobre un sustituto declarado, pintura automotriz sobre plástico, y se entrenó por completo con datos sintéticos generados por un modelo físico de Fresnel-Rouard. Las conclusiones que siguen se apoyan en los Capítulos 4 y 5.

La respuesta a la pregunta de investigación es afirmativa dentro del dominio estudiado, con espesores entre 20 y 100 μm y sustrato plástico. Una sola red, entrenada solo con datos sintéticos, cuenta hasta cinco clases de capas y estima los parámetros de cada una con el error que recoge la Sección 4.1. Más que el acierto global, lo que da confianza es cómo se equivoca el sistema, porque las confusiones de conteo casi siempre caen entre clases vecinas y el desempeño baja justo donde los ecos se solapan y la información se pierde. Que los errores sigan a la física, y no a otra cosa, indica que la red aprendió el problema real y no detalles propios del simulador.

El estudio también deja claro un límite del método. Las capas más finas, por debajo del piso del rango de espesor, marcan un límite de información y no de datos ni de modelo: el sistema no las resuelve y sobre las pinturas reales su espesor se sobreestima. Superarlo pediría información adicional, como una calibración previa, más ancho de banda o una nueva forma de medir. Conviene aclarar, además, que la exactitud en las clases con más capas es algo optimista, porque las interfaces de muy bajo contraste de índice quedan por debajo del ruido y no siempre son resolubles.

Sobre mediciones reales, el sistema funciona de principio a fin y se comparó contra microscopía. Frente a la estratigrafía verdadera obtenida por corte, acierta el número de capas en dos de las cinco pinturas, mientras que el método tradicional no acierta ninguna (Sección 4.2). Esa comparación es cuantitativa, aunque se apoya en apenas cinco muestras, y la confianza de la red no está bien calibrada fuera de lo que vio durante el entrenamiento, como mostró una muestra con polaridad anómala. Los espesores, en cambio, se sobreestiman, porque las capas reales más finas caen por debajo del piso del rango aprendido, lo que deja al descubierto una brecha todavía abierta entre lo sintético y lo real en la estimación del espesor absoluto, que queda como trabajo futuro.

Por último, el sistema principal se eligió por una razón física comprobable, la resolución temporal que preserva su submuestreo. El entrenamiento no es del todo reproducible al detalle sobre el hardware usado, lo que queda como una limitación declarada.

El objetivo de fondo, la estratigrafía de pinturas artísticas, trae diferencias que este trabajo reconoce pero no resuelve. El soporte de una obra (madera, lienzo, yeso) refleja mucho menos que un metal, pero eso, lejos de ser un obstáculo, juega a favor de la separación de capas: cuando el fondo no domina la señal, las reflexiones de las propias capas aportan una fracción mucho mayor del pulso total y resultan más fáciles de aislar. En el caso opuesto, con una base metálica que se comporta casi como un espejo, la señal es esencialmente la del metal más unas correcciones pequeñas y difíciles de separar de las capas. El soporte plástico usado aquí, de reflexión débil, ya está más cerca de esa condición favorable que de una base metálica. Los materiales de una pintura artística, además, dispersan y absorben la señal de forma más irregular (Fukunaga et al., 2007; Guillet et al., 2017). Nada de esto invalida la idea central, generar con física y proponer con la red, porque no depende del tipo de sustrato; el simulador ya admite soportes y materiales más complejos, de modo que adaptarlo es cuestión de configuración y no de rediseño. Lo que sí hace falta es repetir el mismo ciclo de validación, empezando por maquetas con materiales de pintura artística y estratigrafía verdadera conocida.

Referencias

- Abraham, E., Younus, A., Delagnes, J.-C., & Mounaix, P. (2010). Non-invasive investigation of art paintings by terahertz imaging. *Applied Physics A*, 100(3), 585-590. <https://doi.org/10.1007/s00339-010-5642-z>
- Artesani, A., Abate, F., Lamuraglia, R., Baldo, M. A., Menegazzo, F., & Traviglia, A. (2023). Integrated imaging and spectroscopic analysis of painted fresco surfaces using terahertz time-domain technique. *Heritage*, 6(7), 5202-5212. <https://doi.org/10.3390/heritage6070276>
- Artesani, A., Lamuraglia, R., Menegazzo, F., Bonetti, S., & Traviglia, A. (2023). Terahertz time-domain spectroscopy in reflection configuration for inorganic and mineral pigment identification. *Applied Spectroscopy*, 77(1), 74-87. <https://doi.org/10.1177/00037028221133404>
- Born, M., & Wolf, E. (1999). *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (7.^a ed.). Cambridge University Press.
- Busboom, I., Rohde, N., Christmann, S., Feige, V., Haehnel, H., & Tibken, B. (2020). Towards neural network classification of terahertz measurements for determining the number of coating layers. *2020 45th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*, 01-02. <https://doi.org/10.1109/IRMMW-THz46771.2020.9370440>
- Caballero Mendoza, J. A. (2010). *Emisores y detectores de radiación en terahertz acoplados por fibra óptica* [Tesis de maestría]. Centro de Investigaciones en Óptica, A.C. <https://cio.repositorioinstitucional.mx/jspui/handle/1002/580>
- Cai, B., Chen, H., Xu, G., Zhao, H., & Sugihara, O. (2017). Ultra-Broadband THz Antireflective Coating with Polymer Composites. *Polymers*, 9(11), 574. <https://doi.org/10.3390/polym9110574>
- Caruana, R. (1997). Multitask Learning. *Machine Learning*, 28(1), 41-75. <https://doi.org/10.1023/A:1007379606734>
- Chen, Y., Huang, S., & Pickwell-MacPherson, E. (2010). Frequency-wavelet domain deconvolution for terahertz reflection imaging and spectroscopy. *Optics Express*, 18(2), 1177-1190. <https://doi.org/10.1364/OE.18.001177>
- Dong, J., Bianca Jackson, J., Melis, M., Giovanacci, D., Walker, G. C., Locquet, A., Bowen, J. W., & Citrin, D. (2016). Terahertz frequency-wavelet domain deconvolution for stratigraphic and subsurface investigation of art painting. *Optics Express*, 24(23), 26972-26985. <https://doi.org/10.1364/OE.24.026972>
- Dong, J., Locquet, A., Melis, M., & Citrin, D. (2017). Global mapping of stratigraphy of an old-master painting using sparsity-based terahertz reflectometry. *Scientific Reports*, 7(1), 15098. <https://doi.org/10.1038/s41598-017-15069-2>
- Duvillaret, L., Garet, F., & Coutaz, J.-L. (1996). A Reliable Method for Extraction of Material Parameters in Terahertz Time-Domain Spectroscopy. *IEEE Journal of Selected Topics in Quantum Electronics*, 2(3), 739-746. <https://doi.org/10.1109/2944.571775>
- Fukunaga, K., Ogawa, Y., Hayashi, S., & Hosako, I. (2007). Terahertz spectroscopy for art conservation. *IEICE Electronics Express*, 4(8), 258-263. <https://doi.org/10.1587/elex.4.258>
- Gezimati, M., & Singh, G. (2024). Terahertz data extraction and analysis based on deep learning techniques for emerging applications. *IEEE Access*, 12, 21174-21198. <https://doi.org/10.1109/ACCESS.2024.3360930>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Guillet, J.-P., Roux, M., Wang, K., Ma, X., Fauquet, F., Balacey, H., Recur, B., Darracq, F., & Mounaix, P. (2017). Art painting diagnostic before restoration with terahertz and millimeter waves. *Journal of Infrared, Millimeter, and Terahertz Waves*, 38(4), 369-379. <https://doi.org/10.1007/s10762-017-0358-1>
- Ige, A. O., & Sibiya, M. (2024). State-of-the-art in 1D convolutional neural networks: A survey. *IEEE Access*, 12, 144082-144105. <https://doi.org/10.1109/ACCESS.2024.3433513>
- Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, 448-456.

- Jäckel, A., Mach, C., Schwenson, L., Castro-Camus, E., Liebermeister, L., Kohlhaas, R. B., & Koch, M. (2025). Computationally Efficient Multi-Layer Thickness Determination Using Sparse CW THz Spectroscopy. *Optics Express*, 33(7), 14812-14822. <https://doi.org/10.1364/OE.550923>
- Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *3rd International Conference on Learning Representations (ICLR)*.
- Klokkou, N., Gorecki, J., Wilkinson, J. S., & Apostolopoulos, V. (2022). Artificial neural networks for material parameter extraction in terahertz time-domain spectroscopy. *Optics Express*, 30(9), 15583-15595. <https://doi.org/10.1364/OE.454756>
- Krimi, S., Klier, J., Jonuscheit, J., von Freymann, G., Urbansky, R., & Beigang, R. (2016). Highly Accurate Thickness Measurement of Multi-Layered Automotive Paints Using Terahertz Technology. *Applied Physics Letters*, 109(2), 021105. <https://doi.org/10.1063/1.4955407>
- Kubiczek, T., & Balzer, J. C. (2022). Material classification for terahertz images based on neural networks. *IEEE Access*, 10, 88667-88677. <https://doi.org/10.1109/ACCESS.2022.3200473>
- Lambert, F. E., Ornik, J., Staats, N.-A., Jäckel, A., Hernandez-Cardoso, G. G., Taiber, J., Stübling, E.-M., Rudolph, B., Mack, O., Portsteffen, H., et al. (2022). Layer separation mapping and consolidation evaluation of a fifteenth century panel painting using terahertz time-domain imaging. *Scientific Reports*, 12(1), 21038. <https://doi.org/10.1038/s41598-022-25013-8>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- Li, J., Li, Y., Cen, Y., Zhang, C., Luo, T., & Yang, D. (2020). Applications of neural networks for spectrum prediction and inverse design in the terahertz band. *IEEE Photonics Journal*, 12(5), 1-9. <https://doi.org/10.1109/JPHOT.2020.3022053>
- Loshchilov, I., & Hutter, F. (2017). SGDR: Stochastic Gradient Descent with Warm Restarts. *5th International Conference on Learning Representations (ICLR)*.
- Loshchilov, I., & Hutter, F. (2019). Decoupled Weight Decay Regularization. *7th International Conference on Learning Representations (ICLR)*.
- Park, H., & Son, J.-H. (2021). Machine learning techniques for THz imaging and time-domain spectroscopy. *Sensors*, 21(4), 1186. <https://doi.org/10.3390/s21041186>
- Reyes-Reyes, E. S. (2020). *Desarrollo de algoritmos para procesamiento de señales de THz para el análisis de objetos de valor cultural* [Tesis de maestría]. Centro de Investigaciones en Óptica, A.C.
- Reyes-Reyes, E. S., Carriles-Jaimes, R., D'Angelo, E., Nazir, S., Koch-Dandolo, C. L., Kuester, F., Jepsen, P. U., & Castro-Camus, E. (2024). Terahertz time-domain imaging for the examination of gilded wooden artifacts. *Scientific Reports*, 14(1), 6261. <https://doi.org/10.1038/s41598-024-56913-6>
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15(1), 1929-1958.
- van Mechelen, J. L. M., Kuzmenko, A. B., & Merbold, H. (2014). Stratified Dispersive Model for Material Characterization Using Terahertz Time-Domain Spectroscopy. *Optics Letters*, 39(13), 3853-3856. <https://doi.org/10.1364/OL.39.003853>
- Wu, J. (2017). *Introduction to Convolutional Neural Networks* (Technical report). National Key Lab for Novel Software Technology, Nanjing University.