

REVISIÓN DEL ESTADO DEL ARTE SOBRE EL USO DE LA INTELIGENCIA
ARTIFICIAL PARA EL ESTUDIO DE LA CORROSIÓN DE METALES

Daniel Saín Flórez García, Mauricio Meza Delgado

Trabajo de Grado para optar por el título de Ingeniero Metalúrgico

Director

Darío Yesid Peña Ballesteros

PhD en Corrosión

Universidad Industrial De Santander

Facultad De Ingenierías Fisicoquímicas

Escuela De Ingeniería Metalúrgica y Ciencia de los Materiales

Ingeniería Metalúrgica y Ciencia de Materiales

Bucaramanga

2024

Agradecimientos

“Doy las infinitas gracias a mi familia, mi madre, Liliana, padre, Elisain, hermanos, mi nona Magdalena, y mis nonos Q.E.P.D Ana, Eurípides y Alonso, quienes han sido mi pilar en valores y siempre luchando por sacar a adelante mi educación, su apoyo constante y el amor brindado, su sabiduría y fortaleza que, pese a las adversidades, nunca me dejaron desistir, puesto que la etapa universitaria es un camino de obstáculos y cada uno estuvo presente en diferentes etapas de mi vida, a ustedes les debo este logro ”

“A todos mis amigos, compañeros que me dejo la Universidad, futuros colegas ingenieros, pues gracias a ustedes, y los lazos de amistad que nos unen, se logró sacar adelante este título que tanto anhelamos.”

“Al profesor Darío Peña B., quien estuvo presente desde semestres muy tempranos, compartiendo su sabiduría y apoyo, parte importante de mi formación académica,

A la Universidad Industrial de Santander, pues se me dio la oportunidad de ser parte de esta academia, y gran nido de formación integral y ética.”

“A demás profesores con los cuales tuve la oportunidad de compartir y administrativos quienes fueron fuente de ayuda y construcción, este logro no sería posible sin todos ustedes.”

Att. Daniel Saín Flórez García

“En primer lugar les agradezco a mis padres que siempre me han brindado su apoyo incondicional el cual fue crucial para llegar a cumplir mis objetivos personales y académicos. Los cuales con su amor me han impulsado siempre a perseguir mis metas y nunca abandonarlas frente a las adversidades”.

“A cada docente que ha sido parte de mi camino universitario, y a todos ellos les quiero agradecer por compartirme el conocimiento necesario para poder concluir con este arduo camino”.

“Agradecerles a todos mis compañeros y futuros colegas, los cuales muchos de ellos se han convertido en mis amigos y hermanos. Gracias por el tiempo compartido, los trabajos juntos y las historias vividas”

Att: Mauricio Meza Delgado

Tabla de contenido

<i>Introducción</i>	14
1.1. Justificación	14
1.2. Objetivos de la tesis	15
1.2.1. Objetivo general.....	15
1.2.2. Objetivos específicos.....	15
1.3. Metodología	15
1.3.1. Determinación del tema de estudio.....	15
1.3.2. Búsqueda y revisión de documentos relevantes.....	16
1.3.3. Evaluación y clasificación de la información recolectada.....	16
1.3.4. Análisis bibliométrico de la literatura científica.....	16
1.3.5. Interpretación y análisis detallado de la información recopilada.	16
1.3.6. Reconocimiento de progresos realizados y desafíos identificados.	17
1.3.7. Comunicación de los resultados obtenidos y elaboración del informe final.	17
2. Fundamentos de Corrosión	17
2.1. Definición de Corrosión	17
2.2. Principales medios corrosivos	17
2.2.1. Corrosión en medio húmedo.....	17
2.2.2. Corrosión en otros fluidos como sales y metales fundidos.....	18
2.2.3. Corrosión en medio seco.	18
2.3. Clasificación de las formas de corrosión en base a su apariencia	18

2.3.1. Corrosión uniforme (general)..... 19

2.3.2. Corrosión galvánica (bimetálica) 19

2.3.3. Corrosión termo galvánica Corrosión termo galvánica..... 19

2.3.4. Corrosión por rendija (incluida la corrosión por depósito)..... 20

2.3.5. Picadura, corrosión por picadura..... 20

2.3.6. Ataque selectivo, lixiviación selectiva (des aleación)..... 20

2.3.7. Corrosión intergranular (incluida la exfoliación)..... 20

2.3.8. Corrosión por erosión..... 21

2.3.9. Corrosión por cavitación..... 21

2.3.10. Corrosión por frotamiento..... 21

2.3.11. Agrietamiento por corrosión bajo esfuerzos..... 21

2.3.12. Fatiga por corrosión..... 21

3. Inteligencia artificial aplicada a la corrosión..... 22

3.1. Contextualización..... 22

3.2. Modelos actuales de algoritmos..... 23

3.2.1. Artificial Neural Networks (ANN)..... 23

3.2.2. Convolutional neural networks (CNNs)..... 26

3.2.3. Recurrent neural networks (RNNs)..... 27

3.2.4. Long Short-Term Memory (LSTM)..... 27

3.2.5. Linear regression..... 28

3.2.6. Support vector regression (SVR)..... 28

3.2.7. Decision Trees (DT)..... 29

INTELIGENCIA ARTIFICIAL PARA EL ESTUDIO DE LA CORROSIÓN	6
3.2.8. Random Forests (RF).....	29
3.2.9. Particle Swarm Optimization (PSO).....	31
3.2.10. Boosting algorithms.	31
3.2.10.1. XGBoost.....	31
3.2.10.2. LightGBM.....	32
4. Metodologías de la inteligencia artificial (IA) aplicadas.....	32
4.1. Aplicación de CNN y RNN en el monitoreo de la corrosión por Imágenes.....	32
4.2. Aplicación de la Memoria larga a corto plazo (LSTM) y modelo PSO-RF.....	35
4.3. Aplicación de la combinación de múltiples algoritmos de interés.....	37
4.3.1. Caso 1 en la Predicción de la corrosión de oleoductos y gasoductos.	37
4.3.2. Caso 2. Predicción de velocidad de corrosión del acero al carbono.....	41
5. Análisis de resultados y limitaciones de los modelos IA en corrosión.....	43
5.1 Resultados de los métodos en la detección de corrosión por Imágenes.....	43
5.2 Resultados de Memoria larga a corto plazo (LSTM) y modelo PSO-RF.....	46
5.3 Resultados de la variación de múltiples métodos de ML.....	48
6. Conclusiones.....	53
7. Recomendaciones.....	54
Referencias Bibliográficas.....	55
Apéndices.....	60

Lista de tablas

Tabla 1. Algunos métodos de ML empleados en la predicción de la corrosión, en especial en la industria del gas y petróleo.....	38
Tabla 2. Análisis comparativo de algunos algoritmos de inteligencia artificial aplicados en el monitoreo de la corrosión.....	52

Lista de figuras

Figura 1. Principales formas de corrosión reagrupadas por su facilidad de reconocimiento.....	18
Figura 2. Arquitectura de razonamiento basado en casos para la predicción del comportamiento de materiales.....	22
Figura 3. Esquema de un único procesador o neurona en una red neuronal artificial.....	25
Figura 4. Inputs y Outputs de datos para predecir la corrosión a partir de barridos de polarización mediante una red neuronal artificial.....	25
Figura 5. Arquitectura de razonamiento basado en redes neuronales convolucionales (CNN).....	26
Figura 6. Un clasificador base de método ensamble propuesto.....	30
Figura 7. Procedimiento de inspección visual potenciada con ML.....	33
Figura 8. Un ejemplo que muestra la corrosión por pilotaje y las micro deformaciones superficiales.....	34
Figura 9. Diagrama del método propuesto de predicción dinámica de la corrosión para aceros de casco.....	35
Figura 10. Ejemplos de imágenes que la arquitectura de aprendizaje profundo DenseNet-201 no bien predicho.....	45
Figura 11. Rendimiento de los modelos ML expresado por el RMSE.....	48
Figura 12. A) Comparación del rendimiento del mejor modelo para cada modelo ML. Las barras de error obtenidas realizando 20 iteraciones independientes de cada modelo. B) Gráfico de paridad entre las predichas (mm/a) por el modelo RF y los valores experimentales.....	51

Lista de apéndices

A.	C-Scan Imaging para muestras seleccionadas.....	60
B.	Diagrama esquemático de la corrosión en la industria del petróleo y el gas.....	60
C.	Localización geográfica de los 7 puertos marítimos.	61
D.	Mapa de las 4 principales rutas marítimas mundiales.....	61
E.	Simulación numérica del crecimiento de defectos de corrosión..	62
F.	Las vistas macro de las formas características de corrosión de oleoductos y gasoductos.	63
G.	Rendimiento de A) Red neuronal artificial (ANN), B) Máquina de vectores de apoyo Regresión (SVR), C) Random Forest (RF), y D) K Nearest Neighbors (KNN) en datos experimentales de la tasa de corrosión para diferentes valores de los hiperparámetros.....	64
H.	Imágenes de análisis electroquímico (Lab. y simulación).....	65
I.	D-Sight Aircraft Inspection System (DAIS).	66
J.	Imágenes de inspección del ala creadas con DAIS 250C.....	66
K.	Índice de corrosión de las rutas 1, 2, 3, 4 en diferentes estaciones.	67
M.	Comparación entre la velocidad de corrosión prevista (rojo) y la registrada experimentalmente (gris) en función del tiempo para experimentos de dosis única sin corrosión.....	69
N.	Un estudio paramétrico, utilizando un modelo RF para determinar cómo varían los índices de corrosión en función del tiempo con el tipo de inhibidor	70

Glosario

Algunos de los métodos de aprendizaje automático son definidos a continuación:

- **AdaBoost (Adaptive Boosting)**: AdaBoost es un método de aprendizaje conjunto que combina múltiples aprendices débiles (normalmente árboles de decisión) para crear un clasificador fuerte. Asigna pesos más altos a los puntos de datos mal clasificados en cada iteración, forzando al modelo a centrarse en los casos más difíciles de clasificar. [14]
- **ANN (Red neuronal artificial)**: Siguiendo el modelo de las redes neuronales biológicas, las RNA están formadas por nodos interconectados dispuestos en capas. Estas redes pueden aprender patrones y relaciones complejas a partir de los datos, lo que las hace versátiles para tareas como la clasificación, la regresión y el reconocimiento de patrones en diversos ámbitos. [3]
- **CNN (Red Neuronal Convolutiva)**: Las CNN son redes neuronales especializadas diseñadas para procesar datos reticulares estructurados, como las imágenes. Aprovechan las capas convolucionales para aprender automáticamente características jerárquicas a partir de la entrada, lo que las hace muy eficaces en tareas como el reconocimiento de imágenes, la detección de objetos y otras tareas de visión por ordenador. [7] [8]
- **DNN (Red Neuronal Profunda)**: DNN es un tipo de red neuronal artificial con múltiples capas ocultas entre las capas de entrada y salida. Es capaz de aprender intrincados patrones y relaciones en los datos, lo que la hace adecuada para tareas complejas como el reconocimiento de imágenes y el procesamiento del lenguaje natural. [10] [14] Schmidhuber, J. (2015).
- **GAN (Generative Adversarial Network)**: Las GAN son una clase de modelos generativos compuestos por dos redes neuronales que compiten entre sí: un generador y un discriminador. El objetivo del generador es producir datos sintéticos realistas, mientras que el discriminador aprende a distinguir entre muestras reales y falsas. Este proceso de entrenamiento adversarial da lugar a la generación de datos sintéticos de alta calidad, utilizados habitualmente para la síntesis de imágenes, el aumento de datos y otras tareas generativas. [24] Radford, et al. (2015).
- **GBM (Gradient Boosting Machine)**: GBM es una técnica de aprendizaje por conjuntos que construye una serie de árboles de decisión secuencialmente. Cada árbol corrige los errores del anterior, centrándose en minimizar la función de pérdida. GBM es conocido por su alta precisión predictiva. [14]. Chen, T., & Guestrin, C. (2016).
- **GBRT (Gradient Boosted Regression Trees)**: GBRT es otra forma de gradient boosting que se centra específicamente en tareas de regresión. Construye árboles de regresión de forma secuencial, cada uno corrigiendo los errores de sus predecesores, con el objetivo de minimizar el residuo global. [14]. Natekin, A., & Knoll, A. (2013).

- **POS-SSCN (Parallel Optimization Solver for Semi-Supervised Clustering Networks):** POS-SSCN es un método de clustering semisupervisado que utiliza solvers de optimización paralelos. Consiste en optimizar los parámetros de las redes de clustering semisupervisadas para clasificar los puntos de datos en clusters, haciendo uso tanto de datos etiquetados como no etiquetados. [14] Shi, Y., & Wang, X. (2017).
- **RF (Random Forest):** RF es un método de aprendizaje conjunto que construye una multitud de árboles de decisión durante el entrenamiento y emite la moda de las clases (clasificación) o la predicción media (regresión) de los árboles individuales. Es robusto frente al sobreajuste y el ruido en los datos. [7] [14] [22]
- **RNN (Red neuronal recurrente):** A diferencia de las redes neuronales tradicionales, las RNN poseen bucles de retroalimentación que les permiten retener información sobre puntos de datos secuenciales. Esta capacidad hace que las RNN sean idóneas para procesar datos de series temporales, lenguaje natural y otros datos secuenciales en los que el contexto es crucial. [24]
- **SVR (Regresión de vectores de soporte):** SVR es un tipo de algoritmo de regresión que utiliza máquinas de vectores de soporte para realizar tareas de regresión. Encuentra el hiperplano que mejor se ajusta a los datos minimizando el error. Es particularmente eficaz en espacios de alta dimensión. [14]
- **SVR-FFA (Regresión de Vectores de Soporte con Algoritmo Firefly):** SVR-FFA emplea el Algoritmo Firefly para optimizar los parámetros del modelo de Regresión por Vectores de Soporte. El Algoritmo Firefly está inspirado en el comportamiento intermitente de las luciérnagas y se utiliza para tareas de optimización. [14]. Zhang, et al. (2014).
- **SVR-GA (Regresión de vectores de soporte con algoritmo genético):** SVR-GA combina la regresión por vectores de soporte con algoritmos genéticos para optimizar los hiperparámetros del modelo SVR. Los algoritmos genéticos imitan el proceso de selección natural para encontrar los mejores parámetros para el modelo SVR. [14]. Chatterjee, et al. (2012).
- **SVRPSO (Regresión de vectores de soporte con optimización de enjambre de partículas):** SVRPSO es otra técnica de optimización aplicada a SVR. Utiliza la optimización de enjambre de partículas para encontrar los mejores parámetros para el modelo SVR. La Optimización por Enjambre de Partículas simula el comportamiento social de los organismos, donde cada partícula representa una solución potencial. [14]
- **XGBoost (Extreme Gradient Boosting):** XGBoost es una implementación optimizada y altamente eficiente de gradient boosting. Incorpora varias técnicas de regularización para evitar el sobreajuste y es conocido por su velocidad y rendimiento en varias competiciones de aprendizaje automático. [14]. Chen, T., & Guestrin, C. (2016).

Resumen

Título: Revisión del estado del arte sobre el uso de la inteligencia artificial para el estudio de la corrosión de metales*

Autor: 2191136 – Daniel Saín Flórez García, 2182397 – Mauricio Meza Delgado**

Palabras clave: Corrosión, Inteligencia artificial, IA, Machine learning, metales

Descripción: El presente trabajo de grado, titulado “Revisión del estado del arte sobre el uso de la inteligencia artificial para el estudio de la corrosión de metales” tiene como propósito estudiar las distintas metodologías de Inteligencia Artificial (IA), utilizadas para identificar el fenómeno de la corrosión de una manera cualitativa y sus distintas formas. El objetivo general consiste en investigar el estado actual del arte en el ámbito del uso de la inteligencia artificial para analizar y comprender la corrosión de metales. La presente investigación de acuerdo con el enfoque de los objetivos específicos dispuestos busca proporcionar una revisión bibliográfica integral que oriente el desarrollo y la aplicación de nuevas técnicas y metodologías en este campo e identificar los distintos avances, limitaciones y parámetros de la inteligencia artificial aplicada al estudio de la corrosión. Esta revisión, culmina con una comparativa entre las metodologías de inteligencia artificial, destacando el subcampo del Machine learning, donde se subdivide en métodos de aprendizaje supervisado, no supervisado y el aprendizaje de refuerzo, analizando su empleabilidad, eficacia y recomendaciones por parte los autores, así como argumentativas por parte del presente análisis realizado.

*Trabajo de Grado** Universidad Industrial de Santander. Facultad de Ingenierías Físicoquímicas. Escuela de Ingeniería Metalúrgica y Ciencia de los materiales. Ingeniería Metalúrgica.
Director: Darío Yesid Peña Ballesteros. Doctor en Corrosión.

Abstract

Title: Review of the state of the art on the use of artificial intelligence for the study of metal corrosion*.

Author(s): 2191136 - Daniel Sain Florez Garcia, 2182397 - Mauricio Meza Delgado**.

Keywords: Corrosion, Artificial intelligence, AI, Machine learning, metals

Description: The present Bachelor Thesis, entitled "Review of the state of the art on the use of artificial intelligence for the study of metal corrosion" has the purpose of studying the different methodologies of Artificial Intelligence (AI), used to identify the phenomenon of corrosion in a qualitative way and its different forms. The general objective is to investigate the current state of the art in the use of artificial intelligence to analyze and understand the corrosion of metals. This research will be carried out in depth in accordance with the focus of the specific objectives set out, where the aim is to provide a comprehensive bibliographical review to guide the development and application of new techniques and methodologies in this field and to identify the different advances, limitations and parameters of artificial intelligence applied to the study of corrosion. This review culminates with a comparison between artificial intelligence methodologies, highlighting the subfield of machine learning, where it is subdivided into supervised, unsupervised and reinforcement learning methods, analyzing its usability, effectiveness, and recommendations by the authors, as well as argumentative by the present analysis fulfilled.

*Bachelor Thesis** Universidad Industrial de Santander. Faculty of Physicochemical Engineering. School of Metallurgical Engineering and Materials Science. Metallurgical Engineering. Director: Darío Yesid Peña Ballesteros. PhD in Corrosion.

Introducción

1.1. Justificación

La corrosión es un fenómeno natural con alta demanda de estudio, debido a sus efectos devastadores en el deterioro de los materiales en diversas industrias, afectando la integridad de maquinaria, piezas y estructuras. En los últimos años el progreso en el desarrollo de nuevas técnicas de inteligencia artificial, han sido implementadas en la investigación y mejora del procesamiento de datos para la determinación de los distintos parámetros que intervienen en la presencia de la corrosión. Estas técnicas tienen el potencial de revolucionar la forma en que se lleva a cabo el estudio de la corrosión de materiales, permitiendo un análisis más detallado y preciso.

El aprendizaje automatizado o Machine Learning (ML), es un subcampo de la Inteligencia artificial que tiene como fundamento el uso de métodos predictivos, a través de la recopilación y entrada de datos de muestra, proporcionando una visión más amplia en la combinación de variables y resultados más acertados, de acuerdo con la literatura que la fecha se ha estudiado, es fundamental la mejora e investigación para el avance del conocimiento científico de este campo.

Por lo tanto, existe una necesidad inminente de nuevos esfuerzos en la expansión de investigación y desarrollo para mejorar las capacidades de los algoritmos de ML para la predicción de la corrosión y para garantizar su aplicabilidad práctica en las estrategias de control de la corrosión en el mundo real. Abordar estos retos no sólo producirá un posible avance en el conocimiento científico del campo de la corrosión, sino que también contribuirá al desarrollo de soluciones innovadoras y eficaces para moderar el impacto de la corrosión en los metales y las infraestructuras.

1.2. Objetivos de la tesis

1.2.1. Objetivo general

Investigar el estado actual del arte en el ámbito del uso de la inteligencia artificial para analizar y comprender la corrosión de metales.

1.2.2. Objetivos específicos

Proporcionar una revisión bibliográfica integral que oriente el desarrollo y la aplicación de nuevas técnicas y metodologías en este campo.

Identificar los distintos avances, limitaciones y parámetros de la inteligencia artificial aplicada al estudio de la corrosión.

1.3. Metodología

1.3.1. Determinación del tema de estudio.

Establecimiento del tema de interés, mediante la determinación de la problemática a llevar a estudio, por medio de la lectura de fundamentos teóricos relacionados sobre el tema que se realiza en la revisión bibliográfica.

1.3.2. Búsqueda y revisión de documentos relevantes.

A partir del tema seleccionado, se realiza la búsqueda en las bases de datos proporcionadas por la biblioteca de la universidad, teniendo en cuenta la disponibilidad y calidad de la producción científica a recolectar, entre los cuales se incluirán, artículos de investigación científica, libros relacionados con el tema, revisiones bibliográficas, normas, handbooks, y proyectos de grado.

1.3.3. Evaluación y clasificación de la información recolectada.

Durante esta fase, se identifican artículos pertinentes para la tesis y documentos que contienen datos investigativos, ayudando a comprender aspectos complementarios. Luego, se lleva a cabo un proceso de filtrado y clasificación basado en la prioridad, evaluando qué recursos ofrecen la información más relevante para alcanzar los objetivos establecidos. Por ejemplo, se destaca la utilidad de herramientas como Elsevier B.V. Scopus, fundamental para realizar un análisis del estado del arte relacionado con el proyecto en desarrollo.

1.3.4. Análisis bibliométrico de la literatura científica.

Mediante la herramienta Scopus por Elsevier, se realiza el análisis bibliométrico de las estadísticas de publicaciones científicas relacionados con el tema, categorizadas por año, autores respectivos, disciplinas y países.

1.3.5. Interpretación y análisis detallado de la información recopilada.

Se analizan los documentos científicos seleccionados como base para esta revisión, donde se clasifica y selecciona el material relevante para el cumplimiento de los objetivos dispuestos en este proyecto, destacando estudios de corrosión por medios de aprendizaje automático, redes neuronales y algoritmos derivados.

1.3.6. Reconocimiento de progresos realizados y desafíos identificados.

En esta etapa se hace reconocimiento de las metodologías de inteligencia artificial aplicada al estudio de la corrosión, sus hallazgos positivos y avances tecnológicos en este campo, y de igual modo se realizan anotaciones de las limitaciones obtenidas y oportunidades de mejora en el ámbito investigativo.

1.3.7. Comunicación de los resultados obtenidos y elaboración del informe final.

En esta etapa se da paso a la aprobación por parte del director, transcurridas las revisiones pertinentes del proyecto, con respecto al progreso, consistencia en la información redactada y realización del informe final a presentar ante los evaluadores.

2. Fundamentos de Corrosión.

2.1. Definición de Corrosión.

La corrosión se define como el ataque destructivo ocasionado en un material metal por la reacción química o electroquímica con su medio ambiente, el deterioro por causas físicas en el material no se considera corrosión, sin embargo, al haber presencia de este fenómeno, acompañado de deterioro físico, se llegan a utilizar términos como, corrosión- erosión, desgaste por corrosión, o corrosión por contacto.

2.2. Principales medios corrosivos.

2.2.1. Corrosión en medio húmedo.

Ocurre cuando el medio de corrosión es el agua, con presencia de especies disueltas, el líquido presente actúa a modo de electrolito y es un proceso típicamente electroquímico.

2.2.2. Corrosión en otros fluidos como sales y metales fundidos.

Es un fenómeno que ocurre cuando los materiales metálicos son expuestos a ambientes agresivos y elevadas temperaturas, entrando en contacto con sales fundidas a altas temperaturas las cuales pueden ser altamente reactivas y causan corrosión mediante varios mecanismos, comúnmente, oxidación, sulfuración, cloración o carburación, dependiendo de la composición de la sal y del metal.

2.2.3. Corrosión en medio seco.

Se presenta cuando el ambiente corrosivo es un gas seco, este tipo de fenómeno también es comúnmente llamada corrosión química, siendo un claro ejemplo la corrosión a alta temperatura.

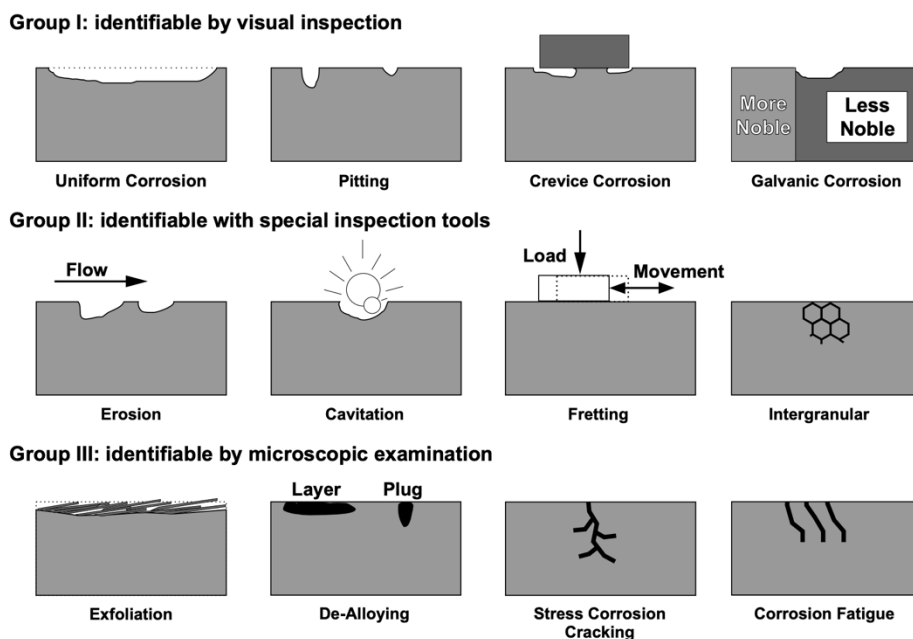
2.3. Clasificación de las formas de corrosión en base a su apariencia.

El fenómeno de la corrosión presenta variaciones respecto a su apariencia física y morfología de la zona afectada, clasificándose en distintitos grupos.

Figura 1

Principales formas de corrosión reagrupadas por su facilidad de reconocimiento.

(Reproducido de Roberge PR. *Handbook of Corrosion Engineering*. McGraw-Hill, 1999.)



2.3.1. Corrosión uniforme (general)

Es el tipo más común de corrosión, donde el metal se corroe uniformemente a lo largo de la superficie debido a la exposición con los agentes corrosivos, tales como el oxígeno o compuestos ácidos.

2.3.2. Corrosión galvánica (bimetálica)

La corrosión galvánica ocurre cuando dos metales similares entran en contacto en presencia de un electrolito, provocando la aceleración de la corrosión en el metal menos noble.

La velocidad de corrosión galvánica depende de diversos factores, tales como la naturaleza de los metales involucrados, la composición del electrolito, el radio de la superficie entre el ánodo y el cátodo, y la conductividad eléctrica del medio. Además, entre mayor sea la diferencia del potencial de electroquímico de los dos metales, más rápido será tu velocidad de corrosión.

Comprender los principios de la corrosión galvánica es esencial para mitigar sus efectos y garantizar la longevidad e integridad de los sistemas y estructuras metálicos en diversas aplicaciones. Algunas medidas de prevención para la corrosión galvánica incluyen:

- Aislar los metales distintos para evitar el contacto directo.
- Utilizar ánodos de sacrificio de un metal más reactivo para proteger el metal primario.
- Aplicar revestimientos protectores o barreras para minimizar el contacto con el electrolito.
- Seleccionar metales con propiedades electroquímicas similares al diseñar estructuras o componentes.

2.3.3. Corrosión termo galvánica Corrosión termo galvánica.

Este tipo de corrosión ocurre cuando los gradientes de temperatura crean diferencias del potencial electroquímico en la superficie del metal, llevando a una corrosión acelerada en zonas localizadas. La corrosión termo galvánica es prevenida por un diseño y medidas apropiadas, que

eviten un calentamiento o enfriamiento desigual, y la formación de puntos calientes. Es importante que el aislamiento sea continuo, en algunos casos esta forma de corrosión puede ser evitada mediante la protección catódica o revestimientos.

2.3.4. Corrosión por rendija (incluida la corrosión por depósito).

La corrosión por rendija o hendiduras se presenta cuando se forman estancamiento de soluciones en huecos estrechos o grietas lo que provoca una corrosión localizada.

2.3.5. Picadura, corrosión por picadura.

La corrosión por picadura involucra la formación de pequeñas picaduras o cráteres sobre la superficie del metal, debido a los ataques localizados provocada por agentes agresivos, tales como iones de cloruro.

Factores que afectan a la corrosión por picaduras:

Este tipo de corrosión es influenciada por múltiples factores, incluyendo el pH y la concentración de cloruros, la velocidad de flujo, la fuerza de gravedad, los iones Cu^{2+} y Fe^{3+} , las propiedades metalúrgicas del metal, la capacidad aislante del óxido, la rugosidad de la superficie, la temperatura, y el contacto galvánico con un material más noble (Einar Bardal, 2004).

2.3.6. Ataque selectivo, lixiviación selectiva (des aleación).

El ataque selectivo o des aleación, hace referencia a la corrosión preferencial de un elemento de la aleación, provocando una menor resistencia a la corrosión en la matriz del material.

2.3.7. Corrosión intergranular (incluida la exfoliación).

La corrosión intergranular ocurre a lo largo de los bordes de grano del metal, usualmente debido a impurezas o sensibilización, causando la debilitación del material. La exfoliación, es una forma de corrosión intergranular donde las capas del material se separan de la superficie.

2.3.8. Corrosión por erosión.

La corrosión por erosión se presenta cuando la combinación de los efectos de corrosión y el desgaste mecánico provocan una pérdida acelerada del material, lo que ocurre usualmente en ambientes de flujo de fluidos a alta velocidad.

2.3.9. Corrosión por cavitación.

La corrosión por cavitación es el resultado de la formación y el colapso de burbujas de vapor en un líquido, lo que provoca daños localizados en las superficies metálicas, normalmente en bombas o hélices.

2.3.10. Corrosión por frotamiento.

La corrosión por frotamiento se produce cuando dos superficies experimentan un movimiento relativo repetido de pequeña amplitud en presencia de ambientes corrosivos, lo que provoca una corrosión acelerada en la interfaz.

2.3.11. Agrietamiento por corrosión bajo esfuerzos.

El agrietamiento por corrosión bajo tensión es una forma de corrosión que se produce bajo tensión de tracción en un entorno corrosivo, lo que provoca la fractura frágil de un material.

2.3.12. Fatiga por corrosión.

La fatiga por corrosión es el efecto combinado de la carga cíclica y la corrosión, que provoca el inicio y la propagación de grietas y, en última instancia, el fallo. Algunos de los mecanismos globales que intervienen en este tipo de corrosión, incluyen:

- La formación de picaduras y surcos de corrosión, que dan lugar a concentradores de esfuerzos. La aceleración mutua de la corrosión, el deterioro mecánico, la absorción y adsorción de hidrogeno del ambiente y la fragilización consiguiente.

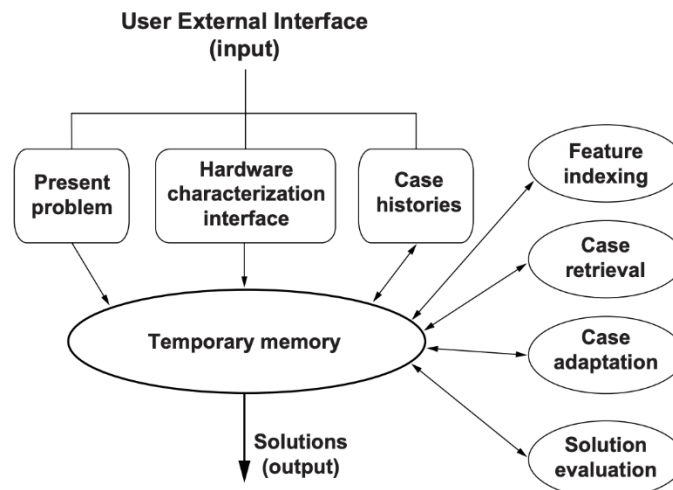
3. Inteligencia artificial aplicada a la corrosión

3.1. Contextualización

Las aplicaciones de la inteligencia artificial, en este caso en el campo de la corrosión, a pesar de que en la última década estas aplicaciones se globalizaron, desde de década de 1980, ya se encontraban estudios mediante el uso de computadores de inteligencia automatizada como destaca, Roberge, Pierre R. (1999), donde señala métodos de modelamiento, predicción y aplicaciones por medio de computer-based learning (CBL) y redes neuronales artificiales (ANN), que para la época representaron un gran avance tecnológico, más sin embargo, también presentaban un consumo de tiempo y costos altos debido a la lentitud de los procesadores y la necesidad de conocimientos especiales de computación y software por parte de los desarrolladores. A continuación, la **figura 2**, explica la arquitectura del razonamiento basada en los modelos de predicción por machine learning.

Figura 2

Arquitectura de razonamiento basado en casos para la predicción del comportamiento de materiales. (Reproducido de Roberge PR. *Handbook of Corrosion Engineering*. New York: McGraw-Hill, 1999.) p.323



3.2. Modelos actuales de algoritmos.

A continuación, se tratarán los principales modelos de aprendizaje automatizado, adicional, algunas de las definiciones de cada método faltante se encuentran en el **glosario**.

3.2.1. *Artificial Neural Networks (ANN)*.

Las redes neuronales artificiales (ANN, sus siglas en inglés) son herramientas que simulan el funcionamiento del cerebro humano, estas se basan en redes de procesos simples, o neuronas, las cuales se conectan y a medida que se le asignan tareas “aprenden” de la información suministrada. Las ANN tienen un gran impacto en la ingeniería metalúrgica, debido a que pueden ser aplicadas en los siguientes procesos:

Predecir el comportamiento de los materiales: Las ANN pueden analizar datos de laboratorio y de campo para predecir la vida útil de armados metálicos, la resistencia a la corrosión, la fatiga y entre otras propiedades.

Diseñar nuevos materiales: Las ANN pueden usarse para identificar nuevas combinaciones de elementos que podrían mejorar las propiedades de los materiales existentes o crear materiales con propiedades completamente nuevas, esto debido a su capacidad de aprendizaje continuo.

Optimizar procesos industriales: Las ANN pueden ayudar a controlar y optimizar procesos industriales como la fundición, la soldadura y el tratamiento térmico.

Desarrollar métodos de prevención de la corrosión: Las ANN pueden usarse para identificar los factores que contribuyen a la corrosión y desarrollar métodos para prevenirla o mitigarla.

Roberge, Pierre R. (1999) presenta las siguientes aplicaciones en las cuales las ANN han desempeñado un alto impacto:

Predicción del riesgo de agrietamiento por corrosión bajo tensión (SCC): Se han utilizado las ANN para analizar el riesgo de SCC en aceros inoxidable en función de la temperatura, la concentración de cloruros y la concentración de oxígeno.

Predicción de la corrosión a partir de barridos de polarización: Las ANN tiene aplicabilidad para predecir la aparición de corrosión general o localizada en diferentes materiales a partir de barridos de polarización potenciodinámica.

Modelado de la corrosión por CO₂: Se ha desarrollado un modelo de "peor caso" de corrosión por CO₂ basado en una ANN que puede predecir la tasa de corrosión en diferentes condiciones.

Predicción de la degradación de materiales de revestimiento: Las ANN permiten detectar la degradación de materiales de revestimiento o recubrimiento no metálicos en función de los resultados de pruebas de laboratorio.

Validación y extrapolación de datos de impedancia electroquímica: Las ANN tienen la capacidad para analizar datos de impedancia electroquímica y predecir la corrosión en condiciones diferentes a las de entrenamiento.

Esta capacidad quedó demostrada en el estudio de Aghaaminiha et al. (2021), en el que se utilizaron modelos de aprendizaje automático, incluidas las ANN, para modelar las velocidades de corrosión dependientes del tiempo del acero al carbono en presencia de inhibidores de la corrosión.

El estudio destacó cómo las ANN podían manejar las relaciones no lineales y los aspectos dependientes del tiempo de los procesos de corrosión, ofreciendo predicciones fiables incluso cuando se extrapolaban más allá de los escenarios de entrenamiento iniciales.

Las **figuras 3 y 4** explican cómo las redes neuronales artificiales (ANN) pueden predecir y monitorear la corrosión. **La figura 3** muestra cómo una neurona en una ANN transforma entradas en salidas mediante funciones de suma ponderada y activación. **La figura 4** amplía este concepto, demostrando cómo las ANN pueden predecir la corrosión utilizando datos electroquímicos complejos.

Figura 3

Esquema de un único procesador o neurona en una red neuronal artificial. (*Reproducido de Roberge PR. Handbook of Corrosion Engineering. New York: McGraw-Hill, 1999.*) p.323

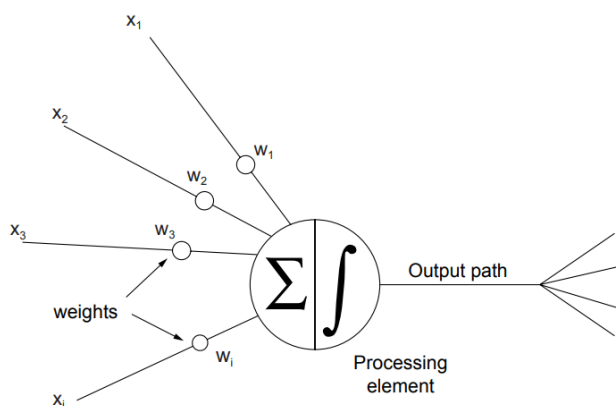


Figura 4

Inputs y Outputs de datos para predecir la corrosión a partir de barridos de polarización mediante una red neuronal artificial. (*Reproducido de Roberge PR. Handbook of Corrosion Engineering. New York: McGraw-Hill, 1999.*) p.323

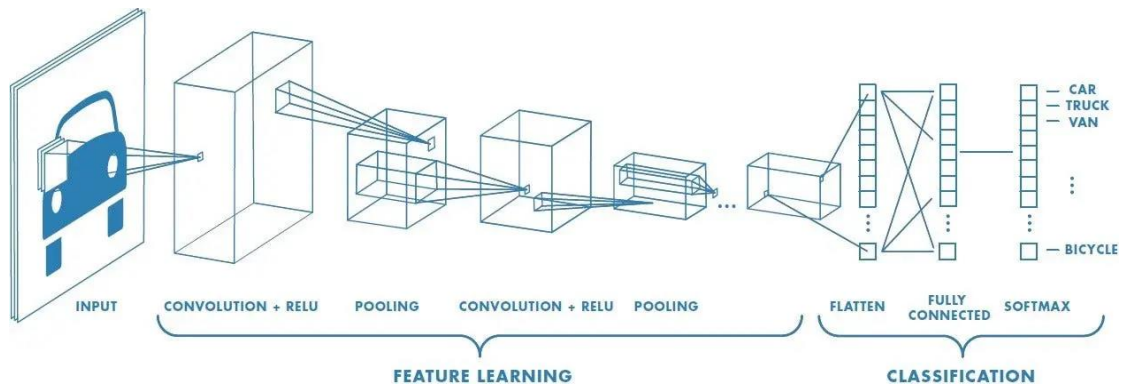
Input parameter	Value of feature
Prepassivation potential	$E_{\text{prot}} - E_{\text{corr}}$
Pitting potential	$E_{\text{pit}} - E_{\text{corr}}$
Hysteresis	+1 = positive 0 = none -1 = negative
Current density at scan reversal	$\mu\text{A}\cdot\text{cm}^{-2}$
Anodic nose	+1 = yes 0 = no
Passive current density	$\mu\text{A}\cdot\text{cm}^{-2}$
Potential at anodic-cathodic transition	$E_{\text{A to C}} - E_{\text{corr}}$
Output parameter	Value of feature
Crevice corrosion predicted	+1 = yes 0 = no
Pitting predicted	+1 = yes 0 = no
Should general corrosion be considered?	+1 = yes 0 = no

3.2.2. Convolutional neural networks (CNNs).

Las redes neuronales convolucionales (CNN) son un tipo de red neuronal artificial diseñada específicamente para procesar datos estructurados en forma de cuadrícula o malla, como las imágenes. Se utilizan ampliamente en tareas de visión por ordenador debido a su capacidad para aprender de forma automática y adaptativa, jerarquías espaciales de características a partir de imágenes de entrada. (Brandoli et al. 2021; Khalaf et al. 2024). Véase **figura 5**.

Figura 5

Arquitectura de razonamiento basado en redes neuronales convolucionales (CNN). (*Reproducido de Sumit Saha. A Guide to Convolutional Neural Networks — the ELI5 way. Data Science & ML. 2018*)



Las redes neuronales convolucionales se distinguen de otras redes neuronales por su rendimiento superior con entradas de señales de imagen, habla o audio. Tienen tres tipos principales de capas, que son:

- Capa convolucional (Convolutional layer)
- Capa de agrupamiento (Pooling layer)
- Capa totalmente conectada (Fully connected (FC) layer)

La capa convolucional es la primera capa de una red convolucional. Mientras que las capas convolucionales pueden ir seguidas de capas convolucionales adicionales o capas de agrupamiento, la capa totalmente conectada es la última capa. Con cada capa, la CNN aumenta su complejidad, identificando mayores porciones de la imagen. Las primeras capas se centran en características simples, como colores y bordes. A medida que los datos de la imagen avanzan por las capas de la CNN, ésta empieza a reconocer elementos o formas más grandes del objeto hasta que finalmente identifica el objeto deseado. (*What Are Convolutional Neural Networks?* / IBM, n.d.)

3.2.3. Recurrent neural networks (RNNs).

Las redes neuronales recurrentes (RNN) son un tipo de red neuronal artificial diseñada para manejar datos secuenciales manteniendo un estado interno o memoria. A diferencia de las redes neuronales feedforward, las RNN tienen conexiones que forman ciclos dirigidos, lo que les permite mostrar dinámicas temporales y captar patrones a lo largo del tiempo. Khalaf et al. (2024).

3.2.4. Long Short-Term Memory (LSTM).

La memoria larga a corto plazo (LSTM) es un tipo de arquitectura de red neuronal recurrente (RNN) diseñada para superar el problema del gradiente evanescente en las RNN tradicionales, lo que les permite capturar dependencias a largo plazo en datos secuenciales.

Las LSTM constan de células de memoria y puertas que controlan el flujo de información dentro de la red. Los componentes clave incluyen un estado de celda que transporta información a través de los pasos temporales, una “puerta de entrada” que regula el flujo de nueva información al estado de celda, una “puerta de olvido” que controla qué información descartar del estado de celda, y una “puerta de salida” que determina la información que debe salir en cada paso temporal. Las LSTM se utilizan ampliamente en tareas de predicción de series temporales, procesamiento

del lenguaje natural, reconocimiento del habla y otras aplicaciones de análisis secuencial de datos. Ji et al. (2024).

3.2.5. *Linear regression.*

“La regresión lineal es un método estadístico utilizado para modelar la relación entre una variable dependiente y una o más variables independientes ajustando una ecuación lineal a los datos observados” (Feng et al., 2023; Hazarika et al., 2021).

En el contexto de los artículos referenciados, la regresión lineal puede emplearse como técnica de modelado predictivo para analizar y predecir la vida a la fatiga por corrosión de materiales como el acero T91 y el acero inoxidable 316 (SS), o como “clasificador” para tareas de clasificación binaria utilizando un enfoque de regresión intuicionista difusa kernel ridge (Feng et al., 2023; Hazarika et al., 2021). El objetivo de este método es encontrar la línea recta (o hiperplano en dimensiones superiores) que mejor se ajuste y que minimice la suma de las diferencias al cuadrado entre los valores observados y predichos.

3.2.6. *Support vector regression (SVR).*

La regresión de vectores de soporte (SVR) es una técnica de aprendizaje automático utilizada para el análisis de regresión. SVR funciona encontrando el hiperplano óptimo que maximiza el margen entre los puntos de datos, al tiempo que permite que algunos puntos caigan dentro del margen o en el lado equivocado del hiperplano.

El objetivo es minimizar el error entre los valores predichos y los valores reales dentro de una tolerancia o ϵ especificada. SVR puede manejar relaciones lineales y no lineales entre variables utilizando diferentes funciones de núcleo para transformar el espacio de entrada en un

espacio de mayor dimensión (Khalaf et al., 2024; Ji & Ye, 2024; Feng et al., 2023; Hazarika et al., 2021).

3.2.7. *Decision Trees (DT)*.

Los árboles de decisión (DT), son una técnica de aprendizaje automático utilizada para tareas de clasificación y regresión. Funcionan particionando recursivamente el espacio de entrada en regiones basadas en valores de características, con el objetivo de minimizar la impureza o maximizar la homogeneidad dentro de cada región.

Cada partición está representada por un nodo de decisión en el árbol, que contiene una condición basada en un valor de característica. El árbol se ramifica en función del resultado de esta condición, dando lugar a nuevas particiones hasta llegar a los nodos hoja, que representan la decisión o predicción final. Los árboles de decisión son intuitivos, interpretables y capaces de manejar tanto datos categóricos como numéricos. A menudo se utilizan como modelos independientes o como parte de la construcción métodos en conjunto, como los Random Forests (RF). (Katuwal et al., 2018).

3.2.8. *Random Forests (RF)*.

Los bosques aleatorios (RF), son una técnica de aprendizaje por conjuntos basada en árboles de decisión que se utiliza principalmente para tareas de clasificación y regresión.

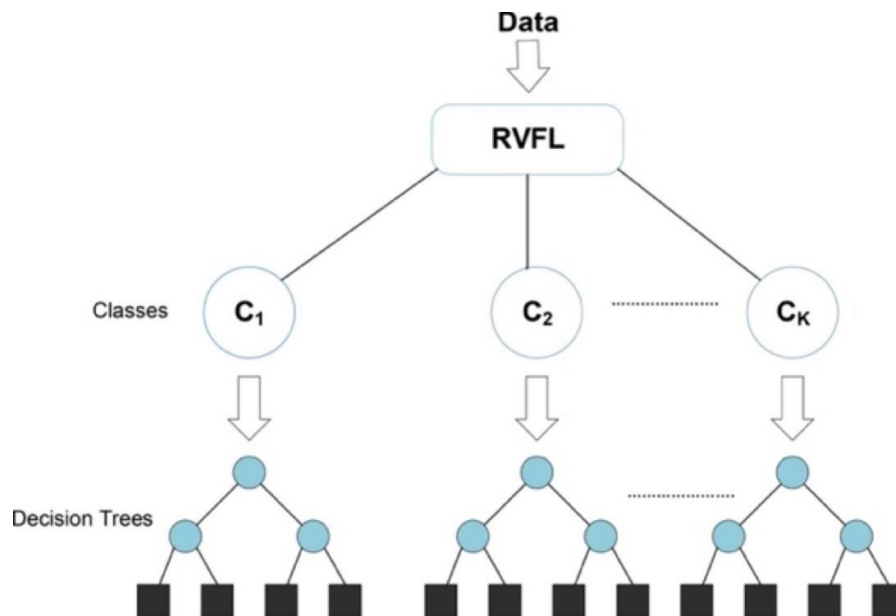
Funcionan construyendo una multitud de árboles de decisión durante el entrenamiento y emitiendo la predicción modal (clasificación) o media (regresión) de los árboles individuales.

Véase fig. 6

Cada “árbol” del “bosque” se entrena con un subconjunto aleatorio de los datos de entrenamiento y un subconjunto aleatorio de las características, lo que ayuda a reducir el sobreajuste y a mejorar el rendimiento de la generalización. Los bosques aleatorios son conocidos por su robustez, escalabilidad y capacidad para manejar eficazmente datos de alta dimensión y grandes conjuntos de datos. Se utilizan ampliamente en varios dominios, incluida la monitorización de la corrosión en la industria del petróleo y el gas (Khalaf et al., 2024), la predicción de las velocidades de corrosión en el acero (Ji & Ye, 2024), la predicción de la vida de fatiga por corrosión del acero (Feng et al., 2023), la detección de la corrosión del fuselaje de las aeronaves (Brandoli et al., 2021) y las tareas de clasificación multiclase (Katuwal et al., 2018).

Figura 6

Un clasificador base de método ensemble propuesto. (Katuwal, R., Suganthan, P. N., & Zhang, L. (2018). *An ensemble of decision trees with random vector functional link networks for multi-class classification.*)



3.2.9. Particle Swarm Optimization (PSO).

La Optimización por Enjambre de Partículas (PSO) es una técnica de optimización computacional inspirada en el comportamiento social de las bandadas de pájaros o los bancos de peces. En la PSO, una población de soluciones potenciales, llamadas partículas, se desplaza por el espacio de búsqueda y cada partícula ajusta su posición en función de su propia experiencia y de la de sus vecinas.

La posición de cada partícula representa una solución candidata al problema de optimización, y la calidad de la solución se evalúa mediante una función de aptitud. Mediante iteraciones, las partículas ajustan sus posiciones basándose en su posición individual mejor conocida (mejor posición local) y en la posición global mejor conocida entre todas las partículas de la población. Este proceso iterativo permite a PSO explorar eficazmente el espacio de búsqueda y converger hacia soluciones óptimas o casi óptimas. Ji et al. (2024).

3.2.10. Boosting algorithms.

Entre los algoritmos utilizados para potenciar los modelos del machine learning, se destacan, XGBoost y LightGBM, son algoritmos de refuerzo de gradiente utilizados habitualmente en el aprendizaje automático para tareas de modelado predictivo. se utilizan ampliamente para tareas de regresión y clasificación en diversos ámbitos debido a su rendimiento superior, escalabilidad y flexibilidad.

3.2.10.1. XGBoost. XGBoost, abreviatura de eXtreme Gradient Boosting, es una implementación eficiente y escalable de las máquinas de gradiente de refuerzo. Funciona construyendo secuencialmente un conjunto de árboles de decisión “débiles”, donde cada árbol subsiguiente corrige los errores cometidos por los anteriores. XGBoost emplea una función

objetivo regularizada para controlar la complejidad del modelo y evitar el sobreajuste. También utiliza un algoritmo novedoso para la división de árboles, que permite una construcción de árboles más rápida y precisa en comparación con los métodos tradicionales de refuerzo por gradiente. Chen, T., & Guestrin, C. (2016).

3.2.10.2. LightGBM. LightGBM (Light Gradient Boosting Machine) es otro marco de refuerzo de gradiente, en este caso desarrollado por Microsoft. Está diseñado para ser más rápido y eficiente en memoria que otras implementaciones de gradient boosting. LightGBM utiliza una novedosa técnica llamada Gradient-based One-Side Sampling (GOSS) para filtrar las instancias de datos con gradientes pequeños durante el proceso de entrenamiento, reduciendo así el coste computacional sin sacrificar el rendimiento del modelo. Además, LightGBM emplea un enfoque basado en histogramas para calcular los puntos de división, lo que mejora aún más la velocidad de entrenamiento. (¡Welcome to LightGBM's Documentation! — LightGBM 4.0.0 Documentation, n.d.)

4. Metodologías de la inteligencia artificial (IA) aplicadas.

4.1. Aplicación de CNN y RNN en el monitoreo de la corrosión por Imágenes.

En el contexto del monitoreo de la corrosión en la industria del petróleo y el gas, se encuentran estudios como los de Kumar, 2022 y Khalaf et al., 2024, donde las CNN pueden utilizarse para analizar imágenes obtenidas en procesos de inspección por medio de Ensayos no destructivos y técnicas C-scan UT-scan (véase **apéndice A**). Estas redes suelen constar de varias capas, incluidas capas convolucionales, capas de agrupamiento y capas totalmente conectadas.

En la fase de entrada, las CNN recogen imágenes de superficies corroídas (véase **apéndice B**) como datos de entrada. Las capas convolucionales extraen características de estas imágenes mediante la aplicación de filtros, que escanean la imagen de entrada para detectar patrones y características relevantes para la corrosión. El diagrama de flujo de la colección de imágenes y asistencia por ML se observa en la **figura 7**.

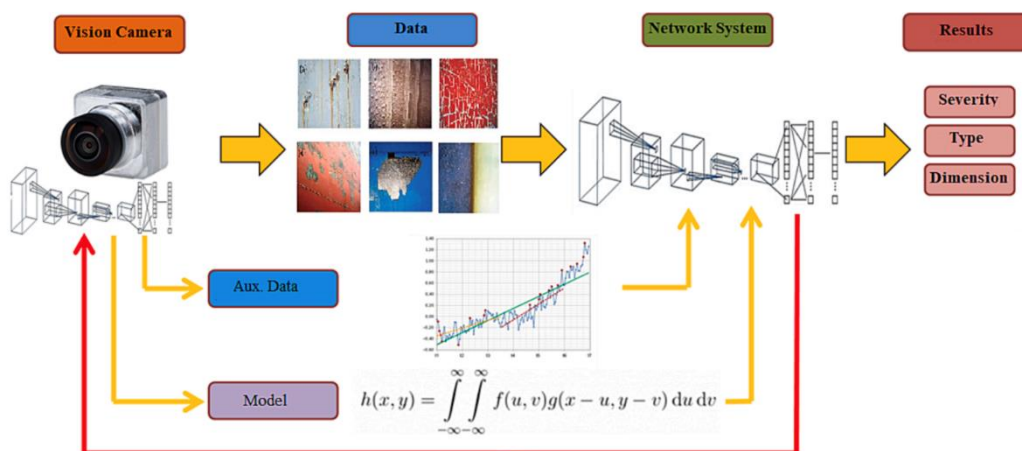
A continuación, las capas de agrupamiento reducen la dimensionalidad de los mapas de características generados por las capas convolucionales, resumiendo la información y conservando las características más importantes.

Por último, las capas totalmente conectadas interpretan las características extraídas y producen predicciones de salida, como la presencia o gravedad de la corrosión en zonas específicas de la imagen.

Al procesar los datos de las imágenes de esta manera, las CNN pueden ayudar a automatizar la detección y supervisión de la corrosión en las infraestructuras de petróleo y gas, proporcionando información valiosa para las tareas de mantenimiento e inspección.

Figura 7

Procedimiento de inspección visual potenciada con ML (*Reproducido de Khalaf et al. (2024). Emerging AI technologies for corrosion monitoring in oil and gas industry*)



En la industria aeronáutica, un estudio realizado por Brandoli et al. (2021), emplea las metodologías de CNN, donde los modelos se entrenan con datos de imágenes etiquetadas, en las que cada imagen se anota con la ubicación y la extensión de la corrosión, véase **figura 8**, permitiendo una detección eficaz y precisa de las regiones de corrosión en las imágenes del fuselaje de los aviones.

Figura 8

Un ejemplo que muestra la corrosión por pilotaje y las micro deformaciones superficiales. (a) Un especialista realiza una inspección visual del fuselaje con una linterna; (b) la foto escaneada del fuselaje muestra microperturbaciones superficiales causadas por la corrosión.

(reproducido de Brandoli et al., 2021 Aircraft Fuselage Corrosion Detection Using Artificial Intelligence)



Redes neuronales recurrentes (RNN):

En el contexto del monitoreo de la corrosión, (Khalaf et al., 2024), las RNN pueden utilizarse para analizar series temporales de datos relacionados con los índices de corrosión u otras variables ambientales. Estas redes son especialmente adecuadas para tareas en las que el orden de los datos de entrada es importante y en las que existen dependencias entre puntos de datos sucesivos. Adicional, en la aplicación de detección de la corrosión del fuselaje de aviones, realizada por Brandoli et al. (2021), las RNN podrían procesar datos secuenciales recogidos de sensores que controlan los niveles de corrosión a lo largo del tiempo.

En cada paso temporal, la RNN toma datos de entrada, como lecturas de sensores o condiciones ambientales, y actualiza su estado interno basándose tanto en la entrada actual como en la información almacenada de pasos temporales anteriores.

Esta naturaleza recurrente permite a las RNN capturar parámetros a largo plazo y patrones temporales en los datos, lo que las hace eficaces para tareas como la predicción de futuros índices de corrosión o la identificación de tendencias indicativas de la progresión de la corrosión.

4.2. Aplicación de la Memoria larga a corto plazo (LSTM) y modelo PSO-RF.

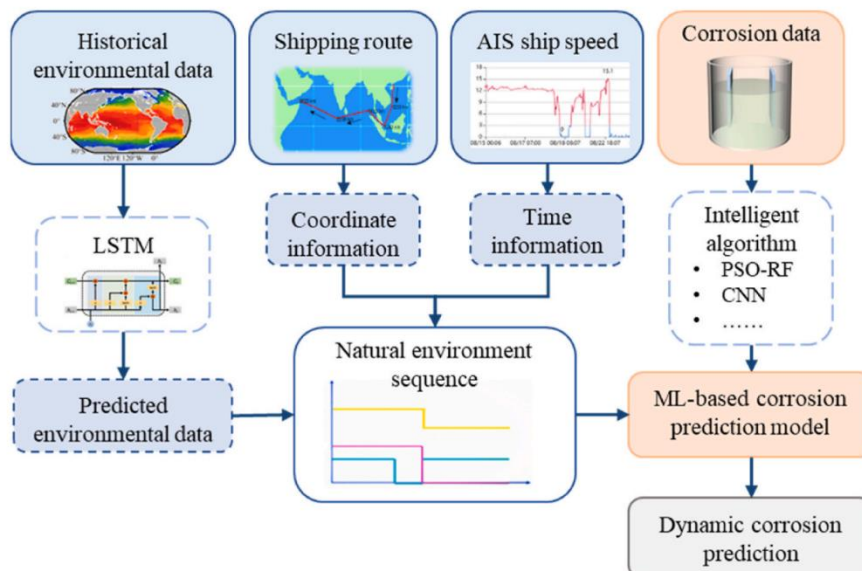
En 2024, una investigación realizada para la predicción del comportamiento frente a la corrosión de cascos de acero en entornos marinos dinámicos se realizó mediante la utilización conjunta de una red LSTM y un modelo PSO-RF, trabajadas conjuntamente. Ji et al. (2024).

El proceso detallado de esta investigación fue adaptado para abordar las complejidades de la predicción del comportamiento de la corrosión en entornos marítimos. Su proceso experimental se detalla en la **Figura 9**.

Figura 9

Diagrama del método propuesto de predicción dinámica de la corrosión para aceros de casco.

(Reproducido por Ji, H., Wang, H., Chen, Q., Ma, X., & Cai, Y. (2024). Corrosion behavior prediction for hull steels under dynamic marine environments by jointly utilizing LSTM network and PSO-RF model. *Ocean Engineering*.)



La metodología destaca por la realización de las siguientes fases:

- **Recopilación y preprocesamiento de datos:**

Recopilación de datos históricos relacionados con las rutas marítimas más empleadas, donde mediante estaciones de control (véase apéndice C y D), las cuales permitieron la medición del comportamiento de la corrosión y las condiciones ambientales en puntos estratégicos, incluidos factores como la temperatura, la humedad, la salinidad, el tiempo de exposición y los índices de corrosión. El preprocesamiento de los datos recopilados implica la limpieza, el filtrado y la normalización para garantizar la coherencia y la fiabilidad. Los valores perdidos pueden tratarse mediante técnicas de imputación, y los valores atípicos se identifican y tratan de manera específica.

- **Ingeniería de características:**

Identificación y selección de características relevantes del conjunto de datos que sean indicativas del comportamiento de la corrosión y de los factores ambientales. Esto puede implicar el conocimiento del dominio y el análisis estadístico para determinar las variables más informativas.

- **Construcción de modelos:**

Red LSTM: Construcción de una red LSTM (Long Short-Term Memory), un tipo de red neuronal recurrente (RNN), para capturar las dependencias temporales en los datos secuenciales de corrosión. La arquitectura LSTM permite al modelo retener la memoria a largo plazo y aprender patrones complejos a lo largo del tiempo.

Modelo PSO-RF: Desarrollo de un modelo predictivo utilizando una combinación de algoritmos de optimización de enjambre de partículas (PSO) y Random Forest (RF). PSO se utiliza para optimizar los hiperparámetros del modelo RF, como el número de árboles y la profundidad máxima, para mejorar su rendimiento predictivo.

- **Integración de modelos:**

Los resultados de la red LSTM y el modelo PSO-RF se combinan para proporcionar una predicción exhaustiva del comportamiento de la corrosión en cascos de acero expuestos a entornos marinos dinámicos. Se pueden emplear técnicas de conjuntos (ensemble learning) o promedios ponderados para integrar las predicciones de ambos modelos de forma efectiva.

- **Evaluación del modelo:**

Evaluación del rendimiento del modelo integrado utilizando métricas apropiadas como la exactitud, la precisión, la recuperación y la puntuación F1. Pueden utilizarse técnicas de validación cruzada para evaluar la solidez del modelo y su capacidad de generalización.

- **Validación y prueba:**

Validación de las capacidades predictivas del modelo mediante pruebas con datos no observados o del mundo real. Este paso tiene como objetivo verificar la fiabilidad y eficacia del modelo en escenarios prácticos de predicción de la corrosión marina.

Combinando los puntos fuertes de las redes LSTM para el modelado temporal y los modelos PSO-RF para la optimización predictiva, la metodología propuesta ofrece un enfoque global para la predicción de la corrosión en entornos marinos dinámicos.

4.3. Aplicación de la combinación de múltiples algoritmos de interés

4.3.1. Caso 1. Predicción de la corrosión de oleoductos y gasoductos.

La metodología descrita en el estudio "Evolution of corrosion prediction models for oil and gas pipelines: From empirical-driven to data-driven" abarca varios pasos destinados a desarrollar modelos sólidos de predicción de la corrosión mediante técnicas de aprendizaje automático. Wang et al. (2023), los métodos y parámetros de entrada y salida empleados se observan en la **tabla 1**.

Durante la fase experimental se realizó el estudio de varios parámetros de corrosión, mediante técnicas mecano-electroquímicas, de manera experimental y por simulación de elementos finitos, como se detalla en el **apéndice E**. las definiciones de cada método se encuentran en el **glosario**.

Tabla 1

Algunos métodos de ML empleados en la predicción de la corrosión, en especial en la industria del gas y petróleo. (adaptado de Wang et al. (2023))

Pre-procesamiento	Método	Parámetros de entrada	Parámetros de salida	Modelo de optimización (métrico)
Recolección de datos	SVR, SVR-GA, SVRPSO, SVR-FFA	Tiempo de exposición, resistividad, contenido de agua, sulfato, bicarbonato, cloruro, pH, potencial tubería/suelo, densidad aparente potencial redox, tipo de revestimiento	profundidad máxima del foso por corrosión	SVR-FFA (RMSE = 0.2909)
	POS-SSCN, DNN, GBM, RF	temperatura, presión de funcionamiento, tasa de producción de gas, tasa de producción de petróleo, tasa de producción de agua, sedimentos básicos y agua, presión parcial de CO ₂ , gravedad específica del gas temperatura, retención de líquido, presión, presión parcial de CO ₂ , velocidad del flujo de líquido, tensión de cizallamiento de la pared, tensión de cizallamiento de la pared, ph	profundidad del defecto de corrosión	POS-SSCN (MSE = 0.0238)
detección de valores atípicos, excluyendo o sustituyendo la información que falta de fuentes de terceros y selección de características si es necesario generar aleatoriamente parámetros de funcionamiento parámetros de funcionamiento	RF, GBRT, AdaBoost, XGBoost	temperatura, retención de líquido, presión presión parcial de CO ₂ , flujo de líquido líquido, tensión de cizallamiento de la pared pared, pH	velocidad de corrosión	XGBoost (RMSE = 0.0228)

Nota: Algunos de los modelos evaluados por el autor, están enunciados en esta tabla, con el fin de analizar los parámetros de entrada necesarios para cada algoritmo en función los parámetros de salida y las métricas de optimización de cada método.

Las fases metodológicas de este estudio se detallan a continuación:

- **Modelo de red neuronal artificial (ANN):**

Se emplearon modelos ANN para captar relaciones complejas y no lineales entre diversos parámetros de entrada y los índices de corrosión. Estos modelos aprovechan las similitudes estructurales y funcionales con las redes neuronales biológicas para aprender patrones de los datos y hacer predicciones. En la investigación, las ANN se entrenaron utilizando parámetros de entrada como el kilometraje, la diferencia de elevación, el ángulo de inclinación, la presión y el número de Reynolds de las tuberías para predecir las tasas de corrosión.

- **Modelo de máquina de vectores de apoyo (SVM):**

Los modelos SVM se utilizaron como algoritmos de aprendizaje supervisado para tareas de clasificación y regresión en la predicción de la corrosión. Las SVM destacan en el manejo de relaciones complejas y no lineales mediante el uso de funciones de núcleo para transformar los datos de entrada en un espacio de mayor dimensión para una separación o regresión eficaz. En este estudio, las SVM se aplicaron para establecer relaciones lineales entre los datos de entrada y de salida, lo que las hace adecuadas para predecir los índices de corrosión en función de diversos factores influyentes.

- **Técnicas de aprendizaje por conjuntos (EL):**

Se emplearon técnicas de aprendizaje por conjuntos (Ensemble learning), como random forest (RF), AdaBoost, gradient boosting regression tree (GBRT) y extreme gradient boosting (XGBoost), para combinar las predicciones de varios modelos y mejorar la precisión de la predicción. Estas técnicas aprovechan la diversidad de los modelos individuales para mitigar los errores y mejorar el rendimiento global.

En concreto, RF combina múltiples árboles de decisión para crear un modelo más robusto y preciso promediando sus predicciones. AdaBoost se centra en mejorar el rendimiento de los “aprendices” débiles mediante el entrenamiento secuencial de modelos en las muestras mal clasificadas de modelos anteriores. GBRT se encarga de entrenar secuencialmente árboles de decisión para minimizar los errores residuales de los modelos anteriores, mejorando gradualmente la precisión de las predicciones. Y, por último, XGBoost realiza una implementación optimizada de gradient boosting que aprovecha los algoritmos de descenso de gradiente para minimizar las funciones de pérdida y mejorar el rendimiento del modelo de forma iterativa.

En la investigación, estas técnicas de aprendizaje por conjuntos se aplicaron para desarrollar modelos de predicción de los índices de corrosión de oleoductos y gasoductos. Al combinar las predicciones de múltiples modelos, los modelos de conjunto lograron un rendimiento superior al de los modelos individuales, mejorando así la precisión y fiabilidad de las predicciones de corrosión.

En general, el estudio utilizó una combinación de RNN, SVM y técnicas de aprendizaje por conjuntos como RF, AdaBoost, GBRT y XGBoost para desarrollar modelos completos de predicción de la corrosión en oleoductos y gasoductos. Estos enfoques de aprendizaje automático permitieron a los investigadores captar relaciones complejas en los datos y realizar predicciones precisas, facilitando así una gestión eficaz de la corrosión y estrategias de mantenimiento en la industria del petróleo y el gas. Algunas de las formas de corrosión en ductos de aprecian en el

apéndice F

4.3.2. Caso 2. Predicción de velocidad de corrosión del acero al carbono

La investigación realizada por Aghaaminiha et al. (2021), “Machine learning modeling of time-dependent corrosion rates of carbon steel in presence of corrosion inhibitors”. Consistió en medir la velocidad de corrosión del acero dulce en soluciones acuosas de CO₂ utilizando la Resistencia de Polarización Lineal (LPR) en diversas condiciones. Los experimentos se llevaron a cabo en cuatro laboratorios con cupones de acero dulce consistentes y dos inhibidores orgánicos (CI-1 y CI-2). [15] Las condiciones variaron en presión de CO₂ (0,5 a 12 bar) y temperatura (80°C a 130°C), con algunos experimentos controlando el pH a 6. Se probaron un total de 25 condiciones, generando 26.855 puntos de datos. Los inhibidores se adicionaron en diferentes dosis y horarios, y algunos experimentos incluyeron un periodo de pre-corrosión. Los datos mostraron que la dosificación secuencial de inhibidores reducía significativamente los índices de corrosión, alcanzando una eficacia de hasta el 99,4%.

Los pasos tomados para el procesamiento de los datos mediante la implementación de machine learning, se enuncian a continuación:

Paso 1: Preprocesamiento de datos

El primer paso consistió en asegurarse de que el conjunto de datos no contenía valores omitidos o incorrectos. Los puntos de datos con valores perdidos se eliminaron o se sustituyeron por la media, la moda o un valor aproximado. Si una variable tenía muchos valores perdidos, se eliminaba del conjunto de datos. Las variables se reescalaron, normalmente a una distribución normal estándar, excepto el tiempo. Las características categóricas se convirtieron en variables ficticias para adaptarlas a los modelos de aprendizaje automático. El preprocesamiento adicional incluyó el ajuste de tiempo a cero en cada momento de inyección del inhibidor de corrosión y la

adición de nuevas características como la concentración previa del inhibidor de corrosión, la velocidad de corrosión inicial y el tipo de prueba realizada.

Paso 2: Ajuste de hiperparámetros

Para optimizar los modelos de aprendizaje automático (ANN, RF, SVR, KNN), se llevó a cabo una búsqueda en cuadrícula sobre un rango de valores de hiperparámetros. El objetivo de este proceso era encontrar el conjunto de hiperparámetros que diera como resultado el menor error cuadrático medio (MSE). Se utilizó un método de validación cruzada quíntuple, en el que los datos se dividieron en cinco subconjuntos. El modelo se entrenó en cuatro subconjuntos y se probó en el quinto, y este procedimiento se repitió cinco veces. Se calculó el MSE medio para identificar los mejores hiperparámetros. Para los modelos de RNA y RF, este proceso implicó 20 iteraciones de validación cruzada quíntuple.

Paso 3: Encontrar el modelo ML óptimo

Una vez ajustados los hiperparámetros, el siguiente paso consistió en seleccionar el mejor modelo de aprendizaje automático. El conjunto de datos se dividió en un conjunto de entrenamiento (75%) y un conjunto de prueba (25%). Los modelos se entrenaron en el conjunto de entrenamiento y su rendimiento se evaluó en el conjunto de pruebas. El modelo óptimo fue el que obtuvo el error cuadrático medio (MSE) más bajo en el conjunto de pruebas. Este método garantizaba que el modelo podía generalizarse bien a datos no observados. Véase **apéndice G**.

Paso 4: Entrenamiento del modelo ML óptimo y comprobación de su rendimiento.

Para la aplicación práctica, el modelo ML óptimo fue entrenado para predecir experimentos completos de tasas de corrosión a lo largo del tiempo, dadas las condiciones ambientales y operativas específicas y los programas de dosificación de inhibidores de corrosión. A diferencia del paso anterior, esta vez el modelo se entrenó utilizando datos de 21 de 25 experimentos, con los

cuatro experimentos restantes completamente ocultos durante el entrenamiento para formar el conjunto de pruebas. Esta estrategia proporcionó una prueba más rigurosa de la capacidad de predicción del modelo. Si un experimento se repetía varias veces y se seleccionaba para el conjunto de pruebas, todas sus réplicas se excluían de los datos de entrenamiento.

Análisis de sensibilidad

El modelo ML entrenado también se utilizó para analizar cómo los cambios en las características de entrada afectaban al comportamiento del sistema. Se realizaron predicciones de las tasas de corrosión variables en el tiempo en función de las variaciones en el tipo de inhibidor, la concentración de inhibidor, la presión parcial de CO₂, la temperatura, el esfuerzo cortante de la pared y el tipo de salmuera. Este análisis permitió comprender mejor la influencia de los distintos factores en los índices de corrosión y ayudó a validar la solidez del modelo.

5. Análisis de resultados y limitaciones de los modelos IA en corrosión.

Los estudios aquí presentados y referencias pertinentes utilizadas permiten realizar un análisis del cual se destacan, el rendimiento, éxito de resultados obtenidos, empleabilidad y, por otro lado, limitaciones y propuestas a futuro.

5.1 Resultados de los métodos en la detección de corrosión por Imágenes.

Los modelos de Machine learning aplicados en detección de la corrosión localizada, realizados por Khalaf et al. (2024), analizan factores como las propiedades de los materiales, las condiciones ambientales y los factores geométricos para predecir y comprender el inicio y la progresión de la corrosión. Estos modelos, que aprovechan algoritmos de IA y enfoques basados en datos, proporcionan información sobre la probabilidad y la gravedad de la corrosión localizada.

El modelado por elementos finitos una vez obtenidos los resultados de laboratorio (véase **apéndice H**), mejoran la fiabilidad de estas predicciones simulando la corrosión por picaduras y teniendo en cuenta la distribución de probabilidad de las picaduras de corrosión. La integración de este enfoque con redes neuronales artificiales (ANN) ayuda a captar la complejidad de la corrosión por picaduras, mejorando la precisión de las predicciones.

Por otro lado, los resultados obtenidos, en el estudio realizado por Brandoli et al. (2021), para el monitoreo de la corrosión en fuselajes aeronáuticos, presentaron un método de redes neuronales convolucionales (CNN) fácil de utilizar para detectar la corrosión en fuselajes de aeronaves utilizando imágenes del sistema de inspección de aeronaves D-Sight (DAIS). (véase **apéndice I**). A diferencia de las inspecciones manuales tradicionales, que pueden llevar mucho tiempo y ser propensas a errores, este método CNN automatizado mejora tanto la precisión como la eficiencia.

Al probar varias configuraciones de redes neuronales, descubrieron que el uso de modelos pre-entrenados como InceptionV3 y DenseNet [8] arrojaba unas impresionantes precisiones del 90,2% y el 92,2%, respectivamente.

A pesar de contar con un conjunto de datos relativamente pequeño debido a los elevados costes de adquisición de imágenes y etiquetado manual, los resultados fueron muy prometedores.

Los investigadores sugieren que aumentar la cantidad de datos etiquetados podría ayudar al modelo a generalizar mejor entre distintos fabricantes de aviones, más allá de Boeing y Airbus.

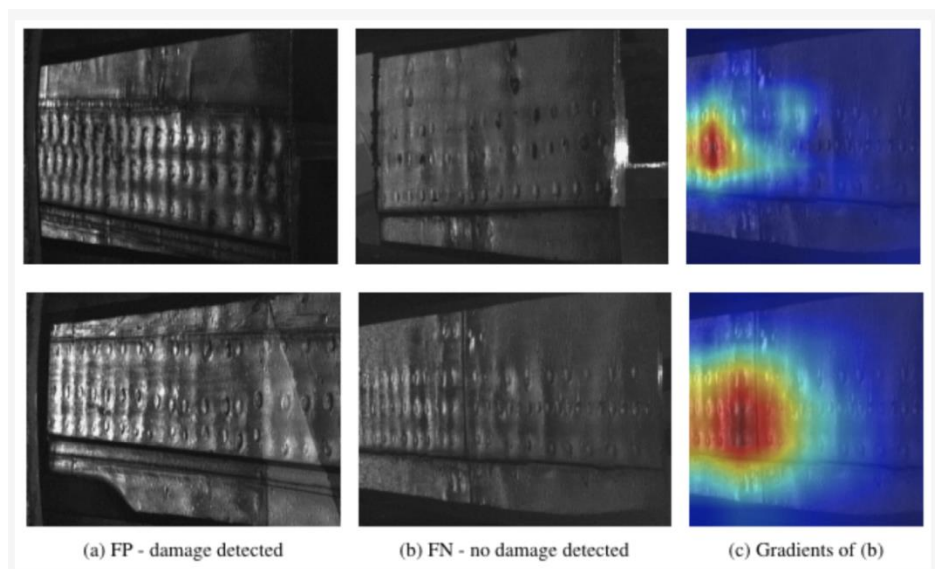
Se estima que, en investigaciones posteriores, se podría incluir la adaptación del modelo a diferentes estructuras y materiales de aeronaves, como la variación del grosor del fuselaje en diferentes modelos de Boeing y helicópteros militares. Algunas de las imágenes obtenidas mediante este método se aprecian en el **apéndice J**.

Este estudio muestra un gran potencial para el uso de redes neuronales profundas (DNN) en la detección de la corrosión en aviones y helicópteros militares, basándose en los resultados positivos obtenidos con datos de aviones comerciales.

La **figura 10**, muestra los resultados obtenidos mediante el dispositivo DenseNet-20, empleado para la captura de imágenes. El aprendizaje profundo (Deep learning), puede transformar significativamente las prácticas de mantenimiento de la aviación. En este estudio, se evaluó el rendimiento de una red neuronal profunda y, concretamente, se descubrió que DenseNet, con 201 capas, era la que mejor funcionaba. Al examinar visualmente los resultados de clasificación errónea, se identificaron casos de falsos positivos (Figura 10a) en los que el modelo predijo incorrectamente la corrosión, y falsos negativos (Figura 10b) en los que no detectó la corrosión real. (*Brandoli et al., 2021*).

Figura 10

Ejemplos de imágenes que la arquitectura de aprendizaje profundo DenseNet-201 no bien predicho. La columna (a) corresponde a imágenes mal clasificadas con corrosión o falsos positivos (FP), mientras que la columna (b) incluye imágenes con corrosión mal clasificada detectada o falsos negativos (FN). En (c), utilizando Grad-CAM para la interpretación visual de los gradientes de DenseNet-201 sobre la imagen de (b). (*Brandoli et al., 2021*)



Adicional, se logra determinar que las imágenes de baja resolución pueden afectar negativamente a los resultados, un problema que se observa en muchos sistemas de IA y visión por computadores, cuando se encuentran con muestras desconocidas o atípicas. Este problema puede reducirse incluyendo muestras de baja resolución durante el entrenamiento del algoritmo, para mejorar la eficacia del modelo frente a nuevas entradas. Sin embargo, los dispositivos DAIS [8] utilizados en este estudio capturan imágenes de alta resolución, lo que ayuda a mantener la precisión.

Al concluir su estudio, los investigadores destacaron los prometedores resultados. Comprobaron que la precisión de sus clasificaciones, sobre todo con la arquitectura óptima, la cual se acercaba a la realizada por creación humana. No obstante, se reconoce la posibilidad de que se produjeran errores humanos, sobre todo cuando las imágenes presentaban grandes similitudes, como las zonas con manchas de corrosión en remaches y juntas. Esta investigación supuso el primer paso hacia la detección automatizada de la corrosión mediante métodos basados en datos. La capacidad de igualar el rendimiento humano en la identificación de la corrosión en imágenes sugiere una vía prometedora para los avances en Inteligencia Artificial.

5.2 Resultados de la aplicación de memoria larga a corto plazo (LSTM) y modelo PSO-RF.

Al analizar los resultados de Ji et al. (2024) se observa un enfoque novedoso que utiliza redes neuronales agrupadas en subespacios (SSCN) y optimización por enjambre de partículas (PSO), el cual estima el crecimiento de la profundidad del defecto de corrosión (CDD) en tuberías. Este método modela el crecimiento de CDD en función del tiempo, ayudando a la gestión de la integridad de las tuberías.

El modelo de optimización de enjambre de partículas y Random Forrest (PSO-RF) demostró una precisión predictiva superior entre los seis modelos evaluados en el conjunto de

datos de entrenamiento, superando a los demás en términos de R^2 y MAE. El modelo PSO-RF también mostró una gran precisión de ajuste, solo superado por el modelo de red neuronal CNN. Esto concuerda con hallazgos anteriores que destacan el gran rendimiento predictivo del algoritmo RF para la corrosión de los cascos de acero. Sin embargo, la precisión de ajuste del modelo PSO-RF, aunque robusta, no es muy superior a la de otros algoritmos que predicen la corrosión en los cascos de acero, basándose en factores medioambientales. Esto se debe a la inclusión de proporciones de elementos de aleación en el modelo, lo que mejora su generalización, permitiendo predicciones a través de varios tipos de acero al carbono en entornos de agua de mar mediante la introducción de la composición de la aleación.

El modelo PSO-RF bien entrenado puede predecir las pérdidas por corrosión del acero al carbono a lo largo de diferentes rutas marítimas y durante varias estaciones del año. Sin embargo, su precisión predictiva depende en gran medida de la calidad de las secuencias ambientales. El modelo se limita a predecir dentro del rango de condiciones probadas en el laboratorio, lo que indica que su principal limitación es su incapacidad para extrapolar más allá de estas condiciones.

Un ejemplo de aplicación del modelo PSO-RF consiste en calcular la secuencia de la corriente de corrosión del acero al carbono a lo largo de una ruta marítima específica. Las diferencias estacionales y geográficas de los factores medioambientales, en particular los niveles de oxígeno disuelto influyen significativamente en las velocidades de corrosión, ya que las mayores concentraciones de oxígeno disuelto producen una corrosión más grave. Los índices de corrosión obtenidos para cada ruta en diferentes estaciones se detallan en el **apéndice K**.

5.3 Resultados de la variación de múltiples métodos de ML

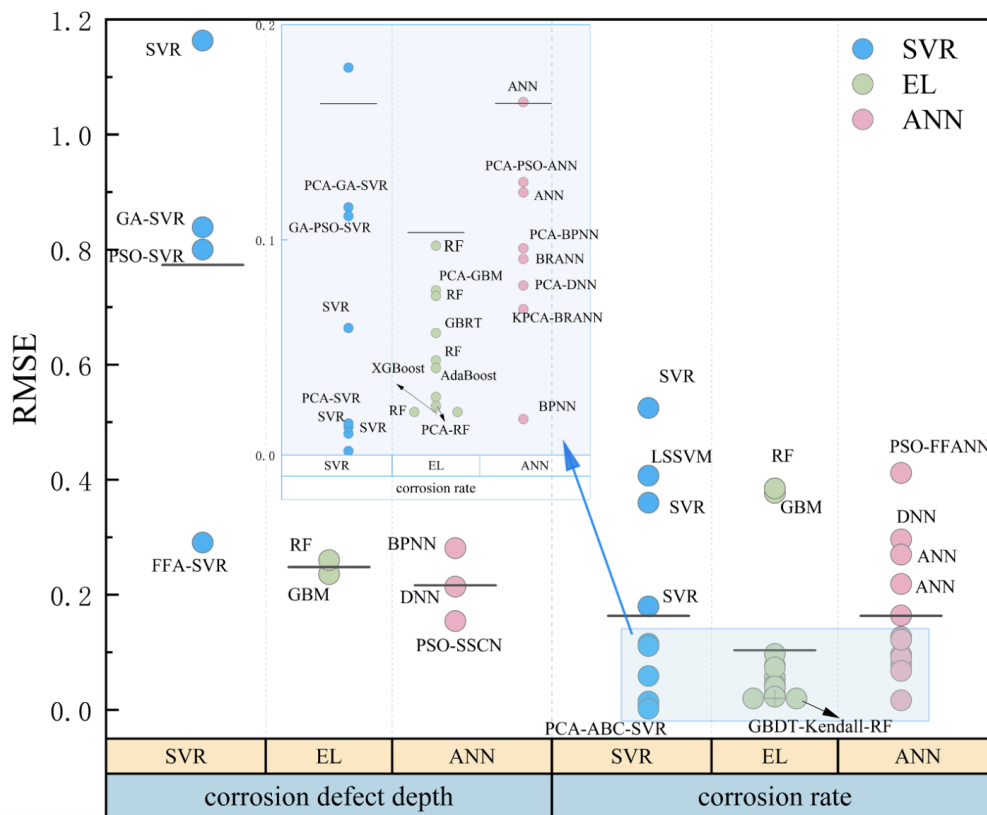
A continuación, en la **figura 11**, se puede detallar una comparativa de rendimientos, de diversos métodos de machine learning, empleados en el estudio realizado por Wang et al. (2023)

Donde se logró obtener resultados en el cálculo de la profundidad del defecto de corrosión, e índice del misma, mediante tres principales modelos base: Support vector regression (SVR), Ensemble learning (EL) y redes neuronales (ANN).

Figura 11

Rendimiento de los modelos ML expresado por el RMSE (Los nombres de los modelos se incluyeron junto a los puntos respectivos, "--" representa el valor RMSE medio de los modelos, El punto superpuesto se desplazará, "+" indica la posición original).

(Reproducido Por Wang et al. (2023) Evolution of corrosion prediction models for oil and gas pipelines: From empirical-driven to data-driven. Engineering Failure Analysis.)



El rendimiento de los distintos modelos varía en función del problema a estudiar y de las características de los datos. Si se analizan las redes neuronales artificiales (ANN) destacan con grandes conjuntos de datos, pero pueden tener problemas con los más pequeños. Los modelos Random Forest (RF) manejan variables continuas y discontinuas sin necesidad de normalizar los datos, lo que los hace versátiles para los problemas de regresión, por otro lado, los modelos SVM (Support Vector Machine) destacan por su velocidad de cálculo y su eficacia cuando se selecciona la función kernel [14] adecuada en función de las propiedades de los datos.

Analizando los demás modelos de redes neuronales, mientras que las CNN destacan en el procesamiento de datos espaciales como las imágenes, las RNN son idóneas para analizar datos secuenciales, lo que las convierte en herramientas valiosas para la supervisión y predicción de la corrosión en diversos contextos industriales.

En algunos casos, la combinación de varios modelos produce mejores resultados que el uso de un único modelo. Por ejemplo, la combinación de RF con el análisis de correlación de Spearman para optimizar las características, seguida del uso de estas características en un modelo SVM, puede reducir los errores, un método conocido como RF-SVM. Las ANN son especialmente adecuadas para crear y combinar múltiples modelos con el fin de lograr los resultados deseados.

En resumen, aprovechar los puntos fuertes de diferentes modelos y combinarlos en enfoques híbridos, como se expone en la investigación realizada por Wang et al. (2023), puede mejorar la precisión y el rendimiento de la predicción, sobre todo en escenarios complejos en los que intervienen diversas características de los datos.

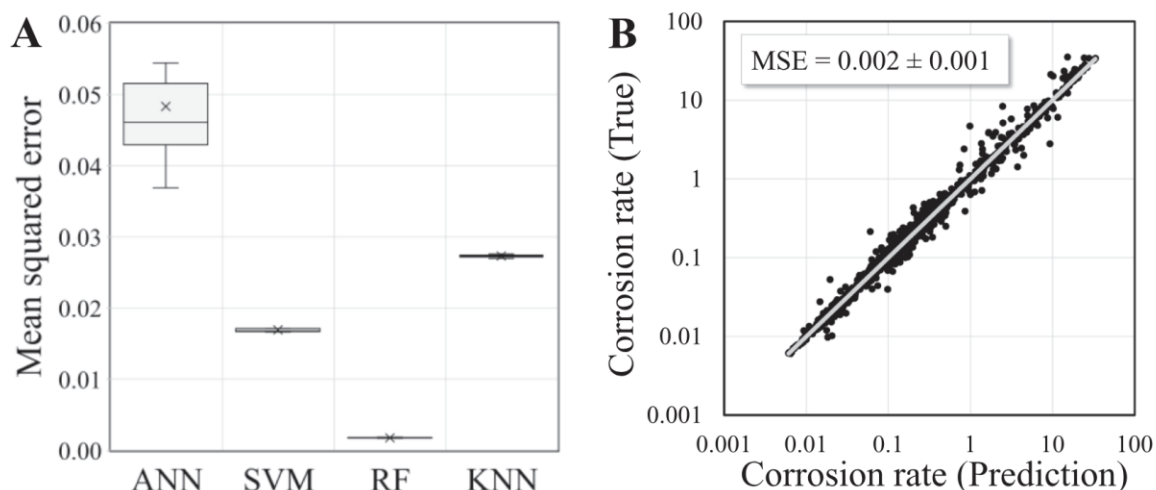
El estudio realizado por Aghaaminiha et al. (2021), evaluó el rendimiento de varios algoritmos de aprendizaje automático (RNA, SVR, RF y KNN) ajustando sus hiperparámetros y realizando una validación cruzada de cinco veces. En el caso de la ANN, la arquitectura óptima resultó ser de 4 capas ocultas con 8 nodos por capa, ya que el error cuadrático medio (ECM) disminuía al aumentar el número de nodos hasta ese punto. Para SVR, el mejor rendimiento se obtuvo con un parámetro de coste de 1000 y un coeficiente de núcleo (γ) de 1, ya que valores más altos de γ mejoraban significativamente el rendimiento. El modelo RF funcionó mejor con 500 árboles de decisión y una fracción máxima de variables de entrada (mFF) fijada en 0,7. El modelo KNN funcionó de forma óptima con K fijado en 3, y se obtuvieron mejores resultados cuando se dio más peso a los vecinos más cercanos. Las simulaciones de las velocidades de corrosión por ML respecto a la obtenida experimental para cada parámetro se detalla en el apéndice M.

El modelo Random Forest (RF) fue identificado como el mejor algoritmo, alcanzando el error cuadrático medio (MSE) más bajo de $0,002 \pm 0,001$. Este modelo fue seleccionado para futuras investigaciones de las velocidades de corrosión a lo largo del tiempo.

Estos resultados se compararon mediante un gráfico de paridad (véase **figura 12**) de las predicciones del modelo RF frente a las tasas de corrosión reales demostró una alta precisión en la predicción de los resultados experimentales, confirmando su eficacia en el modelado de los datos de la tasa de corrosión. RF fue utilizado para el análisis de la velocidad de corrosión bajo diferentes parámetros e inhibidores utilizados en este estudio, los resultados se observan en el **apéndice N**, donde se detallan los índices de corrosión para cada uno de los montajes y la variación de parámetros correspondiente, destacando la presencia de CO₂, cambios de presión, esfuerzos de tensión, y tipos de salmuera.

Figura 12

A) Comparación del rendimiento del mejor modelo para cada modelo ML. Las barras de error obtenidas realizando 20 iteraciones independientes de cada modelo. B) Gráfico de paridad entre las predichas (mm/a) por el modelo RF y los valores experimentales. El valor MSE indicado es la media de 20 iteraciones independientes de entrenamiento seguidas de pruebas en el conjunto de pruebas del modelo RF. pruebas en el conjunto de pruebas del modelo RF. En ambas figuras, las predicciones se realizan conjunto de pruebas, seleccionado aleatoriamente como el 25 de los 26.855 puntos de datos disponibles. *Aghaaminiha et al. (2021)*



A continuación, en la **tabla 2**, se realiza un análisis comparativo de algunos de los principales grupos de métodos de machine learning, donde se detallan los sensores más comúnmente usados, para el cálculo de cada parámetro relacionado con el monitorio de la corrosión que se desea investigar.

Adicional, se detalla el análisis sistemático del método y los factores limitantes que estos presentan, a modo de mejora o recomendaciones.

Tabla 2

Análisis comparativo de algunos algoritmos de inteligencia artificial aplicados en el monitoreo de la corrosión. (Adaptado de 2024. Khalaf et al. *Emerging AI technologies for corrosion monitoring in oil and gas industry: A comprehensive review. Engineering Failure Analysis*)

Algoritmo	Aplicación	Modelo específico de ejemplo	Sensores de	Análisis sistemático	Factores funcionales
Modelo de Regresión	Predicción de velocidades de corrosión	Redes neuronales, ANN, arboles de decisión (DT)	Electroquímicos	-Limitado cuando los datos son complejos o no lineales - Sensibilidad a los valores atípicos	- Fácil de aplicar - Predicción del índice de corrosión - Bueno para relaciones lineales - Puede no manejar bien las no linealidades - Supone una relación Lineal -Puede carecer de precisión para datos de corrosión complejos
Modelos de clasificación	Identificación de tipos de corrosión	Máquinas de vectores soporte (SVM), Redes neuronales, Random Forest (RF)	Resistencia eléctrica	-Limitado cuando se trata con más de dos clases de Corrosión	- Bueno para la clasificación - Evaluación de riesgos - Interpretable- Limitado a binarios o problemas multiclase -Eficaz para distinguir entre dos o más tipos de corrosión.
Aprendizaje profundizado (DP)	Reconocimiento de patrones complejos	Redes neuronales convolucionales (CNN), Redes neuronales recurrentes (RNN), Redes Generativas Adversariales (GAN)	Emisión acústica, inspección óptica, captura de imágenes en alta resolución.	-Requiere un cuidadoso modelo de selección y optimización	- Puede manejar patrones Complejos. - Eficaz para datos de imágenes, requiere grandes conjuntos de datos. - Puede necesitar importantes recursos informáticos. -Extremadamente potente para reconocimiento complejos.

Nota: En la tabla 1 se abordaron los principales grupos de algoritmos empleados en el ML mediante un análisis comparativo.

6. Conclusiones

La presente revisión bibliográfica, enfocada en las aplicaciones de la inteligencia artificial en el ámbito de la corrosión de metales, más en específico, los modelos de machine learning, y sus subcampos, permitieron desarrollar una profundización, de los distintos algoritmos comúnmente usados, por distintos investigadores, a nivel mundial, destacando la identificación de las principales metodologías utilizadas, resultados a obtener y permitió analizar las limitaciones y oportunidades de mejora de los distintos algoritmos.

Las diversas estrategias, mezcla de modelos de algoritmos y trabajo en paralelo con múltiples investigadores y ambientes o tipos de corrosión, permiten a los modelos de machine learning, la mejora constante y ampliar su rango de aplicaciones, lo que genera una cercanía de resultados con altos estándares de calidad, a los obtenidos de manera experimental, y favoreciendo las optimizaciones de técnicas de inspección y clasificación del fenómeno de la corrosión, donde se destacan la cercanía de cálculos de índice de corrosión, el monitoreo de superficies corroídas mediante imágenes en alta definición, potenciadas por medio de inteligencia supervisada y modelos aleatorios, arrojando proyecciones y tratamiento de datos muy eficaz y acertado.

Adicional, se logra concluir, que los modelos de machine learning, han logrado un avance significativo, en la última década, destacando las redes neuronales, modelos de decisión aleatoria y los potenciadores o Boosters, como modelos con alta aplicación de optimización en cuanto a resultados de una manera asertiva y eficaz, más sin embargo, se observan que modelos derivados de los Support Vector Machine, SVM, aun cuentan con limitantes respecto al número de tipos de corrosión que pueden procesar por cargue de datos, y análisis de salida de los mismos, los cuales son un factor clave en la mejora continua de las tecnologías emergentes.

7. Recomendaciones

Las investigaciones futuras, en la detección mediante la aplicación de machine learning, en el campo de la corrosión, deben dar prioridad a varias áreas clave. En primer lugar, el desarrollo de técnicas avanzadas de integración de datos para combinar información de diversas fuentes, puesto que, mejorará la calidad de los datos utilizados para entrenar modelos de IA.

En segundo lugar, la exploración de métodos de aprendizaje profundo (Deep learning), como las redes neuronales convolucionales y recurrentes, junto con el aprendizaje por transferencia, esto puede mejorar la detección de patrones con alta complejidad, en los datos de análisis de corrosión.

En tercer lugar, la integración de nuevos modelos basados en la física con algoritmos de IA generaría la mejora en la interpretabilidad y la precisión de las predicciones, especialmente en situaciones de escasez de datos.

Referencias Bibliográficas

- [1]: R. Winston Revie. Herbert H. Uhlig. (2008). CORROSION AND CORROSION CONTROL An Introduction to Corrosion Science and Engineering. Cuarta Edición. John Wiley & Sons.
- [2]: Einar Bardal. (2004). Corrosion and Protection. Springer.
- [3]: Roberge, Pierre R. (1999). Handbook of Corrosion Engineering. McGraw-Hill Education.
- [4]: Fontana, Mars G., and Roger W. Staehle. Advances in Corrosion Science and Technology. Springer Science & Business Media, 2012.
- [5]: Young, Donald J. High Temperature Corrosion and Materials Applications. Wiley, 1987.
- [6]: ASM Handbook, Volume 13A: Corrosion: Fundamentals, Testing, and Protection. ASM International, 2003.
- [7]: Khalaf, A. H., Xiao, Y., Xu, N., Wu, B., Li, H., Lin, B., Nie, Z., & Tang, J. (2024). Emerging AI technologies for corrosion monitoring in oil and gas industry: A comprehensive review. *Engineering Failure Analysis*, 155, 107735. <https://doi.org/10.1016/j.engfailanal.2023.107735>
- [8]: Brandoli, Bruno, André R. de Geus, Jefferson R. Souza, Gabriel Spadon, Amilcar Soares, Jose F. Rodrigues, Jr., Jerzy Komorowski, and Stan Matwin. (2021). "Aircraft Fuselage Corrosion

Detection Using Artificial Intelligence" Sensors 21, no. 12: 4026.
<https://doi.org/10.3390/s21124026>

[9]: Ji, H., Wang, H., Chen, Q., Ma, X., & Cai, Y. (2024). Corrosion behavior prediction for hull steels under dynamic marine environments by jointly utilizing LSTM network and PSO-RF model. *Ocean Engineering*, 300, 117371. <https://doi.org/10.1016/j.oceaneng.2024.117371>

[10]: Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. arXiv 2017. arXiv preprint arXiv:1610.02357, 8.

[11]: Kumar reddy, G.R., 2022. Hybrid model for detection of corrosion in water pipeline images using CNN and comparing accuracy with SVM. *ECS Trans.* 107
<https://doi.org/10.1149/10701.13861ecst>.

[12]: What are Convolutional Neural Networks? | IBM. (n.d).
<https://www.ibm.com/topics/convolutional-neural-networks>

[13]: Katuwal, R., Suganthan, P. N., & Zhang, L. (2018). An ensemble of decision trees with random vector functional link networks for multi-class classification. *Applied Soft Computing (Print)*, 70, 1146–1153. <https://doi.org/10.1016/j.asoc.2017.09.020>

[14]: Wang, Q., Song, Y., Zhang, X., Dong, L., Xi, Y., Zeng, D., Liu, Q., Zhang, H., Zhang, Z., Yan, R., & Luo, H. (2023). Evolution of corrosion prediction models for oil and gas pipelines: From empirical-driven to data-driven. *Engineering Failure Analysis*, 146, 107097. <https://doi.org/10.1016/j.engfailanal.2023.107097>

[15]: M. Aghaaminiha, R. Mehrani, M. Colahan, B. Brown, M. Singer, S. Nesic, S.M. Vargas, S. Sharma, Machine learning modeling of time-dependent corrosion rates of carbon steel in presence of corrosion inhibitors, *Corros. Sci.* 193 (2021), 109904, <https://doi.org/10.1016/j.corsci.2021.109904>.

[16]: X. Li, L. Zhang, F. Khan, Z. Han, A data-driven corrosion prediction model to support digitization of subsea operations, *Process Saf. Environ. Protect.* 153 (2021) 413–421, <https://doi.org/10.1016/j.psep.2021.07.031>.

[17]: Welcome to LightGBM's documentation! — LightGBM 4.0.0 documentation. (n.d.). <https://lightgbm.readthedocs.io/en/stable/#>

[18]: Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). <https://doi.org/10.1145/2939672.2939785>

[19]: Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural*

Information Processing Systems (Vol. 30). <http://papers.nips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree>

[20]: H. Ji and H. Ye. (2024). Machine learning prediction of corrosion rate of steel in carbonated cementitious mortars. *Engineering Failure Analysis* 155 107735

[21]: Feng, S., Sun, X., Chen, G., Wu, H., & Chen, X. (2023). LBE corrosion fatigue life prediction of T91 steel and 316 SS using machine learning method assisted by symbol regression. *International Journal of Fatigue*, 177, 107962. <https://doi.org/10.1016/j.ijfatigue.2023.107962>

[22]: Hazarika, B. B., Gupta, D., & Borah, P. (2021). An intuitionistic fuzzy kernel ridge regression classifier for binary classification. *Applied Soft Computing*, 112, 107816. <https://doi.org/10.1016/j.asoc.2021.107816>

[23]: Seghier, M. E. a. B., Keshtegar, B., Taleb-Berrouane, M., Abbassi, R., & Nguyen-Thoi, T. (2021). Advanced intelligence frameworks for predicting maximum pitting corrosion depth in oil and gas pipelines. *Process Safety and Environmental Protection*, 147, 818–833. <https://doi.org/10.1016/j.psep.2021.01.008>

[24]: Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*.

[25]: Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

[26]: Chatterjee, A., & Bandopadhyay, S. (2012). Optimization of machining parameters using genetic algorithm and support vector regression model. *International Journal of Computational Intelligence Systems*, 5(5), 990-1001.

[27]: Zhang, H., Sun, J., Hu, H., & Xu, X. (2014). A hybrid firefly algorithm-support vector regression model for annual electricity demand forecasting. *Energy Conversion and Management*, 84, 459-468.

[28]: Shi, Y., & Wang, X. (2017). Parallel optimization algorithms for semi-supervised clustering. *Information Sciences*, 382-383, 192-211.

[29]: Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.

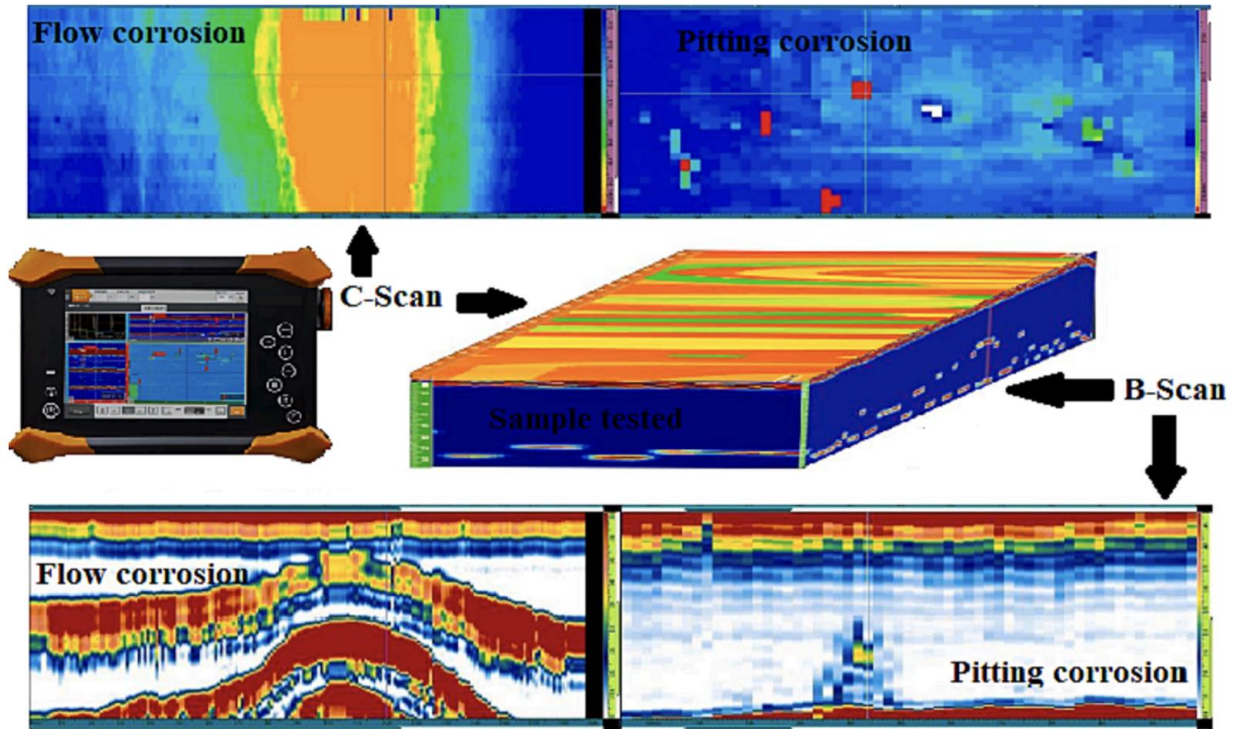
[30]: Srivastava, V., Basu, B. N., & Prabhu, N. (2024). Application of machine learning (ML)-based multi-Classifications to identify corrosion fatigue cracking phenomena in naval steel weldments. *Materials Today Communications*, 108591.

<https://doi.org/10.1016/j.mtcomm.2024.108591>

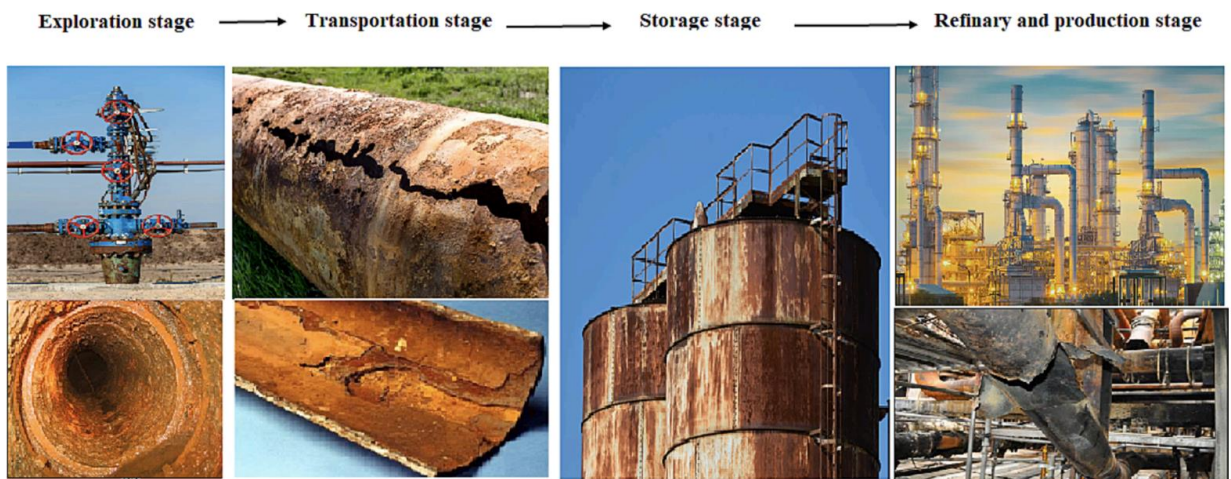
[31]: Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, 7, 21.

Apéndices

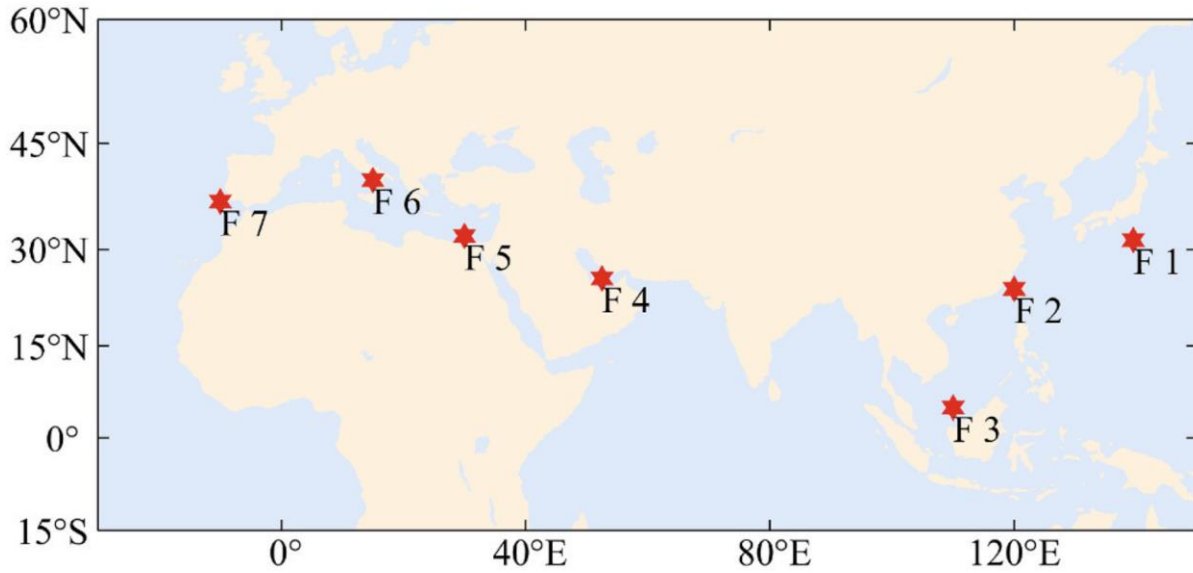
A. C-Scan Imaging para muestras seleccionadas (Parte superior C-scan, parte inferior es B-scan). [7]



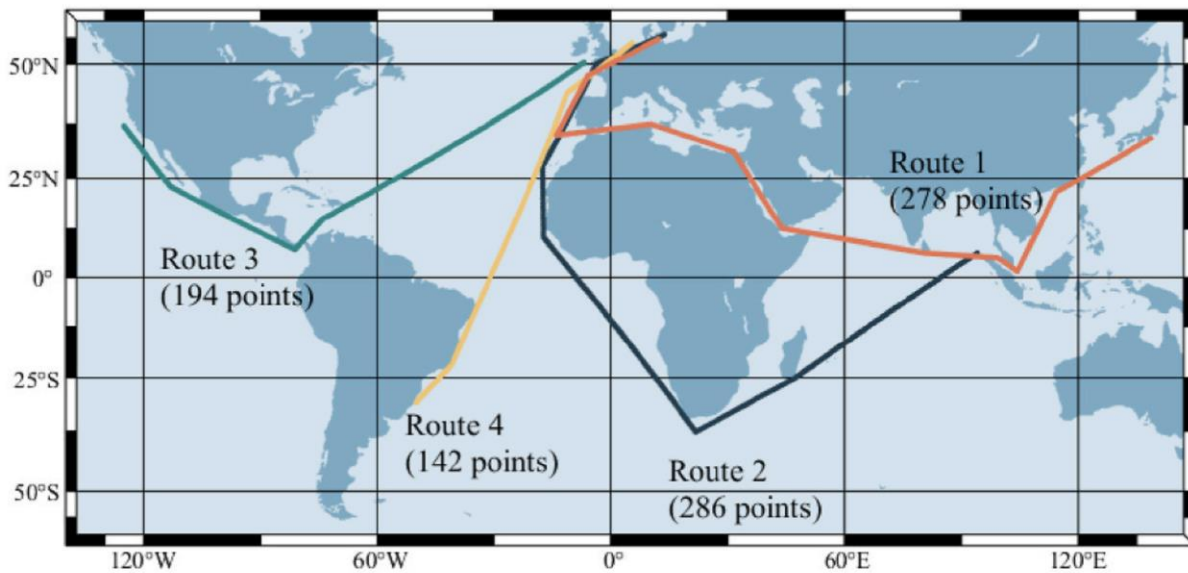
B. Diagrama esquemático de la corrosión en la industria del petróleo y el gas. (Khalaf et al., 2024) [7]



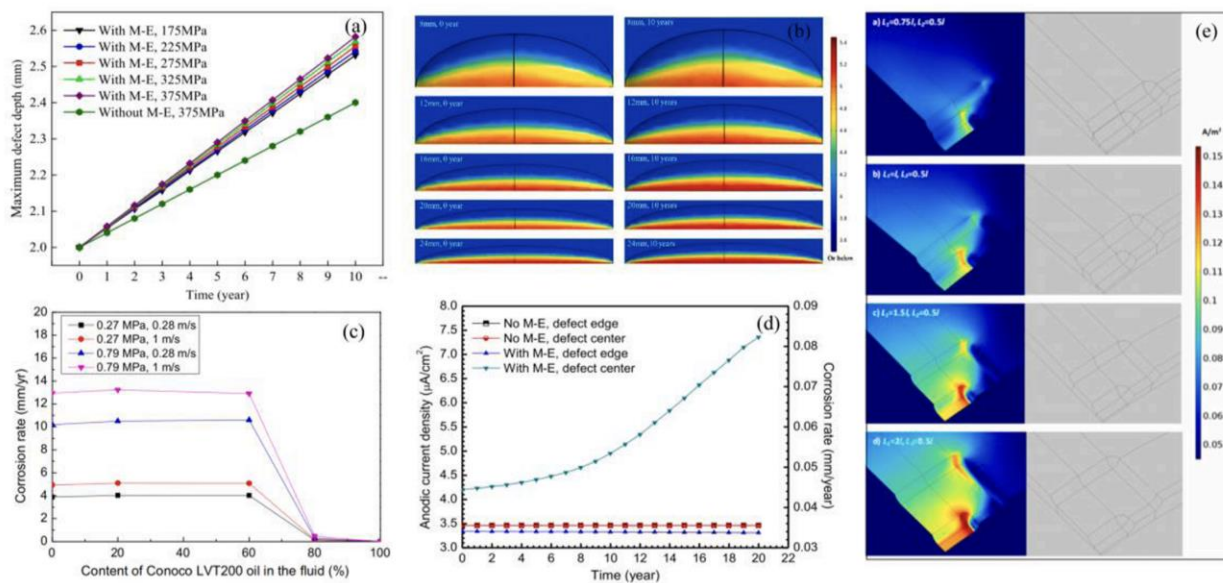
C. Localización geográfica de los 7 puertos marítimos. Ji et al. (2024).



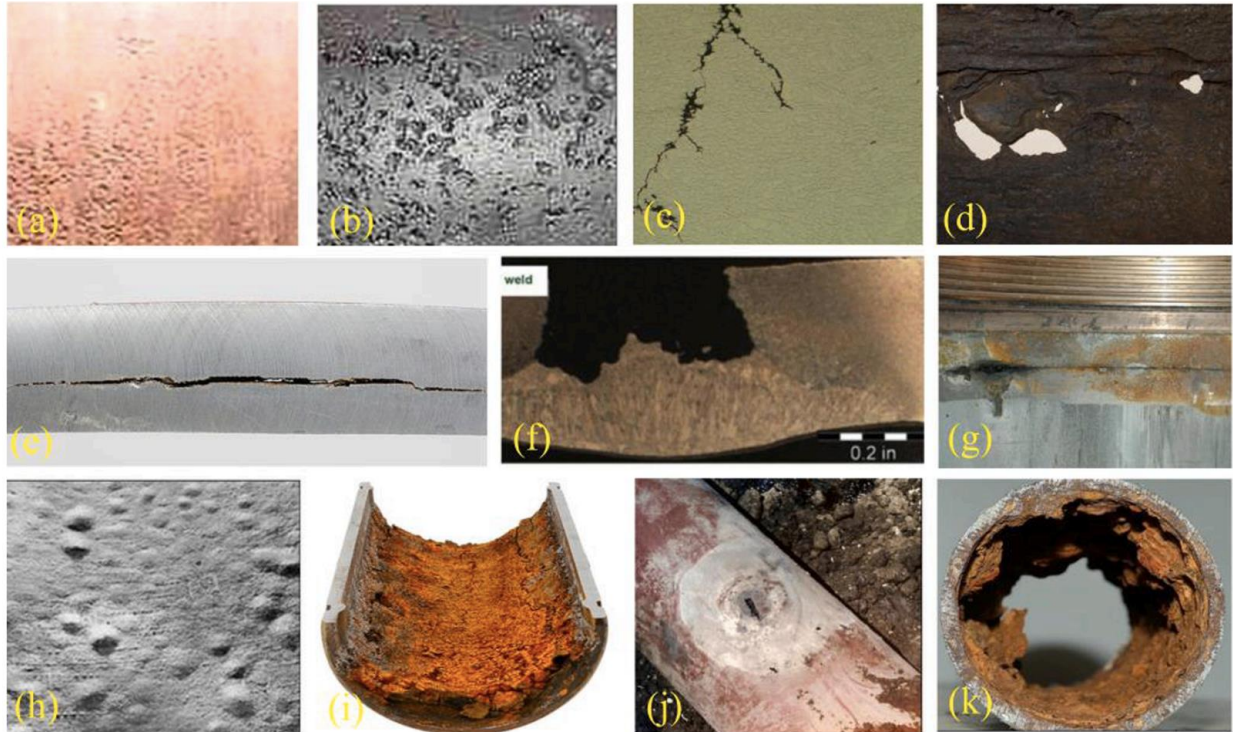
D. Mapa de las 4 principales rutas marítimas mundiales. Ji et al. (2024).



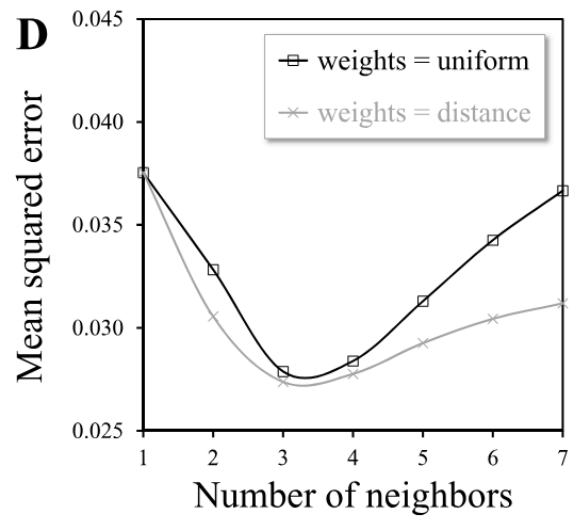
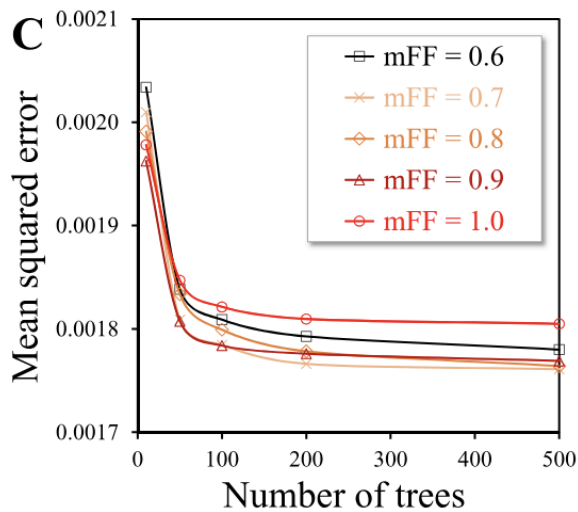
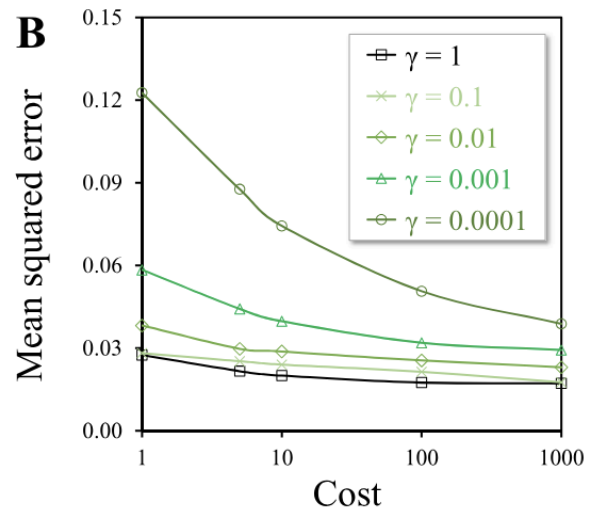
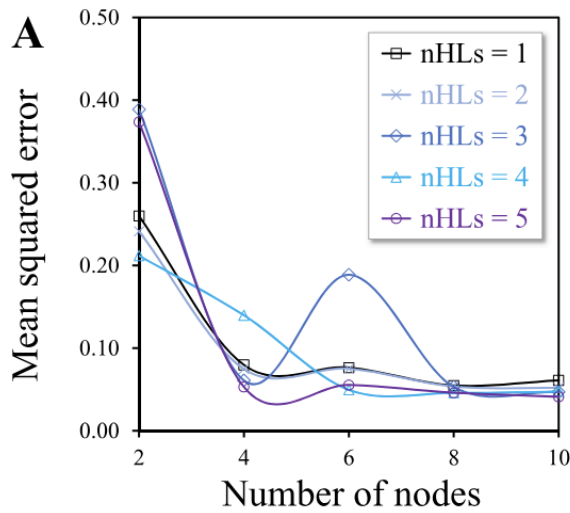
E. Simulación numérica del crecimiento de defectos de corrosión. (a) Variación de la profundidad máxima de los defectos a diferentes tensiones axiales de tracción y presiones internas fijas de 18 MPa, y con o sin considerar las interacciones mecano electroquímicas. presiones internas fijas de 18 MPa, y con o sin considerar las interacciones mecano-electroquímicas. (b) Densidad de corriente anódica en el defecto de corrosión con distintas longitudes iniciales del defecto en un segmento de tubo de acero X100 sometido a una tensión interna fija de 20 MPa y a una tensión de tracción axial de 375 MPa inicialmente y tras 10 años de servicio. (c) Velocidad de corrosión modelada en función del contenido de aceite Conoco LVT200 en el fluido a dos presiones parciales de CO₂ y dos velocidades de flujo, respectivamente. (d) Dependencia temporal de la densidad de corriente anódica (o velocidad de corrosión) en el borde y el centro de un defecto de corrosión con una presión parcial de 20 bar. un defecto de corrosión con una presión operativa de 20 MPa en función del tiempo. (e) Distribuciones de la densidad de corriente anódica en los defectos de corrosión solapados con longitudes variadas del defecto de la capa superior bajo la longitud fija del defecto de la capa inferior de 0,5 l (es decir, 33,87 mm) y la presión interna de 17,65 MPa. de 17,65 MPa. (reproducido de Wang et al. 2023).



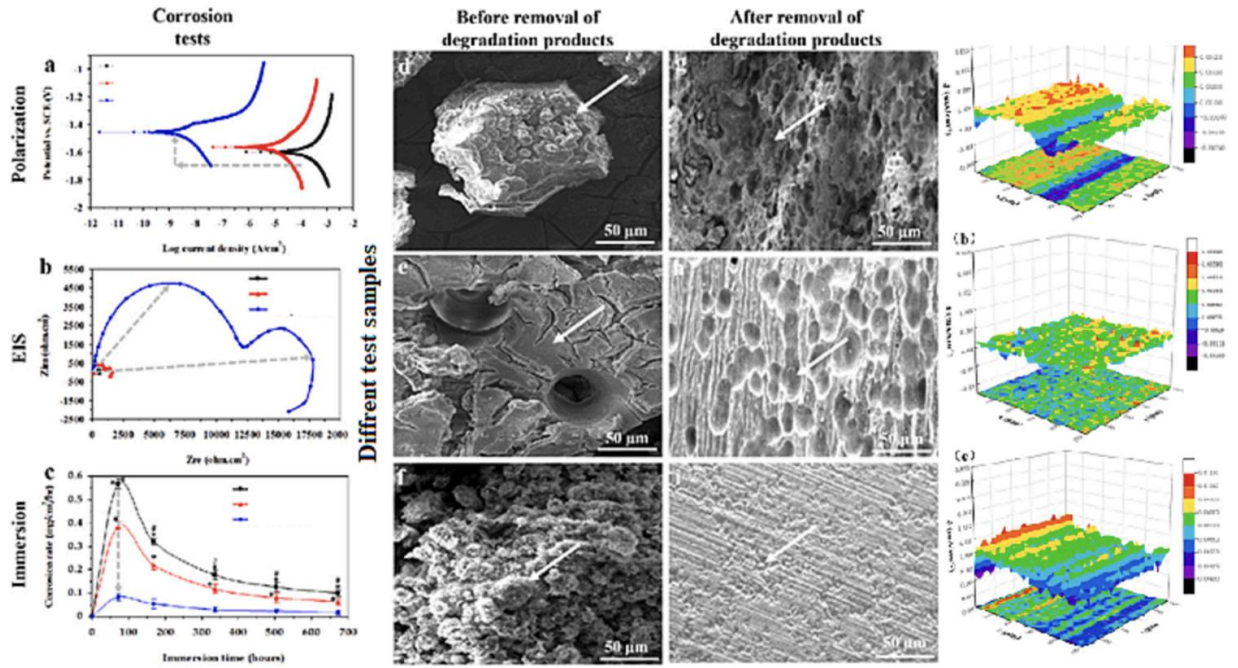
F. Las vistas macro de las formas características de corrosión de oleoductos y gasoductos. (a) Corrosión uniforme. (b) picaduras. (c) SCC. (d) ataque de masa. (e) HIC. (f) PWC. (g) corrosión en grietas. (h) HB. (i) UDC. (j) Daños en el revestimiento - corrosión local por grietas. (k) MIC. (reproducido de Wang et al. 2023).



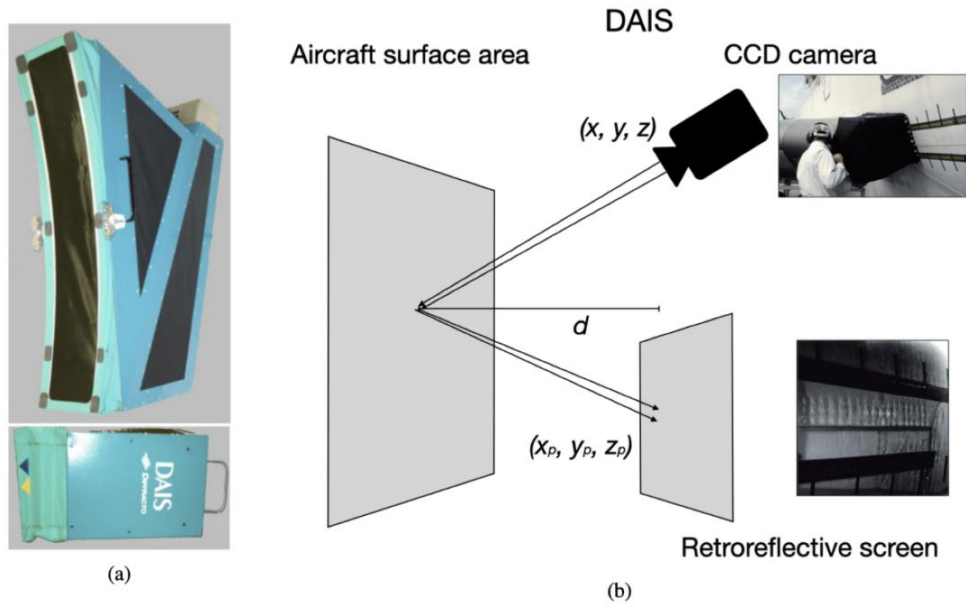
G. Rendimiento de A) Red neuronal artificial (ANN), B) Máquina de vectores de apoyo Regresión (SVR), C) Random Forest (RF), y D) K Nearest Neighbors (KNN) en datos experimentales de la tasa de corrosión para diferentes valores de los hiperparámetros. Cada punto es la media de 20 iteraciones independientes de entrenamiento seguidas de pruebas en el conjunto de pruebas. Las líneas son guías para facilitar la lectura. Aghaaminiha et al. (2021)



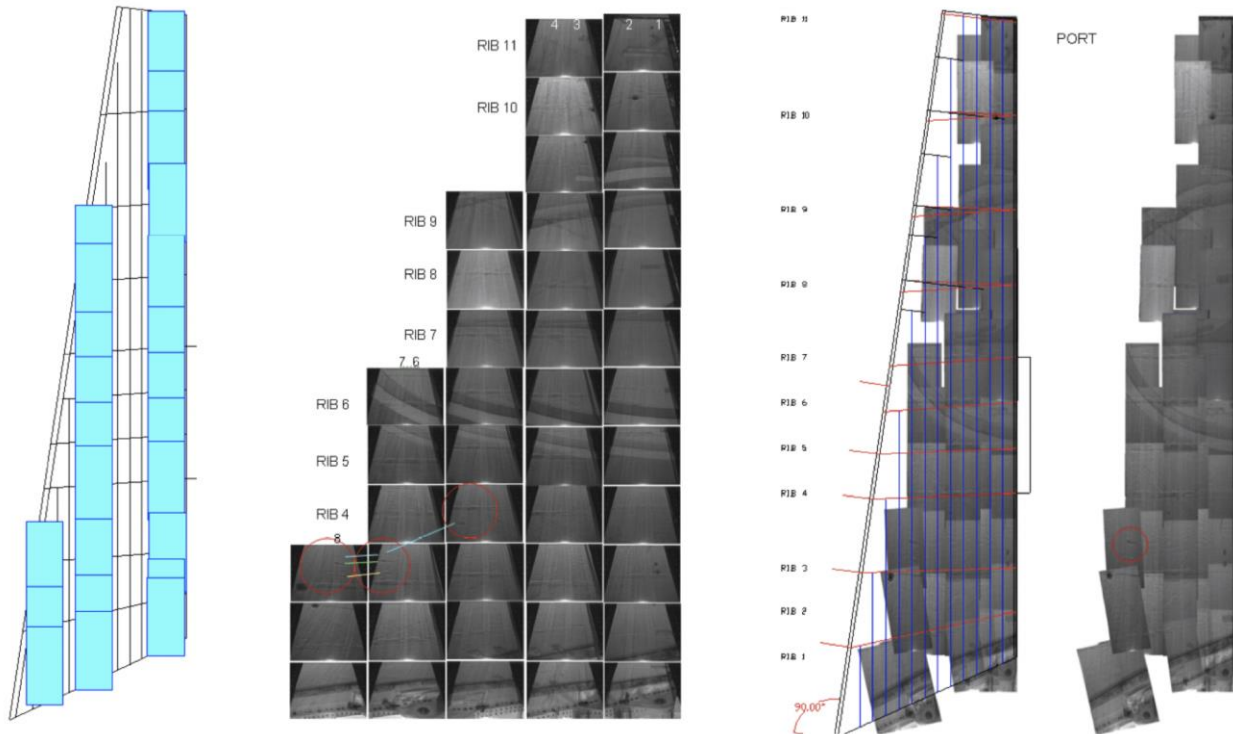
H. Imágenes de análisis electroquímico (Lab. y simulación). (Khalaf et al., 2024)



I. D-Sight Aircraft Inspection System (DAIS). (a) DAIS 250c, and (b) schematic of DAIS functioning. [8] (Brandoli et al., 2021)

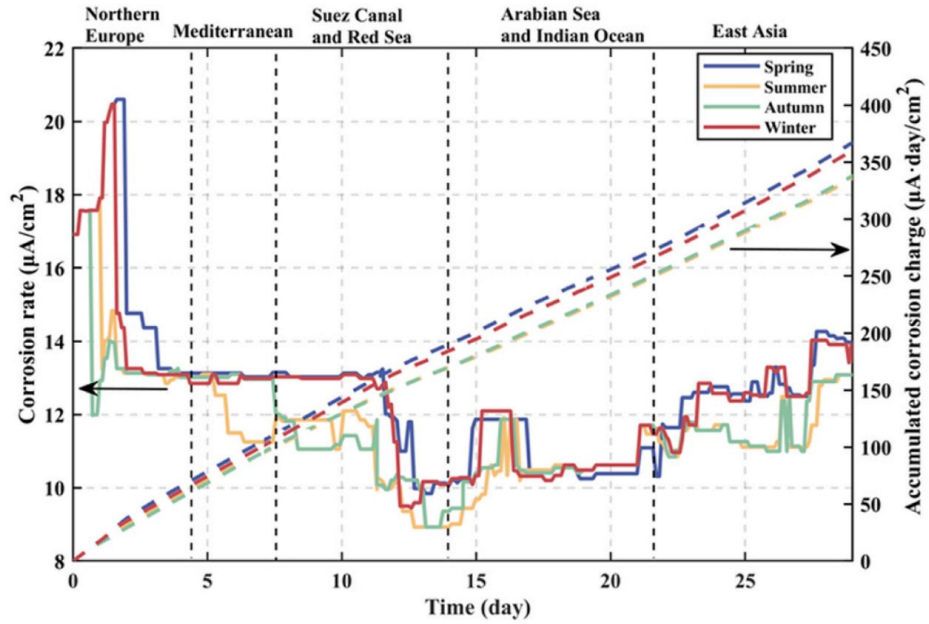


J. Imágenes de inspección del ala creadas con DAIS 250C. Las imágenes se componen como un mosaico para cubrir la totalidad del ala. Cada muestra de imagen se enumera para controlar la inspección. A continuación, un especialista realiza la inspección de cada muestra. [8]

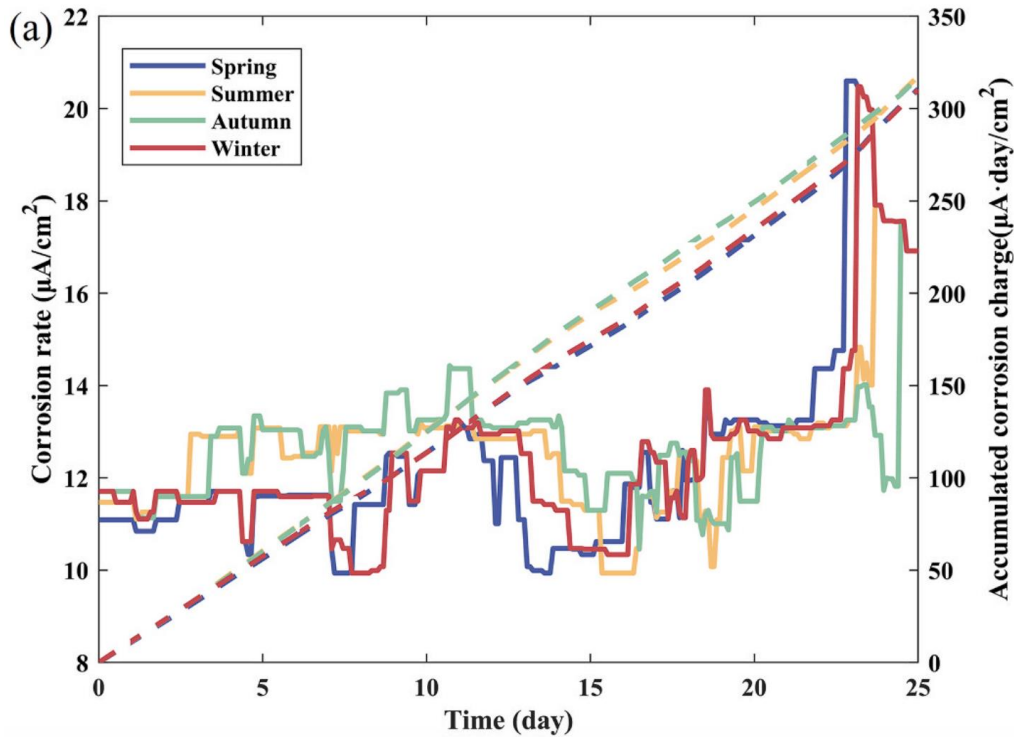


K. Índice de corrosión de las rutas 1, 2, 3, 4 en diferentes estaciones. Ji et al. (2024).

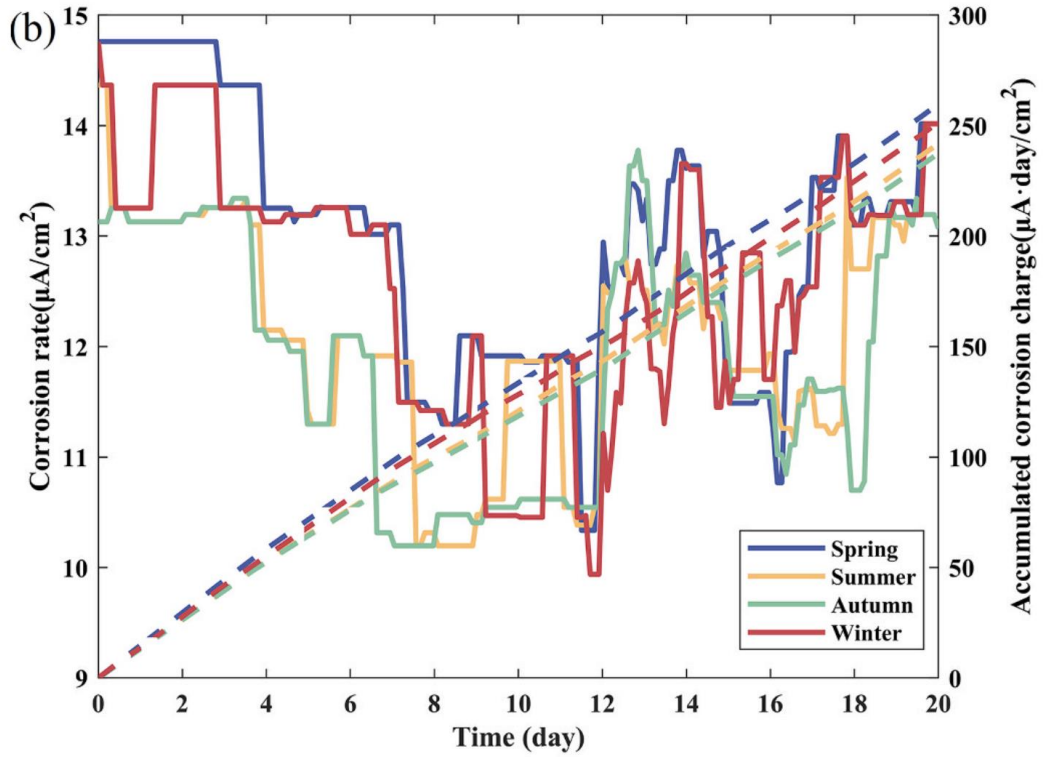
Índice de corrosión de la **Ruta 1** en diferentes estaciones.



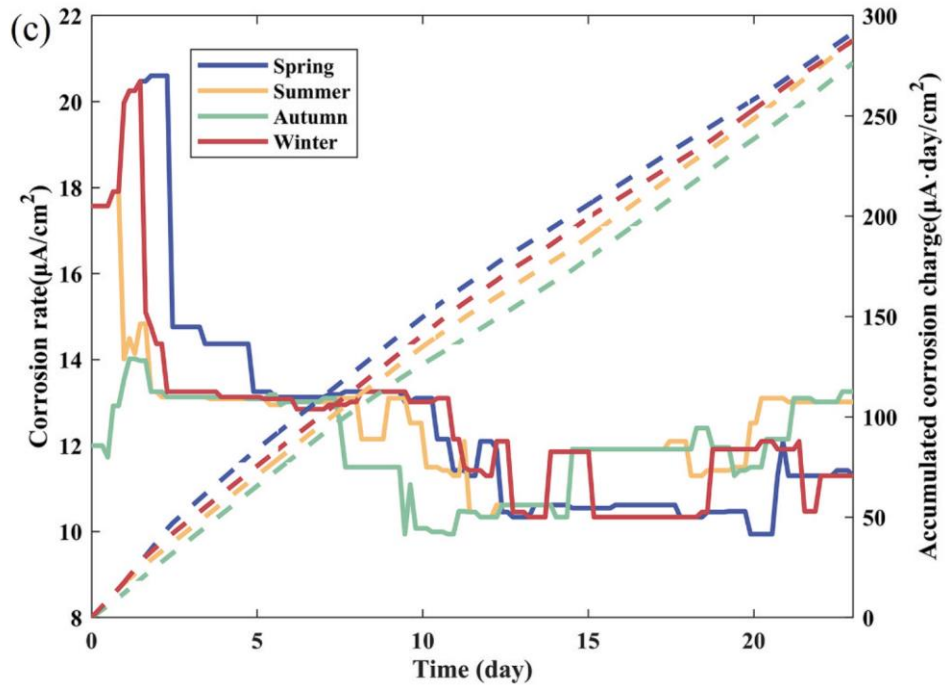
Índice de corrosión de **la Ruta 2** en diferentes estaciones. Ji et al. (2024).



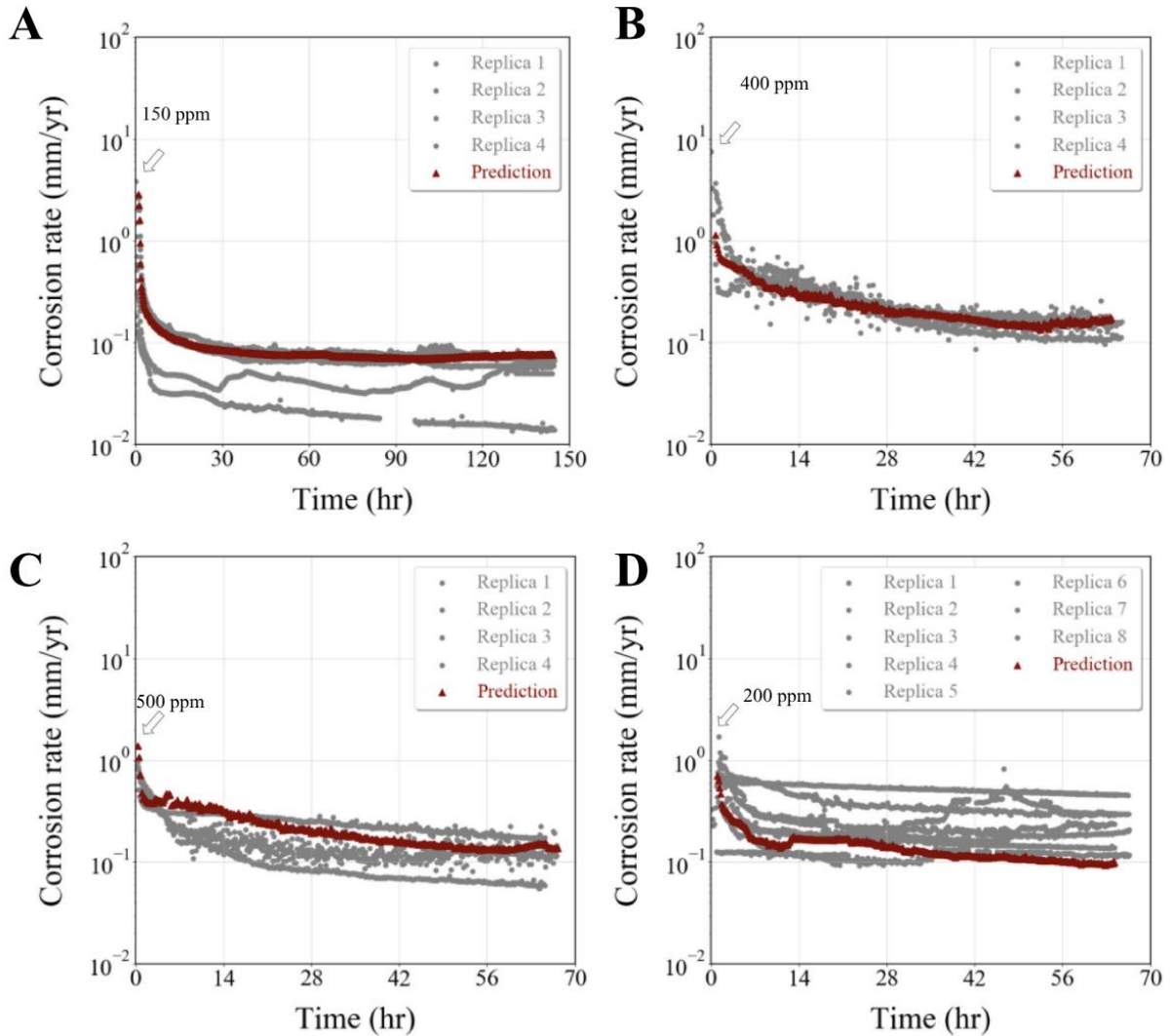
Índice de corrosión de la **Ruta 3** en diferentes estaciones.



Índice de corrosión de la **Ruta 4** en diferentes estaciones.



M. Comparación entre la velocidad de corrosión prevista (rojo) y la registrada experimentalmente (gris) en función del tiempo para experimentos de dosis única sin corrosión. A continuación, se presentan los detalles de los experimentos. Aghaaminiha et al. (2021)



N. Un estudio paramétrico, utilizando un modelo RF para determinar cómo varían los índices de corrosión en función del tiempo con el tipo de inhibidor A) tipo de inhibidor, B) concentración de inhibidor, C) presión parcial de CO₂, D) temperatura, E) tensión F) tipo de salmuera. La condición común en todas las figuras, a menos que se especifique lo contrario es la siguiente: tipo de inhibidor = CI-1, concentración de inhibidor 100 ppm, duración de la exposición ~ 22 h, pCO₂ = 12 bar, T = 130 °C, pH= no controlado, tensión de cizallamiento de la pared = 20 Pa, salmuera fuerza iónica = 0,615 M, tipo de salmuera = A. Aghaaminiha et al. (2021)

