



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Escola Tècnica
Superior d'Enginyeria
Informàtica

Escuela Técnica Superior de Ingeniería Informática
Universidad Politécnica de Valencia

Uso de IA en la detección de Artrosis para rodillas

TRABAJO FIN DE GRADO

Grado en Ingeniería Informática

Autor: Hernández Martínez, Carlos

Tutor: Juan Ciscar, Alfonso

Curso 2024-2025

Resum

Aquí citamos a un datajkkset [3]. Citación paper IEE [5], Dataset [1] aqui citamos un paper [6] otra cita [4]

Quitar imagenes brillantes -> [2]

Paraules clau: Palabras clave en catalán

Resumen

(Resumen en castellano)

Palabras clave: Palabras clave en español

Abstract

(Resumen en inglés)

Key words: Keywords in English

Índice general

Índice general	V
Índice de figuras	VII
Índice de tablas	VII
<hr/>	
1 Introducción	1
1.1 Motivación	1
1.2 Objetivos	2
1.3 Estructura del documento	2
2 Preliminares	3
2.1 Aprendizaje automático	3
2.1.1 Categorías de Aprendizaje Automático	3
2.1.2 Procesamiento de Datos en Aprendizaje Automático	4
2.1.3 Métricas de Evaluación	4
2.1.4 Aplicaciones del Aprendizaje Automático en Imágenes Médicas	4
2.2 Redes neuronales	5
2.3 ML aplicado a CV y tareas biomédicas (MedMNIST)	5
3 Capítulo 1 de contribución	7
3.1 Metodología y Adaptaciones del Modelo	7
3.2 Configuración del Entrenamiento y Optimización	8
3.3 Preprocesamiento y Aumento de Datos	8
3.4 Implementación del Modelo en PyTorch	9
3.5 Resultados y Discusión	10
3.6 Conclusión del Capítulo	11
4 Capítulo 2 de contribución	13
5 Capítulo 3 de contribución	15
6 Conclusiones	17
6.1 Resumen del trabajo realizado	17
6.2 Objetivos alcanzados	17
6.3 Trabajo futuro	17
<hr/>	
Apéndices	
A Configuración del sistema	21
B Otro apéndice	23

Índice de figuras

Índice de tablas

3.1	Matriz de confusión del modelo <i>EfficientNetB5Custom</i> en el conjunto de prueba.	10
3.2	Métricas de evaluación por clase en el conjunto de prueba.	10

CAPÍTULO 1

Introducción

1.1 Motivación

La artritis es una de las enfermedades musculoesqueléticas más prevalentes a nivel mundial y una de las principales causas de discapacidad en adultos mayores. Su diagnóstico y seguimiento se basa tradicionalmente en la evaluación clínica y en la interpretación de imágenes médicas, como radiografías, resonancias magnéticas y tomografías computarizadas. Sin embargo, este proceso suele depender en gran medida de la experiencia del profesional médico, lo que puede generar variabilidad en los diagnósticos y retrasos en la detección temprana de la enfermedad.

En los últimos años, los avances en inteligencia artificial, y en particular en el aprendizaje profundo, han demostrado un gran potencial para mejorar la precisión y la eficiencia en el análisis de imágenes médicas. Las redes neuronales convolucionales (CNN) han sido ampliamente utilizadas en el campo de la visión por computadora para tareas como la detección de patologías en radiografías, la segmentación de tejidos en resonancias magnéticas y la clasificación de niveles de severidad en enfermedades degenerativas.

Este Trabajo de Fin de Grado (TFG) se motiva por la necesidad de desarrollar métodos automáticos y robustos para el análisis de la artritis mediante el uso de técnicas de aprendizaje profundo. En particular, se busca explorar el uso de redes neuronales para la clasificación de imágenes médicas, utilizando bases de datos estandarizadas como *Mendeley dataset* [1]. La aplicación de estos modelos podría no solo optimizar el proceso de diagnóstico, sino también contribuir al desarrollo de herramientas de soporte a la decisión clínica, facilitando una intervención más temprana y personalizada para los pacientes.

La relevancia de este estudio radica en su potencial impacto en la práctica clínica. Un sistema basado en inteligencia artificial podría reducir la carga de trabajo de los especialistas, mejorar la objetividad del diagnóstico y ofrecer segundas opiniones automáticas que complementen la evaluación médica tradicional. Además, el desarrollo de estas tecnologías en el ámbito de la artritis podría sentar un precedente para su aplicación en otras enfermedades musculoesqueléticas, ampliando el alcance del aprendizaje profundo en el campo de la salud.

Además, se plantea la posibilidad de realizar *transfer learning* utilizando modelos pre-entrenados en artritis humana para aplicarlos en el diagnóstico de artritis en gatos. Esta adaptación podría beneficiar la práctica veterinaria, proporcionando herramientas automatizadas para la evaluación de la enfermedad en animales y mejorando la precisión en su diagnóstico.

En este contexto, el presente trabajo busca contribuir al avance del uso de inteligencia artificial en la detección y análisis de la artritis, evaluando diferentes enfoques de redes neuronales y analizando su rendimiento en la clasificación de imágenes médicas. La motivación principal es demostrar la viabilidad y efectividad de estos modelos en un problema biomédico concreto, promoviendo la integración de tecnologías emergentes en el ámbito de la salud.

1.2 Objetivos

1.3 Estructura del documento

CAPÍTULO 2

Preliminares

2.1 Aprendizaje automático

El aprendizaje automático (*Machine Learning, ML*) es una rama de la inteligencia artificial que ha transformado múltiples disciplinas al permitir que los sistemas aprendan y mejoren su desempeño en tareas específicas a partir de la experiencia, sin necesidad de ser programados explícitamente para cada situación. Este enfoque se basa en la construcción de modelos matemáticos capaces de identificar patrones y relaciones en grandes volúmenes de datos, lo que resulta esencial para aplicaciones tan diversas como la visión por computadora, el procesamiento del lenguaje natural, la bioinformática y, en particular, el diagnóstico médico.

En el contexto del análisis de imágenes médicas, el aprendizaje automático permite automatizar procesos de detección, clasificación y segmentación de patologías, facilitando diagnósticos más rápidos y precisos. Por ejemplo, modelos basados en redes neuronales profundas han alcanzado desempeños comparables a los de expertos humanos en la identificación de anomalías en radiografías, resonancias magnéticas y tomografías computarizadas.

2.1.1. Categorías de Aprendizaje Automático

El aprendizaje automático se clasifica en tres paradigmas fundamentales:

- **Aprendizaje supervisado:** Se apoya en conjuntos de datos etiquetados, donde cada entrada cuenta con una salida o etiqueta conocida. Durante el entrenamiento, el modelo aprende a mapear las entradas a las salidas correctas, facilitando la predicción de etiquetas en datos nuevos. Entre los algoritmos más comunes en este paradigma se encuentran la regresión lineal, las máquinas de soporte vectorial (*Support Vector Machines, SVM*) y las redes neuronales profundas.

java Copy

- **Aprendizaje no supervisado:** Aquí, los datos carecen de etiquetas y el objetivo es descubrir patrones o estructuras subyacentes. Técnicas como el clustering (agrupamiento), la reducción de dimensionalidad mediante análisis de componentes principales (PCA) y los modelos generativos permiten identificar relaciones ocultas en los datos y organizar la información de forma significativa.
- **Aprendizaje por refuerzo:** En este enfoque, un agente interactúa con un entorno y aprende a tomar decisiones mediante un sistema de prueba y error, optimizando

sus acciones en función de una función de recompensa. Este paradigma es especialmente útil en problemas de toma de decisiones secuenciales y en la optimización de estrategias en entornos dinámicos.

2.1.2. Procesamiento de Datos en Aprendizaje Automático

El rendimiento y la capacidad de generalización de los modelos dependen en gran medida de la calidad de los datos de entrada. Por ello, el procesamiento de datos es una etapa crítica que involucra varias fases:

- **Preprocesamiento:** Incluye la limpieza, normalización y transformación de los datos para eliminar ruidos, gestionar valores faltantes y unificar las escalas de las variables. En el ámbito de imágenes médicas, esta fase puede implicar la corrección de artefactos, la estandarización de intensidades y la segmentación preliminar de regiones de interés.
- **División del conjunto de datos:** Se segmenta la información en subconjuntos de entrenamiento, validación y prueba. Esta división es crucial para ajustar los hiperparámetros del modelo, prevenir el sobreajuste y evaluar de manera objetiva el desempeño final.
- **Extracción y selección de características:** En ciertos casos, se realiza una selección o generación de características relevantes que potencien la capacidad del modelo para identificar patrones significativos, lo cual es particularmente importante en dominios con alta dimensionalidad o datos complejos.

2.1.3. Métricas de Evaluación

Para medir la efectividad de un modelo de aprendizaje automático, se emplean diversas métricas que varían según el tipo de problema:

- **Para clasificación:** Se utilizan métricas como la precisión (*accuracy*), la precisión (precision), el recall (sensibilidad) y el F1-score, especialmente en contextos con clases desbalanceadas. La curva ROC y el área bajo la curva (AUC-ROC) son también indicadores fundamentales para evaluar la capacidad del modelo de distinguir entre clases.
- **Para regresión:** Se valora el desempeño mediante el error cuadrático medio (MSE) y el error absoluto medio (MAE), que cuantifican la diferencia entre las predicciones del modelo y los valores reales.

2.1.4. Aplicaciones del Aprendizaje Automático en Imágenes Médicas

La aplicación del aprendizaje automático en el análisis de imágenes médicas ha abierto nuevas fronteras en el diagnóstico y seguimiento de diversas patologías. Algunas de las aplicaciones más destacadas incluyen:

- **Detección de enfermedades:** Mediante modelos supervisados, se pueden identificar anomalías en radiografías, resonancias magnéticas y tomografías computarizadas, facilitando diagnósticos tempranos y precisos.

- **Segmentación de tejidos y órganos:** Algoritmos basados en redes neuronales permiten delimitar estructuras anatómicas y patologías en imágenes médicas, lo que es fundamental para planificar tratamientos y cirugías.
- **Clasificación de tumores:** Los modelos de aprendizaje profundo pueden diferenciar entre tumores benignos y malignos, aportando una segunda opinión automatizada que complementa la evaluación clínica.

En este trabajo se explora la aplicación del aprendizaje profundo para la detección de artritis en imágenes de rodillas. La integración de técnicas avanzadas de ML en el análisis de imágenes médicas no solo mejora la precisión diagnóstica, sino que también reduce la variabilidad interobservador, ofreciendo un soporte robusto y automatizado que puede transformar el proceso de diagnóstico en el ámbito clínico. Con el continuo avance en el procesamiento de datos y el desarrollo de modelos más sofisticados, se espera que el aprendizaje automático siga abriendo nuevas posibilidades para la personalización y optimización del tratamiento médico.

2.2 Redes neuronales

2.3 ML aplicado a CV y tareas biomédicas (MedMNIST)

CAPÍTULO 3

Capítulo 1 de contribución

El objetivo principal de este capítulo es replicar y ampliar el estudio realizado en [5] aplicando técnicas avanzadas de aprendizaje profundo al análisis de imágenes de rodillas, utilizando el conjunto de datos proporcionado en [1]. La detección temprana de artrosis resulta crucial para la intervención clínica oportuna, y la automatización del proceso diagnóstico mediante inteligencia artificial ofrece una solución prometedora para reducir la variabilidad en la interpretación de imágenes médicas y aliviar la carga de trabajo en los centros de salud.

La elección de utilizar el modelo **EfficientNetB5** se fundamenta en su reconocido equilibrio entre precisión y eficiencia computacional. Su capacidad para aprender representaciones complejas a partir de imágenes médicas lo convierte en un candidato ideal para abordar los retos inherentes a la clasificación de patologías en rodillas. Además, la replicación del estudio original permite validar la reproducibilidad de los resultados y sienta las bases para futuras investigaciones y mejoras en la metodología.

3.1 Metodología y Adaptaciones del Modelo

Para la replicación del estudio, se parte de un modelo preentrenado de **EfficientNetB5** que se ha adaptado para ajustarse al problema específico de la detección de artrosis en imágenes de rodilla. La principal modificación realizada consiste en la redefinición de la capa de clasificación, adaptándola para discriminar entre las cinco clases presentes en el dataset. La nueva arquitectura de la capa de clasificación es la siguiente:

```
self.efficientnet.classifier = nn.Sequential(  
    nn.Linear(in_features, 256),  
    nn.ReLU(),  
    nn.Linear(256, 256),  
    nn.ReLU(),  
    nn.Linear(256, num_classes),  
    nn.Softmax(dim=1)  
)
```

Esta estructura permite transformar las características extraídas por la red en probabilidades asociadas a cada clase, facilitando una toma de decisiones más precisa durante el proceso de clasificación.

3.2 Configuración del Entrenamiento y Optimización

El proceso de entrenamiento del modelo se ha diseñado cuidadosamente para garantizar la robustez y la capacidad de generalización del sistema. Se ha empleado la función de pérdida **CrossEntropy**, idónea para problemas de clasificación múltiple, complementada con técnicas de regularización que combinan penalizaciones L1 y L2 (coeficiente de 0.001) para mitigar el riesgo de sobreajuste.

Los hiperparámetros críticos definidos para el entrenamiento son los siguientes:

- **Paciencia** (*patience*): 5
- **Factor de reducción del learning rate** (*factor*): 0.1
- **Tasa de aprendizaje inicial** (*lr*): 0.001
- **Betas**: (0.9, 0.999)

Para la optimización se ha optado por el algoritmo **Adam**, reconocido por su adaptabilidad en entornos con datos complejos, mientras que el ajuste dinámico de la tasa de aprendizaje se realiza mediante el scheduler **ReduceLROnPlateau**. Este enfoque permite ajustar la tasa de aprendizaje en función de la evolución de la función de pérdida, contribuyendo a una convergencia más estable y eficiente:

```
optimizer = torch.optim.Adam(self.model.parameters(),
                             lr=self.learning_rate,
                             betas=self.betas)

scheduler = torch.optim.lr_scheduler.ReduceLROnPlateau(optimizer=optimizer,
                                                         factor=self.factor,
                                                         patience=self.patience)
```

3.3 Preprocesamiento y Aumento de Datos

El conjunto de datos empleado, conocido como *Mendeley dataset*[\[3\]](#), se ha dividido en subconjuntos de entrenamiento (70 %), test (20 %) y validación (10 %), asegurando una adecuada representación de cada clase. Para potenciar la capacidad del modelo y mejorar su generalización, se ha aplicado una serie de técnicas de aumento de datos utilizando transformaciones aleatorias. Estas transformaciones incluyen:

- Rotación de imágenes hasta 20 grados.
- Desplazamientos horizontales y verticales de hasta un 20 % del tamaño de la imagen.
- Aplicación de cizallamiento (shear) y zoom.
- Inversión horizontal aleatoria.

El proceso de aumento de datos ha permitido transformar un conjunto original con distribución desequilibrada en un dataset equilibrado, alcanzando aproximadamente 1000 muestras por clase en el conjunto de entrenamiento. La distribución de imágenes es la siguiente:

Clase	0	1	2	3	4
Train (original)	359	333	162	154	144
Validation	51	47	23	22	20
Test	104	97	47	45	42
Train (aumentado)	1000	1000	1000	1000	1000

3.4 Implementación del Modelo en PyTorch

La implementación se ha desarrollado en PyTorch, lo que ha permitido una integración eficiente de módulos preentrenados y una configuración flexible para la experimentación. A modo de ejemplo, se presenta una versión simplificada de la clase `EfficientNetB5Custom`, que incorpora las modificaciones descritas:

```
class EfficientNetB5Custom(nn.Module):
    def __init__(self, num_classes=5, pretrained=True):
        super(EfficientNetB5Custom, self).__init__()
        if pretrained:
            self.efficientnet = models.efficientnet_b5(weights=EfficientNet_B5_Weights.DEFAULT)
        else:
            self.efficientnet = models.efficientnet_b5(weights=None)
        self.name = "EfficientNetB5Custom"
        in_features = self.efficientnet.classifier[1].in_features
        self.efficientnet.classifier = nn.Sequential(
            nn.Linear(in_features, 256),
            nn.ReLU(),
            nn.Linear(256, 256),
            nn.ReLU(),
            nn.Linear(256, num_classes),
            nn.Softmax(dim=1)
        )

    def forward(self, x):
        return self.efficientnet(x)
```

Paralelamente, se ha definido un pipeline de preprocesamiento para las imágenes mediante las siguientes transformaciones:

```
transform = transforms.Compose([
    transforms.Resize((224, 224)),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406],
                        std=[0.229, 0.224, 0.225]),
])
```

La configuración de entrenamiento adicional se ha establecido con los siguientes parámetros:

- **BATCH_SIZE:** 20
- **LEARNING_RATE:** 0.001
- **FACTOR:** 0.001

- Regularización L1: 0.0001
- Regularización L2: 0.0001
- PATIENCE: 5
- BETAS: (0.9, 0.999)

3.5 Resultados y Discusión

Tras el entrenamiento, se obtuvieron los siguientes resultados en el conjunto de prueba:

- Loss: 1.16
- Accuracy (ACC): 0.74
- Área bajo la curva (AUC): 0.92

La **Matriz de Confusión** para las cinco clases (0 a 4) se muestra a continuación:

		Predicción				
		0	1	2	3	4
Clase Real	0	44	4	3	0	0
	1	5	36	6	0	0
	2	3	11	6	3	0
	3	0	3	2	17	0
	4	0	2	1	0	17

Tabla 3.1: Matriz de confusión del modelo *EfficientNetB5Custom* en el conjunto de prueba.

Precision = $TP / (TP + FP)$

Sensibilidad (Recall) = $TP / (TP + FN)$

Especificidad = $TN / (TN + FP)$

F1 = $2 * (Precision * Recall) / (Precision + Recall)$

Clase	Precisión	Sensibilidad	Especificidad	F1
0	0.8461	0.8627	0.9285	0.8543
1	0.7058	0.7659	0.8275	0.6990
2	0.3333	0.2608	0.9140	0.2926
3	0.8500	0.7727	0.9787	0.8095
4	1.0000	0.8500	1.0000	0.9189

Tabla 3.2: Métricas de evaluación por clase en el conjunto de prueba.

En ella se aprecia un desempeño satisfactorio en la mayoría de las clases, con especial eficacia para las clases 0 y 4. Sin embargo, existe cierta confusión entre las clases 1 y 2, lo que sugiere la necesidad de refinar la estrategia de preprocesamiento o ajustar hiperparámetros que permitan una mejor discriminación en futuros experimentos.

La precisión global (0.74) y el AUC (0.92) indican que, a pesar de los falsos positivos y negativos observados, el modelo presenta una capacidad notable para distinguir entre diferentes niveles de artrosis. Con todo, se prevén mejoras adicionales si se profundiza en técnicas de regularización o se evalúan arquitecturas alternativas de redes neuronales.

3.6 Conclusión del Capítulo

En este capítulo se ha descrito el proceso de replicación del estudio de [5] con un modelo **EfficientNetB5** adaptado al diagnóstico de artrosis en imágenes de rodilla. Los resultados muestran un *accuracy* del 74 % y un *AUC* del 92 %, valores prometedores para un primer experimento de replicación. La matriz de confusión permite identificar las clases con mayor dificultad de clasificación, evidenciando la relevancia de un preprocesamiento cuidadoso y una optimización sistemática de los hiperparámetros.

Estos hallazgos establecen las bases para trabajos posteriores, en los que se podrá investigar la influencia de distintos métodos de aumento de datos, arquitecturas de red y enfoques de regularización. Asimismo, el potencial de la inteligencia artificial en el ámbito de la salud se ve reforzado, promoviendo la adopción de estas técnicas como herramienta de apoyo para los profesionales médicos.

CAPÍTULO 4

Capítulo 2 de contribución

CAPÍTULO 5

Capítulo 3 de contribución

CAPÍTULO 6

Conclusiones

6.1 Resumen del trabajo realizado

6.2 Objetivos alcanzados

6.3 Trabajo futuro

Bibliografía

- [1] Pingjun Chen. *Knee Osteoarthritis Severity Grading Dataset*. Ver. V1. Mendeley Data, 2018. DOI: [10.17632/56rmx5bjcr.1](https://doi.org/10.17632/56rmx5bjcr.1).
- [2] Sajid Fardin Dipto y Md. Omaer Faruq Goni. "Classification of X-Ray Images for the Automated Severity Grading of Knee Osteoarthritis by Ensemble Learning Through EfficientNet Architectures with Grad-CAM Visualization". En: *2024 IEEE International Conference on Power, Electrical, Electronics and Industrial Applications (PEEIACON)*. 2024, págs. 108-113. DOI: [10.1109/PEEIACON63629.2024.10800349](https://doi.org/10.1109/PEEIACON63629.2024.10800349).
- [3] Shivanand Gornale y Pooja Patravali. *Digital Knee X-ray Images*. Ver. V1. Mendeley Data, 2020. DOI: [10.17632/t9ndx37v5h.1](https://doi.org/10.17632/t9ndx37v5h.1).
- [4] Neha Sharma et al. "A Comprehensive Review on Knee Osteoarthritis Detection using Medical Imaging and Machine Learning". En: *2024 International Conference on Intelligent Systems for Cybersecurity (ISCS)*. 2024, págs. 1-6. DOI: [10.1109/ISCS61804.2024.10581051](https://doi.org/10.1109/ISCS61804.2024.10581051).
- [5] Shubham Kumar Singh, Kuldeep Chouhan y Arun Prakash Agrawal. "Osteoarthritis Prediction in Knee Joint Using Deep Learning Techniques". En: *2024 27th International Symposium on Wireless Personal Multimedia Communications (WPMC)*. 2024, págs. 1-5. DOI: [10.1109/WPMC63271.2024.10863523](https://doi.org/10.1109/WPMC63271.2024.10863523).
- [6] Elias Vaattovaara et al. "Kellgren-Lawrence Grading of Knee Osteoarthritis using Deep Learning: Diagnostic Performance with External Dataset and Comparison with Four Readers". En: *Osteoarthritis and Cartilage Open* (2025), pág. 100580. ISSN: 2665-9131. DOI: <https://doi.org/10.1016/j.ocarto.2025.100580>. URL: <https://www.sciencedirect.com/science/article/pii/S2665913125000160>.

APÉNDICE A

Configuración del sistema

APÉNDICE B

Otro apéndice
