

# Optimización eficiente de CNNs mediante características precalculadas y entrenamiento ponderado \*

Daniel Vila-Cruz, Verónica Bolón-Canedo y Laura Morán-Fernández

CITIC, Universidade da Coruña, A Coruña, Spain  
d.vila1@udc.es

## Resumen

Los modelos de aprendizaje profundo, en especial las redes convolucionales (CNNs), han logrado gran rendimiento en la clasificación de imágenes médicas, aunque su uso real se ve limitado por el hardware, el consumo energético y los tiempos de entrenamiento. En este trabajo proponemos una estrategia ligera que separa la extracción de características de la optimización del clasificador. Al precalcular las características y adaptar la normalización con un umbral dinámico, reducimos notablemente la carga computacional. Un rediseño del clasificador y un entrenamiento que pondera los casos ambiguos evitan la pérdida de precisión. En los conjuntos de datos Brain Cancer MRI y BreakHis, nuestro enfoque disminuye tiempo y emisiones de CO<sub>2</sub>, manteniendo o superando la precisión del *fine-tuning* tradicional.

## 1. Introducción

Los rápidos avances en aprendizaje profundo han generado modelos cada vez más precisos, pero complejos y costosos. En medicina, el hardware especializado es escaso y esto limita la adopción práctica de estas tecnologías. Aunque técnicas como la poda de pesos o la cuantización reducen el coste computacional, requieren optimizaciones adicionales o pueden degradar el rendimiento. Asimismo, los conjuntos de datos médicos suelen ser pequeños y no contar con las suficientes muestras representativas para cada caso, lo que puede provocar sobreajuste o sesgos al reutilizar modelos preentrenados. Surge así la necesidad de estrategias frugales y conscientes del consumo energético que mantengan la precisión con un menor impacto ambiental.

## 2. Metodología propuesta

El objetivo de la propuesta es maximizar la precisión manteniendo un entrenamiento rápido y de bajo consumo. Para ello, se plantea un flujo de trabajo que desacopla el extractor de características (*backbone*) del clasificador, evitando la retropropagación a través de toda la red. En lugar de ajustar el modelo completo, se obtiene el conjunto de características con el *backbone* preentrenado, y luego se reutilizan en múltiples iteraciones de entrenamiento y validación del clasificador. De esta forma se elimina el coste repetido de procesar las imágenes en cada época. No obstante, al prescindir de la adaptación directa del *backbone*, puede producirse un desajuste estadístico entre dominios. Para solventarlo, se actualizan las estadísticas de las capas de *Batch Normalization* haciendo pasar los datos en modo entrenamiento sin retropropagación, permitiendo que las medias y varianzas internas se ajusten al nuevo dominio. Este proceso se detiene automáticamente cuando el cambio relativo entre medias consecutivas cae por debajo de un umbral predefinido, evitando cálculos innecesarios y manteniendo la estabilidad del modelo.

---

\*Esta investigación ha sido financiada por el Ministerio de Ciencia e Innovación de España (Subvención PID2023-147404OB-I00 / AEI / 10.13039 / 501100011033), junto con "NextGenerationE"/PRTR del Ministerio de Transformación Digital y Función Pública bajo la subvención TSI-100925-2023-1, y por la Xunta de Galicia (Subvenciones ED431G 2023/01 y ED431C 2022/44).

Una vez obtenidas las características, se entrena un clasificador rediseñado, aumentando ligeramente su complejidad, lo que permite mejorar su capacidad de generalización gracias a la normalización intermedia y la activación no lineal. Durante el entrenamiento, se incorpora un mecanismo de ponderación que otorga mayor importancia a las muestras ambiguas, es decir, aquellas cuya diferencia entre las dos probabilidades más altas del *softmax* es menor. Estas instancias reciben un peso ampliado a la hora de calcular la pérdida, lo que orienta el aprendizaje hacia los ejemplos más informativos. Con ello, el clasificador converge más rápido y generaliza mejor sin aumentar el número de iteraciones. En conjunto, la metodología elimina el ajuste completo del modelo, reduce drásticamente el tiempo de entrenamiento y conserva la capacidad de adaptación mediante un ajuste estadístico controlado.

### 3. Resultados

La propuesta se evaluó en los conjuntos de datos Brain Cancer MRI [1] y BreakHis [2]. Los experimentos se realizaron en una GPU NVIDIA RTX 4060, comparando el método propuesto con el ajuste fino convencional en términos de precisión, tiempo y emisiones de CO<sub>2</sub>. Se han realizado pruebas empleando cuatro extractores de características: ResNet18, ResNet50, DenseNet121 y MobileNetV3. En el conjunto Brain Cancer MRI, nuestro método mejora de media en un 0,27 % ( $\pm 3,48\%$ ) la precisión, reduciendo el tiempo de entrenamiento en un 87,76 % ( $\pm 5,62\%$ ) y las emisiones asociadas en un 88,60 % ( $\pm 5,37\%$ ). En el conjunto BreakHis, la mejora de precisión es de un 0,16 % ( $\pm 0,78\%$ ) de media, mientras que las reducciones correspondientes al tiempo de entrenamiento y emisiones asociadas son 89,71 % ( $\pm 3,33\%$ ) y 90,15 % ( $\pm 2,93\%$ ), respectivamente.

También se evaluó el método realizando ajuste de hiperparámetros, un escenario en el que, debido a la posibilidad de reutilizar las características extraídas, es especialmente eficiente. El caso que presenta una menor mejora con respecto al *fine-tuning* tradicional se obtuvo empleando 1000 muestras de entrenamiento, 10 combinaciones de hiperparámetros y el extractor ResNet18, reduciendo el tiempo de entrenamiento 28,29 veces (3,62 minutos frente a 7,67 segundos). El caso de mayor mejora corresponde al experimento realizado empleando 100 000 muestras de entrenamiento, 160 combinaciones de hiperparámetros y el extractor DenseNet121, reduciendo el tiempo de entrenamiento 167,34 veces (11,73 días frente a 1,68 horas).

### 4. Conclusiones

Se presenta una estrategia alternativa al *fine-tuning* tradicional, combinando preprocesamiento de características, adaptación ligera al dominio y entrenamiento ponderado para reducir drásticamente el coste temporal y energético. La eliminación de cálculos redundantes sobre el *backbone* disminuye tanto el tiempo de entrenamiento como las emisiones hasta en un 95 %, manteniendo una precisión competitiva. Destaca también la eficiencia del método en el contexto de ajuste de hiperparámetros, donde cuanto mayor es el número de combinaciones evaluadas, mayor es la reducción obtenida en coste computacional con respecto al *fine-tuning* tradicional. Los resultados demuestran que es posible entrenar clasificadores precisos en entornos con recursos limitados mediante técnicas simples, modulares y reproducibles, abriendo la puerta a un uso más sostenible del aprendizaje profundo en aplicaciones médicas.

## Referencias

- [1] Md Mizanur Rahman. Brain Cancer - MRI dataset, 2024.
- [2] Fabio A. Spanhol, Luiz S. Oliveira, Caroline Petitjean, and Laurent Heutte. A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, 63(7):1455–1462, 2016.