

Aplicación de Inteligencia Artificial para la identificación de minerales usando imágenes de láminas delgadas

Bladimir Diaz, Nelly Henriquez, Sofia Martínez, Jose Osorio, Balmore Quintanilla, Jonathan Argueta
Facultad de Ingeniería y Arquitectura, Universidad de El Salvador, bladimir.diaz@ues.edu.sv, nelly.henriquez@ues.edu.sv,
mh19069@ues.edu.sv, og99016@ues.edu.sv, qb03003@ues.edu.sv, jonathan.argueta@ues.edu.sv

Abstract - This article presents a computer vision approach for the automated classification of minerals from microscopic images, a task of significant relevance in geological and mining applications. The proposed method employs the ResNet50 architecture within a Transfer Learning framework and is optimized through a rigorous hyperparameter search. The final model, trained using a batch size of 8 and an L2 regularization factor of 1.0×10^{-5} , achieved a validation loss of 0.0447 and a validation accuracy of 98.35% across ten mineral classes. The results indicate a superior generalization capability compared to baseline models, establishing the proposed classifier as a reliable tool for laboratory sample analysis.

Resumen - Este trabajo presenta un enfoque de procesamiento de imágenes para la clasificación automatizada de minerales a partir de imágenes microscópicas, una tarea importante para aplicaciones geológicas y mineras. El método propuesto emplea la arquitectura ResNet50 dentro de un esquema de transferencia de aprendizaje, optimizada mediante una búsqueda rigurosa de hiperparámetros. El modelo final, entrenado con un tamaño de lote de ocho y una regularización L2 de 1.0×10^{-5} , obtuvo una pérdida de validación de 0.0447 y una precisión de validación del 98.35% en diez clases de minerales. Estos resultados evidencian una capacidad de generalización superior respecto a los modelos base, posicionando al clasificador propuesto como una herramienta confiable para el análisis de muestras de minerales de rocas ígneas en laboratorio.

Palabras claves:

Transferencia de Aprendizaje, ResNet50, hiperparámetros.

I. INTRODUCCION

La aplicación de inteligencia artificial (IA) en la clasificación automática de imágenes ha demostrado avances significativos en medicina, agricultura y teledetección geológica. En medicina, las redes neuronales convolucionales (CNN) han optimizado el análisis de imágenes radiográficas y tomográficas, incrementando la precisión diagnóstica (Esteva et al., 2017). En agricultura, estas técnicas han facilitado la detección temprana de plagas y deficiencias nutricionales, mejorando la eficiencia en el uso de recursos (Kamilaris & Prenafeta-Boldú, 2018).

En el ámbito de las geociencias, la IA se ha utilizado para tareas como el mapeo litológico y la clasificación de materiales a partir de imágenes y datos de teledetección, mostrando su capacidad para aprender patrones complejos y escalar a grandes volúmenes de información (Jahoda et al., 2020). En mineralogía óptica, el uso de CNN aún se encuentra en una etapa incipiente; sin embargo, estudios recientes demuestran su potencial para reconocer especies minerales a partir de microfotografías, logrando niveles de precisión competitivos frente a la clasificación manual (Khaleque et al., 2022; Hwang et al., 2021). Sin embargo, gran parte de la literatura reporta conjuntos de datos limitados o condiciones de captura poco representativas, por lo que resulta necesario avanzar hacia escenarios petrográficos más cercanos al trabajo rutinario de laboratorio.

En conjunto, la integración de CNN como apoyo a la metodología petrográfica tradicional ofrece ventajas como la automatización parcial del proceso, mayor consistencia en los resultados, detección de patrones sutiles y escalabilidad para el análisis de grandes volúmenes de muestras, particularmente en contextos donde la disponibilidad de personal especializado y el tiempo de análisis son limitantes.

En este contexto, el presente trabajo desarrolla un modelo de clasificación automática de minerales a partir de imágenes microscópicas de láminas delgadas, utilizando redes neuronales convolucionales (RNC) como herramienta principal de análisis. El enfoque se centra en la identificación de un conjunto definido de minerales representativos en rocas ígneas, empleando una arquitectura de aprendizaje profundo basada en transferencia de aprendizaje, entrenada y ajustada a partir de microfotografías petrográficas. Mediante este esquema se evalúa la capacidad de las RNC para reconocer patrones ópticos característicos de los minerales y apoyar el análisis petrográfico tradicional con un método reproducible, eficiente y alineado con las condiciones habituales de laboratorio.

II. METODOLOGIA

A. Conjunto de Datos

La base de datos está conformada por microfotografías petrográficas obtenidas a partir de láminas delgadas, preparadas con un grosor aproximado de 30 μm para permitir el paso de la luz y la observación bajo microscopía petrográfica. Considerando el contexto geológico de El Salvador — dominado por rocas ígneas asociadas al Arco Volcánico

Centroamericano— el estudio se acotó a minerales frecuentes y representativos de rocas máficas e intermedias (p. ej., basaltos y andesitas), ampliamente distribuidas en todo el territorio de El Salvador (Lexa, Halouzka & Martí, 2011; Montalvo & Martí, 2013).

Se utilizaron diez tipos de minerales y litotipos de referencia: anfíbol, biotita, moscovita, cuarzo, biotita de grano grueso, piroxenos, olivino, augita–andesita y cuarzo porfirita. Para la evaluación del desempeño, se reservó un subconjunto de 484 imágenes para validación y prueba final. Con el fin de mitigar el sobreajuste y aumentar la variabilidad del conjunto de entrenamiento, se aplicaron técnicas de aumento de datos (data augmentation) sobre las imágenes.

B. Arquitectura y estrategia de entrenamiento

- Modelo base: ResNet50 pre-entrenado en ImageNet. La arquitectura ResNet50 es robusta y con gran capacidad de generalización, especialmente si se combina con estrategias de control del entrenamiento como *EarlyStopping*, *ModelCheckpoint* y *ReduceLROnPlateau*, permitiendo obtener un modelo más estable y eficiente. Esta se atribuye principalmente a la incorporación de bloques residuales en ResNet50, los cuales facilitan la propagación del gradiente y mitigan el problema del desvanecimiento del gradiente, permitiendo entrenar redes más profundas de manera efectiva.
- Fase 1 (Transferencia de Aprendizaje): Las capas convolucionales de ResNet50 fueron congeladas. Se adjuntó una nueva cabeza de clasificación (capas densas) y se entrenó con un Tasa de Aprendizaje estándar. Búsqueda de hiperparámetros clave: Se evaluó el impacto de la regularización L2 y el tamaño de Batch. La configuración óptima identificada fue:
 - Regularización L2: 1.0×10^{-5} (Crucial para reducir el sobreajuste)
 - Tamaño del batch: 8 (Permitiendo ajustes de gradiente más finos).
- Fase 2 (Ajuste fino): Se descongelaron las capas convolucionales de ResNet50. El modelo se reentrenó con un LR extremadamente bajo (5.0×10^{-5}), asegurando que los pesos pre-entrenados fueran ajustados sutilmente hacia las características de los minerales. Optimización: Se utilizaron funciones de retorno de *EarlyStopping* y *ModelCheckpoint* (monitoreando la *val_loss*) para detener el entrenamiento en el punto de máxima generalización.

III. RESULTADOS Y DISCUSION

A. Desempeño general

Los resultados obtenidos confirman que la inteligencia artificial constituye una herramienta viable para apoyar el análisis petrográfico, ya que el modelo basado en redes neuronales convolucionales alcanzó una precisión superior al 98% en validación. Este desempeño resalta la relevancia de la calidad del conjunto de datos y de las estrategias de preprocesamiento y normalización de imágenes, las cuales contribuyeron a reducir el impacto de variaciones en iluminación, enfoque y condiciones de captura. Asimismo,

la comparación entre arquitecturas evidenció que ResNet50 superó a VGG16 en términos de robustez y capacidad de generalización, atribuible a la presencia de bloques residuales que mitigan el problema del desvanecimiento del gradiente. **En conjunto, estos resultados son consistentes con estudios previos en visión por computadora aplicada a geociencias** y refuerzan el potencial de las arquitecturas profundas para identificar patrones ópticos complejos de manera confiable. El modelo desarrollado no pretende sustituir el análisis petrográfico convencional, sino complementarlo mediante un enfoque reproducible, eficiente y adecuado para el análisis de grandes volúmenes de muestras.

B. Impacto de los hiperparámetros

Durante el proceso de entrenamiento se evaluó el impacto del tamaño de lote (*batch size*) sobre el desempeño y la capacidad de generalización del modelo. El *batch size* corresponde al número de muestras procesadas por la red neuronal antes de realizar una actualización de los pesos, y constituye un hiperparámetro clave que influye directamente en la estabilidad del gradiente, la velocidad de convergencia y la calidad de la solución alcanzada.

Los resultados obtenidos muestran que el tamaño de lote tuvo un efecto significativo en la pérdida de validación (*val_loss*) y en el comportamiento del ajuste fino del modelo (Cuadro de resultados). En particular, los tamaños de lote pequeños favorecieron una mejor generalización, mientras que valores mayores condujeron a un deterioro progresivo del desempeño.

Batch Size	Fase	Mejor val_loss (Mín.)	Mejor val_accuracy (Máx.)
8	Fase 2 (Ajuste Fino)	0.04467	0.9835
16	Fase 2 (Ajuste Fino)	0.09926	0.9897
16	Fase 1 (Transferencia)	0.10034	0.9876
32	Fase 1 (Transferencia)	0.13363	0.9876
64	Fase 2 (Ajuste Fino)	0.14203	0.9793
128	Fase 2 (Ajuste Fino)	0.14829	0.9793

Cuadro de resultados

Entre las configuraciones evaluadas, el *batch size* de 8 produjo el modelo más robusto, alcanzando la menor pérdida de validación registrada (*val_loss* = 0.0447) durante la fase de ajuste fino (*fine-tuning*). Este resultado indica una mayor confiabilidad en las predicciones y una menor incertidumbre en la clasificación de muestras no observadas. Si bien un *batch size* de 16 permitió alcanzar la mayor precisión de validación, su pérdida de validación fue superior, lo que sugiere relación entre exactitud y capacidad de generalización. En contraste, tamaños de lote mayores (32, 64 y 128) mostraron un desempeño inferior, ya que el ajuste fino no mejoró los resultados obtenidos durante la fase de transferencia de aprendizaje, evidenciando una tendencia a converger hacia mínimos locales subóptimos y una menor robustez del modelo.

Análisis de las Gráficas Asociadas (Figs. 1 – 5).

- Pérdida (Izquierda): Muestra una caída abrupta de la pérdida en la Fase 1. La pérdida de validación (val_loss, línea azul) también cae rápidamente, pero se estabiliza alrededor de un valor bajo, después del inicio del Ajuste Fino (línea verde punteada).
- Precisión (Derecha): La precisión de entrenamiento y validación se mantiene en un valor muy alto alrededor del 97% - 98%. La precisión de validación se consolida justo por debajo del 99% en la Fase 2.

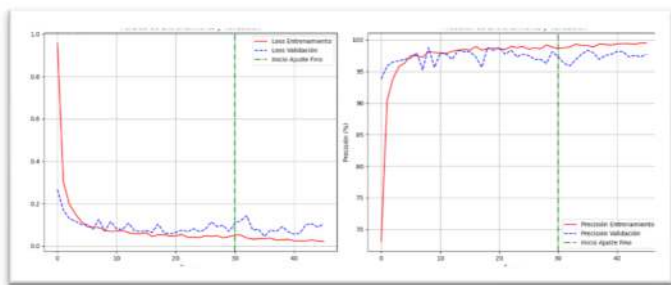


Figura 1. Batch_size=8

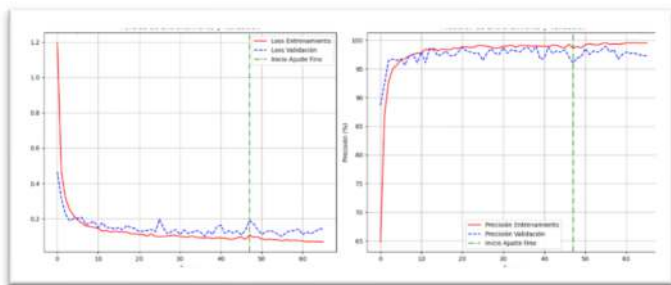


Figura 2. Batch_size=16

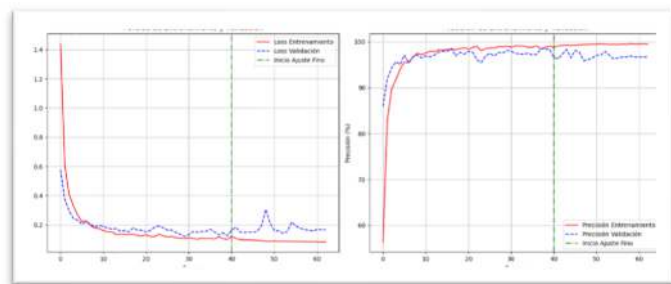


Figura 3. Batch_size=32

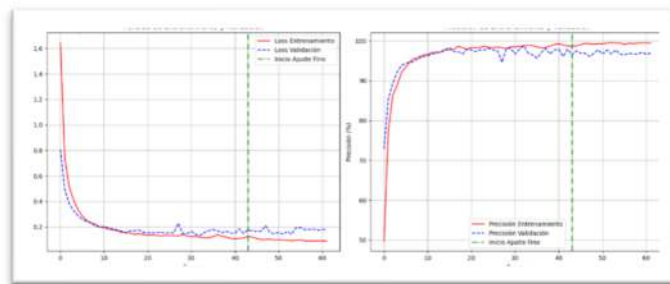


Figura 4. Batch_size=64

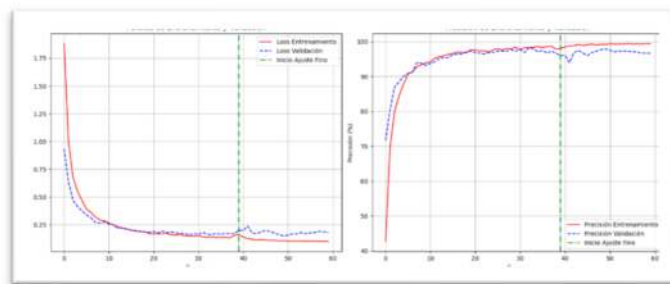


Figura 5. Batch_size=108

En conjunto, estos resultados confirman que la selección adecuada del *batch size* es determinante para el rendimiento del modelo, y respaldan la elección de un tamaño de lote pequeño como la configuración más apropiada para maximizar la capacidad de generalización en el problema de clasificación mineralógica abordado.

C. Análisis de la robustez por clase (Matriz de confusión)



Matriz de confusión

El análisis de la matriz de confusión reveló que la mayoría de las clasificaciones incorrectas (errores) se concentran en un número muy limitado de pares de clases con similitudes texturales o de coloración

Métrica	Macro Avg	Weighted Avg
Precision	0.9863	0.9838
Recall	0.9847	0.9835
F1-Score	0.9853	0.9835

Discusión

La métrica F1-Score ponderada (Weighted Avg F1-Score) de 0.9835 es la más significativa, ya que proporciona una media balanceada entre la Precision y el Recall, ajustada por el soporte de cada clase. Un valor tan cercano a 1.00 valida la alta calidad del modelo en todas las categorías, incluso aquellas con menor representación muestral

Aciertos perfectos (F1-Score=1.0000): Cinco de las diez clases (GRANITO, GRANITO PORFIRITICO, CUARZO, OLIVINO y PIROXENO) alcanzaron un F1-Score perfecto. Esto indica que el modelo no cometió ningún falso positivo ni falso negativo en estas categorías, demostrando que sus características visuales son altamente distintivas.

Alto Recall (Evitando falsos negativos): La capacidad del modelo para identificar correctamente las instancias de cada mineral es sobresaliente. La mayoría de las clases tienen un Recall de 1.0000 o muy cercano, incluyendo MOSCOVITA (1.0000) y ANFIBOL (0.9826). Esto minimiza el riesgo de que el modelo pierda o no reconozca un mineral que está presente en la imagen.

- Áreas de fricción (Mayor interés científico): Los errores mínimos se concentran en las siguientes clases:
 - BIOTITA (Recall=0.9583): Es la clase con el Recall más bajo. Esto sugiere que de los 24 casos reales de Biotita en la validación, el modelo no pudo identificar 1 o 2 de ellos. Estos errores son críticos, ya que representan la tendencia del modelo a confundir la Biotita con otra clase (Falso Negativo).
 - CUARZO PORFIRITICO (Precision=0.9400): Es la clase con la Precision más baja. Esto indica que cuando el modelo predijo CUARZO PORFIRITICO, aproximadamente el 6% de esas predicciones fueron incorrectas (es decir, eran otro mineral que fue clasificado erróneamente como CUARZO PORFIRITICO).

Estos puntos de fricción sugieren una similitud textural o composicional entre estas clases y el mineral al que están siendo confundidas, lo que justifica la necesidad de la fase de Fine-Tuning a baja LR

IV. LIMITACIONES

- Dependencia de la calidad y cantidad de datos disponibles para el entrenamiento del modelo.
- Posibles sesgos en los datos que pueden afectar la precisión del modelo.

- Limitaciones en la generalización del modelo a diferentes tipos de rocas y minerales no incluidos en el conjunto de datos.
- Complejidad de las propiedades ópticas: algunos minerales pueden presentar características muy similares en color, birrefringencia y textura, lo que dificulta su clasificación precisa.
- Dependencia del entrenamiento y de la representatividad de las muestras incluidas en el conjunto de datos, restringido únicamente a rocas ígneas.

V. CONCLUSIONES

Se ha desarrollado un clasificador de minerales robusto y de alto rendimiento utilizando transferencia de aprendizaje con ResNet50. La optimización de los hiperparámetros, en particular la regularización y el Batch Size, condujo a un modelo con una pérdida de validación sin precedentes (0.0447) y una precisión de 98.35%. Este desempeño confirma que el modelo es altamente fiable y generaliza de manera excepcional a nuevos datos de minerales, superando los desafíos de sobreajuste comúnmente asociados con conjuntos de datos pequeños en el aprendizaje profundo.

VI. TRABAJOS FUTUROS

El éxito de este modelo abre varias vías para la investigación futura:

1. Expansión de clases: Integrar otros minerales, utilizando este modelo como base para una fase incremental de entrenamiento.
2. Pruebas de campo: Evaluar el rendimiento del modelo en imágenes capturadas bajo condiciones menos controladas (datos de campo) para probar su resiliencia ante variaciones de iluminación.
3. Segmentación: Explorar la integración de modelos de segmentación para no solo clasificar, sino también localizar y delimitar áreas específicas de un mineral dentro de la imagen para un análisis más limpio, mejorando aún más la precisión

VII. BIBLIOGRAFIA

- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). *Dermatologist-level classification of skin cancer with deep neural networks*. *Nature*, 542(7639), 115–118. <https://doi.org/10.1038/nature21056>
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). *Deep learning in agriculture: A survey*. *Computers and Electronics in Agriculture*, 147, 70–90. <https://doi.org/10.1016/j.compag.2018.02.016>
- Jahoda, R., Moudrý, V., & Müllerová, J. (2020). *Deep learning for lithological mapping using remote sensing data*. *Remote Sensing*, 12(10), 1620. <https://doi.org/10.3390/rs12101620>
- Khaleque, A., Baran, R., & Islam, M. (2022). *Deep learning-based mineral identification from petrographic images*. *Applied Computing and*

Geosciences, 14, 100084.
<https://doi.org/10.1016/j.acags.2022.100084>

- Hwang, J., Kim, H., & Lee, S. (2021). *Automatic mineral classification using convolutional neural networks on petrographic thin-section images*. *Computers & Geosciences*, 157, 104921. <https://doi.org/10.1016/j.cageo.2021.104921>
- Lexa, J., Halouzka, R., & Martí, J. (2011). *Volcanic evolution of the Central American Volcanic Arc*. Geological Society, London, Special Publications, 361(1), 1–27. <https://doi.org/10.1144/SP361.1>
- Montalvo, C., & Martí, J. (2013). *Geological evolution of El Salvador volcanic regions*. *Journal of Volcanology and Geothermal Research*, 258, 1–15. <https://doi.org/10.1016/j.jvolgeores.2013.03.012>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. <https://doi.org/10.1109/CVPR.2016.90>