

Automatic Classification of Bee Pollen Types

Julieta Shammah^{1[0009-0009-6901-5706]}, María Elena Buemi^{2[0009-0008-1076-4035]} and Agustín Sanguinetti^{1,3,4,5[0000-0003-3808-3131]}

¹ Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Biodiversidad y Biología Experimental. Buenos Aires, Argentina.

² Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales, Departamento de Computación. Buenos Aires, Argentina .

³ CONICET – Universidad de Buenos Aires. Instituto de Biodiversidad y Biología Experimental y Aplicada (IBBEA). Buenos Aires, Argentina.

⁴ Universidad Nacional de Entre Ríos. Facultad de Bromatología. Gualeguaychú, Argentina.

⁵ sangos@gmail.com

Abstract. This study assesses the effectiveness of automatically classifying pollen types relevant to apiculture using convolutional neural networks. Using samples from ten bee-foraged plant species from Buenos Aires (Argentina), non-acetolyzed hydrated pollen specimens were prepared and photographed under an optical microscope at a medium magnification ($10\times$ objective lens). A segmentation algorithm was developed to extract individual pollen grain images, generating two independent datasets. Preliminary non-acetolyzed results with the ResNet18 network show 90% accuracy for grayscale images versus 63% for color images. This research aims to optimize the automatic identification of the floral origin of honeys from Buenos Aires using low-complexity equipment.

Keywords: Pollen Classification, Convolutional Neural Networks, Melissopalynology, Automatic Image Extraction, ResNet18.

Clasificación Automática de Tipos Polínicos Apícolas

Resumen. Este trabajo evalúa la eficacia de clasificación automática de tipos polínicos apícolas mediante redes neuronales convolucionales. Utilizando muestras de diez especies de plantas apícolas de Buenos Aires (Argentina) se prepararon y fotografiaron especímenes polínicos hidratados sin acetolizar bajo microscopio óptico con un aumento mediano ($10\times$ lente objetivo). Se desarrolló un algoritmo de segmentación para extraer imágenes individuales de granos de polen, generando dos conjuntos de datos independientes. Los resultados preliminares con la red ResNet18 sin acetólisis muestran una precisión del 90% para imágenes en escala de grises versus 63% para imágenes en color. Esta investigación busca optimizar la identificación automática del origen floral de mieles bonaerenses utilizando equipamiento de baja complejidad.

Palabras clave: Clasificación de Polen, Redes Neuronales Convolucionales, Melisopalnología, Extracción Automática de Imágenes, ResNet18.

1. Introducción

Las propiedades fisico-químicas y organolépticas de la miel dependen de la flora visitada por las abejas que la elaboran. En la miel quedan en suspensión granos polínicos como testigos de las plantas visitadas. Por esta razón, el espectro polínico de cada miel constituye una huella digital de la vegetación circundante a la colmena y es

utilizado al momento de validar y certificar el origen floral de una miel (Loveaux et al., 1978).

Los estudios melisopalínológicos para validar este origen floral demandan de operadores muy calificados que puedan identificar al microscopio los tipos polínicos presentes en las muestras analizadas. Además, estos operadores deben pasar horas realizando las observaciones y conteos manuales de cada uno de los tipos polínicos presentes. Un procedimiento habitual en estos estudios es someter a las muestras de polen a la acetólisis: una digestión química agresiva que barre con toda la materia orgánica a excepción de la cobertura de los granos (Lieux, 1980). Esto facilita la observación de detalles estructurales sin la interferencia del contenido celular o de aceites asociados. Por contrapartida, es un tratamiento que consume tiempo y reactivos y que debe realizarse en estrictas condiciones de seguridad, y que puede eliminar otros caracteres informativos asociados. Otro procedimiento estándar en estudios melisopalínológicos consiste en hidratar los granos de polen en un medio de gelatina-glicerina o solo glicerol. Este es un método simple pero efectivo en el cual se obtienen grandes cantidades de granos de polen sin desintegrar el contenido celular, el cual queda dentro de las paredes del grano de polen y que se expande ligeramente al ser hidratado. Frecuentemente, en algunas especies se puede visualizar una porción de este contenido sobresaliendo por las aberturas del polen, lo cual presenta una herramienta útil para la identificación (Koelzer et al., 2024). En los últimos años se han realizado ensayos con diferentes enfoques de aprendizaje automático para la detección, identificación y clasificación de tipos polínicos de interés económico y ambiental (Viertel & Konig, 2022). La mayor parte de estos ensayos se ha realizado utilizando imágenes de los tipos polínicos obtenidas con lentes objetivo de elevados aumentos ($40\times$ a $100\times$) a diferentes alturas de foco y un posterior apilamiento (*z-stacking*) y procesado (Olsson et al., 2021; Polling et al., 2021). La captura de estas imágenes demanda de equipamiento automatizado de elevado costo y baja disponibilidad en laboratorios convencionales.

El objetivo de esta línea de trabajo es indagar en la performance de modelos de aprendizaje automáticos (en particular, redes neuronales convolucionales o RNC) respecto a la clasificación de muestras palinológicas de importancia apícola hidratadas y no acetolizadas, fotografiadas a aumentos de mediana magnitud y sin apilamiento posterior utilizando equipamiento de microscopía normalmente disponible en laboratorios convencionales. Para ello se tomaron muestras polínicas de flora apícola de Buenos Aires y con ellas se realizaron preparados microscópicos sin acetolizar. Estos preparados microscópicos fueron fotografiados bajo el microscopio óptico de campo claro con una lente objetivo $10\times$.

Este trabajo pretende llenar un vacío preexistente en cuanto a la clasificación automática de tipos polínicos encontrados en mieles bonaerenses utilizando material e instrumentos de laboratorio de baja complejidad.

2. Metodología

En distintas regiones de la provincia de Buenos Aires se tomaron muestras polínicas de diez especies de plantas forrajeadas intensamente por abejas melíferas: *Baccharis* sp., *Brassica oleracea*, *Cirsium vulgare*, *Echium plantagineum*, *Eucalyptus* sp., *Helminthotheca echooides*, *Ligustrum lucidum*, *Lotus* sp., *Solanum sisymbriifolium* y

Trifolium pratense. Para cada especie, se tomaron muestras de dos ejemplares de distinta localidad. Cada muestra consiste en polen reservado en etanol al 70% proveniente de anteras de un ejemplar de una misma especie. De allí se tomó una alícuota, se centrifugó a 3000 rpm durante 6 minutos en un tubo eppendorf de 1,5 ml, se descartó el sobrenadante y se resuspendió el polen en una gota de alcohol al 70%. Se tomó la suspensión de polen con una pipeta y se realizó un preparado microscópico montado sobre gelatina-glicerina teñida con fucsina (Zarlavsky, 2014) sellado con parafina (Fig. 1).

A partir de estos preparados se tomaron imágenes de microscopía de campo claro utilizando un microscopio Zeiss Ax10 con un lente objetivo Zeiss LD A-Plan 421241-9910 de apertura numérica 0.25 y aumento 10× acoplado a una cámara digital Zeiss Axiocam 208 con sensor CMOS color de 3840×2160 pixeles. Se barrió el preparado manualmente tomando aproximadamente veinte fotografías por muestra. Cada una de estas fotografías abarcaba un campo visual equivalente a 1,13×0,63 mm de la muestra, conteniendo al menos diez granos de polen en distintas orientaciones (Fig. 2A).

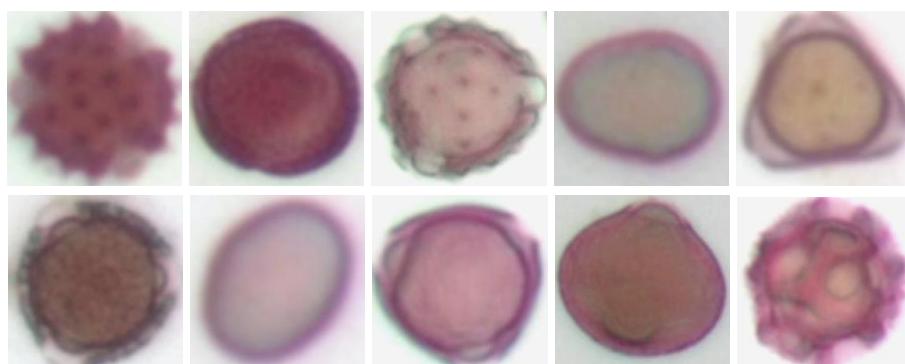


Fig. 1. Granos de polen no acetolizados de *Baccharis* sp., *Brassica oleracea*, *Cirsium vulgare*, *Echium plantagineum*, *Eucalyptus* sp., *Ligustrum lucidum*, *Lotus*, *Solanum sisymbriifolium*, *Trifolium pratense* y *Helminthotheca echiooides*, respectivamente. Los granos se encuentran en su estado hidratado y fueron teñidos con fucsina básica.

2.1 Segmentación y extracción de imágenes

Se trajeron automáticamente imágenes de los granos de polen individuales presentes en cada fotografía, procurando obtener al menos 200 granos por ejemplar (Fig. 2). Para ello, desarrollamos un algoritmo que realiza segmentación y extracción de objetos, basado en una serie de operaciones utilizando la librería Image Processing Toolbox de MATLAB. Partiendo de una imagen a color en formato TIFF, se convirtió la imagen a escala de grises y se la suavizó utilizando un filtro gaussiano con desvío estándar proporcional al tamaño del grano de polen (parámetro ajustable). Para detectar los bordes de cada objeto, se aplicó una máscara de gradiente utilizando el operador "Sobel", obteniendo así una imagen binaria. Para mejorar la definición de los bordes, se realizaron operaciones morfológicas de dilatación y posterior erosión,

usando elementos estructurantes con forma de disco. Luego se llenó la imagen binaria y se filtraron los objetos de acuerdo al tamaño del grano de polen.

Finalmente, se extrajeron los componentes conectados y se utilizaron centroides para obtener cajas delimitadoras alrededor de cada objeto, y de este modo extraerlos de la imagen original. Se guardaron las imágenes individuales en JPG, después de reescalarlas a un tamaño adecuado para la red neuronal a entrenar, en nuestro caso 224×224 píxeles (Fig. 1). Dado que el reescalado iguala el tamaño entre los granos de polen, la información sobre el tamaño de los distintos tipos polínicos no forma parte de nuestro análisis.

Se analizaron las imágenes obtenidas y se descartaron los granos fuera de foco, rotos, superpuestos, colapsados o contaminados. La extracción se realizó de forma automatizada para todas las imágenes tomadas de un mismo ejemplar.

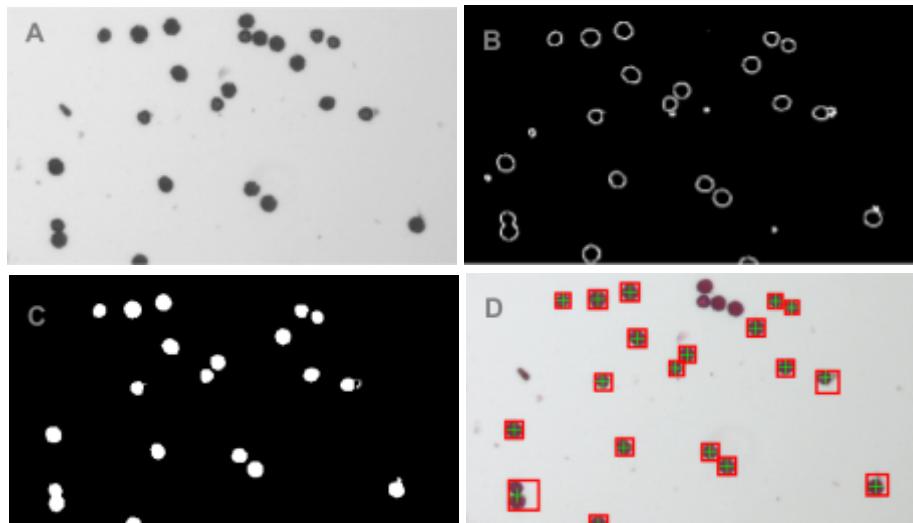


Fig. 2. Esquema de extracción de imágenes. **A.** Conversión a escala de grises y suavizado. **B.** Máscara de gradiente y dilatación. **C.** Imagen binaria luego de, llenar, erosionar y aplicar filtro de tamaño. **D.** Extracción de componentes conectados y recorte de cada grano en la imagen original.

2.2 Generación del conjunto de datos

Para cada especie, se asignaron los granos de polen de un ejemplar al conjunto de entrenamiento y los del otro ejemplar al conjunto de validación (Fig. 3). Esta división por origen del ejemplar, en lugar de una separación aleatoria, permite evaluar de forma más rigurosa la capacidad de generalización del modelo (Olsson et al., 2021). Al enfrentarse a ejemplares completamente nuevos durante la validación, comprobamos que el modelo aprende las características morfológicas distintivas de cada tipo polínico y no las particularidades del contexto microscópico (como patrones de fondo, iluminación o condiciones de preparación específicas de cada muestra).

Para determinar entre los dos ejemplares cuál sería asignado al conjunto de entrenamiento y cuál al de validación, se optó por asignar el ejemplar del cuál se obtuvo la mayor cantidad de imágenes individuales de granos de polen al conjunto de entrenamiento, y utilizar las imágenes obtenidas del otro ejemplar para validar. Para ello se cuantificó la cantidad de imágenes luego de extraer y limpiar los datos, conservando únicamente los granos que se encontraban en buen estado y libres de suciedad, procurando obtener un mínimo de 200 granos en el caso del conjunto de entrenamiento, y un mínimo de 50 granos en el caso del conjunto de validación. En ambos casos se buscó abarcar una amplia variabilidad en cuanto a la orientación de los granos de polen. Cada especie fue considerada como una clase independiente a clasificar por la red neuronal.

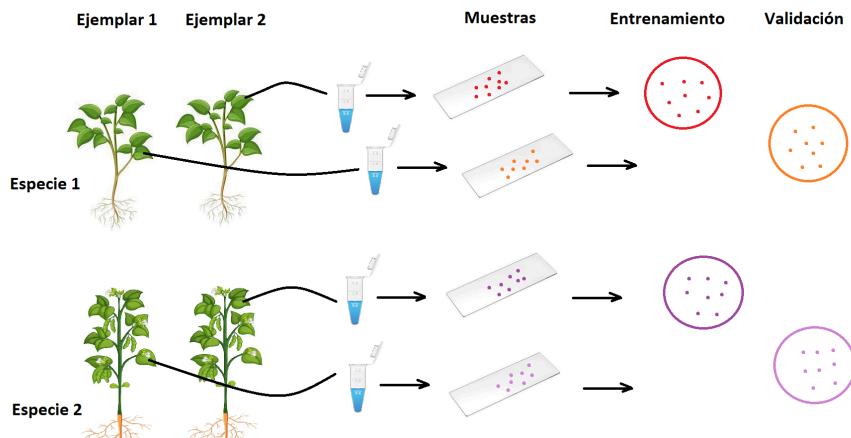


Fig. 3. Esquema de partición de datos. Se muestra la separación por ejemplares para cada especie, donde las muestras del ejemplar 1 se destinan al conjunto de datos de entrenamiento y las del ejemplar 2 al conjunto de datos de validación.

2.3 Entrenamiento de la red neuronal convolucional ResNet18

En una computadora Intel Core i5 con 16 GB RAM y placa de video GPU NVIDIA GeForce RTX con 12 GB VRAM, se corrieron entrenamientos de RNC en Python usando aprendizaje por transferencia con el modelo ResNet18 de la librería PyTorch (He et al., 2016). Esta es una RNC con 18 capas de profundidad pre-entrenada con más de un millón de imágenes. Para ello se utilizaron todas las imágenes de los granos del conjunto de entrenamiento, con un rango de 200 a 1400 imágenes por clase, contando con un total de 6411 imágenes. Se entrenó durante un total de 10 épocas, la tasa de aprendizaje se fijó a 0.0001, y el tamaño de lote fue 32. Durante el entrenamiento, se utilizó la función de pérdida de entropía cruzada para optimizar la clasificación (Terven et al., 2023).

Para abordar el desequilibrio de clases y mejorar la generalización del modelo, implementamos una estrategia de aumento de datos al vuelo combinada con muestreo ponderado (Johnson et al., 2019). Durante el entrenamiento, cada imagen fue

sometida a transformaciones aleatorias únicas (incluyendo rotaciones, reflexiones y variaciones de brillo y contraste dentro de un rango definido) cada vez que era seleccionada a través de diferentes épocas, creando diversas variaciones de las imágenes originales. Empleamos un enfoque de muestreo aleatorio ponderado donde los pesos de las clases se calcularon de manera inversamente proporcional a la frecuencia de clase (1/recuento de clase), asegurando una representación equilibrada entre todas las especies de polen independientemente de su distribución original en el conjunto de datos. Este enfoque combinado busca aumentar la variabilidad de nuestros datos de entrenamiento y evitar que el modelo esté sesgado hacia las clases mayoritarias, con cada clase recibiendo aproximadamente la misma representación a lo largo del proceso de entrenamiento. Luego se usó el conjunto de muestras con el que no se entrenó para validar.

Se realizaron dos entrenamientos independientes, uno usando las imágenes en escala de grises y otro en color. Para adaptar el conjunto de datos al modelo ResNet18, que requiere imágenes RGB, se convirtieron las imágenes originales en color a escala de grises, y luego se replicó el valor de cada píxel en los tres canales (rojo, verde y azul). De este modo se mantuvo la información deseada mientras se cumplía con el formato de entrada requerido por la red neuronal.

2.4 Métricas

El rendimiento de ambas RNC (entrenada con imágenes en escala de grises o en color) se estimó a partir de la validación con imágenes propias del conjunto de validación. Para ello se calcularon las siguientes métricas (Viertel & Konig, 2022). Además, para cada clasificación se realizó una matriz de confusión para indagar entre qué tipos polínicos ocurren las principales confusiones.

$$\text{Exactitud} = \frac{\text{Verdaderos positivos} + \text{Verdaderos negativos}}{\text{Total}} \quad (1)$$

$$\text{Precisión} = \frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{Falsos positivos}} \quad (2)$$

$$\text{Tasa de recuperación} = \frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{Falsos negativos}} \quad (3)$$

$$F1 = \frac{2 * \text{Precisión} * \text{Tasa de recuperación}}{\text{Precisión} + \text{Tasa de recuperación}} \quad (4)$$

Las ecuaciones 2, 3 y 4 al ser estimadores de clasificadores binarios se calcularon para cada tipo polínico en particular y se promediaron entre todos los tipos polínicos para referirse a una clasificación con múltiples categorías como es una muestra polínica multiespecie.

3. Resultados

Los entrenamientos de la ResNet18 demoraron 10 minutos para los sets a color y blanco y negro, con 6411 imágenes de entrenamiento en ambos casos. Para los ensayos con la red ResNet18 se obtuvo una precisión de 63% en el caso de imágenes en color y de 90% en el caso de usar las imágenes en escala de grises.

El entrenamiento de la RNC con el dataset en color mostró una curva de aprendizaje con valores cercanos a 99%, mientras que para la curva de validación se observó una estabilización en un 63% (Fig. 4A). La matriz de confusión (Fig. 5) indica una elevada precisión de clasificación para *Baccharis sp.*, *Cirsium vulgare*, *Echium plantagineum*, *Helminthotheca echioidea*, *Ligustrum lucidum* y *Solanum sisymbriifolium* (mayor a 95%), y una baja precisión para *Brassica oleracea*, *Eucalyptus sp.*, *Lotus sp.* y *Trifolium pratense*. *Lotus sp.* es confundida frecuentemente con *Echium plantagineum* y que hay una confusión entre *Brassica oleracea* y *Trifolium pratense*.

El entrenamiento de la RNC con el dataset en escala de grises mostró una curva de aprendizaje con valores cercanos a 99%, mientras que para la curva de validación se observó una estabilización en un 90% (Fig. 4B). La matriz de confusión (Fig. 6) muestra la capacidad de distinción entre clases, destacando cierta confusión entre *Lotus sp.* y *Echium plantagineum*, la cual puede ser debida a similitudes morfológicas de los granos de polen. La mayoría de las especies presenta una elevada precisión de clasificación (mayor a 85%).

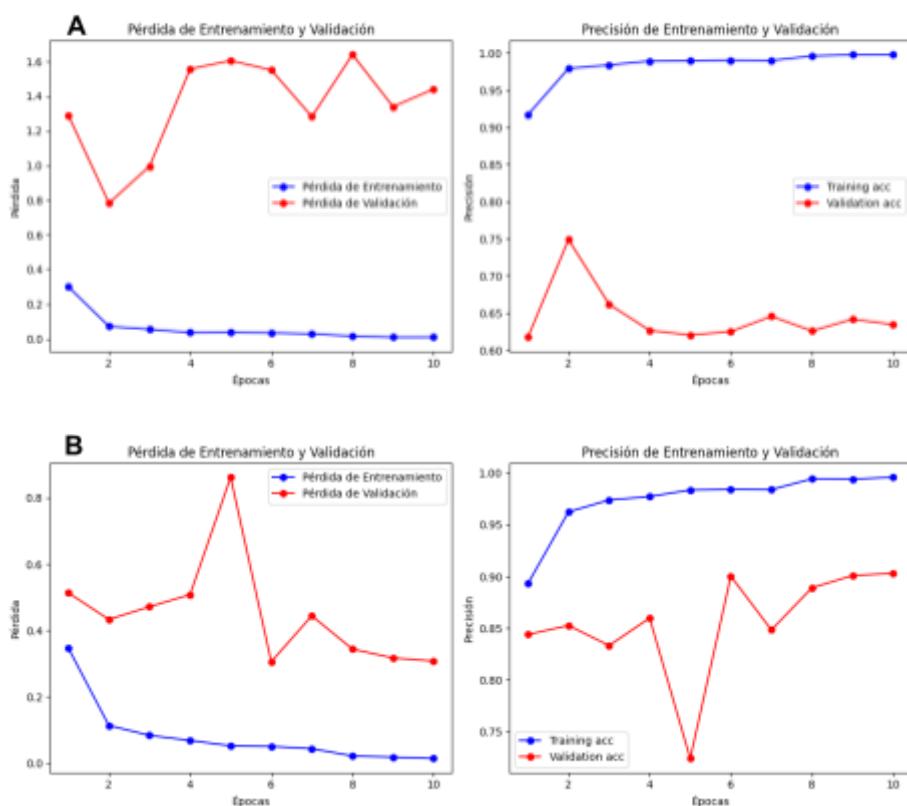


Fig. 4. A. Ensayo ResNet18 con imágenes en color. Precisión: 63%. **B.** Ensayo ResNet18 con imágenes en escala de grises. Precisión: 90%

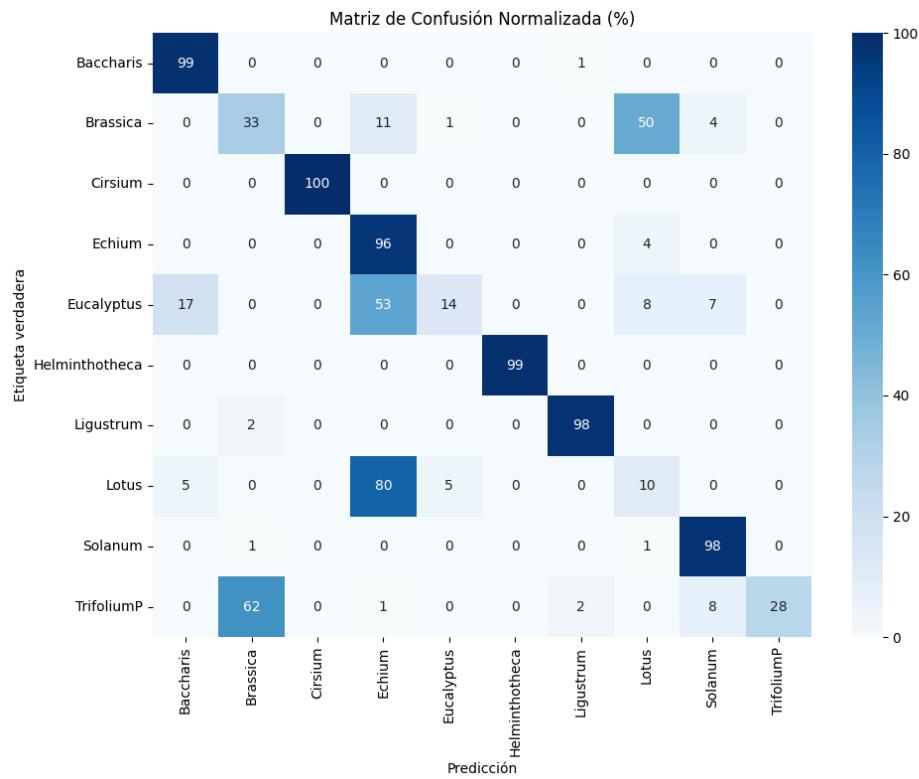


Fig. 5. Matriz de confusión para los tipos polínicos resultante del ensayo en color. Las especies verdaderas se muestran en las columnas, y las especies predichas se muestran en las filas. El color indica la frecuencia relativa en cada celda, de acuerdo a la leyenda insertada a la derecha.

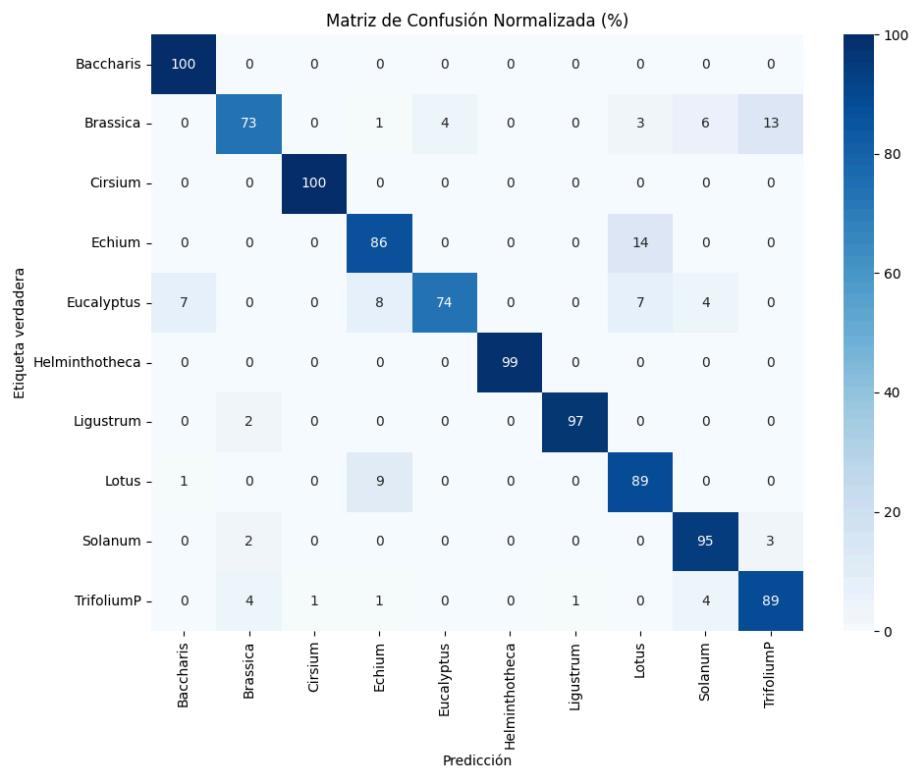


Fig. 6. Matriz de confusión para los tipos polínicos resultante del ensayo en escala de grises.

Tabla 1. Métricas de rendimiento de las RNC.

	Gris		Color	
	Entrenamiento	Validación	Entrenamiento	Validación
Exactitud	0,9982	0,9032	0,9996	0,6347
Precisión	0,9982	0,9100	0,9996	0,6590
Tasa de recuperación	0,9982	0,9032	0,9996	0,6347
F1	0,9982	0,9036	0,9996	0,6001

4. Discusiones

Los resultados obtenidos muestran que las redes neuronales convolucionales, específicamente ResNet18, pueden clasificar eficazmente imágenes de granos de polen hidratados de diez especies de interés apícola de Buenos Aires, Argentina, obtenidas con un *set up* disponible en laboratorios convencionales. La implementación de un enfoque de clasificación automática utilizando aprendizaje por transferencia y técnicas de aumentación y balanceo de datos alcanzó una precisión elevada, particularmente cuando se utilizaron imágenes en escala de grises. A su vez, se puso a prueba la capacidad de generalización del modelo mediante el uso de ejemplares independientes para el entrenamiento y la validación, evitando el sobreajuste a características particulares de los granos de polen con los que entrenó el modelo (Olsson et al., 2021).

La capacidad de clasificación del modelo entrenado con imágenes en escala de grises fue significativamente mayor a la del modelo entrenado con imágenes en color, con una precisión de 90% y 63% respectivamente. Esto puede atribuirse a que la conversión a escala de grises reduce la variabilidad cromática introducida por la tinción con fucsina, que puede presentar intensidades y tonalidades variables entre preparados microscópicos distintos. Al eliminar esta fuente de variación, la cual puede ser poco informativa, el modelo podría concentrarse en características morfológicas más estables como la forma y los patrones de ornamentación (Daood et al., 2016).

En cuanto a las principales confusiones en la clasificación automática de estos tipos polínicos, la confusión observada entre *Lotus* sp. y *Echium plantagineum*, particularmente en el modelo de escala de grises, puede atribuirse a similitudes morfológicas. Ambos tipos polínicos presentan formas similares cuando se visualizan en ciertas orientaciones, lo cual resalta la importancia de incluir diversas orientaciones de los granos en el conjunto de datos de entrenamiento, sobre todo para especies con morfologías similares. Además, dado que ambos granos de polen son los más pequeños del conjunto de clases, al observarlos bajo un microscopio óptico con lente objetivo de 10×, la baja resolución óptica puede dificultar distinguir las diferencias en la ornamentación ya que los detalles finos de la exina pueden no ser completamente visibles (Halbritter et al., 2018).

En el caso del modelo entrenado con imágenes a color, la baja precisión en la clasificación de *Brassica oleracea*, *Eucalyptus* sp., *Lotus* sp. y *Trifolium pratense*, podría estar relacionada con similitudes en la respuesta a la tinción con fucsina o con variaciones en la intensidad de la tinción entre muestras. Las características visuales de nuestras imágenes en color podrían no estar proporcionando suficiente información, o proveyendo demasiado ruido, para que el modelo distinga los granos fácilmente, a diferencia de las imágenes en escala de grises, donde se reduce la complejidad visual de la imagen permitiendo resaltar características que aumentan la capacidad de reconocimiento del modelo (Bui et al., 2016). Aplicar técnicas de procesamiento de imágenes para homogeneizar el conjunto de datos y disminuir la variabilidad debida a las condiciones del contexto microscópico o de iluminación podría mejorar el rendimiento del modelo.

Además, a pesar del elevado rendimiento del modelo en comparación con otras experimentos clasificatorios (Viertel & Konig, 2022), es importante remarcar que el conjunto de datos utilizado comprende diez especies de plantas, lo que representa solo una pequeña fracción de la diversidad floral disponible para las abejas en la región bonaerense. Para desarrollar una herramienta de aplicación práctica en análisis de polen apícola regional sería necesario ampliar la palinoteca aumentando considerablemente el número de especies y el número de ejemplares muestreados por especie.

Nuestro trabajo está basado en imágenes de microscopía de campo claro de granos de polen hidratados teñidos con fucsina (Koelzer et al., 2024), uno de los métodos más rápidos y sencillos para procesar muestras palinológicas. Si bien hay mucho margen de mejora, podemos concluir que utilizando equipamiento de baja complejidad se puede lograr una clasificación automática de tipos polínicos con alta precisión y eficiencia de forma accesible.

Bibliografía

- Bui, Hieu & Lech, Margaret & Cheng, Eva & Neville, Katrina & Burnett, Ian. (2016). Using grayscale images for object recognition with convolutional-recursive neural network. 321–325. 10.1109/CCE.2016.7562656.
- Daood, A., Ribeiro, E., & Bush, M. (2016). Pollen Grain Recognition Using Deep Learning. Lecture Notes in Computer Science, 321–330. doi:10.1007/978-3-319-50835-1_30.
- He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. (2016). "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778.
- Halbritter, Heidemarie & Silvia, Ulrich & Grímsson, Friðgeir & Weber, Martina & Zetter, Reinhard & Hesse, Michael & Buchner, Ralf & Svojtka, Matthias & Frosch-Radivo, Andrea. (2018). Illustrated Pollen Terminology. 10.1007/978-3-319-71365-6.
- Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. Journal of Big Data, 6(1). doi:10.1186/s40537-019-0192-5.
- Koelzer, K., Ribarits, A., & Weber, M. (2024). Comparing the acetolysed and hydrated methods for the pollen analysis of honey. Grana. <https://doi.org/10.1080/00173134.2024.2348651>.
- Lieux MH. (1980). Acetolysis applied to microscopical honey analysis. Grana 19: 57–61.
- Louveaux J, Maurizio A, Vorwohl G. (1978). Methods of Melissopalynology. Bee World 59: 139–157.
- Olsson, O., Karlsson, M., Persson, A. S., Smith, H. G., Varadarajan, V., Yourstone, J., & Stjernman, M. (2021). Efficient, automated, and robust pollen analysis using deep learning. Methods in Ecology and Evolution, 12(5), 850–862. <https://doi.org/10.1111/2041-210X.1357>.
- Polling, M., Li, C., Cao, L., Verbeek, F., de Weger, L. A., Belmonte, J., De Linares, C., Willemse, J., de Boer, H., & Gravendeel, B. (2021). Neural networks for increased accuracy of allergenic pollen monitoring. Scientific Reports, 11(1), 1–11. <https://doi.org/10.1038/s41598-021-90433-x>.
- Terven, Juan & Cordova-Esparza, Diana-Margarita & Ramirez-Pedraza, Alfonzo & Chávez Urbiola, Edgar. (2023). Loss Functions and Metrics in Deep Learning. A Review. 10.48550/arXiv.2307.02694.
- Viertel, P., König, M. (2022). Pattern recognition methodologies for pollen grain image classification: a survey. Machine Vision and Applications 33, 18. <https://doi.org/10.1007/s00138-021-01271>.

Zarlavsky, G. E. (2014). Histología Vegetal: técnicas simples y complejas. Sociedad Argentina de Botánica, Buenos Aires, 198.