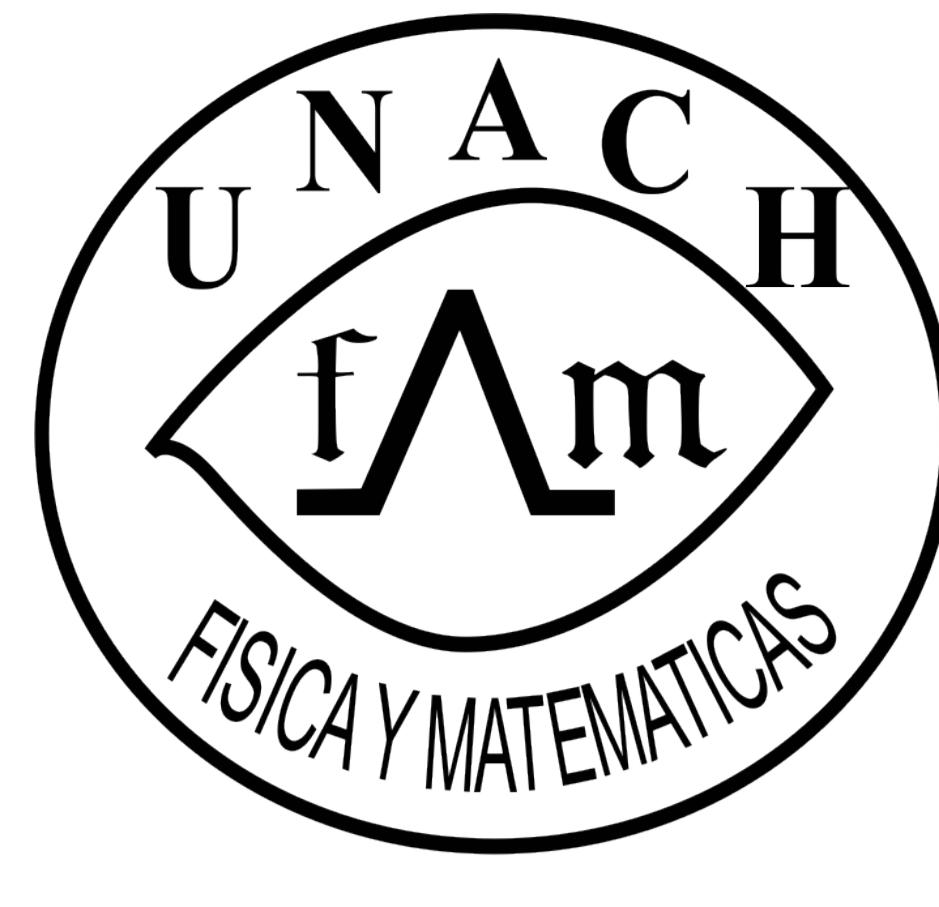


COMPARACIÓN DE MODELOS DE APRENDIZAJE MÁQUINA PARA LA IDENTIFICACIÓN DE PATRONES EN IMÁGENES MONOCROMÁTICAS.

Facultad de Ciencias en Física y Matemáticas
Benemerita Universidad Autónoma de Chiapas (UNACH)
Autor: Mariana Belén Cruz Rodríguez.

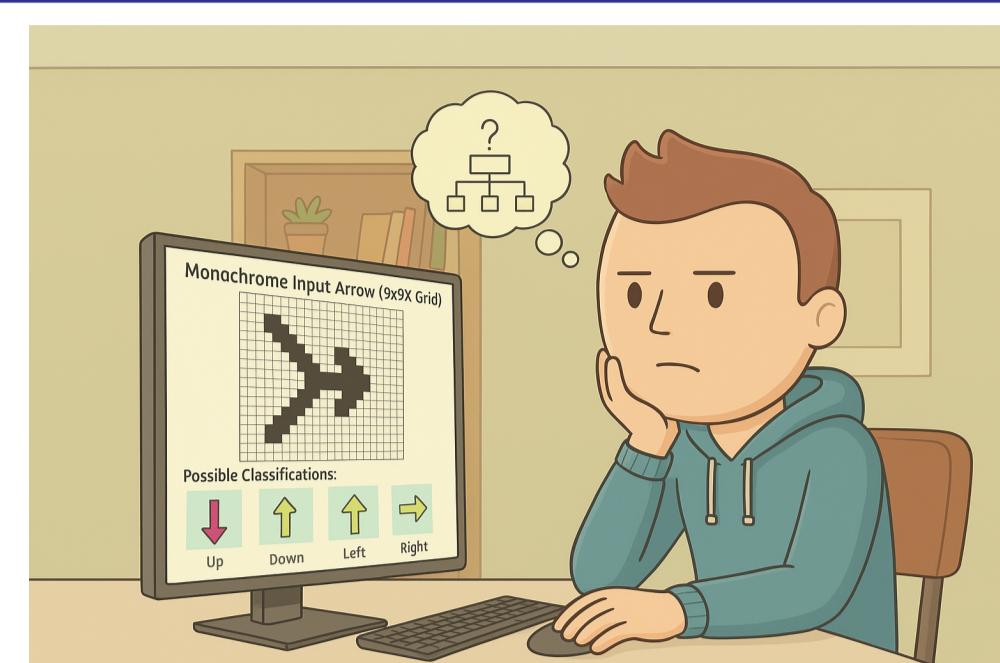


1. Resumen

Este trabajo compara dos enfoques complementarios para la clasificación automática de flechas (arriba, abajo, izquierda y derecha) en imágenes monocromáticas de 9×9 píxeles.

El primero combina Modelos Ocultos de Markov (HMM) con un esquema bootstrap para mejorar la estimación de parámetros y cuantificar la incertidumbre de clasificación.

El segundo emplea Redes Neuronales Convolucionales 2D (CNN) diseñadas para procesar imágenes pequeñas con baja resolución. Se analizan la precisión, estabilidad y requerimientos computacionales de ambos métodos bajo un mismo protocolo experimental.



2. Introducción

Las imágenes, al ser procesadas como matrices, pueden analizarse para identificar patrones, bordes y texturas.

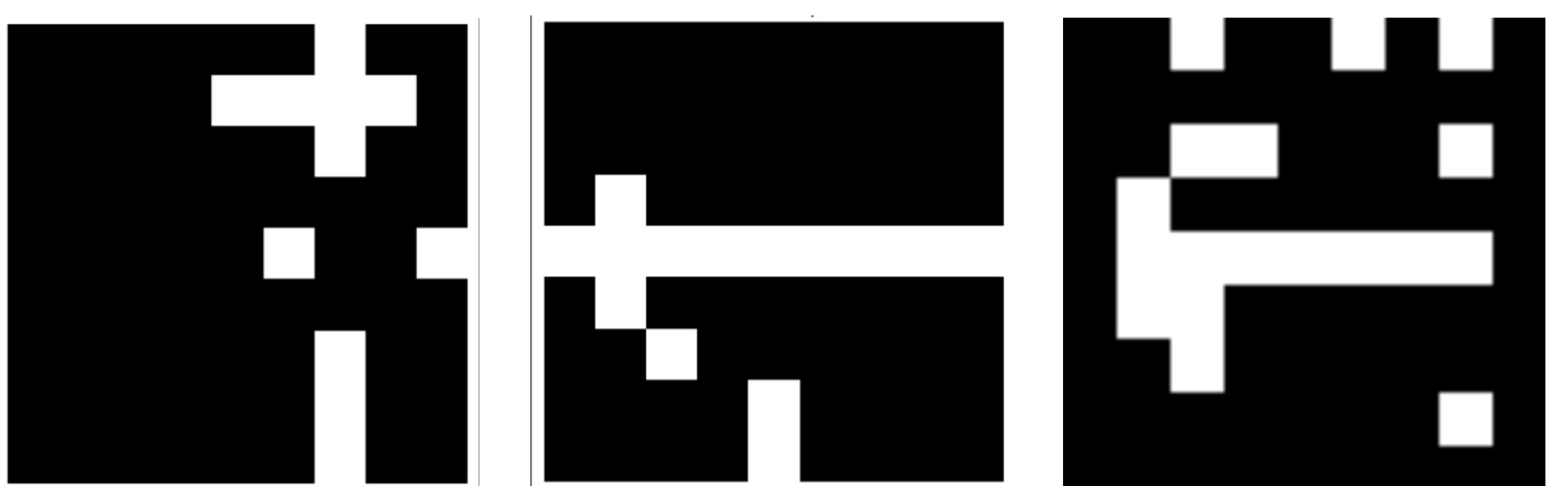
En particular, al analizar imágenes monocromáticas se pueden trasladar estas al campo matemático mediante el uso de matrices binarias (0,1).

Uno de los métodos clásicos para esto es el Modelo Oculto de Markov (HMM), que modela la imagen de forma secuencial. Para robustecerlo, se emplea la técnica de bootstrap, un método de remuestreo que estima la variabilidad y fiabilidad estadística del modelo.

Por otro lado, las Redes Neuronales Convolucionales (CNN) adoptan un enfoque espacial, utilizando filtros que recorren la imagen para aprender representaciones jerárquicas de las características.

3. Objetivo

Determinar cuál enfoque ofrece mejor equilibrio entre interpretabilidad, precisión y costo computacional para la clasificación de imágenes con ruido en diferentes categorías.



4. Metodología

Primero es necesaria la creación de los patrones base, es decir las flechas binarias base $9 \times 9 (\uparrow\downarrow\leftarrow\rightarrow)$, definidas mediante matrices binarias (0,1).

Se genera un dataset que consta de 600 imágenes sintéticas introduciendo un ruido del 5%. De ellas, 500 se usan para entrenamiento y 100 para prueba.

El primer modelo implementado es HMM + Bootstrap, empleando Baum-Welch, 500 remuestreos y decodificación Viterbi.

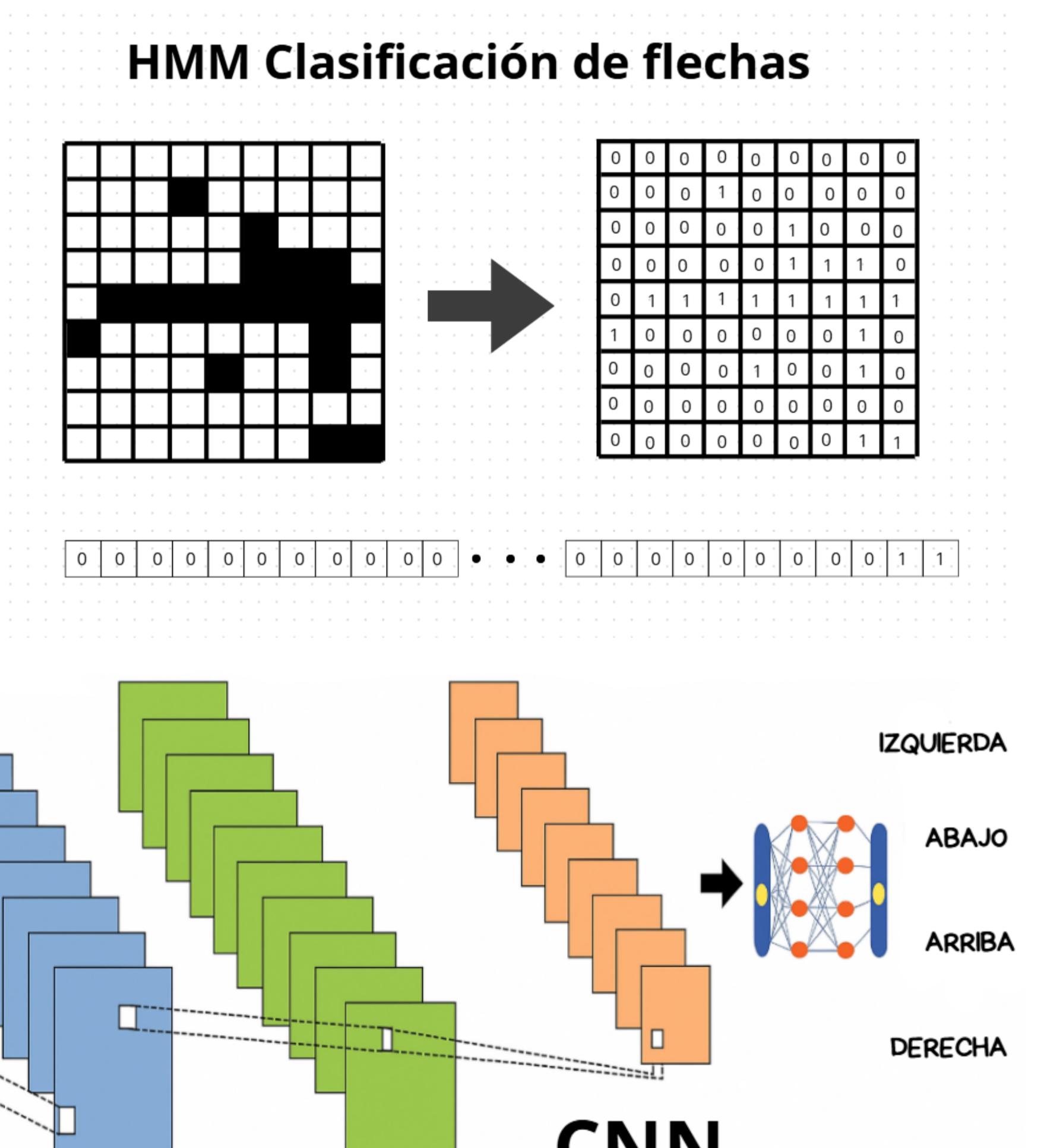
El bootstrap es un método estadístico de remuestreo con reemplazo que permite estimar la variabilidad y la incertidumbre de un estimador o modelo a partir de una única muestra de datos, generando múltiples réplicas del conjunto original para obtener una distribución empírica de los parámetros o métricas de interés.

Por otro lado, el algoritmo de Baum-Welch es un método iterativo de máxima verosimilitud utilizado para entrenar Modelos Ocultos de Markov. Se basa en el algoritmo EM (Expectation-Maximization) y ajusta las probabilidades de transición y emisión del modelo para maximizar la probabilidad de las secuencias observadas.

La CNN 2D se compone de 2 capas convolucionales 3×3 (ReLU (Rectified Linear Unit) + MaxPooling), seguidas de una capa densa + Softmax (4 clases), optimizador Adam (0.001), *early stopping* y 100 épocas.

Las métricas consideradas son:

Accuracy y tiempo de entrenamiento, para el caso de CNN, también se analiza la función de pérdida que es una medida cuantitativa que evalúa la discrepancia entre las salidas predichas y las reales, y cuya minimización orienta el ajuste de los parámetros del modelo durante el entrenamiento.



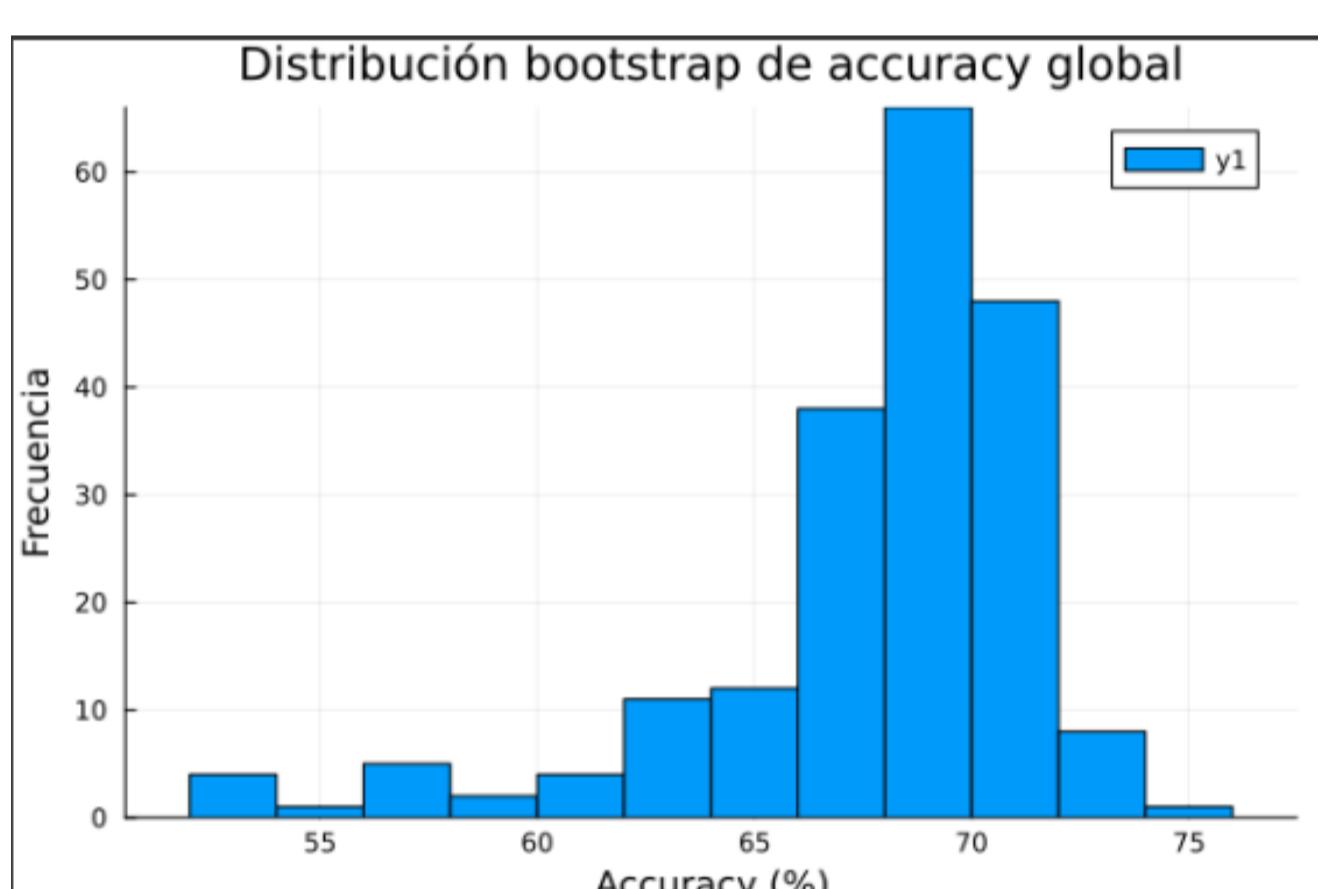
6. Conclusión

De los resultados obtenidos se observa la gran ventaja que presentan las CNN frente a métodos clásicos como el HMM, tanto en precisión como en tiempo de ejecución y recursos computacionales. Mientras que el HMM mostró un desempeño limitado al tratar con la complejidad espacial de las imágenes bidimensionales, la CNN aprovechó de manera eficiente la información espacial y las relaciones locales entre píxeles, logrando una mejor representación de las características relevantes para la clasificación. **Trabajo futuro:** Es importante saber que hay muchos trabajos por realizar, es posible explorar arquitecturas de aún más complejas o modelos híbridos que combinen características secuenciales y espaciales para mejorar el rendimiento sin sacrificar la interpretabilidad, esta vez aplicándolo a imágenes más complejas y en RGB.

5. Resultados

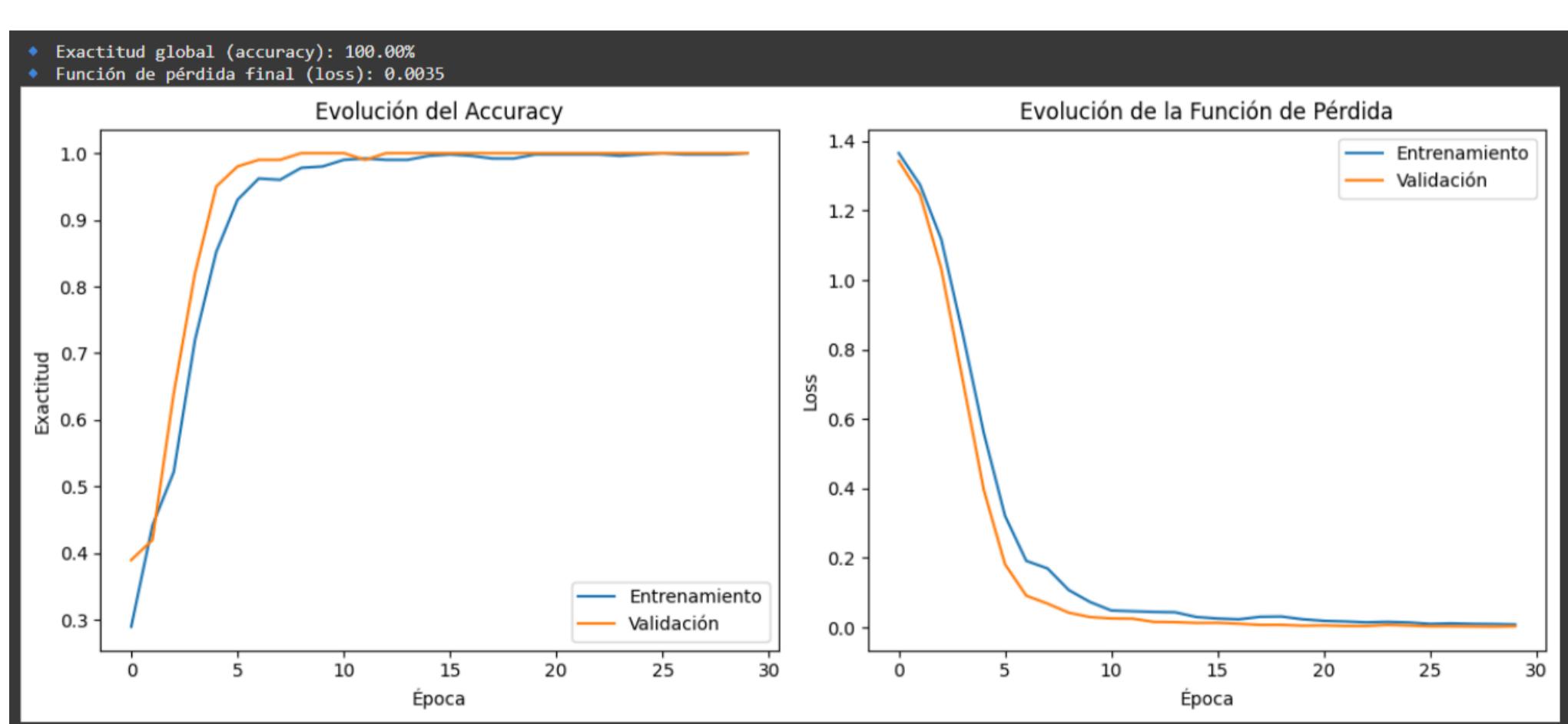
Al realizar la prueba del HMM, el rendimiento del modelo es el siguiente: Accuracy 68%

- Clase 1: correct=54 / total=100 → 54.0%
- Clase 2: correct=38 / total=100 → 38.0%
- Clase 3: correct=89 / total=100 → 89.0%
- Clase 4: correct=91 / total=100 → 91.0%



La implementación tardó considerablemente más debido a los remuestreos del bootstrap, dando así un tiempo aproximado de 15 minutos

Por otro lado, para el modelo CNN se obtuvo una precisión del 100%, con una función de pérdida de 0.0035, logrando el entrenamiento en tan solo 25 segundos



7. Referencias

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 1994.