



El variograma como herramienta de caracterización

| The Variogram as a Tool for Characterization |

| O variograma como ferramenta de caracterização |

 **Katalina Oviedo Rodríguez**

katalina.oviedo.rodriguez@una.cr

Universidad Nacional
Heredia, Costa Rica

 **Alda Carvalho**

alda.carvalho@uab.pt

DCeT, Universidade Aberta Lisboa
Lisboa, Portugal

Recibido: 14 de julio del 2025

Aceptado: 30 de abril del 2026

Resumen: En este artículo se explora el uso del variograma experimental como descriptor para la extracción de características en imágenes. Esta caracterización se evalúa mediante dos modelos de clasificación: una red neuronal completamente conectada (FCNN) y máquinas de vectores soporte (SVM). Para la validación experimental se emplea la base de datos MNIST, ampliamente utilizada como referencia en problemas de clasificación de imágenes. El estudio considera variogramas experimentales direccionales calculados en las orientaciones de 0° , 45° y 90° , así como el variograma omnidireccional como línea base de comparación. Con el fin de explotar la información direccional, se analizan dos estrategias de combinación: una fusión tardía basada en un esquema de votación ponderada de las probabilidades generadas por modelos independientes, y una fusión temprana mediante la concatenación de las características direccionales. Los resultados permiten evaluar la capacidad del variograma para reducir la dimensionalidad del problema y, al mismo tiempo, preservar información espacial relevante para la clasificación.

Palabras Clave: Variograma experimental, MNIST, redes neuronales, votación ponderada.

Abstract: This paper explores the use of the experimental variogram as a descriptor for feature extraction in images. This characterization is evaluated using two classification models: a fully connected neural network (FCNN) and support vector machines (SVM). The MNIST dataset, which is widely used as a benchmark in image classification tasks, is employed for experimental validation. The study considers directional experimental variograms computed at orientations of 0° , 45° , and 90° , as well as the omnidirectional variogram as a baseline for comparison. In order to exploit directional information, two combination strategies are analyzed: a late fusion approach based on a weighted voting scheme applied to the class probability outputs of independent models, and an early fusion approach based on the concatenation of directional feature vectors. The results allow assessing the ability of the variogram to reduce the dimensionality of the problem while preserving spatial information that is relevant for classification.

¹Katalina Oviedo Rodríguez, académica de la Escuela de Matemática de la Universidad Nacional de Costa Rica. Dirección postal: Mercedes Norte, Heredia, Costa Rica. Código postal: 40102. Correo electrónico: katalina.oviedo.rodriguez@una.cr

²Alda Carvalho, professora Auxiliar, Universidade Aberta, Lisboa, Portugal. Dirección postal: Rua Almirante Barroso, n.º 38, 1000-013 Lisboa, Portugal. Código Postal: 1000-013. Correo electrónico: alda.carvalho@uab.pt

Keywords: Experimental variogram, MNIST, neuronal network, weighted voting.

Resumo: Neste artigo investiga-se o uso do variograma experimental como descritor para extração de características em imagens. Essa forma de caracterização é analisada por meio de dois modelos de classificação: uma rede neural totalmente conectada (FCNN) e máquinas de vetores de suporte (SVM). Para a validação experimental, utiliza-se a base de dados MNIST, amplamente reconhecida como referência em tarefas de classificação de imagens. O estudo contempla variogramas experimentais direcionais calculados nas orientações de 0° , 45° e 90° , além do variograma omnidirecional, adotado como linha de base para comparação. Com o objetivo de explorar de maneira mais abrangente a informação direcional, são examinadas duas estratégias de combinação: uma fusão tardia, baseada em um esquema de votação ponderada das probabilidades produzidas por modelos independentes, e uma fusão precoce, realizada por meio da concatenação das características direcionais. Os resultados obtidos permitem avaliar em que medida o variograma contribui para a redução da dimensionalidade do problema, ao mesmo tempo em que preserva informações espaciais relevantes para a tarefa de classificação.

Palavras-chave: Variograma experimental, MNIST, redes neurais, votação ponderada.

1. Introducción

El análisis de imágenes en esta era es una herramienta de gran utilidad en muchas disciplinas, por mencionar algunas: medicina, biometría, geografía y medio ambiente (Chan et al., 2020; Ghorbanzadeh et al., 2022; Gómez-de-Mariscal et al., 2021; Nascimento & de Sousa Bezerra, 2021; Nazarkevych et al., 2021; Shen et al., 2017; Xia et al., 2022). Existen métodos robustos para la clasificación de imágenes, por ejemplo, las redes neuronales convolucionales (CNN por sus siglas en inglés) (Bharadiya, 2023). Las redes neuronales convolucionales están compuestas por diversas capas (entrada, capa convolucional, ReLU, agrupamiento, totalmente conectada, softmax (Bharadiya, 2023; Krishna & Kalluri, 2019)) para realizar la tarea de clasificación. Las CNN son un tipo de aprendizaje profundo, en este tipo de aprendizaje se extraen las características sin necesidad de interpretación explícita, y al mismo tiempo se realiza la clasificación (Krishna & Kalluri, 2019). Es común utilizar CNN para la clasificación de imágenes (Krishna & Kalluri, 2019). Una desventaja de este enfoque es que requiere un alto poder computacional y mucho tiempo de entrenamiento (Krishna & Kalluri, 2019).

Una de las etapas principales en el análisis es el preprocesamiento de las imágenes; esta etapa puede ser crucial para lograr una buena clasificación, pero también para mejorar la eficiencia computacional (Bansal et al., 2023; Maharana et al., 2022). Ahora bien, en el caso de que sea posible obtener las características de alguna otra manera, entonces se puede utilizar una red neuronal completamente conectada (FCNN por sus siglas en inglés) así como también el método de vectores soporte (SVM por sus siglas en inglés), para probar un tipo de características dadas por el variograma experimental de las imágenes.

El variograma es una función matemática de estadística espacial que mide la variabilidad de una distribución de una variable regionalizada (Journel & Huijbregts, 1978). El variograma experimental es una medida real basada en observaciones reales (Journel & Huijbregts, 1978). Con él se puede observar cómo varían los valores a medida que cambia la distancia entre ellos (ver ecuación 1). Es decir, el variograma experimental puede capturar características y estructuras espaciales importantes presentes en los datos (Journel & Huijbregts, 1978).

Por lo tanto, en este trabajo se pretende verificar si el variograma experimental puede realizar la extracción de características de imágenes para luego utilizar una FCNN u otra técnica de aprendizaje automático como el SVM. Existen trabajos donde se ha utilizado el variograma experimental con distintos objetivos, desde la detección de masas cancerígenas hasta la detección de texturas de materiales (Berberoğlu et al., 2010; Ericeira et al., 2013; Yang et al., 2012). Una ventaja que puede tener este enfo-

que es la reducción de la dimensión: por ejemplo, para imágenes de tamaño $n \times n$, se puede obtener un vector de características de tamaño menor o igual a n .

En este trabajo se utilizó la base de datos MNIST (Modified National Institute of Standards and Technology) (LeCun & Cortes, 2010), que contiene imágenes de dígitos manuscritos del 0 al 9. Se consideraron variogramas experimentales direccionales calculados en 0° , 45° y 90° , así como un variograma omnidireccional como referencia. Dado que cada dirección puede capturar distintas estructuras espaciales o perder información relevante, se analizaron estrategias para combinar la información direccional.

En particular, se evaluaron dos esquemas de fusión: una fusión tardía basada en un método de votación ponderada (weighted voting) (Zhou, 2021), donde se combinan las probabilidades generadas por clasificadores independientes, y una fusión temprana mediante la concatenación directa de las características obtenidas a partir de los variogramas direccionales.

Asimismo, se comparó el desempeño de estos enfoques direccionales con un esquema basado en el variograma omnidireccional, utilizado como línea base para evaluar el aporte de la información direccional en la clasificación.

Finalmente, es importante destacar que en este estudio se utilizan imágenes del conjunto MNIST debido a que constituye una base de datos ampliamente conocida y consistente para la evaluación de métodos de clasificación. Si bien las redes neuronales convolucionales han demostrado un alto rendimiento en esta tarea, el objetivo de este trabajo no es competir con dichos enfoques, sino resaltar el valor del variograma experimental como herramienta de caracterización. El interés principal radica en demostrar que, a partir de una representación de baja dimensión basada en la variabilidad espacial, es posible capturar información relevante para la clasificación. Además, el uso del variograma no se limita al análisis de imágenes, sino que puede extenderse a otros contextos donde los datos presentan una distribución espacial.

El objetivo principal de este estudio es evaluar la utilidad del variograma experimental como descriptor de imágenes en tareas de clasificación, utilizando los modelos FCNN y SVM. Como objetivos específicos, se analiza el aporte de la información direccional mediante variogramas experimentales en tres orientaciones principales, se comparan estrategias de fusión temprana y tardía, y se contrasta el desempeño obtenido con un enfoque basado en el variograma omnidireccional.

2. Fundamentos teóricos y computacionales del enfoque propuesto

El enfoque propuesto se sustenta en dos fundamentos teóricos. En primer lugar, se emplean conceptos de estadística espacial, en particular el variograma experimental. En segundo lugar, se aplican técnicas de aprendizaje automático para la caracterización y clasificación del conjunto de datos.

2.1. Variograma Experimental

El variograma es una herramienta de la geoestadística o estadística espacial que permite analizar la relación espacial presente en los datos. Para definir al variograma es necesario introducir el concepto de variable regionalizada. Una variable regionalizada es una variable aleatoria definida en cada una de las regiones consideradas, y se denota como

$$\{Z(s) : s \in D \subset \mathbb{R}^n\}.$$

Donde cada $Z(s)$ es una variable aleatoria. De este modo, el variograma, el cual se denota con γ , se

define como

$$\gamma(\mathbf{h}) = \frac{1}{2} E \left[(Z(s) - Z(s + \mathbf{h}))^2 \right],$$

donde \mathbf{h} es el vector de separación entre los puntos, el cual incorpora tanto la distancia como la dirección. El variograma γ puede ser estimado mediante el variograma experimental, denotado por $\hat{\gamma}$, el cual se define como

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{(i,j) \in N(h)} (Z(i) - Z(j))^2, \quad (1)$$

donde $h \in \mathbb{R}$ y $N(h) = \{(i, j) : \|i - j\| = h\}$ representa el conjunto de pares de índices cuya distancia es h .

En este trabajo, la región D corresponde a una matriz de tamaño 28×28 , donde $Z(i, j)$ representa la intensidad del píxel en escala de grises, con valores entre 0 y 255. El conjunto $N(h)$ puede interpretarse como el conjunto de pares de píxeles dentro de la imagen que se encuentran separados por una distancia h .

Desde un punto de vista práctico, el variograma experimental mide el grado de similitud entre los valores de los píxeles en función de la distancia y la dirección de separación. Valores bajos del variograma indican una alta similitud entre píxeles cercanos, mientras que valores altos reflejan cambios abruptos de intensidad, lo cual suele asociarse con bordes o estructuras relevantes en la imagen. De este modo, el variograma permite capturar información sobre la textura y la organización espacial de los píxeles.

El cálculo del variograma experimental direccional permite analizar la variabilidad espacial a lo largo de orientaciones específicas. En el caso de imágenes, las direcciones de 0° , 45° y 90° corresponden a las orientaciones horizontal, diagonal y vertical, respectivamente. Estas direcciones capturan la mayoría de las estructuras geométricas presentes en los dígitos manuscritos, tales como trazos rectos, inclinados y verticales. La selección de estas tres orientaciones constituye un conjunto mínimo y representativo que permite caracterizar anisotropías (propiedades que cambian según la dirección) espaciales relevantes sin incrementar de forma significativa la complejidad computacional.

2.2. Aprendizaje automático

El aprendizaje automático supervisado se basa en el ajuste de modelos a partir de datos etiquetados, con el objetivo de aprender una función que permita predecir correctamente la clase de nuevas observaciones. En este trabajo se emplean dos modelos ampliamente utilizados en tareas de clasificación: redes neuronales completamente conectadas (FCNN) y máquinas de vectores soporte (SVM), con el fin de evaluar la capacidad discriminativa de las características extraídas mediante el variograma experimental.

Para utilizar modelos supervisados es necesario dividir los datos en subconjuntos disjuntos. En este estudio, los datos se separaron en tres conjuntos: entrenamiento, validación y prueba. El conjunto de entrenamiento se utilizó para el ajuste de los modelos, el de validación para la optimización de hiperparámetros y de los pesos del esquema de votación ponderada, y el conjunto de prueba se mantuvo completamente aislado para la evaluación final del desempeño.

Adicionalmente, sobre el conjunto de entrenamiento se empleó validación cruzada estratificada K-Fold. Este procedimiento permite obtener estimaciones más robustas del desempeño de los modelos, ya que cada observación participa tanto en el entrenamiento como en la validación, reduciendo la dependencia de una partición específica de los datos y mitigando posibles sesgos.

Las redes neuronales completamente conectadas empleadas en este trabajo utilizan la función de activación ReLU (*Rectified Linear Unit*), definida como $f(x) = \max(0, x)$, debido a su capacidad para modelar relaciones no lineales y facilitar el entrenamiento eficiente de redes profundas. Asimismo, se incorpora la técnica de *dropout*, la cual consiste en desactivar aleatoriamente un subconjunto de neuronas durante el entrenamiento con el fin de reducir el sobreajuste y mejorar la capacidad de generalización del modelo.

Las máquinas de vectores soporte (SVM) son modelos de aprendizaje supervisado que buscan construir una frontera de decisión óptima entre las clases, maximizando el margen entre las observaciones de distintas categorías. En problemas multiclase, como el abordado en este trabajo, las SVM se extienden mediante esquemas de descomposición. Además, se emplea una calibración de probabilidades que permite interpretar las salidas del modelo como distribuciones de probabilidad, lo cual resulta fundamental para los esquemas de fusión basados en votación ponderada.

Para la combinación de la información direccional se emplean dos estrategias. La fusión tardía consiste en entrenar modelos independientes para cada dirección y combinar sus salidas mediante un esquema de votación ponderada. Los pesos de esta votación se optimizan utilizando el algoritmo de *Differential Evolution*, un método de optimización global basado en mecanismos evolutivos como mutación, recombinación y selección. Por otro lado, la fusión temprana se realiza mediante la concatenación de los vectores de características correspondientes a las distintas direcciones, generando una representación conjunta que se utiliza como entrada de un único modelo de clasificación.

La evaluación de los modelos de clasificación supervisada se realiza mediante el análisis de la matriz de confusión, la cual resume el número de aciertos y errores de clasificación al comparar las etiquetas reales con las predichas por el modelo (ver Tabla 1). Más específicamente, la matriz de confusión permite identificar cuatro tipos de resultados: verdaderos positivos (VP), falsos positivos (FP), verdaderos negativos (VN) y falsos negativos (FN). A partir de esta matriz se pueden calcular métricas fundamentales para evaluar el desempeño de un clasificador, tales como la exactitud, la precisión, la sensibilidad (recall) y el F1-score.

Tabla 1: Matriz de confusión para clasificación binaria. Elaboración propia.

	<i>Predicho Positivo</i>	<i>Predicho Negativo</i>
<i>Real Positivo</i>	VP	FN
<i>Real Negativo</i>	FP	VN

A partir de la matriz de confusión se definen las siguientes métricas de evaluación:

- **Exactitud (Accuracy):** proporción de predicciones correctas respecto al total de observaciones.

$$\text{Accuracy} = \frac{VP + VN}{VP + VN + FP + FN}$$

- **Precisión (Precision):** proporción de instancias predichas como positivas que son realmente positivas.

$$\text{Precision} = \frac{VP}{VP + FP}$$

- **Sensibilidad o Exhaustividad (Recall):** proporción de instancias positivas correctamente identificadas.

$$\text{Recall} = \frac{VP}{VP + FN}$$

- **F1-score:** media armónica entre precisión y sensibilidad, especialmente útil cuando existe desequilibrio entre clases.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

3. Metodología

Como se mencionó anteriormente, para las pruebas se utilizó la base de datos MNIST. El procedimiento experimental se estructuró en varias etapas: (i) preparación de los datos, (ii) cálculo de variogramas experimentales direccionales en 0° , 45° y 90° , así como del variograma omnidireccional como referencia, (iii) construcción de representaciones de características mediante fusión temprana y fusión tardía, y (iv) aplicación y evaluación de modelos de clasificación supervisada.

3.1. Descripción de los datos

La base de datos MNIST es una de las más utilizadas en el desarrollo y validación de modelos de clasificación de imágenes. MNIST contiene imágenes de dígitos manuscritos del 0 al 9, representadas como matrices de tamaño 28×28 píxeles en escala de grises (véase la Figura 1). El conjunto de datos está compuesto por 60000 imágenes de entrenamiento y 10000 imágenes de prueba.

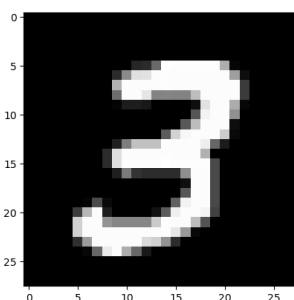


Figura 1: Imagen del dígito tres del conjunto MNIST. Elaboración propia.

MNIST ha sido ampliamente utilizada con métodos como redes neuronales convolucionales (CNN) y máquinas de vectores soporte (SVM), debido a su simplicidad, estandarización y relevancia como problema de referencia (Ahlawat & Choudhary, 2020; Gope et al., 2021; Keerthi, 2022). En este trabajo, los datos fueron cargados utilizando la biblioteca TensorFlow, que facilita el manejo eficiente de grandes volúmenes de datos (TensorFlow, s.f.).

3.2. Cálculo de variogramas experimentales

Para cada imagen del conjunto MNIST se calculó el variograma experimental direccional, es decir, el variograma experimental restringido a una dirección específica (véase la Figura 2). Para ello se utilizó la función **DirectionalVariogram** de la biblioteca *SciKit-GStat* para Python.

En este estudio, cada variograma experimental direccional se representa como un vector unidimensional de tamaño 28, correspondiente al número de retardos (*lags*) considerados. En todos los casos se empleó el estimador clásico de Matheron, con un esquema de binning uniforme y un alcance máximo (*maxlag*) determinado por la extensión espacial de la imagen.

El número de retardos se fijó en 28, en concordancia con la dimensión espacial de las imágenes (28×28 píxeles), lo que permite capturar la variabilidad espacial a distintas escalas sin introducir inestabilidad

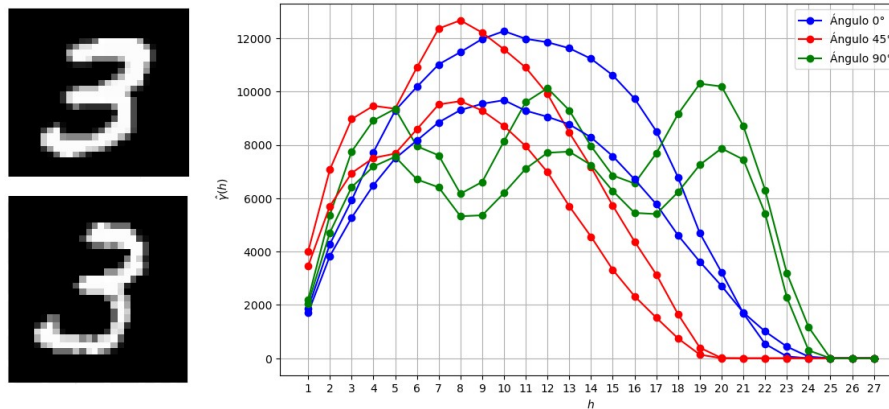


Figura 2: Ejemplo de dos imágenes del dígito tres del conjunto MNIST y sus variogramas experimentales direccionales en 0° , 45° y 90° . Elaboración propia.

numérica asociada a un número excesivo de lags. Para los variogramas direccionales se utilizó una tolerancia angular de 45° , de modo que cada dirección integra pares de píxeles cuya orientación se encuentra dentro de dicho rango, incrementando la robustez estadística del estimador.

Adicionalmente, se calculó el variograma omnidireccional de cada imagen como línea base de comparación frente a las configuraciones direccionales.

De este modo, al emplear el variograma como descriptor, se obtiene una representación de dimensión considerablemente menor que la imagen original. En particular, se utiliza un vector de tamaño 27 en lugar de una matriz de 28×28 píxeles, reduciendo significativamente la dimensionalidad del problema.

3.3. Protocolo experimental y prevención de fugas de información

Con el fin de evitar fugas de información, se definió un protocolo experimental estricto basado en tres subconjuntos disjuntos: entrenamiento, validación y prueba. El conjunto de entrenamiento se utilizó para el ajuste de los modelos, el conjunto de validación para la optimización de hiperparámetros y de los pesos del esquema de votación ponderada, y el conjunto de prueba se mantuvo completamente aislado hasta la evaluación final.

En el caso de los modelos SVM, se aplicó una calibración explícita de probabilidades, ajustada únicamente sobre el conjunto de validación. Para los modelos FCNN, se empleó validación cruzada estratificada K-Fold sobre el conjunto de entrenamiento, junto con un esquema de detención temprana (*early stopping*), con el fin de mitigar el sobreajuste.

Todas las métricas de desempeño se calcularon exclusivamente sobre el conjunto de prueba y se reportan junto con intervalos de confianza al 95 %, estimados mediante el método de *bootstrap*.

3.4. Definición de los conjuntos de entrenamiento y prueba

Para cada imagen de la base de datos se calcularon los variogramas experimentales direccionales en las orientaciones 0° , 45° y 90° . Como resultado, se obtuvieron tres conjuntos de datos independientes, uno por cada dirección, los cuales fueron utilizados para entrenar modelos FCNN y SVM de manera separada.

Además, se construyó una representación conjunta mediante la concatenación de los vectores de características correspondientes a las tres direcciones, dando lugar a un conjunto de datos adicional que

permite evaluar una estrategia de fusión temprana.

3.5. Modelo FCNN

Cada modelo FCNN está compuesto por dos capas densas ocultas de 128 neuronas cada una, con función de activación ReLU. Después de cada capa densa se aplicó un *dropout* con tasa 0.5, lo que implica la desactivación aleatoria del 50% de las neuronas durante el entrenamiento, con el fin de reducir el sobreajuste. Finalmente, la capa de salida está formada por 10 neuronas, correspondientes a las clases de los dígitos, y utiliza la función *softmax* para generar una distribución de probabilidad sobre las clases (véase la Figura 3).

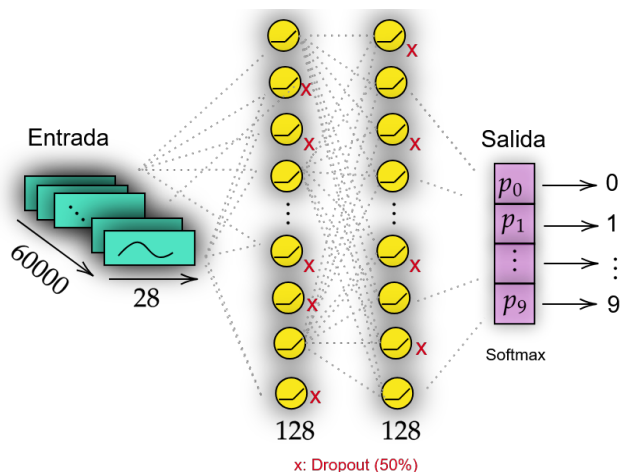


Figura 3: Arquitectura de la red neuronal completamente conectada (FCNN).
Elaboración propia.

3.6. Modelo SVM

Para el método de máquinas de vectores soporte se entrenaron modelos independientes utilizando los variogramas correspondientes a cada dirección (0° , 45° y 90°), así como el conjunto de características concatenadas. Dado que las SVM no producen probabilidades de forma nativa, se aplicó un proceso de calibración posterior que permite obtener distribuciones de probabilidad sobre las clases.

3.7. Fusión temprana, fusión tardía y evaluación

En este trabajo se analizaron dos estrategias de combinación de la información direccional. La fusión temprana se realizó mediante la concatenación de los vectores de características correspondientes a las direcciones 0° , 45° y 90° , generando una única representación que se utilizó como entrada de un modelo FCNN o SVM.

Por otro lado, la fusión tardía se implementó mediante un esquema de votación ponderada (*weighted voting*). En este enfoque, cada modelo entrenado de forma independiente genera un vector de probabilidades asociado a las clases, y dichas probabilidades se combinan mediante un conjunto de pesos $P = (p_1, p_2, p_3)$.

Los pesos del esquema de votación se obtuvieron utilizando el algoritmo de optimización *Differential Evolution*, implementado en el módulo `scipy.optimize`. La función objetivo se definió como la minimización de $1 - \text{accuracy}$, evaluada exclusivamente sobre el conjunto de validación. Los pesos resultantes fueron normalizados de modo que su suma fuese igual a uno.

Finalmente, el desempeño de los modelos se evaluó mediante la exactitud, la precisión, la sensibilidad, el F1-score y la matriz de confusión, utilizando las herramientas provistas por la biblioteca scikit-learn.

4. Resultados y discusión

Una vez entrenados los modelos, calibradas las probabilidades (en el caso de SVM) y obtenidos los pesos del esquema de combinación, se evaluó el desempeño final sobre el conjunto de prueba, compuesto por 10000 imágenes.

En esta sección se reportan resultados para dos estrategias de combinación de la información direccional: (i) **fusión tardía** mediante *votación ponderada* de probabilidades producidas por modelos independientes entrenados en 0° , 45° y 90° ; y (ii) **fusión temprana** mediante la *concatenación* de los descriptores direccionales en un único vector de características.

4.1. Método FCNN

4.1.1. Fusión tardía: ensemble direccional (0° , 45° y 90°)

El modelo FCNN basado en variogramas direccionales y combinado mediante votación ponderada alcanzó una exactitud (*accuracy*) de 0.8741, con un intervalo de confianza al 95 % de [0.8674, 0.8805]. Adicionalmente, se obtuvo un F1-score macro de 0.8714 y un F1-score ponderado de 0.8729, lo cual sugiere un desempeño global balanceado entre clases.

La Tabla 2 resume las métricas por clase (precisión, sensibilidad y F1-score) para el ensemble direccional. Se observa un desempeño particularmente alto en las clases 0 y 1, mientras que las clases 2 y 3 presentan valores inferiores, lo que sugiere que ciertos dígitos comparten patrones espaciales similares desde la perspectiva del variograma.

Tabla 2: Métricas por clase para FCNN con fusión tardía (ensemble direccional 0° , 45° y 90°). Elaboración propia.

Clase	Precisión	Sensibilidad	F1-Score
0	0.8893	0.9673	0.9267
1	0.9533	0.9885	0.9706
2	0.8323	0.7888	0.8100
3	0.8187	0.8317	0.8251
4	0.9220	0.9389	0.9304
5	0.8742	0.7948	0.8326
6	0.8621	0.7568	0.8060
7	0.8554	0.8804	0.8677
8	0.8602	0.9035	0.8813
9	0.8588	0.8682	0.8635

Al analizar la matriz de confusión (Figura 4), los errores más frecuentes corresponden a confusiones entre los dígitos 2, 3 y 5. Este patrón es consistente con similitudes estructurales entre dichos dígitos manuscritos, lo que puede inducir variabilidad espacial comparable en las direcciones evaluadas.

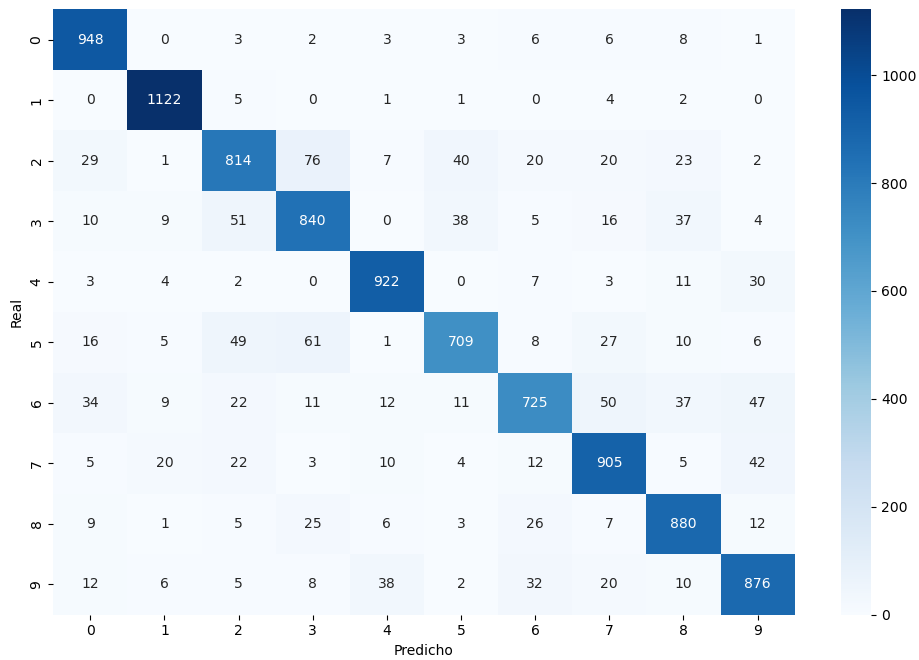


Figura 4: Matriz de confusión del modelo FCNN con ensemble direccional (0°, 45° y 90°) y votación ponderada. Elaboración propia.

4.1.2. Referencia: variograma omnidireccional (O)

Como línea base dentro del marco variográfico, se evaluó una FCNN utilizando exclusivamente el variograma omnidireccional. En este caso, el modelo alcanzó una exactitud de 0.7207 (IC95%: [0.7122, 0.7296]), con F1-score macro de 0.7082 y F1-score ponderado de 0.7133. La matriz de confusión se muestra en la Figura 5.

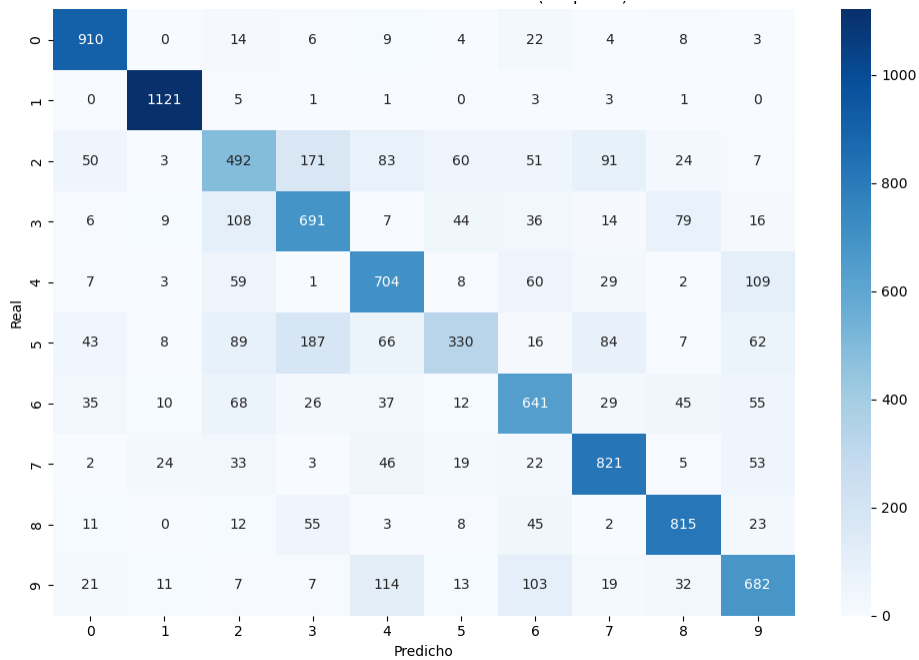


Figura 5: Matriz de confusión del modelo FCNN utilizando el variograma omnidireccional. Elaboración propia.

Estos resultados indican que el variograma omnidireccional, al promediar la variabilidad espacial sin considerar orientación, pierde información discriminativa relevante para algunas clases.

4.1.3. Fusión temprana: concatenación

La estrategia de fusión temprana mediante concatenación de los descriptores direccionales (0/45/90) alcanzó una exactitud de 0.9202 (IC95 %: [0.9147, 0.9256]), con F1-score macro de 0.9190 y F1-score ponderado de 0.9200.

Este resultado sugiere que la concatenación preserva información complementaria entre direcciones de forma más directa que la combinación tardía, aumentando la separabilidad entre clases en el espacio de características.

La Tabla 3 presenta las métricas por clase del modelo FCNN con fusión temprana. En comparación con el ensemble direccional, se observa una mejora sistemática en clases conflictivas (por ejemplo, 2, 3, 5 y 6), lo cual es consistente con una mayor disponibilidad de información direccional al nivel del vector de características.

Tabla 3: Métricas por clase para FCNN con fusión temprana. Elaboración propia.

<i>Clase</i>	<i>Precisión</i>	<i>Sensibilidad</i>	<i>F1-Score</i>
0	0.9638	0.9786	0.9711
1	0.9748	0.9885	0.9816
2	0.8923	0.9070	0.8996
3	0.8921	0.8842	0.8881
4	0.9299	0.9593	0.9444
5	0.8962	0.9002	0.8982
6	0.8959	0.8716	0.8836
7	0.9267	0.9095	0.9180
8	0.9119	0.9138	0.9128
9	0.9070	0.8791	0.8928

4.2. Método SVM

4.2.1. Fusión tardía: ensemble direccional (0°, 45° y 90°)

El clasificador SVM calibrado, combinado mediante votación ponderada, alcanzó una exactitud de 0.8439 (IC95 %: [0.8366, 0.8511]), con F1-score macro de 0.8401 y F1-score ponderado de 0.8422.

La Tabla 4 resume las métricas por clase para el SVM calibrado con ensemble direccional.

La matriz de confusión del SVM (Figura 6) muestra un patrón de errores similar al de FCNN, con confusiones recurrentes entre 2, 3 y 5, lo cual refuerza la hipótesis de similitudes espaciales capturadas por el variograma en estas clases.

4.2.2. Resultados con el variograma omnidireccional (O)

Al utilizar exclusivamente el variograma omnidireccional, el desempeño disminuyó de manera notable, alcanzando una exactitud de 0.6829 (IC95 %: [0.6742, 0.6922]) y un F1-score macro de 0.6682. Esto confirma que la direccionalidad aporta información esencial para mejorar la discriminación en el espacio de variogramas.

Tabla 4: Métricas por clase para SVM calibrado con fusión tardía (ensemble direccional 0° , 45° y 90°). Elaboración propia.

Clase	Precisión	Sensibilidad	F1-Score
0	0.8786	0.9531	0.9143
1	0.9350	0.9885	0.9610
2	0.8043	0.7568	0.7798
3	0.7614	0.8089	0.7844
4	0.8828	0.9053	0.8939
5	0.8687	0.7119	0.7825
6	0.8149	0.7171	0.7629
7	0.8279	0.8473	0.8375
8	0.8215	0.8881	0.8535
9	0.8322	0.8305	0.8313

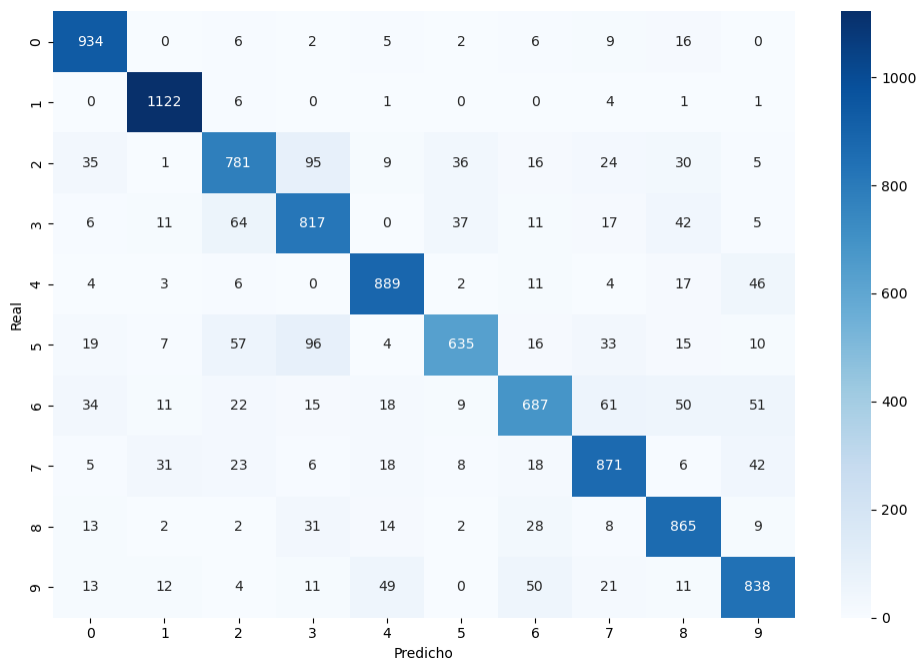


Figura 6: Matriz de confusión del modelo SVM calibrado con ensemble direccional y votación ponderada. Elaboración propia.

4.2.3. Fusión temprana: concatenación (0/45/90)

La estrategia de fusión temprana mediante concatenación también fue evaluada con SVM calibrado. En este caso se obtuvo una exactitud de 0.9200 (IC95 %: [0.9148, 0.9255]), con F1-score macro de 0.9189 y F1-score ponderado de 0.9200.

La Tabla 5 presenta las métricas por clase del SVM calibrado con fusión temprana. En términos generales, el enfoque de concatenación mejora consistentemente la sensibilidad y el F1-score en clases previamente conflictivas bajo fusión tardía (por ejemplo, 2, 3 y 5).

4.3. Comentario final

La Tabla 6 resume el desempeño global de los modelos evaluados. El uso de direccionalidad mejora de forma consistente el rendimiento frente al enfoque omnidireccional, lo que confirma que las

Tabla 5: Métricas por clase para SVM calibrado con fusión temprana (0/45/90).
Elaboración propia.

<i>Clase</i>	<i>Precisión</i>	<i>Sensibilidad</i>	<i>F1-Score</i>
0	0.9766	0.9796	0.9781
1	0.9706	0.9877	0.9790
2	0.9023	0.9215	0.9118
3	0.9207	0.8970	0.9087
4	0.9440	0.9440	0.9440
5	0.9138	0.8913	0.9024
6	0.8536	0.8706	0.8620
7	0.9259	0.8988	0.9121
8	0.8947	0.9158	0.9051
9	0.8900	0.8821	0.8860

orientaciones 0° , 45° y 90° capturan información espacial complementaria.

Además, la concatenación alcanzó los mejores resultados globales tanto en FCNN como en SVM, lo cual sugiere que combinar las direcciones a nivel de características conserva información.

Respecto a baselines basados en píxeles o análisis en componentes principales (ACP), estos enfoques han sido ampliamente estudiados para MNIST y persiguen un objetivo distinto, pues operan sobre representaciones de alta dimensión o arquitecturas más complejas. En este trabajo el interés principal es aislar y evaluar el aporte del variograma experimental como descriptor espacial de baja dimensión, y cuantificar el efecto de la direccionalidad y de la estrategia de fusión empleada.

Tabla 6: Resumen comparativo de desempeño y tiempo. Elaboración propia.

<i>Modelo</i>	<i>Representación</i>	<i>Accuracy</i>	<i>F1 macro</i>	<i>Tiempo (s)</i>
FCNN	Fusión tardía (Ensemble 0/45/90)	0.8741	0.8714	3755.90
FCNN	Omnidireccional (O)	0.7207	0.7082	676.50
FCNN	Fusión temprana (concatenación 0/45/90)	0.9202	0.9190	490.44
SVM	Fusión tardía (Ensemble 0/45/90)	0.8439	0.8401	1831.35
SVM	Omnidireccional (O)	0.6829	0.6682	325.62
SVM	Fusión temprana (concatenación 0/45/90)	0.9200	0.9189	262.84

5. Conclusiones

En este trabajo se evaluó el uso del variograma experimental como descriptor de baja dimensión para la clasificación de imágenes, considerando variogramas direccionales (0° , 45° y 90°) y el variograma omnidireccional como referencia, y empleando dos modelos de clasificación: una red neuronal

completamente conectada y una máquina de vectores soporte calibrada. Los resultados confirman que el variograma permite capturar información espacial relevante aun cuando se trabaja con una representación significativamente reducida respecto a la imagen original.

Al comparar las estrategias de combinación, se observó que la fusión temprana mediante la concatenación de los variogramas direccionales produjo el mejor desempeño global en ambos clasificadores, con valores de exactitud cercanos a 0.92 y F1-scores macro y ponderados igualmente alrededor de 0.92. Este hallazgo sugiere que integrar explícitamente, a nivel de características, la información direccional preserva patrones espaciales complementarios que resultan altamente discriminativos para la tarea de clasificación.

En el caso de la fusión tardía basada en votación ponderada de probabilidades (ensemble direccional), los resultados también evidencian que incorporar direccionalidad mejora de manera consistente el desempeño frente al variograma omnidireccional. En particular, el esquema de votación ponderada superó al enfoque omnidireccional en términos de accuracy y F1-score tanto para FCNN como para SVM, lo que confirma que el promedio omnidireccional tiende a perder información útil para discriminar ciertas estructuras presentes en los dígitos manuscritos.

Desde el punto de vista computacional, el SVM calibrado ofreció una alternativa más eficiente que la FCNN, manteniendo un rendimiento competitivo en los esquemas direccionales. No obstante, en la configuración de fusión temprana, ambos clasificadores alcanzaron desempeños muy similares, lo cual refuerza la idea de que la calidad del descriptor (variograma direccional concatenado) domina el rendimiento final más que la complejidad del clasificador empleado.

El análisis por clase mostró un desempeño particularmente alto y estable para los dígitos 0 y 1, mientras que clases como 2 y 3 tendieron a presentar mayores confusiones. Este comportamiento puede atribuirse a similitudes morfológicas y de variabilidad espacial entre ciertos dígitos, que generan firmas variográficas parcialmente superpuestas en las direcciones consideradas.

En conjunto, los resultados confirman que el variograma experimental constituye un caracterizador adecuado para imágenes y, en particular, una herramienta útil para reducir dimensionalidad sin perder completamente la estructura espacial relevante. Dado que el variograma preserva información sobre la variabilidad media en función de la distancia, su aplicación resulta especialmente atractiva en dominios donde la estructura espacial es crítica, como imagen médica, tomografía, análisis de texturas y otras áreas afines.

Finalmente, este trabajo representa un primer paso hacia la aplicación del variograma como descriptor en escenarios más complejos. Como trabajo futuro, se propone extender este enfoque a conjuntos de datos médicos e imágenes clínicas, con el objetivo de evaluar su capacidad de generalización y su utilidad en contextos de apoyo al diagnóstico, así como comparar su desempeño con otras técnicas de aprendizaje automático y descriptores de mayor complejidad.

Contribución de las personas autoras: Ambas autoras contribuyeron de manera equitativa en todas las partes del trabajo, colaborando en el desarrollo, la investigación y la redacción de cada sección.

Accesibilidad de los datos: La base de datos utilizada es la MNIST. Sin embargo, las bases de datos ya depuradas y con el cálculo de los variogramas, además del código utilizados se puede hallar en https://drive.google.com/drive/folders/1I74O9y3kg-WYUab5-Mj3n-XwEKS3rDhw?usp=share_link.

Declaración de uso de inteligencia artificial: Para la elaboración de este artículo, se emplearon herramientas de inteligencia artificial para optimizar los códigos de Python utilizados en el análisis de datos y para mejorar aspectos de la redacción del documento.

Referencias

- Ahlawat, S., & Choudhary, A. (2020). Hybrid CNN-SVM classifier for handwritten digit recognition. *Procedia Computer Science*, 167, 2554-2560. <https://doi.org/10.1016/j.procs.2020.03.309>
- Bansal, M., Kumar, M., Sachdeva, M., & Mittal, A. (2023). Transfer learning for image classification using VGG19. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-021-03488-z>
- Berberoğlu, S. Ü., et al. (2010). Utilizing image texture to detect land-cover change. *International Journal of Remote Sensing*, 31(11), 2793-2815. <https://doi.org/10.1080/01431160903111077>
- Bharadiya, J. (2023). Convolutional neural networks for image classification. *International Journal of Innovative Science and Research Technology*, 8(5), 673-677. <https://doi.org/10.5281/zenodo.7952031>
- Chan, H. P., Samala, R. K., Hadjiiski, L. M., & Zhou, C. (2020). Deep learning in medical image analysis. En *Deep Learning in Medical Image Analysis: Challenges and Applications* (pp. 3-21). Springer. https://doi.org/10.1007/978-3-030-33128-3_1
- Ericeira, D. R., et al. (2013). Detection of masses based on asymmetric regions of digital bilateral mammograms. *Computers in Biology and Medicine*, 43(8), 987-999. <https://doi.org/10.1016/j.combiomed.2013.04.019>
- Ghorbanzadeh, O., Shahabi, H., Crivellari, A., Homayouni, S., Blaschke, T., & Ghamisi, P. (2022). Landslide detection using deep learning and object-based image analysis. *Landslides*, 19(4), 929-939. <https://doi.org/10.1007/s10346-021-01843-x>
- Gómez-de-Mariscal, E., et al. (2021). DeepImageJ: A user-friendly environment to run deep learning models in ImageJ. *Nature Methods*, 18(10), 1192-1195. <https://doi.org/10.1038/s41592-021-01262-9>
- Gope, B., et al. (2021). Handwritten digits identification using MNIST database via machine learning models. *IOP Conference Series: Materials Science and Engineering*, 1022(1). <https://doi.org/10.1088/1757-899X/1022/1/012108>
- Journel, A. G., & Huijbregts, C. J. (1978). *Mining Geostatistics*. Academic Press.
- Keerthi, T. (2022). MNIST handwritten digit recognition using machine learning. *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*. <https://doi.org/10.1109/ICACITE53722.2022.9823806>
- Krishna, S. T., & Kalluri, H. K. (2019). Deep learning and transfer learning approaches for image classification. *International Journal of Recent Technology and Engineering*, 7(5S4), 427-432.
- LeCun, Y., & Cortes, C. (2010). MNIST handwritten digit database.
- Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1), 91-99. <https://doi.org/10.1016/j.gltp.2022.04.020>
- Nascimento, I. G. B., & de Sousa Bezerra, A. F. (2021). Tecnologias digitais e metodologias ativas (TDMA) no ensino do processamento e análise de imagens de interesse médico-relato de

experiência. *Brazilian Journal of Development*, 7(4), 38340-38346. <https://doi.org/10.34117/bjdv7n4-343>

Nazarkevych, M., Voznyi, Y., Hrytsyk, V., Klyujnyk, I., Havrysh, B., & Lotoshynska, N. (2021). Identification of biometric images by machine learning. *2021 IEEE 12th International Conference on Electronics and Information Technologies (ELIT)*, 95-98. <https://doi.org/10.1109/ELIT53502.2021.9501064>

Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 221-248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>

TensorFlow. (s.f.). MNIST dataset.

Xia, H., Liu, Z., Efremochkina, M., Liu, X., & Lin, C. (2022). Study on city digital twin technologies for sustainable smart city design. *Sustainable Cities and Society*, 84, 104009. <https://doi.org/10.1016/j.scs.2022.104009>

Yang, J., Zhuang, Y., & Wu, F. (2012). ESVC-based extraction and segmentation of texture features. *Computers & Geosciences*, 49, 238-247. <https://doi.org/10.1016/j.cageo.2012.06.007>

Zhou, Z.-H. (2021). *Machine Learning*. Springer Nature. <https://doi.org/10.1007/978-981-15-1967-3>