

# Deep learning aplicado al reconocimiento de señales de tránsito

Andrés Orlando López Henao

Universidad de Antioquia

aorlando.lopez@udea.edu.co

**Abstract—** This initiative aims to develop a machine learning model that can recognize traffic signs present in an image. Its main objective is to offer assistance to visually impaired people in the future. One possible application scenario would be the use of glasses equipped with a camera that captures images of the environment. The machine learning model would be able to analyze these images in real time and recognize the traffic signs present, providing useful information and facilitating the mobility of visually impaired people in urban environments and road trips by giving greater autonomy and safety to users while moving around the city.

**Keywords:** Traffic sign recognition, Machine learning, Deep learning, Convolutional Neural Network.

## I. CONTEXTO DE LA APLICACIÓN

Esta iniciativa tiene como propósito desarrollar un modelo de aprendizaje automático que pueda reconocer las señales de tránsito presentes en una imagen. Su objetivo principal es ofrecer asistencia a personas con discapacidad visual en el futuro. Un posible escenario de aplicación sería el uso de gafas equipadas con una cámara que capture imágenes del entorno. El modelo de machine learning sería capaz de analizar estas imágenes en tiempo real y reconocer las señales de tráfico presentes, proporcionando información útil y facilitando la movilidad de las personas con limitaciones visuales en entornos urbanos y viajes por carretera otorgando mayor autonomía y seguridad a los usuarios mientras se desplazan por la ciudad.

## II. OBJETIVO DE MACHINE LEARNING

El objetivo de machine learning es predecir las señales de tránsito que están presentes en una imagen.

## III. DATASET

A continuación se describen las características del dataset.

A. *Tipo de datos:* Imágenes de señales de tránsito.

B. *Número de datos:* 73139.

C. *Tamaño en disco:* 61.2 megabytes.

D. *Distribución de las clases:* 43 clases.

TABLA 1. DISTRIBUCIÓN DE LAS CLASES

Id	Nombre	Cantidad
0	Speed limit (20km/h)	451
1	Speed limit (30km/h)	4920

2	Speed limit (50km/h)	2610
3	Speed limit (60km/h)	2670
4	Speed limit (70km/h)	3750
5	Speed limit (80km/h)	3510
6	End of speed limit (80km/h)	780
7	Speed limit (100km/h)	2730
8	Speed limit (120km/h)	2670
9	No passing	2790
10	No passing for vehicles over 3.5 metric tons	3810
11	Right-of-way at the next intersection	2490
12	Priority road	3990
13	Yield	4080
14	Stop	1470
15	No vehicles	1170
16	Vehicles over 3.5 metric tons prohibited	780
17	No entry	2100
18	General caution	2280
19	Dangerous curve to the left	390
20	Dangerous curve to the right	660
21	Double curve	600
22	Bumpy road	720
23	Slippery road	960
24	Road narrows on the right	510
25	Road work	2850
26	Traffic signals	1140
27	Pedestrians	450
28	Children crossing	1020
29	Bicycles crossing	510
30	Beware of ice/snow	840
31	Wild animals crossing	1470
32	End of all speed and passing limits	450
33	Turn right ahead	1288
34	Turn left ahead	780
35	Ahead only	2280
36	Go straight or right	720
37	Go straight or left	390
38	Keep right	3930
39	Keep left	570
40	Roundabout mandatory	660
41	End of no passing	450
42	End of no passing by vehicles over 3.5 metric tons	450

#### IV. MÉTRICAS DE DESEMPEÑO

A continuación en la tabla 2 se describen las métricas de machine learning que se aplicarán al modelo.

TABLA 2. MÉTRICAS DE MACHINE LEARNING

Métrica	Definición
<i>Exactitud (Accuracy)</i>	Proporción de predicciones correctas respecto al total de predicciones.
<i>Precisión</i>	Proporción de verdaderos positivos respecto a todos los positivos predichos.
<i>Sensibilidad (Recall)</i>	Proporción de verdaderos positivos respecto a todos los positivos reales.
<i>F1-Score</i>	Media armónica de precisión y sensibilidad, proporciona un balance entre ambas.
<i>ROC-AUC</i>	Medición de la capacidad discriminativa del modelo.
<i>Matriz de Confusión</i>	Representación de la distribución de las predicciones correctas e incorrectas del modelo.
<i>Entropía Cruzada</i>	Medición de la diferencia entre dos distribuciones de probabilidad, utilizada como función de pérdida en problemas de clasificación.

En la tabla 3 se describen las métricas de negocio que se aplicarán al modelo.

TABLA 3. MÉTRICAS DE NEGOCIO

Métrica	Definición
<i>Tiempo de inferencia</i>	Tiempo necesario para que el modelo realice una predicción en una imagen en tiempo real.
<i>Tasa de Error</i>	Proporción de veces que el modelo no identifica correctamente una señal de tránsito en comparación con el total de imágenes procesadas.

#### V. REFERENCIAS Y TRABAJOS PREVIOS

Se han aplicado diversas técnicas de aprendizaje automático a la detección y reconocimiento de señales de tráfico, con resultados prometedores. En el artículo Review. Machine learning techniques for traffic sign detection (Mukhometzianov & Wang, 2017), el autor proporcionó una revisión exhaustiva de estos métodos, clasificándolos en aprendizaje automático general y redes neuronales.

En el trabajo Extreme learning machine based traffic sign detection (Huang et al., 2014) se propone un método eficiente para la detección de señales de tráfico. Este método utiliza principalmente clasificadores Extreme learning machine (ELM) de 5 categorías que fueron entrenados usando características del Histograma de Gradiente Orientado (HOG) y usa clasificadores ELM de 9 categorías que fueron entrenados usando características HOG e histograma de color. Los resultados experimentales han demostrado que este método obtiene un 98,60% en términos del área bajo la curva (AUC) y cuesta 6,5 segundos en promedio para detectar una imagen.

En el artículo Traffic Sign Recognition System using CNN (Agrawal et al., 2023), los autores utilizan un modelo de aprendizaje profundo basado en redes neuronales convolucionales (CNN) descrito en la figura 1 para entrenar y clasificar las señales basándose en el aprendizaje a partir de las imágenes de entrenamiento.

Este modelo alcanza una precisión del 93% y una sensibilidad del 92%.

Model: "sequential"	
Layer (type)	Output Shape
conv2d (Conv2D)	(None, 26, 26, 32)
conv2d_1 (Conv2D)	(None, 22, 22, 32)
max_pooling2d (MaxPooling2D)	(None, 11, 11, 32)
dropout (Dropout)	(None, 11, 11, 32)
conv2d_2 (Conv2D)	(None, 9, 9, 64)
conv2d_3 (Conv2D)	(None, 7, 7, 64)
max_pooling2d_1 (MaxPooling2D)	(None, 3, 3, 64)
dropout_1 (Dropout)	(None, 3, 3, 64)
flatten (Flatten)	(None, 576)
dense (Dense)	(None, 256)
dropout_2 (Dropout)	(None, 256)
dense_1 (Dense)	(None, 43)

FIG. 1. MODELO CNN (HUANG)

En el trabajo A Real-Time Traffic Sign Detection and Recognition System on Hybrid Dataset using CNN (Bhatt et al., 2022), los autores emplean el modelo CNN de la figura 2. Se utiliza un conjunto de datos híbrido que combina el conjunto de datos estándar de reconocimiento de señales de tráfico alemán con un conjunto de datos adicional creado a partir de señales de tráfico indias, alcanzando una exactitud del 95.42%.

Model: "sequential\_1"

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 28, 28, 60)	1560
conv2d_2 (Conv2D)	(None, 24, 24, 60)	90060
max_pooling2d_1 (MaxPooling2D)	(None, 12, 12, 60)	0
conv2d_3 (Conv2D)	(None, 10, 10, 30)	16230
conv2d_4 (Conv2D)	(None, 8, 8, 30)	8130
max_pooling2d_2 (MaxPooling2D)	(None, 4, 4, 30)	0
dropout_1 (Dropout)	(None, 4, 4, 30)	0
flatten_1 (Flatten)	(None, 480)	0
dense_1 (Dense)	(None, 500)	240500
dropout_2 (Dropout)	(None, 500)	0
dense_2 (Dense)	(None, 102)	51102

FIG. 2. MODELO CNN (AGRAWAL)

El artículo Computer Vision System for Speed Limit Traffic Sign Recognition (Salimullina & Budanov, 2022) se centra en el desarrollo de un sistema portátil para el reconocimiento general de señales de límite de velocidad en las carreteras. El proceso de reconocimiento se llevó a cabo utilizando un modelo CNN alcanzando un tiempo medio de reconocimiento de aproximadamente 5 segundos.

En la investigación Traffic Sign Recognition with Faster RCNN and RPN for Advanced Driver Assistance Systems (K R et al., 2021) utilizan una Faster Region-based Convolutional Neural Network más una Red de Propuesta de Regiones (RPN) como se muestra en la figura 3. A continuación, se utiliza el algoritmo Random

Forest para realizar la clasificación y la regresión en el conjunto de datos dado. Este modelo proporciona una precisión cercana a 100.

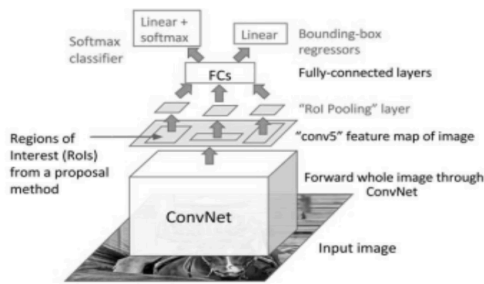


FIG. 3. MODELO DE ATENCIÓN (K R)

En el artículo Enhancing Road Safety through Innovative Traffic Sign Detection and Recognition with YOLOv5 (Ashwath et al., 2024) los autores utilizan el algoritmo YOLOv5 ("You Only Look Once") para la detección de objetos, dividiendo las imágenes en una cuadrícula para localizarlos con eficacia.

El trabajo An Efficient Automatic Traffic Sign Detection and Recognition Method for Smartphones (Shih et al., 2017) propone un nuevo algoritmo basado en el algoritmo de Regiones Extremas Máximamente Estables (MSER) para detectar con precisión todas las señales de tráfico candidatas en tiempo real. A continuación, se extrae el vector de características de cada región candidata mediante el algoritmo de HOG. Para reconocer la señal de tráfico, el algoritmo propuesto se diseña combinando un clasificador de máquina de vectores lineales de soporte (LSVM) con un proceso de votación para mejorar la tasa de reconocimiento de señales de tráfico en entornos reales. Los resultados experimentales muestran que el sistema propuesto alcanza una tasa de precisión del 96% en el reconocimiento de señales de tráfico alemán (GTSRB).

En la tabla 4 se presentan los resultados arrojados por los diferentes enfoques en términos de exactitud, precisión, sensibilidad y tiempo de predicción de cada uno de los modelos como base de comparación en trabajos futuros.

TABLA 4. RESULTADOS

Autor	Enfoque	Exactitud	Precisión	Sensibilidad	Tiempo
(Agrawal et al., 2023)	CNN	-	93%	92%	-
(Bhatt et al., 2022)	CNN	95.42%	-	-	-
(Salimullina & Budanov, 2022)	CNN	98.90%	-	-	5 s
(K R et al., 2021)	Faster RCNN+ RPN	-	97.70%	99.20%	-
(Ashwath et al., 2024)	YOLOv5	90.25%	-	-	-
(Shih et al., 2017)	LSVM	-	96.12%	94.48%	28 ms

REFERENCES

Agrawal, S. C., Tripathi, R. K., Gupta, A., & Sahu, M. (2023). Traffic Sign Recognition System using CNN. *2023 World Conference on Communication & Computing (WCONF)*, 1-5. <https://doi.org/10.1109/WCONF58270.2023.10235024>

Ashwath, S., T, C., M, A. S., S, S., & D R, S. (2024). Enhancing Road Safety through Innovative Traffic Sign Detection and Recognition with YOLOv5. *2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, 1-4. <https://doi.org/10.1109/IITCEE59897.2024.10467719>

Bhatt, N., Laldas, P., & Lobo, V. B. (2022). A Real-Time Traffic Sign Detection and Recognition System on Hybrid Dataset using CNN. *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, 1354-1358. <https://doi.org/10.1109/ICCES54183.2022.9835954>

Huang, Z., Yu, Y., Ye, S., & Liu, H. (2014). Extreme learning machine based traffic sign detection. *2014 International Conference on Multisensor Fusion and Information Integration for Intelligent Systems (MFI)*, 1-6. <https://doi.org/10.1109/MFI.2014.6997672>

K R, S., M, D., & R S, V. D. (2021). Traffic Sign Recognition with Faster RCNN and RPN for Advanced Driver Assistance Systems. *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 1-8. <https://doi.org/10.1109/ICECA52323.2021.9675996>

Salimullina, A. D., & Budanov, D. O. (2022). Computer Vision System for Speed Limit Traffic Sign Recognition. *2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, 415-418. <https://doi.org/10.1109/ElConRus54750.2022.9755744>

Shih, P.-C., Tsai, C.-Y., & Hsu, C.-F. (2017). An efficient automatic traffic sign detection and recognition method for smartphones. *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 1-5. <https://doi.org/10.1109/CISP-BMEI.2017.8301993>