

Reconocimiento y detección de señales de tráfico mexicanas de tipo preventivas, restrictivas e informativas aplicando YOLOv4

Daniela Bolaños-Flores¹, Hamurabi Gamboa-Rosales¹,
José M. Celaya-Padilla¹, Tania A. Ramirez-del Real^{2,3}

¹ Universidad Autónoma de Zacatecas,
México

² Consejo Nacional de Ciencia y Tecnología,
México

³ Centro de Investigación en Ciencias de
Información Geoespacial,
México

tramirez@centrogeo.edu.mx, bolanosfloresdaniela@gmail.com

Resumen. Una variedad de factores a lo largo del camino pueden poner en peligro la seguridad de los conductores o peatones y dar lugar a accidentes de alto impacto durante la conducción, por lo que las señales de tránsito son elementos esenciales que brindan información sobre el estado del camino durante el viaje. Técnicas novedosas permiten el desarrollo de herramientas que ayudan a reconocer y clasificar objetos de interés. En este trabajo en particular, se emplean para diseñar un modelo de reconocimiento y detección de señales de tránsito mexicanas, utilizando técnicas de aprendizaje computacional. Se recolectó un conjunto de datos enfocado en señales de tránsito en 5 ciudades diferentes de la república mexicana dentro de las principales vías urbanas. La base de datos contiene un total de 2,160 elementos de carretera divididos en 14 clases diferentes para el entrenamiento y validación de algoritmos. Además, la recopilación de datos tuvo lugar en diversas condiciones climáticas. La detección automática se realiza mediante una red neuronal convolucional denominada You Only Look Once (YOLOv4), que presenta un porcentaje mean Average Precision (mAP) mayor al 95 %.

Palabras clave: Reconocimiento y detección, señales de tráfico, aprendizaje computacional.

Recognition and Detection of Preventive, Restrictive and Informative Mexican Traffic Signs Applying YOLOv4 Architecture

Abstract. A variety of factors along the road can jeopardize the safety of drivers or pedestrians and lead to high-impact accidents while driving. Therefore, traffic signs are essential elements, providing information about the state of road

during travel. Novel techniques allow for the development of tools that help to recognize and classify objects of interest. This work uses computational learning techniques to design a recognition and detection model for Mexican traffic signs. A dataset focused on traffic signs was collected in 5 different cities of the Mexican Republic within the main urban roads. The database contains 2,160 road elements divided into 14 classes for algorithm training and validation—furthermore, data collection takes place in various weather conditions. The automatic detection uses a convolutional neural network called You Only Look Once (YOLOv4), yielding a mean Average Precision (mAP) percentage more significant than 95%.

Keywords: Recognition and detection, traffic signs, machine learning.

1. Introducción

En la actualidad, el desarrollo de la tecnología ha aumentado de manera considerable, por lo que para la sociedad y para diversas áreas de la ingeniería ha sido de gran aportación, como la interpretación de información visual en máquinas, haciendo uso de técnicas computacionales para el reconocimiento en imágenes, obteniendo resultados prometedores en diferentes campos de aplicación.

En el área automotriz se han implementado arquitecturas basadas en inteligencia artificial, donde se proponen y desarrollan modelos inteligentes de reconocimiento y detección de elementos viales, posteriormente se pueden utilizar en autos de conducción inteligente, permitiendo al piloto considerar señales o elementos que se encuentran en el transcurso de carreteras o zonas urbanas con la finalidad de advertir señales que no haya atendido. Estos vehículos incluyen funciones de piloto automático que permiten al usuario identificar señales de alto y semáforos a cierta distancia, siempre y cuando esté bajo una supervisión de manera activa [24].

Por otra parte, adquirir una de estas unidades suele ser costoso, además, las señales que identifica y reconoce son enfocadas al estándar estadounidense, asiático y europeo, generando una limitación en el área donde se implementa. Con frecuencia la falta de seguridad vial dentro del territorio mexicano se debe a ciertos factores, como la ausencia de elementos de señalización en zonas urbanas o carreteras trayendo consigo consecuencias como accidentes catastróficos [23].

También, en algunas zonas, las señales sufren modificaciones por diversas causas climáticas o vandalizaciones, por lo que la conducción es más complicada para el individuo. La literatura técnico científica reporta modelos inteligentes asociados al análisis de señales viales o de tráfico dentro de algún país en particular, lo cual presenta una brecha, ya que los modelos propuestos no alcanzan la generalización, debido a la implementación con conjuntos de datos específicos a la localidad donde se formularon [17].

Lo anterior, ha permitido la propuesta de diversos trabajos de investigación enfocados al reconocimiento y detección de señales de tráfico basados en técnicas de aprendizaje automático [1, 8, 9, 12, 27], obteniendo resultados con buen desempeño al evaluarlos, el área de oportunidad puede establecerse por la existencia del sesgo de la información, debido a la región de adquisición de datos, que generalmente pertenecen a

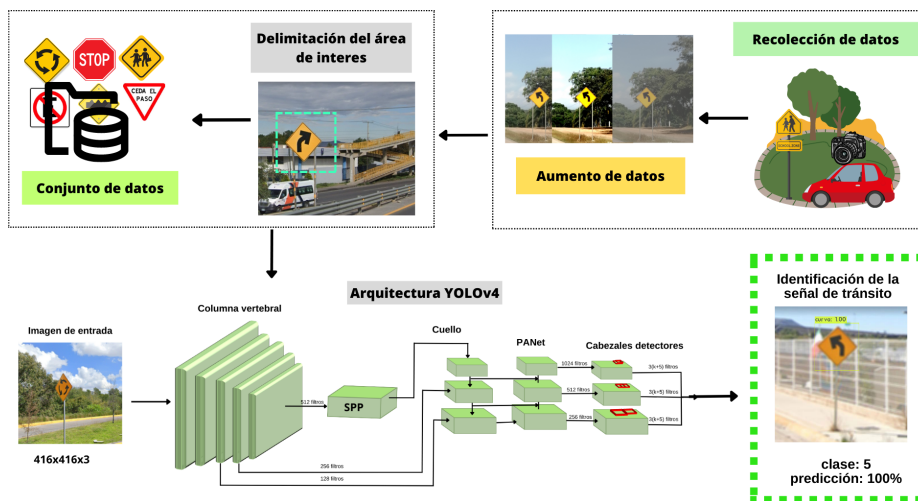


Fig. 1. Diagrama general para el desarrollo del modelo.

señales extranjeras, por lo que complica la implementación para el territorio mexicano, ya que existen diferencias entre las señales mexicanas y las extranjeras, como lo son el diseño, el color, la simbología, entre otras.

Considerando las problemáticas previamente expuestas, surge la necesidad de proponer una base de datos que contenga señales de tránsito en las principales zonas urbanas de diversas ciudades de la región. Este conjunto de datos se utilizará posteriormente para implementar un modelo capaz de identificar y reconocer las señales preventivas y restrictivas en el territorio mexicano.

El presente trabajo está organizado de la siguiente forma, en la Sección 2 se encuentran los principales trabajos relacionados, en la Sección 3 se explican los materiales y métodos enfocados al conjunto y procesamiento de datos, además la arquitectura de YOLOv4 y el proceso para el desarrollo del modelo. A continuación, en la Sección 4 se localiza la experimentación y resultados presentados, además de la evaluación y desempeño general del algoritmo. Finalmente, en la Sección 5, se concluye y proponen los trabajos futuros.

2. Trabajos relacionados

Dentro de la literatura existe una extensa cantidad de investigaciones acerca de Reconocimiento de Señales de Tráfico (TSR, por sus siglas en inglés) y Detección de Señales de Tráfico (TSD, por sus siglas en inglés); donde la mayoría de los autores utilizan un conjunto de datos públicos con una variedad de clases y un extenso número de imágenes, principalmente se pueden mencionar los siguientes:

- High Resolution Remote Sensing Detection(HRRSD) [28]: Este grupo contiene 13 categorías donde solo el 11 % pertenece a señales de tráfico asiáticas.

Tabla 1. Tipos de señales de tránsito.

Número de secuencia	Tipo	Total
Señales Restrictivas		
1	Ceda el paso	69
2	Límite de velocidad	422
3	Prohibido estacionarse	408
4	Alto	173
Señales Preventivas		
5	Cruce peatonal	355
6	Curva	135
7	Glorieta	68
8	Incorporación al tránsito	156
9	Salida	140
10	Reductor de velocidad	144
Señales Informativas		
11	Parada de autobus	20
12	Zona de discapacitados	35
13	Zona de taxis	15
14	Gasolinera	20
Total		2160

- Tsinghua-Tencent 100K (TT100K) [29]:Este conjunto contiene 128 clases donde solo 30,000 están enfocados a señales de tránsito asiáticas.
- German Traffic Sign Recognition Benchmark (GTSRB) [10] : Este conjunto cuenta con 43 categorías y una cantidad aproximada de 51,839 imágenes con elementos viales alemanes.

Derivado de los conjuntos de datos anteriores, surgen diversos trabajos con propuestas para mejorar el desempeño en la detección de señales viales, a continuación se describen algunos de los planteamientos y resultados sobresalientes más actuales.

Liu et al. [16] desarrollaron un modelo inteligente de TSR y TSD. Dentro de la fase de prueba del sistema, se utilizó un conjunto de datos denominado HRRSD donde contiene aproximadamente 21,761 imágenes con categorías como vehículos, avión, puentes, entre otras, es importante resaltar que los datos no están enfocados a señales de tránsito, sino, es utilizada para evaluar la capacidad del algoritmo propuesto, obteniendo un porcentaje de 85.5 % de mean Average Precision (mAP).

En los trabajos de Liu et al. y Li et al. [14, 13] utilizaron la base de datos asiática denominada TT100K para la creación del modelo de reconocimiento o detección. El principal enfoque en el trabajo de Liu et al. [14] es el reconocimiento de señales de tráfico TSR, de la misma forma, los autores utilizaron 10,267 imágenes del total del conjunto de datos TT100K, para desarrollar un modelo que permita reconocer señales de proporción limitada, borrosas y complejas dentro de un ámbito natural, por medio de una arquitectura modificada obteniendo resultados por arriba del 86 % en mAP, una precisión de 87.45 % y una sensibilidad de 79.65 %.

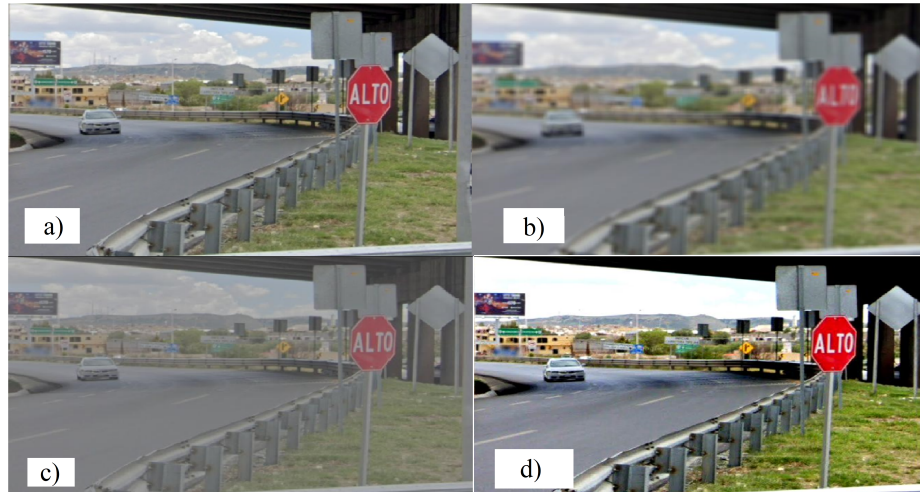


Fig. 2. Demostración de aumento de datos para la clase alto. a) imagen original, b) imagen en desenfoque, c) imagen con baja saturación y d) imagen con alta saturación.

Li et al. [13] realizaron reconocimiento y detección de señales de tráfico (TSR y TSD), utilizando 5,857 imágenes del total del conjunto de TT100K para la fase de entrenamiento y prueba, desarrollando un modelo que reconoce y clasifica señales con dígitos numéricos, dentro de un recorrido en automóvil, utilizando una red neuronal basada en aprendizaje profundo, modificada y propuesta por los creadores, presentando en un desempeño del 72 % medido en mAP.

Posteriormente, los trabajos de Cao et al., Vennelakanti et al., Barodi et al. y Yao et al. [6, 25, 2, 26], hacen uso del conjunto de datos alemanes denominados GTSRB para evaluar el desempeño del modelo propuesto. Los estudios mencionados también utilizan imágenes de diversos países como lo son de China, Nueva Zelanda, India, México y Marruecos.

Particularmente, Cao et al. [6] realizaron reconocimiento y detección de señales de tráfico, por medio de la arquitectura LeNet-5 CNN y otra herramienta modificada para mejorar la seguridad del individuo al momento de conducir, permitiendo tener un asistente en la conducción de vehículos, su desempeño logró un mAP de 99.75 %.

Vennelakanti et al. [25] utilizaron 21,383 imágenes del total del conjunto de datos German Traffic Sign Recognition Benchmark (GTSRB), donde propusieron un modelo con una arquitectura denominada CNN-Ensemble, teniendo la finalidad de reconocer y detectar señales circulares y triangulares pertenecientes a Bélgica y Alemania, obteniendo una tasa de reconocimiento del 99.75 %.

Posteriormente, Ramkumar et al. [20] proponen una arquitectura de elaboración propia contando con 500 capas ocultas, la cual permite realizar reconocimiento y clasificación, utilizando el total de datos del conjunto de datos alemán para el entrenamiento del modelo, presentando resultados con una exactitud de 98.64 %. De igual manera, para la investigación de Barodi et al. [2] proponen una de arquitectura de redes neuronales convolucionales, donde utilizan distintas técnicas de visión e



Fig. 3. Anotación de imágenes por medio de LabelImg.

implementación de inteligencia artificial aplicado a tráfico en tiempo real, presentando resultados del 98 % de precisión, una sensibilidad del 98 % y un valor F1 del 98 %. El estudio de Yao et al. [26] propone un modelo inteligente para reconocer señales asiáticas, su evaluación se realiza en diversos conjuntos de datos.

Para el primer modelo se utiliza la base de datos de GTSDB, donde se utiliza la cantidad de 15,734 imágenes, la arquitectura empleada es una combinación de YOLOv4 Tiny [11], de igual manera un método de fusión de características basado en una Red de Pirámide de Características Adaptativas (AFPN, por sus siglas en inglés) [11] y un Bloque de Campo Receptivo (RFB, por sus siglas en inglés), el cual consiste en poner escalas de campos receptivos mejorando la capacidad de extracción de características para la red neuronal convolucional [15], logrando el 93.5 % y 82.4 % de precisión y sensibilidad, asimismo, presentan porcentajes de valor F1 de 87.6 % y mAP de 86.8 %. De lo anterior es importante destacar la falta de datos con señales en el territorio Mexicano.

3. Materiales y métodos

En la Figura 1 se presenta la metodología usada para el desarrollo del modelo de reconocimiento y detección de señales de tráfico utilizando la arquitectura de una red neuronal convolucional designada como YOLOv4.

De entrada, se realizó una recolección de datos de dos maneras, en la primera se utilizó una cámara de video dentro de un vehículo haciendo una recorrido por la ciudad, para la segunda forma se utilizaron herramientas de Google que permiten la visualización de rutas en calles de zonas urbanas y carreteras.

Una vez terminada la recopilación, se procedió a aplicar la técnica de aumento de datos para transformar las imágenes con vista a diferentes cambios climáticos, tales como días muy soleados simulando ciudades que se encuentran cerca del mar, neblinosos suponiendo a los lugares donde existe más humedad en el ambiente y por último, días donde la temperatura es constante.

Tabla 2. Especificación del etiquetado de imagen.

Clase	Coordenada en x	Coordenada en y	Ancho	Alto
Ceda el paso	0.0.153802	0.282593	0.113688	0.192693
Prohibido estacionarse	0.545817	0.315544	0.122814	0.267192

Una vez completa la tarea anterior, se procede a delimitar la zona de interés, es decir, por medio de herramientas de etiquetado se procede a delimitar donde se encuentran las señales de tránsito, con el fin de proponer el conjunto de datos para entrenamiento, validación y prueba.

El siguiente paso es el entrenamiento del modelo computacional, donde se utiliza una red pre-entrenada con una arquitectura utilizada para predicciones, por medio de supresión máxima, asimismo, la primera fase de la red neuronal convolucional es conocida como CSPDarknet53 y el resto de la estructura contiene diversas capas de extracción de características [22].

El bloque de Agrupación de Pirámide Espacial (SPP, por sus siglas en inglés), está conectado desde las capas principales hasta las capas de extracción de características de igual manera que la Red de agregación de rutas (PANet, por sus siglas en inglés), la cual extrae características en diferentes etapas del procesamiento, por último, se encuentra la capa de detección donde realiza una combinación de capas en varios niveles ayudando a detectar objetos de diferentes tamaños. Cabe mencionar que la función de activación utilizada es la Mish [22].

3.1. Recolección de datos

En la primera etapa se hizo la recolección de señales de tráfico verticales, las cuales se dividen en preventivas, restrictivas e informativas. Este trabajo se enfoca en señales restrictivas y preventivas, debido a que las informativas varían dependiendo el área de la región. Las imágenes recolectadas fueron adquiridas de las ciudades de Puebla de Zaragoza, Veracruz, Zacatecas, Monterrey y Xalapa, dentro de las avenidas principales, parques, áreas verdes, mercados y áreas urbanas.

Por otra parte, para la captura de las imágenes utilizadas en el entrenamiento del modelo computacional de reconocimiento y detección fue de dos maneras diferentes, la primera opción fue por medio de una herramienta de Google Maps denominada Google Street View, la cual permite la visualización de calles y avenidas de manera virtual.

Mientras, que la segunda opción fue dentro las calles principales de la capital de Zacatecas, con la ayuda de una cámara de video (GoPro Session 5). Obteniendo así las siguientes categorías representadas en la Tabla 1.

El total de las imágenes recolectadas fueron 1,966 y el 28.5 % de estas representan señales restrictivas, 43 % pertenecen a señales preventivas y el 28.5 % son señales informativas. Además, como las señales de tránsito recolectadas se encuentran en un entorno no controlado se presentan aspectos fotográficos como cambio de iluminación, diferentes ángulos y enfoque variable.

Tabla 3. Conjunto de datos de entrenamiento, validación y prueba.

	Clase	Entrenamiento	Validación	Prueba	Total
1	Alto	145	20	8	173
2	Prohibido estacionarse	350	41	17	408
3	Ceda el paso	57	9	3	69
4	Limite de velocidad	351	44	27	422
5	Cruce peatonal	281	57	17	355
6	Curva	112	16	7	135
7	Glorieta	54	12	2	68
8	Incorporación al transito	87	59	10	156
9	Salida	109	26	5	140
10	Reductor de velocidad	124	16	4	144
11	Parada de autobus	12	6	2	20
12	Zona de discapacitados	23	10	2	35
13	Gasolinera	12	6	2	20
14	Zona de taxis	11	2	2	15
Total		1728	324	108	2160

3.2. Aumento de datos

El aumento de datos consiste en un conjunto de técnicas que permite ampliar artificialmente la cantidad de información a partir de datos [7], realizando modificaciones a la imagen. Por lo que, para que exista variedad dentro del conjunto de datos recolectado se procederá a modificar la saturación en un rango de menor a mayor y la visibilidad realizando difuminado, con el fin de simular diferentes condiciones climáticas como días con lluvia o neblina, y condiciones de iluminación como nublado y soleado. Teniendo como ventaja de que el conjunto se adapte más en línea con escenas reales dentro de la vialidad.

3.3. Cuadro delimitador en la región de interés

Para realizar la detección del objeto de interés en el interior de la imagen se procede a realizar el etiquetado por medio de un cuadro delimitador conocido como BoundingBox, la delimitación se realiza mediante la herramienta, indicando donde se encuentra la señal de tránsito dentro de la ilustración, con el fin de comparar la etiqueta a predecir contra la del modelo.

La Figura 3 representa la anotación que se realizó por medio de la interfaz de etiquetado LabelImg con una imagen del conjunto de datos propuesto, donde se puede observar dos cuadros delimitadores, uno perteneciente a la clase de ceda el paso y otro a la señal de prohibido estacionarse. Una vez concluido, se procede a guardar en formato tipo YOLO, este formato guarda las coordenadas de la región etiquetada, es decir, las coordenadas x y y demuestran el centro del cuadro delimitador en el que se encuentra el objeto, además del “ancho” y “alto” del cuadro que se debe detectar.

De igual modo, en la Tabla 2 se observa un ejemplo de las coordenadas almacenadas mencionadas anteriormente y correspondientes a las clases de ceda el paso y prohibido estacionarse.

Tabla 4. Parámetros del entorno de experimentación.

Sistema operativo	Ubuntu 9.4.0
Tipo de GPU	NVIDIA Tesla T4
RAM	25.5
Lenguaje de programación	Python 3.8.10
Versión de CUDA	11.6

3.4. Preparación de imágenes de entrenamiento y prueba

Para la preparación de imágenes de entrenamiento y prueba se realizó una división del grupo total de imágenes en una proporción de 80:15:5, correspondiente a un 80 % para entrenamiento, 15 % para validación y un 5 % de prueba, como lo menciona [21] el set de prueba nos permite visualizar si el modelo generado realiza reconocimiento y detección con un desempeño aceptable.

En este trabajo la métrica para medir el rendimiento del modelo es mean Average Precision (mAP). Esta métrica permite evaluar la precisión con la cual se detecta el área que ocupa el objeto, para el cálculo de esta primero se mide la precisión promedio para cada clase, posteriormente se calcula la media de las precisiones obtenidas [19].

De igual modo, la distribución en la que se encuentran las imágenes de entrenamiento validación y prueba se pueden observar en la Tabla 3 el conjunto de datos está conformado por 14 clases, donde están incluidas los tres tipos de señalamientos verticales, de igual manera, se muestra la cantidad equivalente de imágenes para el proceso de entrenamiento, validación y prueba.

4. Experimentación y resultados

4.1. Ambiente de experimentación

Para llevar a cabo el experimento de reconocimiento y detección de señales de tránsito implementando la arquitectura de YOLOv4 y basándose en el conjunto de datos propuesto, se utilizó el entorno de codificación de Google Colaboratory, debido a que contiene un soporte de GPU y dependencias necesarias [3]. En la Tabla 4 se muestra los parámetros que se utilizaron para el entrenamiento del modelo de reconocimiento y detección.

4.2. Configuración de la red y generación del modelo

Una vez separado y preparado el conjunto de datos, se deberán establecer dentro de Google Drive, debido a que el entorno experimental de Google Colab requiere acceso a los datos de esa plataforma. Posterior a esto, se procede a adaptar los parámetros establecidos mostrados en la Tabla 4, debido a que se requiere la GPU ya que se necesitan dependencias que posibiliten utilizar el procesamiento gráfico como el instrumento de CUDA [18], teniendo una mejor optimización de costo computacional para el desarrollo del modelo.

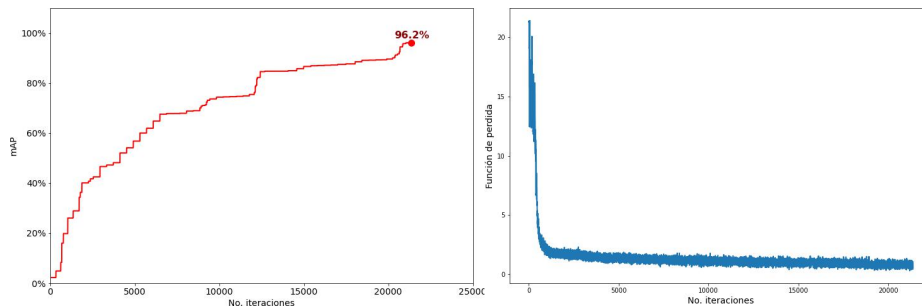


Fig. 4. Rendimiento promedio de la red YOLOv4 en el reconocimiento de señales de tránsito.

Además, del cambio de entorno de ejecución se requieren herramientas que contribuyan a la visión artificial, como OpenCV[5]. Posteriormente, se realiza la implementación de la arquitectura a utilizar desde el repositorio de los autores [4] y se procede a descargar los pesos pre-entrenados para adaptarlos dentro del entrenamiento.

4.3. Resultados del entrenamiento

Para llevar a cabo el entrenamiento del modelo, se concretó por medio de 1,728 imágenes donde están incluidas las 14 clases existentes de señales de tránsito de tipo restrictivas, preventivas e informativas y sus variantes. En la Figura 4, se observa el rendimiento promedio del modelo, donde alcanzó un porcentaje mayor del 95 % de mAP en la detección del objeto de estudio durante 21,371 iteraciones en un aproximado de 10 horas.

Una vez terminada la ejecución del entrenamiento, se procede a evaluar la predicción del modelo desarrollado por medio del algoritmo YOLOv4, para este proceso se utilizó la partición de 108 imágenes pertenecientes al conjunto de prueba.

Asimismo, como demostración de este procedimiento algunas de las pruebas de experimentación se muestran en la Figura 5, donde se puede observar que las 7 imágenes presentadas contienen una o dos señales de tránsito diferentes tales como ceda el paso, prohibido estacionarse, curva, glorieta, cruce de peatones, alto, incorporación y límite de velocidad.

Además de mostrar la delimitación de la detección realizada por el modelo, representada en color morado y por otra parte, en color amarillo, el resultado esperado ideal o mejor conocido como ground truth, el cual se generó mediante la interfaz de etiquetado. Por último, se observa la diferencia que existe entre ambos cuadros delimitadores.

Es relevante destacar que el algoritmo implementado en este estudio recorre la imagen de entrada hasta el punto de identificar el objeto de interés y, posteriormente delimitarlo de acuerdo a su respectiva clase. Los resultados presentados demuestran una alta eficiencia en el reconocimiento de diferentes tipos de señales de tránsito, lo que se traduce a un excelente desempeño del modelo.



Fig. 5. Demostración de la posición real del objeto y la detección del algoritmo propuesto.

5. Conclusiones y trabajos futuros

En el presente trabajo, se hizo uso de una red neuronal convolucional denominada You Only Look Once (YOLOv4), la cual permite realizar un modelo para reconocer y detectar elementos viales dentro de las zonas urbanas y carreteras de tipo restrictivas, preventivas e informativas.

Asimismo, se han aplicado diferentes técnicas de aumento de datos para alimentar el conjunto de datos propuesto para este estudio y tener variabilidad y robustez al momento de desarrollar el modelo. Este enfoque se utiliza comúnmente dentro del aprendizaje computacional mejorando significativamente la capacidad del modelo al tener nuevos escenarios.

La Figura 4 muestra el comportamiento del modelo en términos de la métrica establecida (mAP), las cuales presentan un excelente desempeño con un porcentaje mayor al 95%. Además, el ejemplo de la Figura 5 ilustra el reconocimiento exitoso y preciso en imágenes con distintas señales de tránsito, donde se puede observar al delimitar la posición exacta dentro de la imagen. Estos resultados resaltan la capacidad del modelo para abordar eficientemente la tarea de reconocimiento y detección de señales de tráfico en diferentes entornos complejos.

Finalmente, se propone como trabajo futuro el incremento del conjunto de datos utilizado en este estudio, con el propósito de abarcar una mayor diversidad de señalética urbana presente en el territorio mexicano. Además, se plantea la implementación de nuevas arquitecturas para poder comparar y evaluar el desempeño en relación al modelo propuesto.

Agradecimientos. Los autores agradecen al Programa Nacional de Posgrados de Calidad (PNPC) del Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico a través de la convocatoria Becas Nacional (Tradicional) 2021-2.

Referencias

1. Almutairy, F., Alshaabi, T., Nelson, J., Wshah, S.: ARTS: Automotive repository of traffic signs for the United States. In: IEEE Transactions on Intelligent Transportation Systems, vol. 22, pp. 457–465 (2021) doi: 10.1109/tits.2019.2958486
2. Barodi, A., Bajit, A., Zemmouri, A., Benbrahim, M., Tamtaoui, A.: Improved deep learning performance for real-time traffic sign detection and recognition applicable to intelligent transportation systems. International Journal of Advanced Computer Science and Applications, vol. 13, no. 5 (2022)
3. Bisong, E.: Building machine learning and deep learning models on google cloud platform, Apress (2019) doi: 10.1007/978-1-4842-4470-8
4. Bochkovskiy, A., Wang, C. Y., Mark-Liao, H. Y.: Yolov4: Optimal speed and accuracy of object detection (2020) doi: 10.48550/arXiv.2004.10934
5. Bradski, G.: The OpenCV library. Dr Dobb's Journal of Software Tools, vol. 25, no. 11, pp. 120–125 (2000)
6. Cao, J., Song, C., Peng, S., Xiao, F., Song, S.: Improved traffic sign detection and recognition algorithm for intelligent vehicles. Sensors, vol. 19, no. 18, pp. 4021 (2019) doi: 10.3390/s19184021
7. Dilmegani, C.: What is data augmentation? techniques y examples in 2023 (2022) research. aimultiple.com/data-augmentation/
8. Dubey, U., Chaurasiya, R. K.: Efficient traffic sign recognition using CLAHE-based image enhancement and ResNet CNN architectures. International Journal of Cognitive Informatics and Natural Intelligence, vol. 15, no. 4, pp. 1–19 (2021) doi: 10.4018/IJCINI.295811
9. Gao, X., Chen, L., Wang, K., Xiong, X., Wang, H., Li, Y.: Improved traffic sign detection algorithm based on Faster R-CNN. Applied Sciences, vol. 12, no. 18, pp. 8948 (2022) doi: 10.3390/app12188948
10. Houben, S., Stallkamp, J., Salmen, J., Schlipsing, M., Igel, C.: Detection of traffic signs in real-world images: The german traffic sign detection benchmark. In: International Joint Conference on Neural Networks, pp. 1–8 (2013) doi: 10.1109/ijcnn.2013.6706807
11. Khokhlov, I., Davydenko, E., Osokin, I., Ryakin, I., Babaev, A., Litvinenko, V., Gorbachev, R.: Tiny-YOLO object detection supplemented with geometrical data. In: IEEE 91st Vehicular Technology Conference, pp. 1–5 (2020) doi: 10.1109/vtc2020-spring48590.2020.9128749
12. Lahmyed, R., Ansari, M. E., Kerkaou, Z.: Automatic road sign detection and recognition based on neural network. Soft Computing, vol. 26, no. 4, pp. 1743–1764 (2022) doi: 10.1007/s00500-021-06726-w
13. Li, Z., Chen, M., He, Y., Xie, L., Su, H.: An efficient framework for detection and recognition of numerical traffic signs. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 2235–2239 (2022) doi: 10.1109/icassp43922.2022.9747406
14. Liu, S., Cai, T., Tang, X., Zhang, Y., Wang, C.: Visual recognition of traffic signs in natural scenes based on improved RetinaNet. Entropy, vol. 24, no. 1, pp. 112 (2022) doi: 10.3390/e24010112
15. Liu, S., Huang, D., Wang, Y.: Receptive field block net for accurate and fast object detection (2017) doi: 10.48550/ARXIV.1711.07767
16. Liu, Y., Shi, G., Li, Y., Zhao, Z.: M-YOLO: Traffic sign detection algorithm applicable to complex scenarios. Symmetry, vol. 14, no. 5, pp. 952 (2022) doi: 10.3390/sym14050952
17. Narejo, S., Talpur, S., Memon, M., Rahoo, A.: An automated system for traffic sign recognition using convolutional neural network. 3c Tecnología: Glosas de Innovación Aplicadas a la PYME, vol. 9, no. 1, pp. 119–135 (2020)

18. NVIDIA, Vingelmann, P., Fitzek, F. H.: Cuda, release: 10.2.89 (2020) developer.nvidia.com/cuda-toolkit
19. Padilla, R., Passos, W. L., Dias, T., Netto, S. L., da Silva, E. A. B.: A comparative analysis of object detection metrics with a companion open-source toolkit. *Electronics*, vol. 10, no. 3, pp. 279 (2021) doi: 10.3390/electronics10030279
20. Ramkumar, S., Ganapathy, R., Sridhar, K.: Traffic sign detection and recognition using CNN. *ECS Transactions*, vol. 107, no. 1, pp. 17447–17455 (2022) doi: 10.1149/10701.17447ecst
21. Recuero-de-los Santos, P.: Datos de entrenamiento vs datos de test (2022) [empresas.blogthinkbig.com/datos-entrenamiento-vs-datos-de-test/](https://blogthinkbig.com/datos-entrenamiento-vs-datos-de-test/)
22. Roszyk, K., Nowicki, M. R., Skrzypczyński, P.: Adopting the YOLOv4 architecture for low-latency multispectral pedestrian detection in autonomous driving. *Sensors*, vol. 22, no. 3, pp. 1082 (2022) doi: 10.3390/s22031082
23. Ruiz-Rivero, M. S., Astorga-Bustillos, F. R., Villa-Herrera, J. E.: La importancia de las señales de tránsito en las vías terrestres. *FINGUACH Revista de Investigación Científica de la Facultad de Ingeniería de la Universidad Autónoma de Chihuahua*, vol. 5, no. 17, pp. 10–11 (2018)
24. Support, T.: Piloto automático y capacidad de conducción autónoma total (2020) www.tesla.com/es/_ES/support/autopilot-and-full-self-driving-capability
25. Vennelakanti, A., Shreya, S., Rajendran, R., Sarkar, D., Muddegowda, D., Hanagal, P.: Traffic sign detection and recognition using a CNN ensemble. In: *IEEE International Conference on Consumer Electronics (ICCE)* (2019) doi: 10.1109/icce.2019.8662019
26. Yao, Y., Han, L., Du, C., Xu, X., Jiang, X.: Traffic sign detection algorithm based on improved YOLOv4-tiny. *Signal Processing: Image Communication*, vol. 107, pp. 116783 (2022) doi: 10.1016/j.image.2022.116783
27. Zhang, H., Zhao, J.: Traffic sign detection and recognition based on deep learning. *Engineering Letters*, vol. 30, no. 2 (2022)
28. Zhang, Y., Yuan, Y., Feng, Y., Lu, X.: Hierarchical and robust convolutional neural network for very high-resolution remote sensing object detection. In: *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, pp. 5535–5548 (2019) doi: 10.1109/TGRS.2019.2900302
29. Zhu, Z., Liang, D., Zhang, S., Huang, X., Li, B., Hu, S.: Traffic-sign detection and classification in the wild. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2110–2118 (2016) doi: 10.1109/CVPR.2016.232