

Modelo de clasificación de imágenes de llantas con YOLO11

Brandon Nahú Nava-Martínez, Juan Carlos Elizondo-Leal

Facultad de Ingeniería y Ciencias, Universidad Autónoma de Tamaulipas, Ciudad Victoria 87000, México

Resumen—En este estudio se implementó un algoritmo basado en YOLO (You Only Look Once) para la detección de defectos en llantas. Se utilizó el modelo YOLOv11n y se entrenó con un dataset específico de imágenes de llantas. Los resultados de la matriz de confusión y la curva de confianza F1 demuestran la eficacia del modelo en la clasificación correcta de llantas en buen estado y llantas pinchadas.

I. INTRODUCCIÓN

La detección temprana y precisa de defectos en llantas es fundamental para garantizar la seguridad en las calles y carreteras al prevenir accidentes. Las llantas en mal estado pueden representar un riesgo significativo para los conductores y otros usuarios de la vía, por lo que es vital desarrollar métodos eficaces para su detección. Con el avance de las tecnologías de inteligencia artificial y el aprendizaje profundo, se han explorado diversos enfoques para abordar este desafío, siendo YOLO uno de los más prometedores.

YOLO es un modelo de detección de objetos de vanguardia que ha demostrado una alta precisión y eficiencia en diversas aplicaciones. Su capacidad para realizar detección en tiempo real lo convierte en una opción ideal para identificación de defectos en llantas. Este estudio se centra en el desarrollo y la evaluación de un modelo basado en YOLO para la detección de llantas en buen estado y llantas pinchadas. Utilizando un dataset específico de imágenes de llantas, se entrenó y se evaluó el modelo para determinar su precisión, recall y F1-score, métricas que son cruciales para la evaluación del rendimiento de modelos de clasificación.

El objetivo principal de este estudio es demostrar la eficacia del modelo YOLOv11n en la clasificación de llantas y proponer mejoras basadas en los resultados obtenidos. Los apartados siguientes describen a detalle la metodología utilizada, los resultados alcanzados y las implicaciones de hallazgos.

II. METODOLOGÍA

A. Algoritmo de Detección

Se utilizó el modelo YOLOv11n para la detección de defectos en llantas. Este modelo es conocido por su capacidad de realizar detecciones rápidas y precisas. El dataset utilizado incluye imágenes etiquetadas de llantas en buen estado y llantas pinchadas, lo que permite al modelo aprender a

diferenciar entre ambas categorías.

B. Preprocesamiento de Datos

El preprocesamiento de datos es una etapa crucial en la implementación de modelos de aprendizaje automático. En este estudio, se utilizó un conjunto de datos etiquetados en **Roboflow** que incluye imágenes de llantas en diferentes condiciones. Las imágenes fueron redimensionadas a 640x640 píxeles y normalizadas para mejorar la eficiencia del entrenamiento del YOLO.

C. Modelo de Clasificación

YOLOv11n es la última versión de la serie de modelos YOLO (You Only Look Once) desarrollada por ultralytics. Es un modelo de detección de objetos en tiempo real que se ha mejorado significativamente en comparación con sus predecesores. YOLOv11 se caracteriza por su alta eficiencia computacional, flexibilidad y capacidad para realizar diversas tareas de visión por computadora, como la detección de objetos, segmentación de instancias y clasificación de imágenes.

YOLO utiliza una red neuronal convolucional (CNN) para procesar las imágenes de entrada. La imagen se divide en una cuadrícula de $S \times S$ celdas. Cada celda de la cuadrícula es responsable de predecir un número fijo de “bounding boxes” (cajas delimitadoras) y sus respectivos puntajes de confianza. Estos puntajes reflejan la certeza del modelo sobre la presencia de un objeto y la precisión de la caja delimitadora. Cada celda produce varias bounding boxes con valores de confianza asociado. La confianza es una medida de cuán segura está la celda de que la caja contiene un objeto y de cuán precisa es la predicción de la ubicación de la caja.

Además de predecir las bounding boxes, cada celda predice las probabilidades de clase para los objetos. Estas probabilidades indican la probabilidad de que el objeto dentro de la caja pertenezca a una clase específica. La puntuación de predicción se calcula multiplicando la probabilidad de clase por puntaje de confianza de la bounding box. Esto permite determinar las cajas con mayor probabilidad de contener objetos relevantes. Para eliminar predicciones redundantes, YOLO aplica la técnica de supresión de no máximos. Esta técnica filtra las cajas con puntuaciones bajas y mantiene solo las predicciones más confiables.

D. Entrenamiento del Modelo

* Revista Argentina de Trabajos Estudiantiles. Patrocinada por la IEEE.

Una vez que tenemos el dataset etiquetado, se dividioo en Train (88%), Valid (8%), Test (4%). Al dividirlo de esta forma el modelo tiene un mejor desarrollo y evaluación. Cada una de estas divisiones tiene un objetivo:

- Train: este conjunto de datos ayuda al modelo a minimizar el error en las predicciones aprendiendo a identificar patrones y características relevantes en los datos.
- Valid: El conjunto de validación se usa para ajustar los hiperparámetros del modelo y prevenir el sobreajuste.
- Test: Este conjunto se utiliza para evaluar el rendimiento final del modelo. Para ello, se utilizan imágenes diferentes a Train y Valid.

El modelo se entrenó utilizando los parámetros que se muestran en la tabla I.

TABLA I
PARÁMETROS PARA EL ENTRENAMIENTO DEL MODELO

Parámetro	Valor
Model Weights	YOLO11-nano
Dataset	341 img (Custom Dataset)
Input Image Size	640x640px
Epochs	100
Learning Rate	0.001 (Inicial)

Durante el entrenamiento, se optimizaron las funciones de pérdida para la clasificación de objetos y la regresión de los cuadros delimitadores. La tasa de aprendizaje se ajustó de forma dinámica para mejorar la capacidad del modelo en la detección precisa del estado de las llantas.

E. Métricas de evaluación

Principalmente se utilizan las siguientes métricas:

- Mean Average Precision (mAP): Esta metrica es usada para evaluar modelos de detección de objetos. La mAP se obtiene como el promedio del Average Precision (AP) para cada clase, que se calcular como el área bajo la curva de precisión/recall (PR).
- Box_loss: Esta métrica mide qué tan precisas son las cajas delimitadoras (bounding boxes) predichas por el modelo en comparación con las cajas reales. Las bounding boxes son las cajas que el modelo utiliza para rodear y localizar los objetos en una imagen. Un valor de box_loss bajo indica que el modelo está prediciendo cajas más cercanas a la ubicación real de los objetos.
- CIs_loss: Esta métrica evalúa la precisión en la clasificación de los objetos detectados dentro de las cajas. Después de que el modelo detecta una caja alrededor de un objeto, necesita clasificar correctamente de qué objeto se trata (por ejemplo, si es una "rueda pinchada" o una "rueda buena").
- Dfl_loss: Es una métrica relacionada con la precisión en el ajuste de los bordes de las cajas. La pérdida focal de distribución (DFL, por sus siglas en inglés) extiende la pérdida focal de etiquetas discretas a continuas, lo que permite la optimización exitosa de representaciones

mejoradas que combinan estimación de calidad y predicción de clases. Esto permite una descripción precisa de la distribución flexible en datos reales eliminando el riesgo de inconsistencia. Por lo tanto, un dfl_loss bajo indica que el modelo está ajustando bien los bordes de las cajas.

III. RESULTADOS

A. Matriz de Confusión Normalizada

Para generar el modelo de clasificación se utilizaron los hiperparámetros de la Tabla II. El rendimiento del modelo se evaluó utilizando las métricas Accuracy, Precision, Recall, F1-score. Estas métricas se definen como:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$F1_score = \frac{2 \times Precision \times Recall}{Recall + Precision} \quad (4)$$

Donde TP denota el número de verdaderos positivos, TN denota el número de verdaderos negativos, FN denota el número de falsos negativos, FP denota el número de falsos positivos [1].

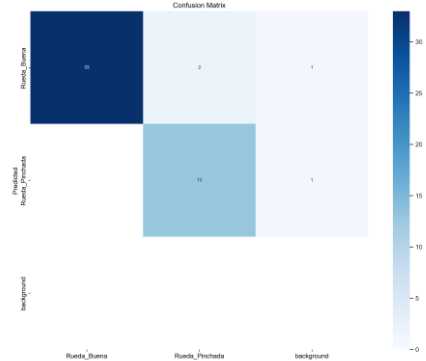
TABLA II
HIPERPARÁMETROS

Parámetro	Valor
Epochs	100
Image size	640
Workers	4

En la tabla III podemos observar las métricas obtenidas utilizando el modelo propuesto.

TABLA III
RESULTADOS DE LA CLASIFICACIÓN

Class	Precision	ReCall	F1-score
Rueda pinchada	1	0.9705	0.985
Rueda buena	0.928	1	0.962



En la figura 2 se muestra la matriz de confusión normalizada, la cual permite analizar el desempeño del modelo en la clasificación de las clases. Esta métrica escala los valores de la matriz de confusión entre 0 y 1.

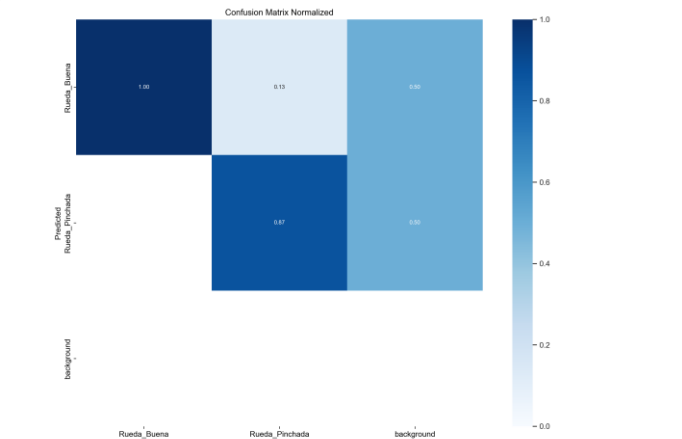


Fig. 2. Matriz de confusión normalizada.

Se tomo en cuenta también la gráfica mostrada en la curva F1-Confidence (fig. 3) la cual representa la relación entre la confianza que un modelo de aprendizaje automático tiene en una predicción y la precisión de esa predicción, medida por el F1_score.

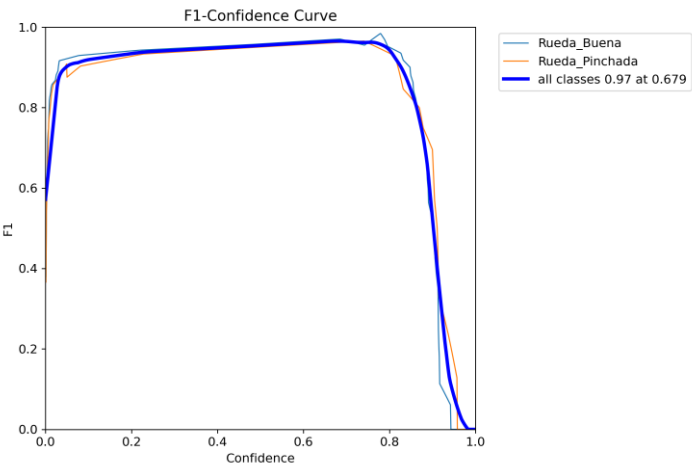


Fig. 3. Curva F1-Confidence; el eje X representa la confianza que el modelo tiene en su predicción y el eje Y representa el F1_score (un valor 1 indica una precisión perfecta).

La curva resultante del modelo entrenado sugiere un desempeño sólido en la tarea de clasificación de ruedas. Presenta altos valores de F1, lo que refleja una notable precisión y recall en clasificación. Esto quiere decir que el modelo está identificando correctamente ambas clases.

Como se puede observar en la figura 3, la curva muestra estabilidad en varios niveles de confianza,

lo que sugiere un rendimiento constante del modelo a medida que aumenta la confianza. Esto garantiza que el modelo mantenga su precisión incluso cuando sus predicciones sean menos seguras.

De igual forma, la curva Precision-Confidence (fig. 4) nos proporciona una visión más detallada de como el modelo entrenado asigna confianzas a sus predicciones.

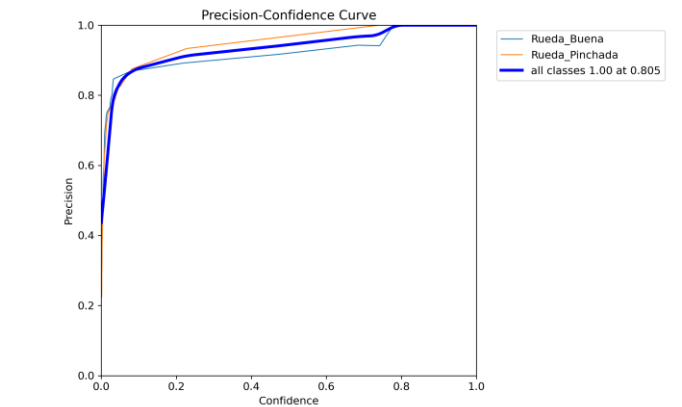


Fig. 4. Curva Precision-Confidence.

Al analizar la grafica de la curva de Precision-Confidence (fig. 4) observamos que cuenta con una precisión alta a medida que aumenta la confianza.

De acuerdo con las métricas, el modelo puede distinguir de manera confiable entre llantas pinchadas y normales. El modelo se probó en entornos reales con fotos y videos como se muestra en la figura 5 y 6.



Fig. 5. Prueba en entorno real del modelo.



Fig. 6. Prueba en entorno real del modelo.

IV. CONCLUSIÓN

El uso de YOLO para la detección de llantas pinchadas ha demostrado ser una técnica eficaz. Los resultados de la matriz de confusión y la curva F1 confirman que el modelo tiene un alto rendimiento en cada clasificación de llantas en diversas condiciones. Futuras investigaciones se centrarán en mejorar la precisión del modelo y su aplicación en entornos industriales.

REFERENCIAS

- [1] J. D. Filoteo -Razo, J. C. Elizondo -Leal, J. R. Martinez -Angulo, J. H. Barron - Zambrano, A. Díaz -Manriquez, V. P. Saldivar -Alonso, J. M. Estudillo -Ayala, and R. Rojas -Laguna, "Non -Invasive Optoelectronic System for Color -Change Detection in Oranges to Predict Ripening by Using Artificial Neural Networks," IEEE Photonics Journal, vol. 15, no. 5, pp. 1 -10, 2023.