

Detección de antiespacios urbanos usando YOLO: Caso de estudio Mexicali

Jorge Adrián Martínez López, Héctor De la Torre Gutiérrez,
Francisco Javier Hernández López

Centro de Investigación en Matemáticas,
Unidad Aguascalientes,
México

{adrian.martinez,hector.delatorre,fcoj23}@cimat.mx

Resumen. El estudio de los antiespacios urbanos ha traído conclusiones interesantes en el campo de la dinámica urbana, pero el número de investigaciones son escasos puesto que la identificación y localización de estos antiespacios resulta ser una tarea laboriosa y complicada. Es por lo anterior que es indispensable contar con una herramienta que facilite a los tomadores de decisiones la detección de estos antiespacios. En este trabajo, se entrenan los modelos de detección de objetos YOLOv4 y YOLOv5, usando imágenes satelitales de alta resolución, para detectar vacíos urbanos y espacios abandonados. Los resultados muestran una precisión promedio del 0.66 y 0.71 utilizando YOLOv4 y YOLOv5 respectivamente. Sin embargo, la detección de cada antiespacio por separado muestra un buen desempeño en la detección de vacíos urbanos pero un desempeño pobre en la detección de espacios abandonados, lo que abre las puertas a futuras investigaciones con el objetivo de mejorar el desempeño de estos modelos.

Palabras clave: Antiespacios urbanos, detección de objetos, deep learning, YOLOv4, YOLOv5, redes neuronales, dinámica urbana.

Urban Antispaces Detection Using YOLO: Case Study Mexicali

Abstract. The urban antispaces study has brought interesting conclusions in the urban dynamics area, but the number of investigations is scarce since identifying and locating these antispaces is laborious and complex. For this reason, a tool for decision-makers is necessary to facilitate the detection of these antispaces. In this work, YOLOv4 and YOLOv5 object detection models are trained using high-resolution satellite images to detect urban empty and abandoned spaces. The results show an average precision of 0.66 and 0.71 for YOLOv4 and YOLOv5, respectively. However, the individual detection of each antispaces shows a good performance in detecting urban empty spaces but a poor performance in detecting abandoned spaces, which depicts further research aimed at improving the performance of these models.

Keywords: Urban antispaces, object detection, deep learning, YOLOv4, YOLOv5, neural networks, urban dynamics.

1. Introducción

De acuerdo con [6], un antiespacio puede asociarse a espacios perdidos, indecisos o desprovistos de función dentro de un espacio urbano. Estos se pueden clasificar en tres grandes grupos (ver Figura 1):

- **Vacío urbano:** Tiene como principal característica que es un espacio sin edificar y sin ningún uso.
- **Espacio abandonado:** Son aquellas edificaciones ya sea terminadas o sin terminar que se encuentran abandonadas y que se han degradado con el paso del tiempo.
- **Remanente urbano:** Son aquellos espacios que se derivan de un proyecto urbano de gran escala y los cuales al finalizar el proyecto dejan de cumplir un propósito.

Los vacíos urbanos han mostrado tener relevancia en el estudio de la dinámica urbana, pues en ciudades fronterizas de México se ha encontrado que su distribución está asociada con la escala de la ciudad, siendo las ciudades de menor escala las que presentan una distribución más homogénea [7]. Por otro lado, las ciudades de mayor escala presentan una correlación espacial entre los vacíos urbanos, la pobreza y las condiciones de seguridad.

Esto provoca que la población con más recursos busque habitar sectores más favorables rechazando estas zonas de la ciudad, ocasionando grandes diferencias en la morfología urbana, en concreto presentando un desequilibrio entre el espacio construido y no construido, propiciando la segregación social [8].

En ciudades fronterizas, como Mexicali, se tiene una alta presencia de vacíos urbanos, los cuales se podrían explicar debido a los altos flujos migratorios e intercambio de mercancía característicos de este tipo de ciudades [2], también por el gran crecimiento de la mancha urbana debido al *boom* económico surgido del Tratado de Libre Comercio de América del Norte (TLCAN) [5]. Además de los hallazgos que han sido descubiertos con respecto a los vacíos urbanos, de acuerdo con [6], la identificación y análisis de antiespacios urbanos puede traer los siguientes beneficios:

- Pueden constituir una valiosa fuente de información que puede dar pie a proyectos de rehabilitación y regeneración urbana, al ser áreas de oportunidad para la reconfiguración de las estructuras urbanas.
- En el ámbito económico, se puede estudiar su influencia en el valor del suelo y plusvalía de la zona.
- En el ámbito legal, permitiría profundizar en estudios para determinar los mecanismos mediante los cuales el Estado podría tomar el control de aquellos vacíos urbanos o espacios abandonados de propiedad privada, que con el pasar de los años se han quedado sin un propietario legítimo que pueda reclamar los derechos de propiedad.

Por todo lo anterior, se tiene interés en el estudio de los antiespacios, en particular los vacíos urbanos, pero debido a que la tarea de identificación no es sencilla, estos se ven limitados. Así pues, surge la necesidad de desarrollar un modelo que facilite la identificación de los antiespacios urbanos sin intervención directa humana.

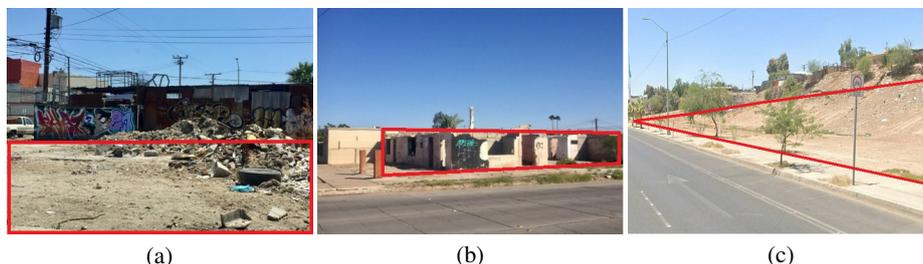


Fig. 1. Ejemplos de los tres antiespacios, marcados dentro del polígono rojo. (a) Vacío urbano. (c) Espacio abandonado. (d) Remanente urbano.

Con el avance de las técnicas de visión computacional, teledetección y aprendizaje supervisado, se abre la interrogante de si es posible utilizar estas técnicas para detectar de manera autónoma los vacíos urbanos. Se espera que utilizando imágenes satelitales de alta resolución y un modelo de detección de objetos como You Only Look Once v4 y v5 (YOLOv4 y YOLOv5) sea posible detectar los antiespacios urbanos.

El resto del manuscrito se organiza como se menciona a continuación. En la sección 2 se exponen las principales técnicas de deep learning y los detectores de objetos más populares así como su uso en trabajos relacionados con la teledetección urbana. En la sección 3 se describe la metodología aplicada en la realización de este trabajo. Los resultados se muestran y discuten en la sección 4. Por último, en la sección 5, se presentan las conclusiones y recomendaciones para trabajo futuro.

2. Marco teórico

La teledetección urbana de objetos ha avanzado a la par del desarrollo y mejora de sensores, la fotogrametría y las técnicas de procesamiento de imágenes. Así también, la teledetección urbano de objetos se ha visto potencializada por la incorporación de técnicas de deep learning y las redes neuronales convolucionales y recurrentes (CNN y RNN, respectivamente, por sus siglas en inglés)[9].

Generalmente, los modelos de detección se clasifican en dos grandes categorías: a dos etapas y a una etapa. Los detectores a una etapa más utilizados son YOLO y Single-Shot Detector (SSD), mientras que los de dos etapas son Faster y Mask Region Based Convolutional Neural Networks (Faster R-CNN y Mask R-CNN) y Region Based Fully Convolutional Networks (R-FCN) [15].

Para este tipo de tareas, los modelos que utilizan imágenes satelitales han mostrado tener buena capacidad de detección. En [16] se utiliza un modelo de segmentación semántica para la identificación de espacios vacantes en áreas urbanas utilizando DeepLabv3 e imágenes satelitales de alta resolución.

Se entrenó el modelo dividiendo las ciudades en bloques parcialmente traslapados. De acuerdo con los resultados, el modelo en general consigue una precisión superior al 90 %. Por otro lado, muchos investigadores optan por una fusión de datos multimodal para mejorar la habilidad de detección [13].

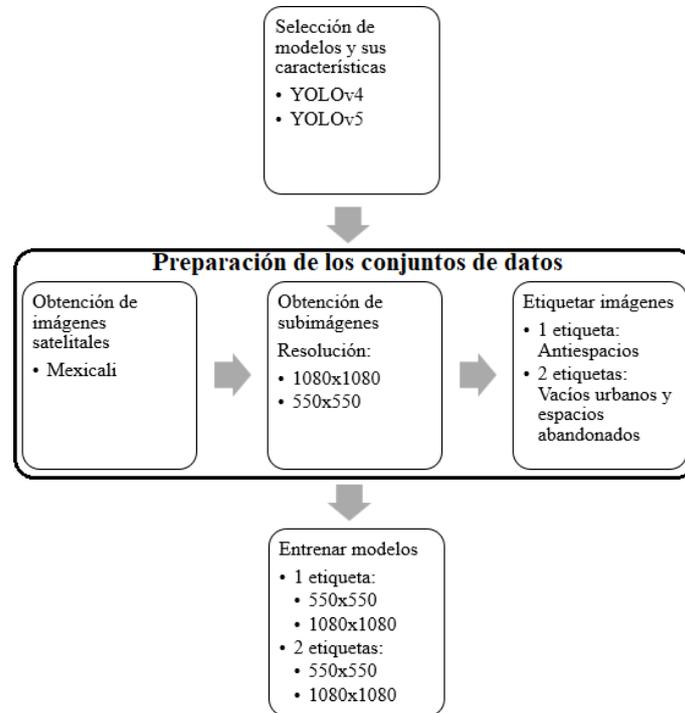


Fig. 2. Diagrama general de la metodología propuesta.

Para la segmentación y detección de áreas funcionales urbanas, se han utilizado imágenes satelitales de alta resolución, puntos sociales de interés e imágenes de alta resolución de las luces nocturnas [4]; además se ha propuesto un modelo de deep learning basado en CNN para imágenes satelitales fusionado con una red con memoria de largo y corto plazo (LSTM, por sus siglas en inglés) para datos de social sensing [1].

También, [3] propusieron un modelo para identificar asentamientos urbanos informales utilizando deep learning y datos provenientes de imágenes satelitales e imágenes de calles (Street-view). En [14] se utilizaron técnicas de deep learning para detectar edificios recientemente construidos que no han sido registrados en el catastro, se utilizaron datos de distintas fuentes con la finalidad de tener información en dos momentos en el tiempo, los cuales fueron entrenados utilizando una CNN.

De los modelos de detección de objetos YOLOv4 y YOLOv5 destacan por su desempeño y flexibilidad de implementación [10]. YOLOv4 destacó por mejorar en gran medida al modelo YOLOv3 al agregar CSPDarknet53 como backbone, Spatial Pyramid Pooling (SPP) y Path Aggregation Network (PAN) para aumentar el campo receptivo [20].

YOLOv5 de igual manera utiliza como backbone CSPDarknet53 y una capa de enfoque al inicio de la red y PAN para aumentar el campo receptivo, aunque en general su arquitectura no se considera tan innovadora comparada con la de YOLOv4, pero ofrece una gran flexibilidad para su implementación [10].



Fig. 3. Imágen satelital de la ciudad de Mexicali.

Además, YOLOv5 ofrece cinco tamaños de modelo distintos lo que permite ajustarse a las necesidades de la tarea a realizar [12], además de utilizar buenas técnicas de aumentación de datos: escalado, ajuste del espacio de color y aumentación de mosaico, este último está también presente en YOLOv4.

Comparaciones entre las distintas versiones de YOLO contra otros detectores muestran que el primero tiene un muy buen desempeño y velocidad [21], y se ha observado una ligera superioridad de YOLOv5 contra YOLOv4 para la detección autónoma de sitios de aterrizaje [17], la detección de aeronaves utilizando imágenes satelitales [11] y en el conteo de multitudes en tiempo real [19].

Pero, no se tiene una comparación entre YOLOv4 y YOLOv5 específicamente en la tarea de detección de antiespacios urbanos utilizando imágenes satelitales, por lo que es de interés investigar cuál de estos dos es mejor para esta tarea.

3. Metodología

En esta sección se explicará los pasos llevados a cabo para la realización de este trabajo. De manera resumida (ver Figura 2), el primer paso consistió en elegir las características de los modelos de YOLOv4 y YOLOv5 a utilizar. Después, con la meta de observar el efecto que tienen las resoluciones de las imágenes en el desempeño de estos modelos se prepararon dos conjuntos de datos, uno con imágenes de resolución de 1080x1080 y el otro con resolución de 550x550.

Una vez definidas las resoluciones, se obtuvieron las imágenes satelitales, las cuales fueron utilizadas para obtener las subimágenes que se etiquetaron con la localización de los antiespacios y finalmente se prepararon los conjuntos de datos para su entrenamiento con los modelos de YOLOv4 y YOLOv5.

Tabla 1. Número de imágenes destinadas al entrenamiento y validación para los dos conjuntos de datos.

Resolución	Entrenamiento	Validación	Total
550×550	780	334	1114
1080×1080	785	337	1122

3.1. Modelo detección de objetos

YOLOv4: Para la implementación de este modelo se ajustaron los siguientes hiperparámetros: 6,000 épocas; tamaño de batch y mini-batch de 64 y 16 respectivamente; tamaño de la red de 896 para las imágenes de resolución de 1080x1080 y 544 para las imágenes de 550x550; tasa de aprendizaje de 0.001; momentum y decaimiento de pesos (weight decay) de 0.949 y 0.0005 respectivamente. Se utilizaron 9 cajas de anclaje, el tamaño de estas cajas se calcula utilizando k-medias con el objetivo de maximizar la intersección sobre la unión (IoU).

YOLOv5: Para su implementación primero es necesario elegir el tamaño del modelo, este ofrece cinco tamaños distintos del modelo: nano, pequeño, mediano, grande y extragrande; para este trabajo se seleccionó el tamaño mediano, el cual ofrece un buen desempeño además de que se adapta a los recursos de cómputo disponibles.

El entrenamiento se ajustó con 1000 épocas y un tamaño de batch de 32. Se eligió el conjunto de hiperparámetros para una alta aumentación de los datos, una tasa de aprendizaje de 0.01; momentum y decaimiento de pesos (weight decay) de 0.937 y 0.0005 respectivamente, los cuales son los predeterminados en YOLOv5.

3.2. Obtención de imágenes

Se cuenta con datos relacionados con la zona de exploración y la localización de los antespacios urbanos en la ciudad de Mexicali, estos están en formato .shp. Estos datos derivaban de una investigación donde se analizó la distribución geoespacial de los antespacios urbanos en ciudades fronterizas mexicanas [6].

Se utilizó el software SAS Planet versión 211230.10225 para obtener dos imágenes satelitales de alta resolución de tres canales (RGB) de la zona de Mexicali, (ver Figura 3), con resolución espacial de 0.25m/píxel y 0.5m/píxel para las imágenes de resolución 1080x1080 y 550x550 respectivamente.

Se obtuvieron las subimágenes a partir de estas dos imágenes de la zona. Partiendo de la idea presentada en [16], cada subimagen tiene un traslape con el objetivo de evitar en la medida de lo posible que un antespacio se encuentre entre dos imágenes, permitiendo que al menos en una imagen el objeto se encuentre completo.

Una vez obtenidas las imágenes se dividieron las imágenes en un conjunto de entrenamiento y uno de validación a una razón de 70-30. Las características del conjunto de datos con las imágenes de alta resolución son:

- Resolución: 1080x1080 píxeles con tres canales (RGB).
- Resolución espacial: 0.25m/píxel.
- Traslape de las imágenes de 80 píxeles en todas las direcciones.

Tabla 2. Número de objetos (antiespacios urbanos) de cada clase de antiespacios en Mexicali.

	Vacíos	Abandonados	Remanentes
Número de objetos	6282	1524	123

Las características del conjunto de datos con imágenes de menor resolución son:

- Resolución: 550x550 píxeles con tres canales (RGB).
- Resolución espacial: 0.5m/píxel.
- Traslape de las imágenes de 50 píxeles en todas las direcciones.

El número de imágenes obtenidas, así como la cantidad destinada al conjunto de entrenamiento y validación, se muestran en la Tabla 1.

3.3. Etiquetado de imágenes

Se crearon las etiquetas en formato YOLO para cada imagen a partir de los archivos .shp con la localización de los antiespacios, se consideraron solo dos clases: vacío urbano y espacio abandonado. Se decidió omitir los remanentes urbanos puesto que el número de estos es muy pequeño (ver Tabla 2) y su localización no resulta ser de interés para tomadores de decisiones en temas relacionados a la dinámica urbana.

Para las pruebas enfocadas en la comparación del desempeño de las distintas resoluciones y versiones de YOLO, se utilizó una sola etiqueta de antiespacios que englobaba los vacíos urbanos y los espacios abandonados. Por otro lado, se utilizaron dos etiquetas, cada una por las dos clases de antiespacios seleccionados, para comparar el desempeño que se tenía en cada clase. En la Figura 4 se muestra a manera de ejemplo como se observaría una imagen etiquetada con los dos antiespacios.

4. Resultados

4.1. Especificaciones del equipo

Los modelos se entrenaron en un equipo con las siguientes especificaciones:

- CPU: Intel Xeon Gold 5222 @3.8Ghz
- GPU: NVIDIA QUADRO RTX 8000 con 48GB GDDR6
- RAM: 48GB DDR4

Se entrenó YOLOv4 en el framework de Darknet, mientras que YOLOv5 se entrenó en el framework de PyTorch. El equipo utilizaba el sistema operativo Ubuntu 20.04.5.

4.2. Métricas de evaluación utilizadas

Se comparó el rendimiento entre ambos modelos utilizando las métricas que estos modelos arrojan al final de su entrenamiento, las cuales son *precision*, *recall*, la puntuación $F1$, la precisión promedio (AP) y la media de la precisión promedio (mAP) con un umbral de la intersección sobre la unión (IoU) de 0.50. A continuación se detalla el cálculo de estas métricas de acuerdo con [18]. Primero es necesario definir los siguientes conceptos básicos:

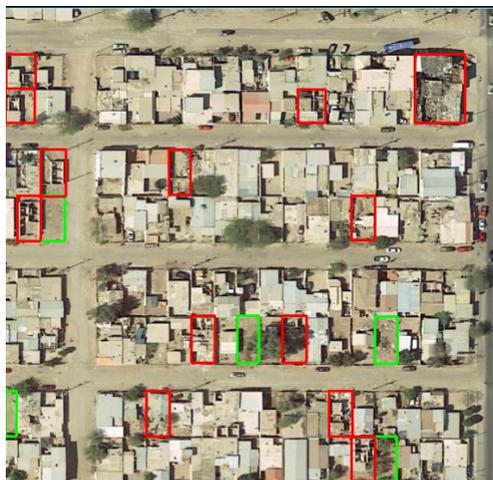


Fig. 4. Etiquetado de los antiespacios. Vacíos urbanos (color verde) y Espacios abandonados (color rojo).

- Verdadero positivo (VP): Una detección correcta de una etiqueta verdadera.
- Falso negativo (FN): Una etiqueta verdadera no detectada.
- Falso positivo (FP): Detección incorrecta de parte del modelo de un objeto no existente.

Es esencial tener un criterio para decidir si el modelo clasifica correcta o incorrectamente cada etiqueta verdadera (antiespacio) dentro de una imagen. El método más común es el IoU.

Considerando que cada etiqueta verdadera tiene una “área” o “caja” asociada B_v y el modelo realiza una predicción de esta “caja” B_p . Entonces el IoU mide el área de traslape de B_v y B_p y lo divide sobre la unión de estas dos, de la siguiente manera:

$$IoU = \frac{\text{area}(B_p \cap B_v)}{\text{area}(B_p \cup B_v)}. \quad (1)$$

Así pues, se puede elegir un umbral u , tal que si $IoU \geq u$ la predicción se considera correcta y si $IoU < u$ la predicción se considera incorrecta. Esto nos permite conocer el número total de VP, FN y FP, para calcular las métricas de precisión P y recall R las cuales se definen de la siguiente manera:

$$P = \frac{VP}{VP + FP}, \quad (2)$$

$$R = \frac{VP}{TP + FN}, \quad (3)$$

donde P la capacidad del modelo de detectar únicamente etiquetas verdaderas, R la capacidad del modelo de detectar todas las etiquetas verdaderas y la puntuación $F1$ es la media armónica de las dos métricas anteriores:

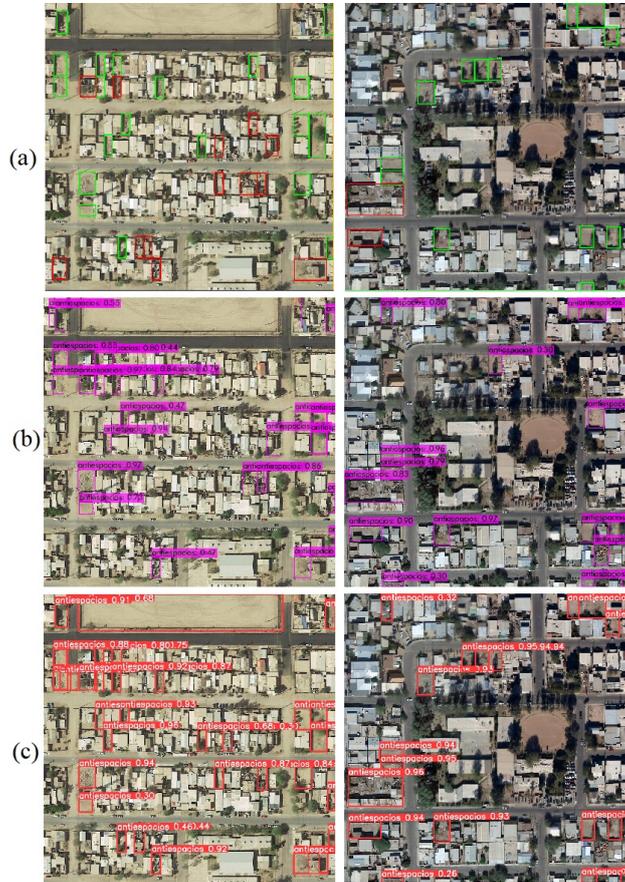


Fig. 5. Detección de antiespacios: (a) Etiquetas verdaderas, en rojo espacios abandonados, en verde vacíos urbanos; (b) YOLOv4; (c) YOLOv5.

$$F_1 = 2 \frac{P * R}{P + R}. \quad (4)$$

La precisión promedio (AP) se calcula por cada clase y es una aproximación del área bajo la curva de la curva $P \times R$. Esta métrica indica la capacidad de los modelos de detección de objetos de mantener altos niveles de P y R , en [18] se detallan las distintas formas de cómo calcularlo. Finalmente, la media de la precisión promedio (mAP) es la media aritmética de las AP , y se define como sigue:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i, \quad (5)$$

Siendo N la cantidad total de clases. Como ya se mencionó, en este trabajo se utilizará un umbral del IoU de 0.5, por lo tanto, se utilizará la notación $AP@50$ y $mAP@50$.

Tabla 3. Resultados comparando YOLOv4 y YOLOv5 entrenado con los dos conjuntos de datos (550x550 y 1080x1080).

Métrica	v4 550	v4 1080	v5 550	v5 1080
<i>P</i>	0.76	0.73	0.67	0.71
<i>R</i>	0.61	0.67	0.64	0.69
<i>F1</i>	0.68	0.7	0.65	0.7
AP@50	0.627	0.664	0.69	0.705

4.3. Comparación entre YOLOv4 y YOLOv5

Los resultados de ambos modelos con los dos conjuntos de imágenes de distinta resolución se muestran en la Tabla 3. Con respecto al conjunto de datos, con imágenes de resolución de 550x550, observamos que YOLOv5 obtiene un mejor AP y *R*, pero pierde ante YOLOv4 en *P* y el puntaje *F1*. Por otro lado, en el conjunto de datos con imágenes de resolución 1080x1080, de igual forma, YOLOv5 obtiene un mejor AP y *R*, pero se queda un poco atrás en *P*, aunque ambos obtienen el mismo puntaje *F1*.

En la Figura 5 se muestra un ejemplo de los objetos detectados por los modelos de YOLOv4 y YOLOv5 entrenados con el conjunto de datos con resolución 1080x1080 píxeles. Se aprecia en este caso que la salida obtenida con YOLOv5 detecta una mayor cantidad de los antespacios verdaderos, pero en ambos modelos se aprecian algunos falsos positivos. En cuanto a la comparación de resultados de acuerdo con la resolución de las imágenes, resulta claro que una más alta resolución conlleva mejores resultados.

En YOLOv4, una mayor resolución mejora en buena medida el AP y el *R*, pero se redujo *P*, aunque a pesar de lo anterior, se obtiene un mejor puntaje *F1*. Lo anterior indica que el modelo entrenado con imágenes de mayor resolución obtiene menos inferencias exitosas, pero a cambio se identifica una mayor cantidad de los antespacios.

En cuanto a YOLOv5, una mayor resolución tuvo como consecuencia una mejora en todas las métricas, especialmente en *P*, *R* y el puntaje *F1*, mientras que se obtuvo una mejora menor en el AP. En cuanto a la tarea de identificación de antespacios, es de mayor interés detectar la mayor cantidad de antespacios verdaderos (mayor *R*) que detectar una menor cantidad de éstos, pero procurando una alta certeza de que los que identificó son verdaderos (mayor *P*).

Lo anterior sucede pues es más sencillo y rápido para una persona descartar los objetos detectados por el modelo (falsos positivos), que realizar la tarea de analizar las imágenes en la búsqueda de los antespacios que el modelo no fue capaz de detectar (falsos negativos). Por lo anterior, el modelo de YOLOv5 entrenado con imágenes de resolución de 1080x1080 es el de mayor utilidad.

4.4. Comparación entre la precisión promedio en cada clase

Con el fin de observar el desempeño en la identificación de vacíos urbanos y espacios abandonados por separado, se entrenaron modelos en YOLOv5 con estas dos clases y con los dos conjuntos de datos de imágenes de distinta resolución. Los resultados se muestran en la Tabla 4. Se vuelve a observar una mejora significativa en todos los resultados al tener una mayor resolución en las imágenes.

Tabla 4. Resultados del desempeño de las dos clases de antiespacios utilizados: vacíos urbanos y espacios abandonados.

Resolución	mAP@50	AP Vacíos	AP Abandonados
550	0.563	0.718	0.408
1080	0.613	0.741	0.485

Por otro lado, observamos un buen desempeño para los vacíos urbanos, llegando a tener una AP de 0.741 en la resolución de 1080x1080; pero se tiene un muy bajo desempeño para los espacios abandonados, llegando apenas a una AP de 0.485.

El problema de identificación de espacios abandonados radica en la complejidad y heterogeneidad de estos mismos al ser observados con solo imágenes satelitales, pues muchos de estos son construcciones terminadas pero que no han sido habitadas, lo que al ser vistas de forma aérea tienen características muy similares a cualquier otra construcción que no está abandonada (véase Figura 6).

Lo anterior implica que difícilmente se podría mejorar la detección de estos antiespacios usando solamente imágenes satelitales y, que sería necesario el uso de otras fuentes de datos, así generando modelo multimodal como lo han hecho otros autores [4, 1, 3, 14] en otros contextos. A pesar de lo anterior, se ha visto un mayor interés en los vacíos urbanos [7, 8] y estos modelos muestran un buen desempeño identificándolos.

5. Conclusiones y trabajo futuro

En este trabajo se buscaba detectar de manera autónoma la ubicación de antiespacios urbanos utilizando imágenes satelitales y un modelo de detección de objetos como YOLO. Después de una exploración inicial de los datos, se observó un número muy bajo de remanentes urbanos en el conjunto de datos; por lo que, aunado al poco interés que estos pueden tener en su estudio, se decidió por no integrarlos, estudiándose solamente los vacíos urbanos y espacios abandonados.

Se utilizó YOLOv4 y YOLOv5 para precisar si es posible detectar estos dos antiespacios y, al mismo tiempo determinar cuál de estos dos modelos tiene un mejor desempeño.

Se entrenaron los modelos de YOLOv4 y YOLOv5 utilizando dos conjuntos de datos con imágenes de la ciudad fronteriza de Mexicali, uno con imágenes a una resolución de 550x550 y el otro a una resolución 1080x1080 y, englobando a los dos antiespacios en una sola clase.

Los resultados mostraron que YOLOv5 obtuvo un mejor desempeño, principalmente utilizando las imágenes de mayor resolución, logrando una precisión promedio del 0.705, contra 0.664 obtenido en YOLOv4. Utilizando al modelo de YOLOv5 como base, se entrenó con los mismos conjuntos de datos pero ahora cada antiespacio tenía su propia clase, esto con el objetivo de observar el desempeño que el modelo tendría en la detección de cada uno de los dos antiespacios.

Se logró un buen desempeño en la detección de vacíos urbanos, pero pobre desempeño con los espacios abandonados. Se concluyó que la posible causa del pobre desempeño es la gran cantidad de espacios abandonados que observados desde una perspectiva aérea tienen características no muy distintas a las de cualquier otra construcción no abandonada.



Fig. 6. Ejemplos de espacios abandonados (en rojo) donde no se presentan características que los permitan diferenciar de las demás estructuras.

Este trabajo abre las puertas a trabajos futuros. Por un lado, si bien se exploró el efecto que tiene la resolución de las imágenes en el desempeño final del modelo, es necesario seguir explorando y utilizar otras configuraciones de YOLO con la meta de determinar las mejores condiciones que aseguren un mejor desempeño en la detección de los antiespacios.

También es de interés explorar el uso de otros tipos de detectores como SDD, o el uso de modelos de segmentación semántica como deeplabv3 el cuál demostró un buen desempeño en una tarea similar [16].

Por otro lado, el uso de imágenes satelitales es suficiente para tener un buen desempeño en la detección de vacíos urbanos, no se da el mismo caso para la detección de espacios abandonados, lo que invita al uso de otros tipos de datos además de las imágenes satelitales que nos permitan mejorar ese desempeño [13].

Finalmente, es de interés el desarrollo de una herramienta que permita la identificación de antiespacios urbanos, lo que facilitará su estudio en distintos ámbitos y podrá ofrecer información valiosa en beneficio de la rehabilitación y planeación urbana.

Referencias

1. Cao, R., Tu, W., Yang, C., Li, Q., Liu, J., Zhu, J., Zhang, Q., Li, Q., Qiu, G.: Deep learning-based remote and social sensing data fusion for urban region function recognition. *Journal of Photogrammetry and Remote Sensing*, vol. 163, pp. 82–97 (2020) doi: 10.1016/j.isprsjprs.2020.02.014
2. Ceniceros, B., Ettinger, C.: Paisaje urbano desde la frontera Juárez-El Paso. Mapeando manifestaciones de arte urbano desde el bordo. *Revista Latinoamericana de Estudios Urbanos Regionales*, vol. 46, pp. 181–201 (2020) doi: 10.4067/S0250-71612020000100181
3. Chen, B., Feng, Q., Niu, B., Yan, F., Gao, B., Yang, J., Gong, J., Liu, J.: Multi-modal fusion of satellite and street-view images for urban village classification based on a dual-branch deep neural network. *International Journal of Applied Earth Observation and Geoinformation*, vol. 109 (2022) doi: 10.1016/j.jag.2022.102794

4. Chen, S., Zhang, H., Yang, H.: Urban functional zone recognition integrating multisource geographic data. *Remote Sensing*, vol. 13, pp. 4732 (2021) doi: 10.3390/rs13234732
5. Coubes, M. L.: Evolución del empleo fronterizo en los noventa: Efectos del TLCAN y de la devaluación sobre la estructura ocupacional. *Frontera Norte*, vol. 15, pp. 33–64 (2017)
6. Curzio, C.: Análisis sobre la distribución geoespacial del anti-espacio urbano; estudio enfocado en ciudades fronterizas del norte de México (2021)
7. Curzio, C., de la Torre, H.: Análisis geoespacial sobre la distribución de los vacíos urbanos localizados en nuevo laredo, tamaulipas. *DECUMANUS*, vol. 6 (2021) doi: 10.20983/decumanus.2021.1.3
8. Curzio, C., de la Torre, H.: Vacíos urbanos y desigualdad socioeconómica: Temas que convergen en la frontera norte de México. *Frontera Norte*, vol. 33, pp. 1–26 (2021) doi: 10.33679/rfn.v1i1.2174
9. Gong, J., Liu, C., Huang, X.: Advances in urban information extraction from high-resolution remote sensing imagery. *Science China Earth Sciences*, vol. 63, pp. 463–475 (2020) doi: 10.1007/s11430-019-9547-x
10. Jiang, P., Ergu, D., Liu, F., Cai, Y., Ma, B.: A review of YOLO algorithm developments. *Procedia Computer Science*, vol. 199, pp. 1066–1073 (2022) doi: 10.1016/j.procs.2022.01.135
11. Jindal, M., Raj, N., Saranya, P., Sundarabalan, V.: Aircraft detection from remote sensing images using YOLOv5 architecture. In: 6th International Conference on Devices, Circuits and Systems, pp. 332–336 (2022) doi: 10.1109/ICDCS54290.2022.9780777
12. Jocher, G.: YOLOv5 by Ultralytics (2020)
13. Li, J., Hong, D., Gao, L., Yao, J., Zheng, K., Zhang, B., Chanussot, J.: Deep learning in multimodal remote sensing data fusion: A comprehensive review. *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, pp. 102926 (2022) doi: 10.1016/j.jag.2022.102926
14. Li, Q., Taubenböck, H., Shi, Y., Auer, S., Roschlaub, R., Glock, C., Kruspe, A., Zhu, X. X.: Identification of undocumented buildings in cadastral data using remote sensing: Construction period, morphology, and landscape. *International Journal of Applied Earth Observation and Geoinformation*, vol. 112 (2022) doi: 10.1016/j.jag.2022.102909
15. Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., Pietikäinen, M.: Deep learning for generic object detection: A survey. *International Journal of Computer Vision*, vol. 128, pp. 261–318 (2020) doi: 10.1007/s11263-019-01247-4
16. Mao, L., Zheng, Z., Meng, X., Zhou, Y., Zhao, P., Yang, Z., Long, Y.: Large-scale automatic identification of urban vacant land using semantic segmentation of high-resolution remote sensing images. *Landscape and Urban Planning*, vol. 222, pp. 104384 (2022) doi: 10.1016/j.landurbplan.2022.104384
17. Nepal, U., Eslamiat, H.: Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs. *Sensors*, vol. 22, pp. 464 (2022) doi: 10.3390/s22020464
18. Padilla, R., Netto, S. L., da Silva, E. A. B.: A survey on performance metrics for object-detection algorithms. pp. 237–242 (2020) doi: 10.1109/IWSSIP48289.2020.9145130
19. Ranjan, A., Pathare, N., Dhavale, S., Kumar, S.: Performance analysis of YOLO algorithms for real-time crowd counting. In: 2nd Asian Conference on Innovation in Technology, pp. 1–8 (2022) doi: 10.1109/ASIANCON55314.2022.9909018
20. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection (2015) doi: 10.48550/arXiv.1506.02640
21. Sultana, F., Sufian, A., Dutta, P.: A review of object detection models based on convolutional neural network (2019) doi: 10.1007/978-981-15-4288-6_1