IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 3, September 2024, pp. 2703~2712
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i3.pp2703-2712

An innovative approach for detecting buildings and construction anomalies in Zenata City

Maryem Ait Moulay[1], Adil Salbi[1], Issam Bouganssa[1], Mohamed-Salim Masmoudi[2], Abdelali Lasfar[1]
[1]Laboratory Systems Analysis and Information Processing and Industrial Management, High School of Technology Salé, Mohammadia School of Engineering, Mohammed V University, Rabat, Morocco
[2]Nouspo IT, Casablanca, Morocco

Article Info

Article history:

Received Jan 7, 2024
Revised Feb 25, 2024
Accepted Mar 5, 2024

Keywords:

Deep learning
Drone
Smart city
Urban development
You only look once version 8

ABSTRACT

Rapid urban development in Morocco has led to increased construction activities and significant environmental concerns. Recently Zenata City has undergone significant urban development, marking a crucial step in its trajectory toward a modern smart city. As a part of this growth, our research incorporates an innovative method within the you only look once version 8 (YOLOv8) model, representing a significant advance over conventional methods. The YOLO algorithm has been updated with new features and improvements that infuse our work with a dash of innovation. YOLOv8 integration improves construction and irregular construction detection accuracy beyond what is possible with traditional applications. We trained our algorithm using orthophoto captured by DJI MATRICE 300 RTK drone split into georeferenced tiles and annotated using LabelImg software. Through this process, we were able to create a solid 742 image dataset for training, testing, and validation purposes related to construction. Utilizing drone imagery and the YOLOv8 object detection algorithm, buildings and construction irregularities are detected with high accuracy after 300 training epochs on Kaggle's GPU P100. Insights for early detection and effective building site management are provided by this all-encompassing strategy, which supports Zenata City's sustainable urban growth.

Corresponding Author:

Maryem Ait Moulay
Laboratory Systems Analysis and Information Processing and Industrial Management
High School of Technology Salé, Mohammadia School of Engineering, Mohammed V University
Rabat, Morocco
Email: maryem_aitmoulay2@um5.ac.ma

1. INTRODUCTION

Morocco as a dynamically developing country in North Africa especially the Casablanca metropolitan area has experienced significant economic growth, rapid urbanization, and a significant increase in population since the 20th century. However, this development has resulted in several problems and obstacles that require special attention, such difficulties include increasing infrastructure problems, problems of access to adequate housing, and environmental concerns [1]. The fact that Zenata City is a vital part of this metro area makes these issues even more urgent [2]. To meet these challenges an effective detection must be developed, to explore the application of object detection algorithms, specifically you only look once version 8 (YOLOv8) to regular and irregular constructions in Zenata City utilizing aerial images captured by DJI Matrice 300 RTK drone.

Artificial intelligence is a field of computer science examining the simulation of intelligent behaviour in computers, permitting them to carry out tasks like learning, reasoning, and self-correction that normally require human intelligence [3]. Machine learning is a subset of artificial intelligence concentrated on developing algorithms that allow computers to learn from data and make predictions or decisions without getting specifically programmed [4]. Lately, advancements in machine learning and pattern recognition have made it possible to automatically extract information from huge volumes of data. Deep learning, a subset of machine learning, is mainly reason for this. Deep learning employs multilayer neural networks to characterize the most representative and discriminative features hierarchically [5].

In this context, convolutional neural networks (CNNs) one of the most frequently employed network architectures in deep learning techniques, have gradually replaced traditionally created feature engineering in image analysis because of their improved ability to identify high-level features. Based on CNN object detection algorithms, there are two kinds. The majority of modern high-performance object-detection frameworks are built using the YOLO and region based convolutional neural network (R-CNN) algorithm series [6]. In several kinds of study fields, R-CNN-based object-detection frameworks have demonstrated excellent accuracy with a slow detection time, though. The YOLO algorithm series handles image recognition as a case of regression using a simple cascade model in order to address the algorithm's performance problem [6]. Since deep learning relies largely on data, the efficacy and accuracy of object detection algorithms vary significantly based on the input data. Several studies have explored this approach, demonstrating promising results. For instance, [7] implemented the proposal algorithm for building detection, incorporating in-depth feature extraction and adaptive super pixel shredding. During training and testing of data collection, they have been used a modified simple linear iterative clustering (SLIC) for image segmentation. Two high-resolution images datasets from the New Zealand region using CNN for building detection produced outstanding results, with remarkable accuracy and an average F1 score of 98.83%. Also, most research employed YOLO-based object detection on a range of private and a limited number of public datasets. Such as Al-Selwi *et al.* [8] reached precise face mask detection and counting using YOLOv5 on Jetson Nano and NVIDIA giga texel shader extreme, showing successful outcomes for applications in different settings such as offices and schools. Benallal and Tayeb [9] used YOLO in the CNN system which provided a precision of 91% to detect road defects. In this context, the purpose of this study was to develop an adequate method for using YOLOv8 to detect regular and irregular construction from drone imagery, as well as to increase accuracy to help urban development management within Zenata City.

2. AREA OF STUDY

In this study, we chose the area of Ain Sebaa/Zenata, a city located in the northern region of Morocco. Geographically, it is situated in the central part of the country, not far from the Atlantic coast. The city benefits from its strategic location, being near a major economic center, making it an ideal hub for urban development and investments as shown in Figure 1.

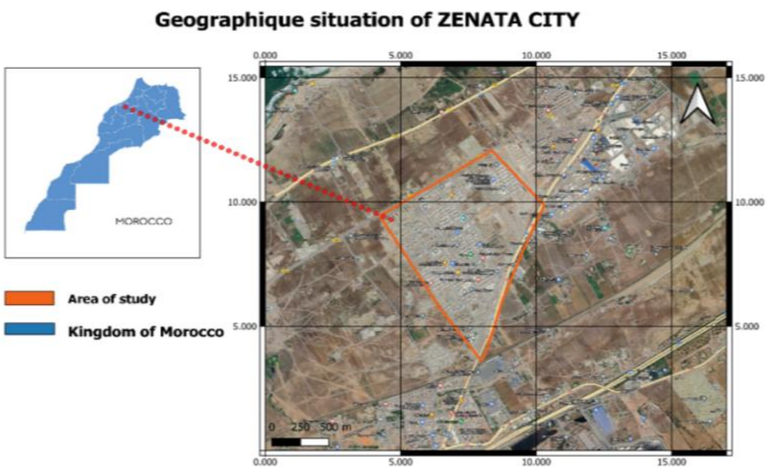

Figure 1. Study area Zenata Eco-City

The Zenata study area is a major urban project that forms part of the national strategy for the creation of new cities. Zenata Eco-City, which is currently under construction, is enthusiastically heralded as an urban model for Africa. Even before the city's completion, the study of this study area highlights the research and policy learning processes used to legitimize this emerging model. This is a crucial area of study for understanding urban development processes, the challenges of urban policy mobility, and the narratives linked to the anticipated export of an urban model under construction [10].

3. METHOD
3.1. Model training
This research work opted to use the Kaggle platform with a P100 GPU accelerator due to its considerable advantages in terms of efficiency and processing power. By using a P100 GPU, machine learning and data modeling tasks can be significantly accelerated, which reduces the time needed to finish complex analyses [11]. Access to the required computing resources is made easier with Kaggle's practical working environment, which also offers a strong infrastructure that enables the quick execution of resource-intensive algorithms. In addition to speeding up modelling iterations, using a P100 GPU on Kaggle allows us to process large datasets more quickly, which enhances the quality and accuracy of the research findings.

3.2. DJI matrice 300 RTK
Choosing the data source has become a crucial step in the preparation of our data to guarantee the best possible final data quality and risk management-always a complex issue when flying in overpopulated urban areas. To achieve that, we used the DJI matrice 300 RTK drone for our study. This system guarantees flight autonomy of more than 45 minutes (the manufacturer claims 55 minutes), allowing for the optimization of the flight project, including the selection of large areas without involuntarily breaking the work. The DJI matrice 300 RTK is equipped with advanced safety features such as intelligent automatic return to home (ARH) technologies, obstacle detection, altitude limits, and geofencing, all of which are already in operation. Figure 2 shows some details of the drone. To maintain functionality if the primary return-to-home unit fails, it also has redundant versions of important systems, including the battery, barometer, compass, and inertial measurement unit (IMU). In addition, the primary circuit boards incorporate communication circuits and a backup power source to ensure dependability.

Figure 2. Snapshot of the drone's technical features

Thus, to facilitate simultaneous control by both pilots and separate the pilot's flight operations from the video operator's photo capture, the radio remote control is also doubled. This preserves visual line-of-sight (VLOS) mode while enabling both pilots to travel independently over great distances. Finally, to ensure maximum safety, the drone is equipped with a flight terminator and a parachute.

3.3. Detection model
3.3.1. YOLO algorithm
Computer vision algorithms and CNN have shown great potential for enhancing object detection [12], they can complete a variety of classification and bounding box regression tasks and are very beneficial things at learning image features [13]. Sevezral studies have explored these approaches, demonstrating promising results. Using the InceptionV3 and MobileNet architectures [14] presents a cost-efficient automated identification solution for food products in smart fridges, with real-time test results validating the system's performance. Additionally, Benradi *et al.* [15] presents a successful example of using CNN for object detection, they employed CNNS to discriminate lung diseases and COVID-19 from chest x-ray images, Arkin *et al.* [16] also present a comparison of CNN and transformer-based methods for object detection in computer vision, revealing transformer's potential and identifying prospects.

Focusing on the detection of construction and construction anomalies, this computer vision is creating a new approach to automatic site monitoring and inspection. Against this backdrop, the University of Washington's YOLO object detection and image segmentation model was developed by Redmon *et al.* [17]. YOLO models have been widely used in remote sensing, such as in studies by [18], [19]. YOLO surpasses many other object detection algorithms such as R-CNN and dimension-based partitioning and merging (DPM)

[17]. To enhance YOLO's accuracy and efficiency, changes have been made to its cost functions and architecture since its initial release in 2015. Using YOLO, an image is split into a grid of smaller regions, and for every object found within those regions, predictions are made for bounding boxes and class probabilities.

Confidence in YOLO is defined as Pr (Object) × intersection over union (IOU) [18], where Pr (object) is the likelihood that an object exists and IOU is the IOU, or the region where inference and ground truth overlap. Five predictions are produced by each grid cell (x, y, w, h, and a confidence score). Every grid also generates p conditional class probabilities, which are represented by the notation Pr (Class|Object). The process for obtaining class-specific confidence scores for each box during the test phase is shown in (1):

$$Class_i \times IOU_{pred}^{truth} = \{Class_i | object \times object \times IOU_{pred}^{truth}\} \quad (1)$$

The last layers forecast the bounding box coordinates as well as the associated class probabilities. The bounding boxes are then adjusted so that they fall between 0 and 1. Except for the final layer, which uses a linear activation function, all subsequent layers increase non-linearity by utilizing the leaky rectified linear activation function, as shown in (2):

$$f(x) = \begin{cases} x, if\ (x > 0) \\ 0.1x, otherwise \end{cases} \quad (2)$$

Released in 2016, YOLOv2 added batch normalization, anchor boxes, and dimension clusters to improve upon the original model [20]. Like this, YOLOv3 was introduced in 2018 and improved the model's performance by using focal loss, a feature pyramid, and a more effective backbone network [21]. Numerous models followed the original YOLO model, such as YOLOv4, YOLOv5, YOLOv6, and YOLOv7. Ultralytics release YOLOv8 in 2023 [22]. Figure 2 displays YOLOv8's intricate architecture.

3.3.2. YOLOv8 detection

YOLOv8 was released by Redmon and Farhadi [22], the company that created YOLOv5. The most recent version, YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), and YOLOv8x (extra-large), introduced five distinct scaled versions. With a broad range of capabilities, YOLOv8 can manage any kind of vision tasks, like tracking, segmentation, object detection, pose estimation, and classification. YOLOv8 is structured around the head, neck, and backbone to predict outcomes, extract features, and merge multiple features [22]. The structure of the YOLOv8 network is illustrated in Figure 3.

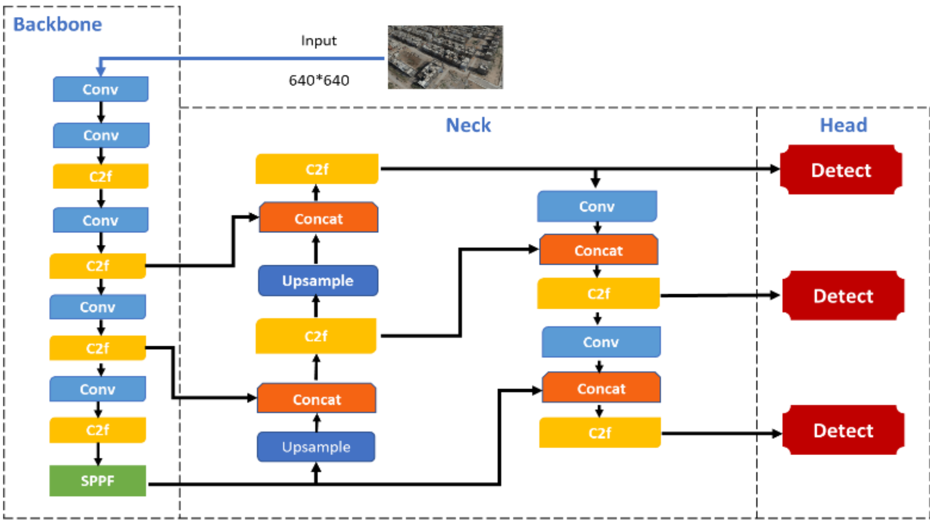

Figure 3. YOLOv8 network structure diagram

YOLOv8 is built on a similar architecture to YOLOv5, but it includes changes to the cross stage partial (CSP) layer, which is known as the cross-stage partial bottleneck with two convolutions (C2f) module. To enhance detection accuracy, this C2f module integrates contextual information with high-level features. The CSP architecture reduces the computational work of the model while improving CNN's learning capability. Split and concat are used to connect the two Conv modules and one BottleNeck to create the C2f module [23].

The backbone park's remaining portion is identical to YOLOv5's. The last layer of the backbone uses the SPPF module, the detailed modules in the YOLOv8 network are shown in Figure 4.

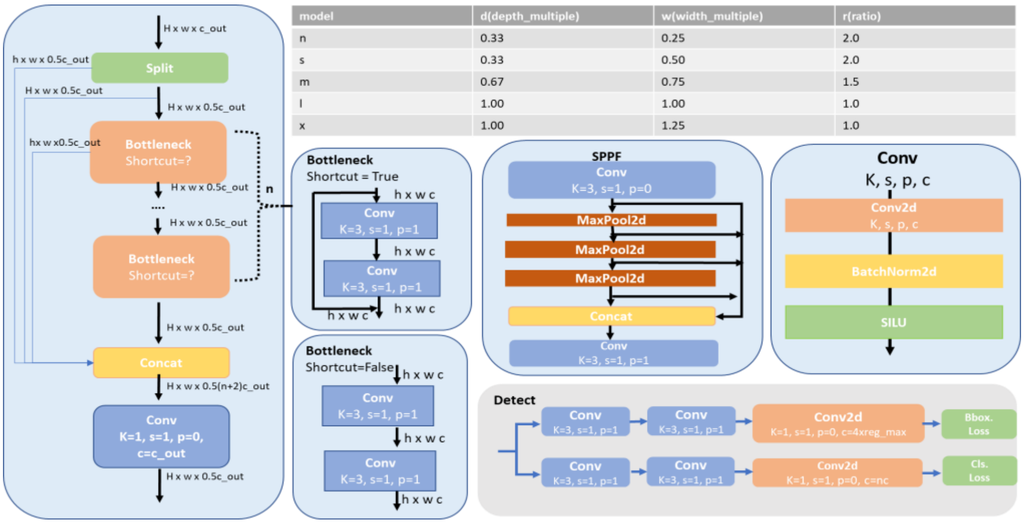

Figure 4. YOLOv8 network detail structure diagram

YOLOv8 employs an anchor-free model with a decoupled head to manage objectness, classification, and regression tasks as separate branches. This strategic layout allows each branch to focus entirely on its assigned task, resulting in improved overall model accuracy. The activation functions used in the output layer of YOLOv8 are important for their respective purposes.

In the output layer, they used the sigmoid function as the activation function for the objectness score, representing the probability that the bounding box contains an object. It uses the softmax function for the class probabilities, representing the objects' probabilities belonging to each possible class [22], YOLOv8's architecture embodies the most recent developments in deep learning-based object detection. It is a great option for several applications due to its speed, accuracy, and efficiency, such as our anomaly detection task in Zenata's construction monitoring system.

3.4. Methodology

As seen in Figure 5, the study employs four primary steps to identify regular and irregular constructions. The first step in the method is to gather constructions image captured by drone and divide it into small images. After that, the data is prepared by labeling the images used for training and validation and separating it into sets for testing, validation, and training. Next, the YOLOv8 model is trained and tested using the matching image sets. Finally, a thorough examination and presentation of the obtained data are made.

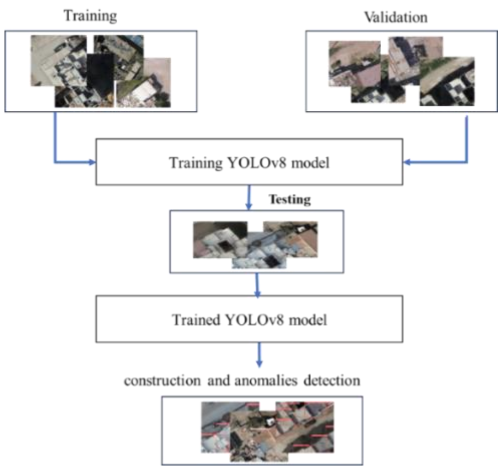

Figure 5. Methodology for buildings and anomalies detection using YOLOv8 in this study

3.5. Data creation

The data collected for this research consisted of high-resolution orthophotos of the city of Zenata, captured using the matrice 300 RTK drone. The drone imagery provided detailed and georeferenced aerial views of the construction sites and urban landscape, enabling a comprehensive analysis of the city's development. To effectively handle the large-scale orthophotos, we employed the rasterio library, a powerful geospatial Python package to divide the orthophotos into smaller, manageable pieces of size 1024×1024 pixels. This subdivision allowed us to create a dataset comprising 742 georeferenced images, each representing various construction scenarios within Zenata as shown in Figure 6.

Figure 6. Training of Zenata City anomaly dataset detection using drone imagery and YOLOv8

All the images were annotated using LabelImg software to precisely annotate the images as part of the data preparation process. The annotations were divided into two different classes: class 0, which remained for regular construction constructions, and class 1, which showed irregular constructions as indicated in Table 1. The annotated images were then carefully divided into three subsets: the training set (63%), the validation set (15%), and the testing set (22%). The training set made up 63% of the dataset.

Table 1. Datasets used in this study

Image size	Training images	Validation images	Testing images	Total
1024*1024	467	111	164	742

4. RESULTS AND DISCUSSION

4.1. Evaluation metrics

Five indicators were used for assessing the performance of the YOLOV8-based approach for detecting regular and irregular construction: recall, precision, F1 score, mean average precision (mAP), detection speed (FPS), and mAP of small targets [mAP (small)]. Precision and recall are defined as follows in (3) and (4) by integrating false positive (FP), true positive (TP), false negative (FN), and true negative (TN) [24].

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

When comparing different models, the average precision at IOU=0.5, or AP50, is a crucial performance metric. By considering precision and recall scores at a particular IOU threshold of 0.5, it offers a useful and succinct metric for evaluating object detection models' overall efficacy without favoring any one class over another. AP50 is frequently used to evaluate the model's object detection accuracy in various scenarios and is essential for comparing various algorithms or architectures [25]. The mAP can be calculated as in the (5):

$$AP = \int_0^1 \text{Precision}(\text{Recall})d\text{Recall}, mAP = \frac{1}{N}\sum_{i=0}^{N} AP \quad (5)$$

N represents the total number of classes present in the dataset. F1 curve the weighted harmonic mean of precision and recall for a classifier is called the F-measure, and it takes =1 (F1 score). It indicates that the highest level of testing confidence, can be defined by (6):

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

4.2. Performance metrics

A variety of important metrics, especially precision, recall, and mAP50, have been employed to evaluate the model's performance. These metrics offer significant details about how well the model works and how accurate it is at recognizing both specific and general anomalies in the construction sites. All details are shown in Table 2.

Table 2. Performance metrics and model details

Performance metrics	Model
Number of epochs	300
Training time	1.600 hours
GPU used	P100
Model size	22.5MB
Precision of constructions	0.918
Recall for constructions	0.83
Map50 for construction	0.957
Precision for general anomalies	0.804
Recall for general anomalies	0.745
mAP50 for general anomalies	0.801

4.3. Visual evaluation

Using a dataset of images taken from Zenata construction sites, the YOLOv8 model was evaluated. A subset of the 97 images that were included in the evaluation are displayed in Figure 7. The model exhibited remarkable efficacy in identifying irregular constructions and regular constructions throughout these images. The bounding boxes precisely pinpointed the anomalies and constructions that were found, demonstrating the high precision of the model. The yellow boxes in the figure, however, show the few instances in which the model made small mistakes. Even with these small errors, the YOLOv8 model demonstrated how well it could visually identify different kinds of construction anomalies and regular constructions. Additional innovations that could potentially improve the model's performance and make it even more capable at identifying anomalies of constructions and construction sites include adding higher resolution images and expanding the dataset.

Figure 7. regular and irregular constructions detection

Figure 8 displays the model's overall analysis result, along with the precision (mAP) measurement, recall, box loss, and training loss. The model clearly advances over the course of the 300 epochs. The training loss decreases as the number of epochs rises, while the mAP value keeps rising. Number of epochs rises. the number of epochs represented by the horizontal axis and the corresponding loss, precision, or recall values represented by the vertical axis.

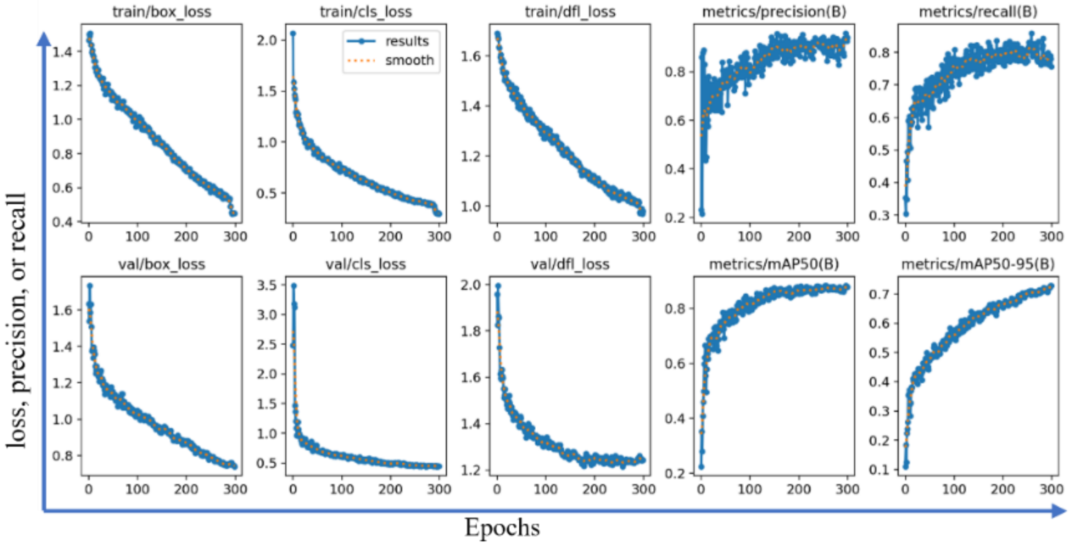

Figure 8. Various evaluation parameters

4.6. Discussion

Our research focused on object detection specially on the detection of regular and irregular constructions, with the help of Kaggle GPU P100 and matrice 300 RTK drone imagery, the model was able to achieve an impressive mAP50 score of 87.9%, which is comparable to other studies using various applications. Table 3 represents a comparison between studies focused on buildings detection using different approaches based on deep learning. Our YOLOv8 model demonstrated perfect accuracy in the context of building detection, achieving an F1 score of 92.27 for both regular buildings. The detection of irregular constructions demonstrated an impressive F1 score of 77.16, highlighting the model's accuracy in real-world settings, because there were relatively fewer examples in Zenata City.

Table 3. Comparing the suggested model's performance against other deep learning methods

Reference	Data	Method used	Performance		
			Precision (%)	Recall (%)	F1 score (%)
Remote sensing image building detection method based on Mask R-CNN [26]	DOTA dataset	Improved mask R-CNN algorithm	---	---	81.8
	Remote sensing image of disaster area	Improved mask R-CNN algorithm	---	---	72.2
Detection of specific building in remote sensing images [27]	Worldview data of Tumshuk City	YOLO-S-CIOU	60	100	---
		YOLOv3	0	0	---
	Google map of Yantai City	YOLO-S-CIOU	50	55.6	---
		YOLOv3	10	5.6	---
Identifying damaged buildings in aerial images [28]	Post-earthquake images	YOLOv4	89	82	65.42
		YOLOv4+ResNext	91	86	72.19
		YOLOv4+ResNext+Focal EIOU Loss	94	89	87.25
Proposed approach	Drone imagery	Yolov8	Regular constructions: 91.8 Irregular constructions 80.4	Regular constructions: 92.7 Irregular constructions 74.5	Regular constructions: 92.27 Irregular constructions 77.16

5. CONCLUSION

The goal of this research is to aid in the management of urban development by detecting the constructions and irregular constructions in Zenata City from drone imagery using the YOLOv8 CNN-based object detection algorithm. The results were reviewed using both integrated test images and metrics to examine the performance. The model training process was significantly sped up with the help of the Kaggle platform using the accelerator P100 GPU, demonstrating its remarkable computing power and efficiency. In future research, we intend to refine our methodology to detect the courtyard anomalies inside of construction by using the YOLOv8 algorithm to strengthen the model's ability to support urban planning and development management efforts.

REFERENCES

[1] Z. Daief, A. Borja, L. Joulami, M. Azzi, A. Fahde, and H. Bazairi, "Assessing benthic ecological status of urban sandy beaches (Northeast Atlantic, Morocco) using M-AMBI," *Ecological Indicators*, vol. 46, pp. 586–595, 2014, doi: 10.1016/j.ecolind.2014.07.023.

[2] J. Hdia *et al.*, "Evaluation of the physicochemical and bacteriological study of the three sites in Ain Sebaa-Zenata region after the installation of the anti-pollution station of the East Coast of Grand Casablanca," *Annual Research & Review in Biology*, pp. 1–11, 2019, doi: 10.9734/arrb/2019/v34i130143.

[3] S. Doroudi, "The intertwined histories of artificial intelligence and education," *International Journal of Artificial Intelligence in Education*, vol. 33, no. 4, pp. 885–928, 2023, doi: 10.1007/s40593-022-00313-2.

[4] I. H. Sarker, "Ai-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems," *SN Computer Science*, vol. 3, no. 2, 2022, doi: 10.1007/s42979-022-01043-x.

[5] W. Fang, Q. Xue, L. Shen, and V. S. Sheng, "Survey on the application of deep learning in extreme weather prediction," *Atmosphere*, vol. 12, no. 6, 2021, doi: 10.3390/atmos12060661.

[6] W. Wu *et al.*, "Application of local fully convolutional neural network combined with YOLO v5 algorithm in small target detection of remote sensing image," *PLoS ONE*, vol. 16, no. 10, 2021, doi: 10.1371/journal.pone.0259283.

[7] A. Benchabana, M. K. Kholladi, R. Bensaci, and B. Khaldi, "Building detection in high-resolution remote sensing images by enhancing superpixel segmentation and classification using deep learning approaches," *Buildings*, vol. 13, no. 7, 2023, doi: 10.3390/buildings13071649.

[8] H. F. Al-Selwi, N. Hassan, H. B. A. Ghani, N. A. B. A. Hamzah, and A. B. A. Aziz, "Face mask detection and counting using you only look once algorithm with Jetson Nano and NVDIA giga texel shader extreme," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 3, pp. 1169–1177, 2023, doi: 10.11591/ijai.v12.i3.pp1169-1177.

[9] M. A. Benallal and M. S. Tayeb, "An image-based convolutional neural network system for road defects detection," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 2, pp. 577–584, 2023, doi: 10.11591/ijai.v12.i2.pp577-584.

[10] L. C. -Roy and S. Moser, "'Fast urban model-making': constructing Moroccan urban expertise through Zenata Eco-City," *Urban Geography*, vol. 44, no. 9, pp. 2016–2036, 2023, doi: 10.1080/02723638.2022.2131261.

[11] C. S. Bojer and J. P. Meldgaard, "Kaggle forecasting competitions: an overlooked learning opportunity," *International Journal of Forecasting*, vol. 37, no. 2, pp. 587–603, 2021, doi: 10.1016/j.ijforecast.2020.07.007.

[12] J. Ding, J. Zhang, Z. Zhan, X. Tang, and X. Wang, "A precision efficient method for collapsed building detection in post-earthquake UAV images based on the improved NMS algorithm and faster R-CNN," *Remote Sensing*, vol. 14, no. 3, 2022, doi: 10.3390/rs14030663.

[13] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object detection with deep learning: a review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212–3232, 2019, doi: 10.1109/TNNLS.2018.2876865.

[14] P. Jain, P. Chawla, M. Masud, S. Mahajan, and A. K. Pandit, "Automated identification algorithm using cnn for computer vision in smart refrigerators," *Computers, Materials and Continua*, vol. 71, no. 2, pp. 3337–3353, 2022, doi: 10.32604/cmc.2022.023053.

[15] H. Benradi, I. Bouganssa, A. Chater, and A. Lasfar, "Discriminative approach lung diseases and COVID-19 from chest x-ray images using convolutional neural networks: a promising approach for accurate diagnosis," *International journal of online and biomedical engineering*, vol. 19, no. 14, pp. 131–141, 2023, doi: 10.3991/ijoe.v19i14.42725.

[16] E. Arkin, N. Yadikar, X. Xu, A. Aysa, and K. Ubul, "A survey: object detection methods from CNN to transformer," *Multimedia Tools and Applications*, vol. 82, no. 14, pp. 21353–21383, 2023, doi: 10.1007/s11042-022-13801-3.

[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem, pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.

[18] U. Alganci, M. Soydas, and E. Sertel, "Comparative research on deep learning approaches for airplane detection from very high-resolution satellite images," *Remote Sensing*, vol. 12, no. 3, 2020, doi: 10.3390/rs12030458.

[19] C. Liu, F. Meng, Z. Zhu, and L. Zhou, "Object detection of UAV aerial image based on YOLOv8," *Frontiers in Computing and Intelligent Systems*, vol. 5, no. 3, pp. 46–50, 2023, doi: 10.54097/fcis.v5i3.13852.

[20] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," *30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, pp. 6517–6525, 2017, doi: 10.1109/CVPR.2017.690.

[21] L. Deng, H. Li, H. Liu, and J. Gu, "A lightweight YOLOv3 algorithm used for safety helmet detection," *Scientific Reports*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-15272-w.

[22] J. Terven, D. M. C. -Esparza, and J. A. R. -González, "A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680–1716, 2023, doi: 10.3390/make5040083.

[23] S. Zhao, J. Zheng, S. Sun, and L. Zhang, "An improved YOLO algorithm for fast and accurate underwater object detection," *Symmetry*, vol. 14, no. 8, 2022, doi: 10.3390/sym14081669.

[24] H. Yi, B. Liu, B. Zhao, and E. Liu, "Small object detection algorithm based on improved YOLOv8 for remote sensing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 1734–1747, 2024, doi: 10.1109/JSTARS.2023.3339235.

[25] L. Cheng, J. Li, P. Duan, and M. Wang, "A small attentional YOLO model for landslide detection from satellite remote sensing images," *Landslides*, vol. 18, no. 8, pp. 2751–2765, 2021, doi: 10.1007/s10346-021-01694-6.

[26] Q. Han, Q. Yin, X. Zheng, and Z. Chen, "Remote sensing image building detection method based on Mask R-CNN," *Complex and Intelligent Systems*, vol. 8, no. 3, pp. 1847–1855, 2022, doi: 10.1007/s40747-021-00322-z.

[27] J. Gao, Y. Chen, Y. Wei, and J. Li, "Detection of specific building in remote sensing images using a novel yolo-s-ciou model. case: gas station identification," *Sensors*, vol. 21, no. 4, pp. 1–19, 2021, doi: 10.3390/s21041375.

[28] L. Shi *et al.*, "Identifying damaged buildings in aerial images using the object detection method," *Remote Sensing*, vol. 13, no. 21, 2021, doi: 10.3390/rs13214213.

BIOGRAPHIES OF AUTHOR

Maryem Ait Moulay is a Ph.D. student at the Laboratory Systems Analysis, Information Processing, and Industrial Management, Mohammadia School of Engineering (EMI), Mohammed V University of Rabat, Morocco. Passionate about artificial intelligence. Her focus is on exploring innovative applications in artificial intelligence, computer vision, and information processing. She can be contacted at email: maryem_aitmoulay2@um5.ac.ma.

Adil Salbi was born in Errachidia, Morocco. He defended his thesis in 2017 in electrical engineering and embedded systems. He is an assistant professor at the Higher School of Technology of Salé (ESTS) at Mohammed V University Morocco and a member of the LASTIMI laboratory (EMI). His research interests include embedded systems and industrial computing as well as the concept of connected objects (IoT) and artificial intelligence (AI). He can be contacted at email: adilsalbi00@gmail.com.

Issam Bouganssa is a professor at Mohammed V University in Rabat since 2019, active in embedded systems and industrial computing. Ph.D. in science and technology for engineers specializing in automation and embedded systems at the LASTIMI Laboratory, Mohammed V University of Rabat. He can be contacted at email: issam.bouganssa@gmail.com.

Mohamed-Salim Masmoudi is a founder of NousPo, active in artificial intelligence and computer science. Master of science in computer science at the Savoie Mont Blanc University. He can be contacted at email: salim@nouspo.com.

Abdelali Lasfar is a professor of Higher Education at Mohammed V Agdal University, Salé Higher School of Technology, Morocco. His research focuses on compression methods, indexing by image content and image at the LASTIMI Laboratory. He can be contacted at email: abdelali.lasfar@est.um5.ac.ma.