Automatic Detection and Counting of Urban Housing and Settlement in Depok City, Indonesia: An Object-Based Deep Learning Model on Optical Satellite Imageries and Points of Interests

A Pindarwati[1]**, A W Wijayanto**[2,*]

[1] Department of Computer Science, Pamulang University, Tangerang Selatan, Indonesia
[2] Department of Statistical Computing, Politeknik Statistika STIS, Jakarta, Indonesia

*Corresponding author's e-mail: ariewahyu@stis.ac.id

Abstract. Detecting urban housing and settlements has a substantial position in decision-making problems such as monitoring housing and development, not to mention the widely required urban mapping application. One of the most important goals in the United Nations Sustainable Development Goals (SDGs) is to improve urban living conditions globally by 2030. We propose an automatic detection of urban housing and settlements on remote sensing satellite imagery data using object detection-based deep learning using semantic segmentation and the potential availability of remote sensing datasets at high spatial resolutions, Open Street Map (OSM) geolocation point of interest dataset, and Sentinel-2 optical satellite imagery data. The detection model using Mask Region-based Convolutional Neural Networks (Mask R-CNN) is implemented in Depok City, Indonesia. These regions were chosen because it is the second most populous suburb in Indonesia and the tenth most populous globally and, making it challenging to extract building features from satellite imagery. This model categorizes dense, moderate, and sparse conditions and has a promising result of an average precision of 100% and an F1-score of 67% with evaluation performance metrics only considering points associated with buildings, not building boundaries or the intersection over union (IoU). The model performance has been compared to ground check results of field surveys, and it performs best in sparse conditions. Our findings offer the potential implementation of the model for fast and accurate monitoring of housing, settlement, and regional planning in urban areas.

1. Introduction

The challenge of accurate detection and counting of urban housing is crucial in decision-making processes [1], such as city planning and development, urban mapping and management, and population estimation. Monitoring house developments and urbanization is a critical task, as it helps in identifying and detecting illegal buildings [2]. This process serves as a valuable tool for ensuring that urban growth is sustainable and complies with the necessary regulations. The detection of urban houses and settlements is of utmost importance as it forms the foundation for conducting field census by (Badan Pusat Statistik) BPS Indonesia.

Urban housing and settlements play a crucial role in shaping the development of cities and regions, influencing various decision-making processes such as urban planning, infrastructure development, and resource allocation. Monitoring and mapping these urban features have become essential in the context of achieving sustainable urban development, a key objective outlined in the United Nations Sustainable Development Goals (SDGs) to improve global urban living conditions by 2030 [3]. Method for capturing characteristics of land area with remote sensing electromagnetic recordings obtained through satellite [4]. In this regard, the accurate detection and classification of urban housing and settlements from remote sensing satellite imagery data have gained significant attention. Satellite imagery is highly potential to be utilized as an alternative source of data mapping urban settlements.

Our study focuses on Depok City, located in Indonesia. Depok, renowned as the second most populous suburb in Indonesia [5] and the tenth most populous globally [6], boasts a vibrant demographic landscape. With a population tally of 2,123,349 individuals, distinctly comprising 1,071,173 males and 1,052,176 females, the city's urban fabric resonates with vitality, translating into an impressive density of 10,621.5 people per km2 [5]. Depok City holds the distinction of being the second most populous suburb in Indonesia and ranks among the top ten most populous globally. The city's rapid population growth and urban expansion have posed unique challenges in accurately delineating and categorizing urban housing and settlements from satellite imagery. The high population density and diverse building patterns necessitate innovative approaches for effective mapping and monitoring.

A robust and reliable framework for statistical analysis must be established to enhance the accuracy and effectiveness of the census process. Field mapping adds a heavy workload, cost, and time pressure to the process, making it imperative to find efficient methods to address the lack of productivity and improve the overall effectiveness of the process. This is a crucial task as it has numerous applications in urban planning and development [7], but the traditional manual approach is time-consuming and resource-intensive [8]. The survey method has several limitations, including high costs and personnel requirements, limited spatial and temporal resolution, and difficulty obtaining timely and accurate data for certain applications.

Table 1. Comparison of Field Mapping and Satellite Imagery for Urban Settlement Mapping

Aspects	Field Mapping	Satellite Imagery
Time and Energy	**Large**	Small
Operational Cost	**Very Large**	More affordable
Data Update Frequency	**Long Time**	Faster
Possibility of Error in Geotagging Results	Low	**High**

There is an opportunity to combine satellite imagery with field mapping for mapping urban settlements. By leveraging the power of satellite imagery and integrating it with traditional field mapping techniques, there is a huge potential to revolutionize the way urban settlements are mapped. The minimal time and energy required for satellite imagery, combined with its affordability and frequent data updates, makes it an ideal solution for mapping in the modern world. To address this challenge, we proposed a study to automatically detect and count urban housing footprints using deep learning techniques and remote sensing datasets. The methods can be used to obtain more frequent and detailed data at a lower cost. The United Nations' goal for sustainable development requires ensuring everyone has access to adequate, secure, and affordable housing and basic services, and improving living conditions in slum areas by 2030 [9]. Method for capturing the characteristics of a land area with remote sensing electromagnetic recordings obtained through satellite [10]. Satellite imagery has a high potential to be utilized as an alternative source of data in mapping urban settlements.

Our main contribution in this study is in terms of ground-check model performance validation using field surveys. While most of the existing studies rely solely on benchmark datasets, we validate our model's performance against ground truth data obtained from field surveys. This ensures that the model's outputs are directly aligned with real-world urban features. By comparing the model's results to on-ground observations, we provide a more accurate assessment of its practical utility. According to research in 2022 conducted building footprint extraction and counting on very high-resolution satellite imagery using object detection deep learning framework with methods CNN in Paris, France, and Khartoum, Sudan [11].

Based on this background, we intend to detect and count buildings in satellite imagery using Sentinel-2 by integrating the Semantic Segmentation deep learning-based object detection architecture. We have also conducted experiments on the model parameters to assess the effectiveness of the model's performance. Figure 1 illustrates the research framework for these objectives. The main contributions of this paper are as follows: i) The proposed method of object-based deep learning is trained and used for building detection from satellite imagery, considering the various characteristics of different regions. Different areas also exhibit different building characteristics tendencies, posing a challenge in itself for detecting building objects on satellite imagery. ii) The proposed method can also be applied to other cities and has the potential for rapid and accurate estimation, mapping, and updating of landmarks for urban housing monitoring and settlement, particularly in metropolitan areas.

Table 2. Main Contribution of Our Study to the Existing Related Studies

No	Relevant Studies	Methods	Data Source	Data Type	Accessibility	Location
1	Nurkarim & Wijayanto, 2022 [11]	Convolutional Neural Networks (CNN) of YOLO architecture	World View-3	Very High-Resolution Satelite Imagery	Paid	Paris, France, Khartoum , Sudan Region.
2	Ye, et al, 2019 [12]	CNN	Aerial imagery datasets.	Very High-Resolution remote sensing data	Paid	Massachu setts, USA
3	Han, et al, 2021 [13]	Mask R-CNN	Drone	Very High Resolution	Paid	China
4	This study	Object Based Deep Learning (Mask R-CNN with Segmentation)	OSM and Sentinel -2 satellite	Medium Resolution and Point of Interest	Free	Depok, Indonesia

Nurkarim and Wijayanto, 2022 introduced an object detection deep learning framework for extracting the building footprint and counting them using very high-resolution satellite imagery [11]. Ye, et al, 2019 introduce a novel fully convolutional network (FCN) called RFA-UNet, which utilizes attention-based re-weighting to extract buildings from aerial imagery, bridging the semantic gap between features and achieving comparable and improved performance compared to other state-of-the-art models for building extraction [12]. Han, et al, propose a building extraction method combining traditional digital image processing and convolutional neural networks to quickly and accurately identify buildings in disaster areas, improving detection accuracy and reducing computational time compared to the R-CNN algorithm [13]. We are the first to introduce an object based detection of urban settlement in Depok, Indonesia using medium resolution and point of interest data.

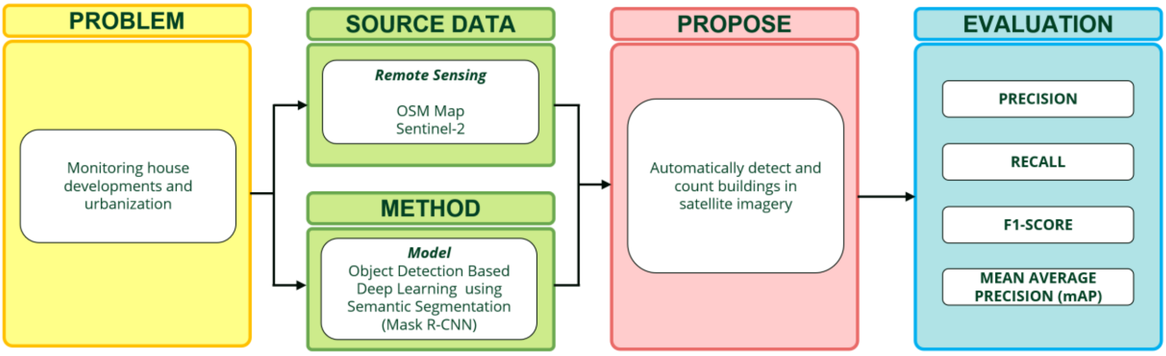


Figure 1. Proposed Research framework

2. Method

2.1. Data Used

This study was conducted in Depok, West Java, Indonesia, and it involved the categorization of three distinct areas: The dense area, characterized by very small buildings and narrow roads, is located on Jalan Anggrek 5-6 Pengasinan, Depok, West Java, covering an area of approximately 0.19 km². The moderate condition area features buildings with medium to small sizes and relatively wide roads. It is located in the Telaga Golf Cluster Bali, Sawangan, Depok, West Java. The sparse area is characterized by large buildings and wide roads and is situated on Jalan Boulevard, Telaga Golf, Sawangan, Depok, West Java

Table 3. The characteristic features a unique area data set

Area Types	Location	Range	Characteristic
Sparse	Jalan Anggrek 5-6, Pengasinan, Depok	0,19	Very small building, Narrow Road
Moderate	Telaga Golf Cluster Bali, Sawangan, Depok	0.2	Small to medium size building, Relatively wide road
Dense	Jalan Bulevard, Telaga Golf, Sawangan, Depok	0.14	Large size Building, Wide road

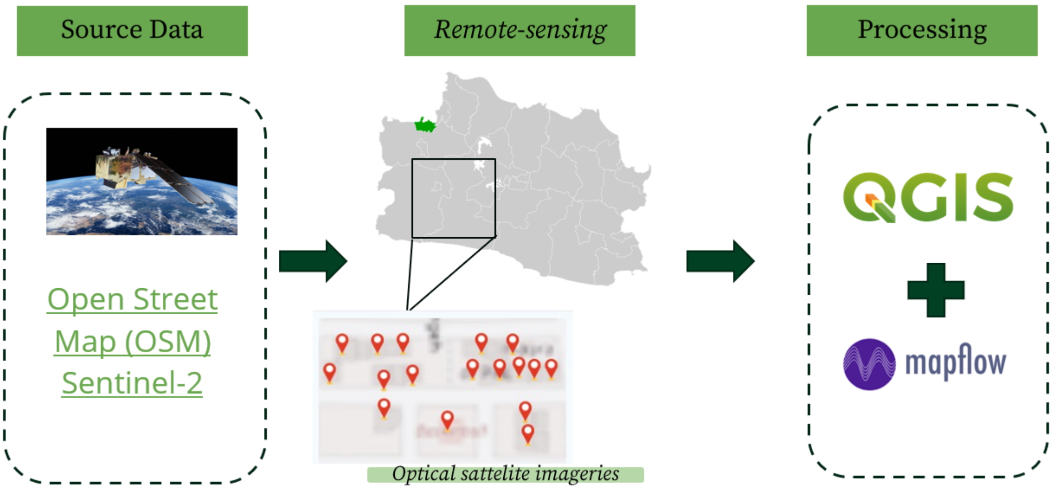


Figure 2. Phase of Data Collection

In this study we use Open Street Map (OSM) Map Dataset, Sentinel-2 satellite imagery. OpenStreetMap (OSM) has developed as a global project and community to produce and maintain a free and editable database and a map of the world based on volunteer mappers' contributions [14]. With about 7.5 billion data points (nodes) generated by approximately 1.8 million individuals as of March 2022 [15], it is possibly the most accomplished example of a crowdsourced geoinformation project and the notion of donated geographic information. OpenStreetMap uses Sentinel-2 imagery. The resolution is sufficient to trace major roads and land use features [16]. Sentinel data is already listed as available to OpenStreetMap. The imagery is very recent, updated between every month to every two weeks, and available just a few days after the sensing date.

The Open Street Map dataset contains Sentinel-2 imagery and per-pixel semantic label masks. The Open Street Map dataset provides a ready-for-use dataset consisting of satellite images of constant size from the Sentinel-2 platform and per-pixel label images using new land use categories derived from OpenStreetMap [16]. In this study, we used Sentinel-2 imagery acquired in May 2023 and Open Street Map data collected on May 16, 2023. Sentinel-2 is a multi-spectral wide-swath, medium-resolution imaging mission designed to support Copernicus Land Monitoring studies that include the monitoring of plants, soil, and water cover in addition to observing aqueducts and coastal areas. The Sentinel-2 Multispectral Instrument (MSI) samples 13 spectral bands during 5 days [17], including four bands at 10 meters, six bands at 20 meters, and three bands at 60 meters spatial resolution (European Space Agency, n.d.). The Sentinel-2 images utilized in this study have a resolution of 10 meters/S2(10) and were rescaled to 30 meters/S2(30).

Table 4. Sentinel-2 MSI Level-2A instrument spectral band specifications

Band	Description	Resolution (m)	Wavelength (nm)
B1	Aerosols	60	443.9 (S2A) / 442.3 (S2B)
B2	Blue	10	443.9 (S2A) / 442.3 (S2B)
B3	Green	10	443.9 (S2A) / 442.3 (S2B)
B4	Red	10	443.9 (S2A) / 442.3 (S2B)
B5	Red Edge 1	20	443.9 (S2A) / 442.3 (S2B)
B6	Red Edge 2	20	443.9 (S2A) / 442.3 (S2B)
B7	Red Edge 3	20	443.9 (S2A) / 442.3 (S2B)
B8	NIR	10	443.9 (S2A) / 442.3 (S2B)
B8A	Red Edge 4	20	443.9 (S2A) / 442.3 (S2B)
B9	Water vapor	60	443.9 (S2A) / 442.3 (S2B)
B10	Cirrus	60	443.9 (S2A) / 442.3 (S2B)
B11	SWIR 1	20	443.9 (S2A) / 442.3 (S2B)
B12	SWIR 2	20	443.9 (S2A) / 442.3 (S2B)

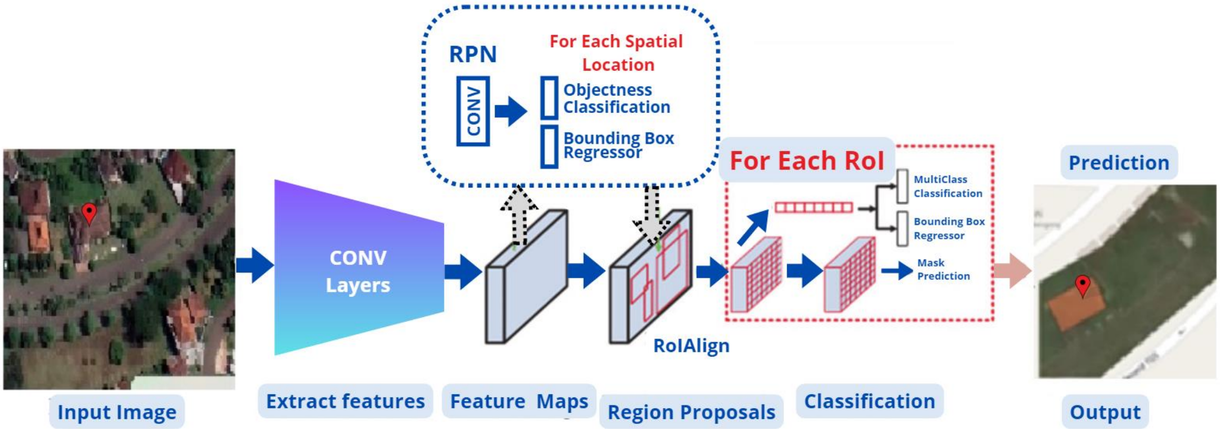


Figure 3. The architecture of Mask R-CNN applied in this study

Mapflow.ai uses a deep learning method called Mask R-CNN for building detection using segmentation [18]. Mask R-CNN is a state-of-the-art deep learning model that combines object detection and instance segmentation. It is an extension of the popular Faster R-CNN model, which adds a branch to the network that outputs a binary mask that delineates the object of interest. This allows for pixel-level segmentation of objects in an image, which is particularly useful for detecting buildings in satellite imagery. The Mask R-CNN model used by Mapflow.ai is trained on a large dataset of satellite imagery and building annotations. During training, the model learns to identify the features that distinguish buildings from other objects in the image, such as roads, trees, and water bodies. Once trained, the model can be used to detect buildings in new satellite images with high accuracy. Overall, the use of deep learning and specifically the Mask R-CNN model allows Mapflow.ai to perform building detection with high accuracy and efficiency, which is useful for a variety of applications such as urban planning, disaster response, and environmental monitoring. For the visualization result we use Quantum GIS (QGIS) software.

2.2. Evaluation

In object-based detection, the commonly used evaluation criteria include Recall, Precision, F1 Score [17]. However, the performance metrics calculation only considered the points associated with the buildings and did not consider building boundaries (Intersection over Union). For evaluation performance metrics calculation does not consider building boundaries/ IoU (Intersection of Union), only the points associated with the building are considered. We used overall accuracy to evaluate the global performance of the methods. In addition, the F1-score of the positive (building) class) were used to evaluate classification performance. In evaluating object detection models, there are several evaluation metrics used, including:

$$Precision = \frac{True\ Positives}{(True\ Positives + False\ Positives)} \tag{1}$$

$$Recall = \frac{True\ Positives}{(True\ Positives + False\ Negatives)} \tag{2}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{(Precision + Recall)} \tag{3}$$

$$Mean\ Average\ Precision\ (mAP) = \frac{1}{N}\sum_{i=1}^{N}\int_0^1 P_{i(r)dr} \tag{4}$$

True positive, *false positive*, and *false negative* can be represented through the following confusion matrix:

Table 5. Confusion Matrix

Confusion Matrix		*Predicted Class*	
		Positive	*Negative*
Actual Class	*Positive*	True Postive	False Negative
	Negative	False Positive	True Negative

Figure 4. Illustrations of *True positive*, *false positive*, and *false negative*

True positives can occur when the model accurately predicts the presence of a building object. False positives occur when the model predicts an object to be a building, but the detected object is not actually a building. Meanwhile, false negatives happen when a building object is not detected as a building, even though it truly is a building. In this study, evaluation performance metrics only considering points associated with buildings, not building boundaries or the intersection over union (IoU).

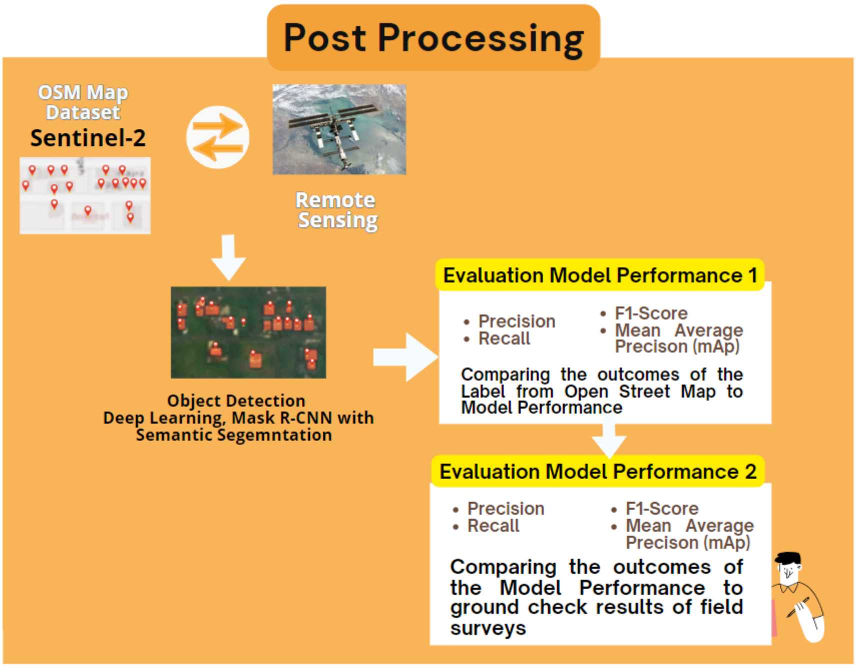
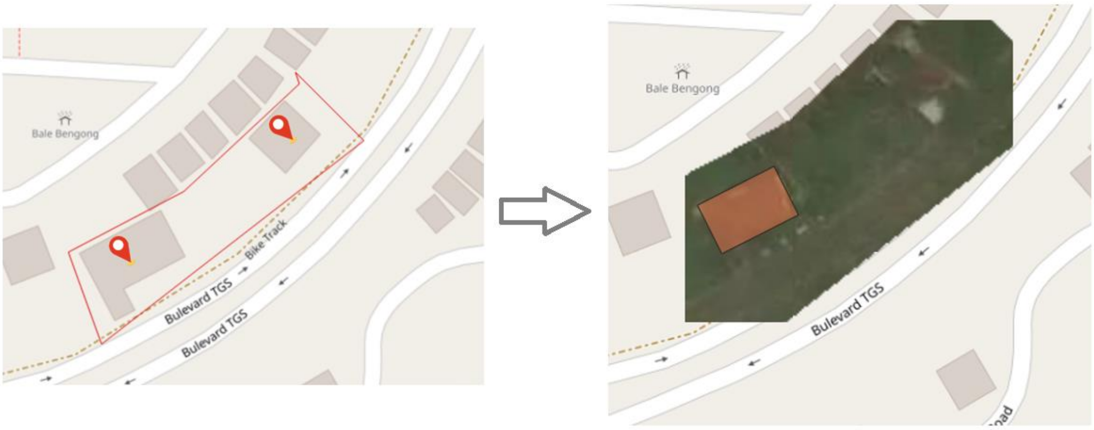


Figure 5. Evaluation of Model Performance in Post Processing

The two techniques used in this study were used twice to evaluate the model's performance. In the first, the results of the labels from OpenStreetMap were compared to model performance by calculating precision, recall, F1 score, and Mean Average Precision. The second method involved comparing the model performance results to the field survey ground-check results by calculating the same evaluation metrics.

3. Results

Using the ground truth labels extracted from OpenStreetMap (OSM) point of interest data and the predictions generated by our model using Mask R-CNN semantic segmentation-based object detection hrough deep learning techniques. These comparisons are visually depicted in Figure 3 below.



a) Sparse Condition: Jalan Bulevard Telaga Golf Sawangan, Depok, Jawa Barat (Range: 0,14 km)

b) Moderate Condition: Telaga Golf Cluster Bali, Sawangan, Depok, Jawa Barat (Range: 0,20 km)

c) Dense Condition, Jalan Anggrek 5-6, Pengasinan, Depok, Jawa Barat (0,19 km)

Figure 6. Implementation model on data test.

3.1. Evaluation Results

After the model successfully predicts all three areas, namely sparse, moderate, and dense, it is subsequently compared with the results obtained from a field survey. The test results, comparing the model's predictions to the field survey data based on specific points, are presented in Table 6 below.

Table 6. The number of buildings compared between two methods

Building Type	(Σ) Amount building OSM	(Σ) Building Mapflow	(Σ) Building Groundcheck of Field Surveys
Sparse	2	1	2
Moderate	89	42	89
Dense	37	19	37

a) Sparse Condition b) Moderate Condition c) Dense Condition

Figure 7. Results of Conditions field surveys

Table 7. Summary of Model's performance in Comparing Two Methods, the Label Open Street Map data and the Field Survey

Area Types	Label Precision (%)	Label Recall (%)	Label F1-Score (%)	Label mAP (%)	Field Survey Precision (%)	Field Survey Recall (%)	Field Survey F1-Score (%)	Field Survey mAP (%)
Sparse	100%	50%	67%	100%	100%	50%	67%	100%
Moderate	90%	47%	64%	66.58%	100%	47%	64%	66.58%
Dense	100%	46%	63%	100%	100%	45%	63%	100%

The results showed promising performance of the model compared to two methods, which are Label Data Open Street Map and Field Survey, with a precision of 100%, recall of 50%, mean average precision of 100%, and F1-score of 67%. However, the performance metrics calculation only considered the points associated with the buildings and did not consider building boundaries (Intersection over Union). The result provides a promising solution for the fast and accurate evaluation of urban housing and settlement monitoring and city planning, particularly in metropolitan areas. By improving the parameter model determination using IoU/ Intersection of Union (overlapping area) and incorporating more diverse data sets and data augmentation techniques, the accuracy of the model can be further improved.

4. Discussions

Current developments in Earth observation science and satellite image analysis have created numerous opportunities for precise and effective monitoring of our planet's geographical features [32-36]. The present study has introduced an innovative approach for automating the detection of urban housing and settlements using object-based deep learning techniques. The model's performance was evaluated in the context of central Depok, Jawa Barat, Indonesia, with a focus on average precision and F1 score as evaluation metrics. However, it's worth noting that the evaluation metrics do not take into account building boundaries through Intersection over Union (IoU) calculation; instead, they only consider points associated with buildings. This choice may influence the overall assessment of the model's accuracy and localization capabilities.

The evaluation outcomes shed light on the complexity of the urban environment in Depok. The city's unique combination of building types, densities, and layouts poses inherent challenges for automated detection. These challenges were reflected in the varying model performance across different areas. Notably, the Boulevard Road and Telaga Golf area demonstrated the model's best performance, particularly in sparse conditions within an elite area. This result suggests that the model's accuracy might be influenced by the specific characteristics of the urban landscape, such as the presence of larger, more distinguishable buildings in upscale neighborhoods.

The results showed promising performance, with a precision of 100%, recall of 50%, mean average precision of 100%, and F1-score of 67%. However, the performance metrics calculation only considered the points associated with the buildings and did not consider building boundaries (Intersection over Union). Our results offer a practical implementation for rapid and accurate estimation of urban housing monitoring and settlement, particularly in metropolitan areas. The result provides a promising solution for the fast and accurate evaluation of urban housing and settlement monitoring and city planning, particularly in metropolitan areas. This aligns with the United Nations' objective to enhance urban living standards worldwide by 2030. By improving the parameter model determination using IoU/ Intersection of Union (overlapping area) and incorporating more diverse data sets and data augmentation techniques, the accuracy of the model can be further improved.

PERFOR-MANCE	DENSELY	MODERATE	SPARSE
Precision	100 %	**90%**	100%
Recall	47 %	**46 %**	50%
F1-score	57%	**56 %**	67%
MAP	100%	**66.58 %**	100%

Figure 8. Model Evaluation per Region/ Area

The model predicts object with the best accuracy of 67% and Mean Average Precision (MAP) of 100%

Furthermore, the model's performance was evaluated against human-labeled ground truth data, acquired through field surveys. This validation process underscores the importance of cross-referencing automated results with human expertise, especially in complex urban settings. While the model's accuracy is promising, there may still be limitations in its ability to accurately identify all building instances, particularly in densely populated areas or regions with irregular building layouts.

The choice to prioritize points associated with buildings over precise building boundaries raises considerations about the model's potential real-world applications. While this approach simplifies evaluation metrics, it might not fully represent the model's practical utility for tasks that require precise boundary delineation, such as urban planning and infrastructure development. Incorporating IoU calculations or considering building boundaries explicitly could provide a more comprehensive assessment of the model's accuracy and suitability for various applications. The results highlight both the potential and challenges of applying such techniques in a diverse and densely populated urban environment like Depok. The model's performance in different conditions and areas underscores the importance of tailoring detection strategies to the specific characteristics of urban landscapes. By acknowledging the limitations of the evaluation metrics and considering building boundaries, future iterations of the model can be refined for more accurate and comprehensive results, ultimately facilitating improved urban planning and development.

5. Conclusion

This project proposes an automatic detection of urban housing and settlements in satellite imagery using object-based deep learning with a promising result of an average precision of 100% and an F1 score of 67% for the central area in Depok, Jawa Barat, Indonesia. With evaluation performance metrics calculation does not consider building boundaries/ IoU (Intersection of Union), only the points associated with the building are considered.

The evaluation of the model showed that the density and type of buildings in Depok, Jawa Barat are challenging and have an impact on the performance of the model. The best performance was achieved in Boulevard Road, Telaga Golf (sparse conditions in an elite area). The model's performance provides promising results, compared to the performance labeled with the field surveys by human evaluators. Our findings offer a potential practical application for fast and accurate evaluation of urban housing and settlement monitoring and city planning, especially in metropolitan areas. This aligns with the United Nations' goal to improve living conditions in urban areas globally by 2030.

For future research, it can be improved by enhancing the Parameter Model Determination, specifically IoU (Intersection over Union), which measures the overlapping area divided by the combined area. This improvement aims to enhance the accuracy of the model. The method also offers potential for further enhancement through the addition of more diverse datasets and the incorporation of data augmentation techniques to boost model performance. Similar research can be implemented in other provinces, and the model developed has the potential to be used for detecting and locating buildings for mapping and updating landmarks, including buildings and infrastructure.

References

[1] Luan, X., Huang, X., Guo, W., & Luan, S. 2018. *A CNN-Based Approach for Automatic Building Detection and Recognition of Roof Types Using a Single Aerial Image*. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 86(5-6), 235-248.

[2] Benali, A., Dermeche, H., Belhadj, S., Adnane, A., & Hanifi Elhachemi Amar, R. 2014. *Buildings extraction of very high spatial resolution satellite images*. 2014 International Conference on Multimedia Computing and Systems (ICMCS). doi: 10.1109/icmcs.2014.6911229

[3] United Nations. 2022. The Sustainable Development Goals Report 2022. Retrieved from https://unstats.un.org/SDGS/REPORT/2022

[4] Campbell, J. B., & Wynne, R. H. 2011. *Introduction to Remote Sensing* (5th ed.). The Guilford Press

[5] Badan Pusat Statistik, "Kota Depok Dalam Angka 2023," Katalog-BPS 1102001.3276, Jakarta, 2023. (5)

[6] Badan Pusat Statistik, "Population and Gender Distribution by Regency/City in West Java Province, 2015." [Online]. Available: https://jabar.bps.go.id/linkTableDinamis/view/id/12 (6)

[7] Huang, X., & Zhang, L. 2012. *Morphological building/Shadow index for building* extraction from high-resolution imagery over urban areas. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5(1), 161-172. doi: 10.1109/jstars.2011.2168195

[8] Wen, Q., Jiang, K., Wang, W., Liu, Q., Guo, Q., Li, L., & Wang, P. 2019. *Automatic building extraction from Google Earth images under complex backgrounds based on deep instance segmentation network.* Sensors, 19(2), 333. doi: 10.3390/s19020333

[9] Putra, Y. C., & Wijayanto, A. W. 2023. *Automatic detection and counting of oil palm trees using remote sensing and object-based deep learning.* Remote Sensing Applications: Society and Environment, 29, 100914.

[10] Han, Q., Yin, Q., Zheng, X., & Chen, Z. 2021. *Remote sensing image building detection method based on mask R-CNN.* Complex & Intelligent Systems, 8(3), 1847-1855. doi: 10.1007/s40747-021-00322-z

[11] Nurkarim, W., & Wijayanto, A. W. 2022. *Building footprint extraction and counting on very high-resolution satellite imagery using object detection deep learning framework*. Earth Science Informatics, 10(4), 895-907. doi: 10.1007/s12145-022-00895-4

[12] Zhang, L., Wu, J., Fan, Y., Gao, H., & Shao, Y. 2020. An efficient building extraction method from high spatial resolution remote sensing images based on improved mask R-CNN. Sensors, 20(5), 1465. doi: 10.3390/s200514655

[13] Z. Ye *et al.*. 2019. "Building extraction from very high resolution aerial imagery using joint attention deep neural network, Remote Sensing, vol. 11, no. 24, p. 2970. doi:10.3390/rs11242970

[14] OpenStreetMap Statistics," OpenStreetMap, OpenStreetMap Foundation. [Online]. Available: https://web.archive.org/web/20210813123456/https://www.openstreetmap.org. [Accessed: 2 September 2023]
[15] "Gearing up for third Sentinel-2 satellite," ESA, 9 Agustus 2021. [Online]. Available: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Gearing_up_for_third_Sentinel-2_satellite. [Accesed: 9 Agustus 2023].
[16] T. D. T. Saadi and A. W. Wijayanto, "Machine Learning Applied to Sentinel-2 and Landsat-8 Multispectral and Medium-Resolution Satellite Imagery for the Detection of Rice Production Areas in Nganjuk, East Java, Indonesia," International Journal of Remote Sensing and Earth Sciences, vol. 18, no. 1, pp. 19-32, Jun. 2021.
[17] N. Johnson, W. Treible, and D. Crispell, "OpenSentinelMap: A large-scale land use dataset using OpenStreetMap and sentinel-2 imagery," 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022. doi:10.1109/cvprw56347.2022.00139
[18] Mapflow.ai website: [Online]. Available: https://mapflow.ai/ [Accessed: 2 Setember 2023]
[19] Long Wen, Yu Cheng, Yi Fang, Xinyu Li, "A comprehensive survey of oriented object detection in remote sensing images," Expert Systems with Applications, vol. 224, 2023, article 119960, ISSN 0957-4174, doi: 10.1016/j.eswa.2023.119960.
[20] Tian D, Han Y, Wang B, Guan T, Gu H, Wei W (2021) Review of object instance segmentation based on deep learning. J Electron Imaging 31. https:// doi. org/ 10. 1117/1. JEI. 31.4. 041205
[21] Wijayanto AW, Afira N, Nurkarim W (2022) Machine Learning proaches using Satellite Data for Oil Palm Area Detection in Pekanbaru City, Riau. in 2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)84–89 (IEEE). https://https:// doi. org/ 10. 1109/ CyberneticsCom5 5287.2022.9865301)
[22] Ren Y, Zhu C, Xiao S (2018) Small Object Detection in Optical RemoteSensing Images via Modified Faster R-CNN. Appl Sci 8:813 Rongshuang F, Yang C, Qiheng X, Jingxue W (2019) A high-resolution remote sensing image building extraction method based on deep learning. Acta Geodaetica Et Cartographica Sinica 48:34
[23] Nurmasari Y, Wijayanto AW (2021) Oil Palm Plantation Detection in Indonesia Using Sentinel-2 and Landsat-8 Optical Satellite Imagery (Case Study: Rokan Hulu Regency, Riau Province). International Journal of Remote Sensing and Earth Science (IJReSES) 18:1–18. https:// doi. org/ 10.30536/j. ijres es. 2021. v18. a3537
[24] Liu L et al (2020) Deep Learning for Generic Object Detection: A Survey. Int J Comput vis 128:261–318
[25] Liu L et al (2021) Learning region-guided scale-aware feature selection for object detection. Neural Comput Appl 33:6389–6403
[26] 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, p. 1375, 2021
[27] E. Nixon, "Human Vulnerability and Climate Change: An assessment of Greater Vancouver's human vulnerability to sea level rise in 2100," Environmental Science, 2008.
[28] N. M. Tarmizi and A. Manan, "A Concept of Urban Poverty Area Identification Using Spatial Correlation Studies on High Resolution Satellite Imagery," in Environmental Science, 2014.
[29] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2961-2969. DOI: 10.1109/ICCV.2017.322
[30] R. Girshick, "Fast R-CNN," in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1440-1448. DOI: 10.1109/ICCV.2015.169
[31] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in Advances in Neural Information Processing Systems, 2015, pp. 91-99. DOI: 10.1109/ICCV.2015.169

[32] Wijayanto AW, Triscowati DW, Marsuhandi AH 2020 Maize Field Area Detection in East Java, Indonesia: An Integrated Multispectral Remote Sensing and Machine Learning Approach. 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE).
[33] Putra YC, Wijayanto AW, Chulafak GA 2022 Oil palm trees detection and counting on Microsoft Bing Maps Very High Resolution (VHR) satellite imagery and Unmanned Aerial Vehicles (UAV) data using image processing thresholding approach, Ecological Informatics, 101878, 2022, Elsevier
[34] Utami NA, Wijayanto AW, Pramana S, Astuti ET 2023 Spatially granular poverty index (SGPI) for urban poverty mapping in Jakarta metropolitan area (JMA): a remote sensing satellite imageries and geospatial big data approach, Earth Science Informatics, 2023, Springer
[35] Kurniawan R, Saputra AMW, Wijayanto AW, Caesarendra W 2022 Eco-environment vulnerability assessment using remote sensing approach in East Kalimantan, Indonesia, Remote Sensing Applications: Society and Environment, 27, 100791, Elsevier
[36] Wijayanto, AW, Putri SR 2022 Estimating Rice Production using Machine Learning Models on Multitemporal Landsat-8 Satellite Images (Case Study: Ngawi Regency, East Java, Indonesia), Proceedings of the 2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), 2022