



IMPROVING SEGMENTATION PERFORMANCE AND OBJECT IDENTIFICATION IN THERMAL ROOF IMAGERY USING ARTIFICIAL INTELLIGENCE

**Liu Yuexiang¹, Farazeira Mohd Raslim^{2*}, Nurul Sakina Mokhtar Azizi³,
Muhammad Hafeez Abdul Nasir⁴**

^{1,3,4}*School of Housing, Building and Planning,
UNIVERSITI SAINS MALAYSIA*

²*Civil Engineering Programme, Faculty of Engineering,
UNIVERSITI MALAYSIA SABAH*

Abstract

The segmentation performance for boundary delineation of rooftop areas and the identification of objects in thermal images is often inaccurate due to low contrast and thermal noise. Improving both boundary delineation and object identification is essential to support further analyses. Therefore, the research objective of this paper is to identify roof-level elements in thermal images and improve boundary delineation accuracy for segmentation. This research employed an experimental methodology, examining 100 thermal images from a case study done at Universiti Sains Malaysia. Thermal images were captured using an unmanned aerial vehicle (UAV) and subsequently processed by frame extraction, similarity-based filtering, and manual polygon-based annotation. The study uses an artificial intelligence (AI) framework to execute the analysis. The findings showed that the AI framework was able to identify roof related objects, such as solar panels, roof equipment, and roofs. The proposed Mask R-CNN model achieved a segmentation mAP (AP averaged over IoU 0.50–0.95) of 77.2%, an AP50 of 95.1%, and an AP75 of 90.2%, indicating strong segmentation performance in the present case-study setting. Compared with the benchmark studies reviewed in this paper, which reported performance levels ranging from 9.4% to 80.8% depending on the task and evaluation metric, the present study showed a markedly stronger reported segmentation result. The delineation boundary of the roof elements was also improved, leading to more accurate assessments of the roof's condition and better identification of potential maintenance needs.

Keywords: Thermal imaging, Image segmentation, Boundary delineation, Rooftop inspection, Building assessment

² Corresponding author. Email: faraziera@ums.edu.my

INTRODUCTION

Thermal infrared imaging has been widely applied in building inspection and energy-related assessment because it reveals surface temperature patterns associated with building conditions, heat loss, and rooftop components relevant to building operation (Mastrucci et al., 2021). In rooftop analysis, thermal images can provide useful information on roof surfaces, solar panels, and roof-mounted equipment, which are closely related to building operation and energy performance. However, for thermal images to be used more systematically in building analysis, these rooftop elements must first be identified and separated with sufficient spatial clarity.

This makes segmentation performance in thermal image boundary delineation an important issue. In rooftop thermal imagery, the usefulness of the image does not depend only on whether an object is visible, but also on whether its boundary can be identified clearly enough for further interpretation, comparison, and analysis. Accurate segmentation is therefore important because boundaries between rooftop components define the regions of interest for later assessment (Khan & Bhattacharjee, 2021). However, unlike RGB images, thermal images often contain weak texture contrast and gradual temperature transitions. As a result, boundaries between adjacent rooftop elements may appear thermally continuous rather than visually distinct, which makes both object identification and boundary delineation more difficult.

A further challenge is that progress in rooftop thermal-image segmentation is limited not only by the nature of thermal imagery itself, but also by the availability of appropriate data. Building-oriented thermal datasets remain limited (Balaras & Argiriou, 2002), and many existing thermal datasets are designed for surveillance or autonomous driving rather than for individual buildings or rooftop-level semantic components (Choi et al., 2022). This affects both model development and the reliability of performance interpretation in real building contexts. In other words, segmentation performance in rooftop thermal imagery is shaped not only by model capability, but also by whether the available data are sufficiently representative of rooftop conditions.

To address this problem, the present study adopts a UAV-based rooftop thermal case study and applies an artificial intelligence framework to examine segmentation performance in rooftop thermal image boundary delineation. Rather than treating the dataset only as an end product, the study uses it as an empirical basis to evaluate whether roof-level elements can be identified and segmented effectively in a rooftop thermal-image setting. In particular, the study focuses on roof, solar panel, and roof equipment as rooftop semantic classes, and examines how segmentation performance can support clearer rooftop interpretation in a building-related context.

This focus is also relevant to built-environment applications because rooftop thermal information is useful only when it can be interpreted and organised with reasonable spatial consistency. In institutional building settings, unclear rooftop boundaries can weaken the consistency of inspection records, reduce the comparability of condition observations across building components, and limit the usefulness of thermal imagery for maintenance review, inspection prioritisation, and asset-related documentation. Recent studies have similarly highlighted the importance of building condition assessment, digital built-environment workflows, and cost- and facility-related decision-making (Mohd Nazri et al., 2025; Che Haron & Muhd Zafir, 2025). In this sense, the present study is not only about applying image segmentation to thermal imagery, but also about improving the reliability of rooftop information used in building-related interpretation and decision support.

LITERATURE REVIEW

Thermal Imaging Applications in Building Assessment

Balaras and Argiriou (2002) demonstrated that infrared thermography has long been used in building diagnostics and energy-related assessment. Nardi et al. (2016) demonstrated that thermographic methods can also be applied to quantitative building-performance evaluation, such as U-value assessment. Chicherin et al. (2023) demonstrated that aerial thermography can support large-scale building energy audits by revealing insulation-related deficiencies across residential areas. Kim et al. (2023), Pedranzini et al. (2024), and Rosti et al. (2025) further demonstrated that thermal imaging has been extended to moisture detection, air leakage assessment, and building-envelope monitoring.

However, Mondal et al. (2025) demonstrated that automation in image-based analysis is still constrained by limited training data, while many building-related thermal studies still rely heavily on expert interpretation. These studies showed that thermal imaging is valuable in building assessment, but they also showed that practical interpretation still depends on how clearly building elements can be separated and understood in the image.

Adoption of artificial intelligence in segmentation performance

Hernandez et al. (2020), Naiman et al. (2022), Peng et al. (2021), and Xu et al. (2023) demonstrated that traditional thermal-image segmentation methods, such as thresholding, edge detection, region growing, and clustering, can be used to separate thermal regions under controlled conditions. However, Acorsi et al. (2020) demonstrated that aerial thermal imagery often contains weak contrast and variable image quality, which makes segmentation more difficult in real scenes.

Batchuluun et al. (2020) demonstrated that deep learning has improved thermal-image analysis by enabling stronger feature learning than traditional

methods. Hou et al. (2021) demonstrated that Mask R-CNN can be applied to building-component segmentation in multichannel aerial imagery. However, Hou et al. (2021) also showed that segmentation performance still varies with scene complexity and object type. Together, these studies demonstrated that AI has strong potential for segmentation improvement, but its application to rooftop thermal imagery remains limited and still requires further validation in building-oriented settings.

Although thermal segmentation techniques have advanced substantially, most existing work has been developed in controlled or non-building contexts. Their applicability to rooftop thermal imagery remains uncertain because building elements often present weak, overlapping, and thermally ambiguous boundaries (Acorsi et al., 2020).

Existing Thermal Image Datasets

Chu et al. (2024) demonstrated that large thermal datasets have been developed for pedestrian detection, while Dai et al. (2020) demonstrated that thermal datasets are also widely used in autonomous driving research. Gou et al. (2023) demonstrated that some thermal datasets have been developed for defect-related studies, but these datasets are generally small and task-specific. These studies showed that publicly available thermal datasets mainly focus on transportation, surveillance, or laboratory-based defect detection rather than rooftop-level building analysis.

Evaluation of Segmentation Performance

Hou et al. (2021) demonstrated that segmentation performance in building-related imagery can be evaluated using precision-based measures. Mayer et al. (2021) demonstrated that rooftop thermal detection may still show very low recall when the target is difficult to identify. Mayer et al. (2023) demonstrated that stronger deep learning models can improve rooftop thermal detection performance, especially for larger rooftop targets. Park et al. (2021) demonstrated that F1 score can be used to evaluate thermal segmentation performance in building-envelope studies.

RESEARCH METHODOLOGY

This study adopted a case-based methodology to evaluate segmentation performance in rooftop thermal boundary delineation. Figure 1 shows the workflow, which includes UAV thermal-image acquisition, frame extraction, similarity-based filtering, annotation, dataset preparation, and Mask R-CNN training. Python with OpenCV was used for frame extraction and filtering, LabelMe was used for polygon-based annotation, and Detectron2 with a ResNet-101 backbone was used for model training and evaluation.

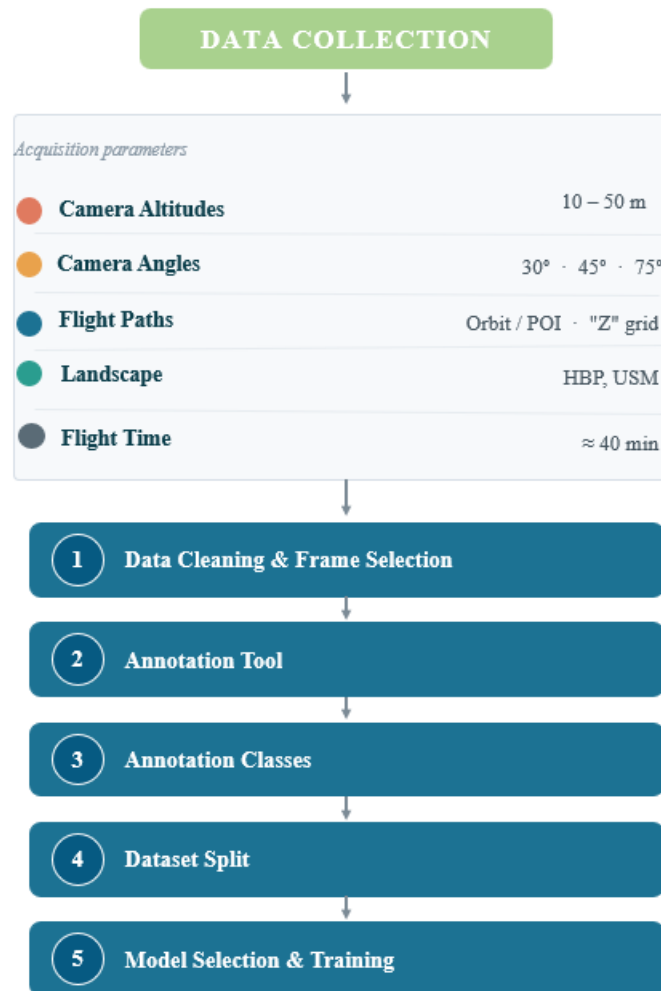


Figure 1: Workflow of UAV-based thermal dataset construction and annotation

The case study was conducted on Building E49 at the School of Housing, Building and Planning, Universiti Sains Malaysia (USM), with an estimated roof area of approximately $1649.31m^2$. The rooftop was suitable for thermal-image acquisition because its main roof-level elements could be captured with limited occlusion from surrounding structures and vegetation.



Figure 2: A general aerial view of the HBP, USM

Data Collection

Thermal images of the building were collected using an unmanned aerial vehicle (UAV) equipped with an infrared camera. The UAV platform enables flexible observation of building scenes from multiple viewpoints and heights, allowing comprehensive coverage of building structures and surrounding environments. Compared with fixed observation stations, UAV-based acquisition provides greater adaptability to complex building layouts while maintaining consistent thermal sensing capability. Thermal data were collected using a UAV-mounted infrared camera, as shown in Figure 3.



Figure 3: UAV platform used for rooftop thermal image acquisition

During each flight mission, thermal data were recorded in video format (MP4) rather than as isolated still images. This approach ensures continuous spatial coverage and captures dynamic variations in viewing angles and scene composition.

The recorded thermal videos were subsequently decomposed into individual image frames using a custom frame extraction program. Each frame represents a single thermal snapshot corresponding to a specific UAV position and viewing angle. This process generated a large pool of candidate thermal images, forming the raw dataset for further processing.

Data Cleaning and Frame Selection

Ejaz et al. (2012) and Zhang et al. (2020) demonstrated that key-frame extraction can reduce redundancy while retaining representative image content. Garrido et al. (2021) showed that thermographic preprocessing can support the learning process, while Deng et al. (2020) demonstrated that grayscale infrared images can be used for thermal-region analysis and feature extraction. Batchuluun et al. (2020) also showed that transformed thermal-image representations can support subsequent image analysis. Following this general approach, consecutive thermal frames were filtered by similarity and converted to gray-scale representations, and only frames with visual dissimilarity above 5% were retained. This resulted in a refined dataset of 100 thermal images.

Annotation Tool

After frame extraction and similarity-based filtering, the selected thermal images were manually annotated using the LabelMe (Dobbie et al., 2021) annotation tool. LabelMe was chosen due to its polygon-based annotation capability, which allows precise delineation of irregular and fine-grained object boundaries commonly observed in rooftop thermal imagery.

Polygon annotations are particularly suitable for thermal images, where object contours are often blurred and do not align with sharp pixel-level edges. Polygon-based annotation was selected because thermal boundaries in rooftop imagery are often blurred and transitional. This allowed the annotation process to reflect the practical difficulty of boundary delineation rather than imposing artificially sharp boundaries.

Annotation Classes

The dataset was annotated with a focus on roof-level semantic components that are critical for rooftop thermal inspection and boundary analysis. Each thermal image was annotated using the following labels of semantic classes: (1) solar panel, (2) roof, and (3) roof equipment. These semantic classes of annotation were selected because these objects in the thermal image are often adjacent and exhibit similar thermal patterns at their boundaries, making it difficult to clearly delineate the boundaries. This creates a challenging boundary-sensitive segmentation scenario that is appropriate for examining limitations in rooftop thermal image delineation. The images were annotated by three annotators using

shared labelling conventions for the three classes (roof, solar panel, roof equipment), with attention to the ambiguous contact regions between them. All labels were then cross-checked in a review pass, where inconsistent boundaries or class assignments were re-examined and corrected to ensure a consistent standard across the dataset.

Dataset Split

Hou et al. (2021) demonstrated that separated data subsets can support building-component segmentation training and evaluation. Following this general approach, the 100 annotated images were randomly split into training (72), validation (13), and testing (15) subsets (Figure 4). The model was trained on the training subset, and the reported metrics were obtained on the validation subset during training; independent testing is left as future work.

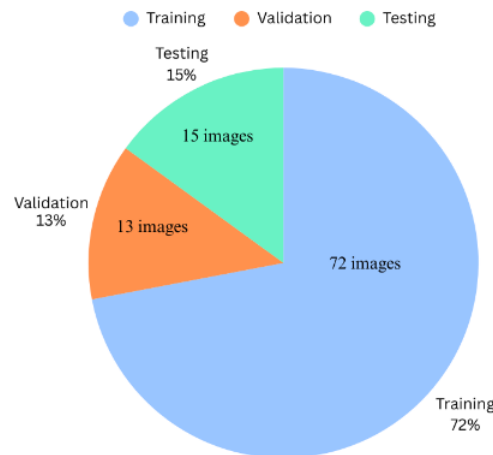


Figure 4: Distribution of the dataset into training, validation, and testing subsets

Model selection and training

Mask R-CNN was selected as the segmentation model in this study because it can perform object detection and instance segmentation simultaneously, which is suitable for separating the roof, solar panel, and roof equipment as distinct rooftop objects. The selection of Mask R-CNN was informed by the original work of He et al. (2017), who demonstrated its effectiveness for instance segmentation, and by Hou et al. (2021), who applied Mask R-CNN to building-component segmentation from aerial RGB–thermal imagery. Following this general line of research, Mask R-CNN was adopted in the present study to evaluate whether

rooftop elements in UAV thermal images could be identified and segmented effectively at the object level.

A ResNet-101 backbone was used in the present study because deeper residual networks can extract stronger hierarchical image features while remaining easier to optimize through residual learning. This choice was informed by He et al. (2016), who showed that deep residual networks improve visual recognition performance, and by He et al. (2017), who evaluated Mask R-CNN with ResNet backbones of 50 and 101 layers. In the present study, ResNet-101 was selected to provide richer feature extraction for rooftop thermal images, where rooftop elements may differ in size, shape, and thermal appearance.

The training process was set to 60,000 iterations. This follows the general practice of using fixed training schedules in instance-segmentation baselines, as seen in standard Detectron2 model-zoo configurations. However, the exact value of 60,000 iterations in the present study was chosen empirically for the current dataset rather than copied directly from a single previous study. It was retained because the model had already learned stably by the later stage of training, and further training was not expected to improve the result significantly. Therefore, 60,000 iterations were considered sufficient for the present case-study setting.

ANALYSIS AND DISCUSSION

Benchmark comparison and overall model performance

To make the model performance easier to understand, benchmark results from related thermal-image studies were first summarised in Table 1. Hou et al. (2021) showed that Mask R-CNN can be used to segment building-related components from aerial RGB–thermal imagery, but the reported precision was 39.70%, which indicates that segmentation in aerial thermal-related scenes is still not an easy task. Mayer et al. (2021) reported an average recall of 9.4% for rooftop thermal bridge detection, showing that many rooftop thermal targets were still missed in their early study. Mayer et al. (2023) later reported an average recall above 50% for large rooftop thermal bridge objects, which suggests that stronger pretrained models can improve rooftop thermal detection. Park et al. (2021) reported an F1 score of 80.8% for wall thermal anomaly detection, showing that segmentation can also support building-envelope thermal analysis beyond rooftop studies. Together, these studies show that thermal-image-based building analysis is feasible, but performance still varies greatly depending on the target object, image source, and evaluation metric. Hou et al. (2021), Mayer et al. (2021), Mayer et al. (2023), and Park et al. (2021) provide a useful benchmark context for the present study.

Against this benchmark context, the present study achieved a segmentation mAP of 77.2%, an AP50 of 95.1%, and an AP75 of 90.2% for roof,

solar panel, and roof equipment in UAV rooftop thermal images. The mAP averages performance across IoU 0.50–0.95, AP50 reflects the lenient 0.50-overlap standard, and AP75 the stricter 0.75 standard. Because the studies in Table 1 used different datasets, targets, and metrics, Table 1 provides contextual reference rather than a same-dataset baseline comparison.

Table 1: Benchmark comparison with related studies

Study	Main target	Image type	Model/method	Reported metric	Reported performance
Hou et al. (2021)	Building components and outdoor scenes	Aerial RGB–thermal imagery	Mask R-CNN	Precision	39.70%
Mayer et al. (2021)	Rooftop thermal bridges	Aerial thermal images	Neural-network detector	Average Recall	9.4%
Mayer et al. (2023)	Rooftop thermal bridges	Aerial thermal	Mask CNN / Swin-T	Average Recall	Above 50%
Park et al. (2021)	Wall thermal anomalies	Thermal visible building envelope images	CNN-based segmentation	F1 score	80.8%
Present study	Roof, solar panel, roof equipment	UAV rooftop thermal images	Mask R-CNN	AP50 AP75 mAP	95.1% 90.2% 77.2%

Training behaviour and model reliability

Hou et al. (2021) showed that Mask R-CNN can support building-component segmentation, whereas the present study further showed strong classification and mask-level performance in rooftop thermal imagery. As reported in Table 2, the proposed model achieved a Fast R-CNN classification accuracy of 99.85%, a foreground classification accuracy of 98.88%, and a mask accuracy of 98.99%. The corresponding error rates were also low, with a Fast R-CNN false negative of 1.03%, a mask false negative of 0.78%, and a mask false positive of 1.41%. These results showed that the model not only detected rooftop objects accurately, but also produced masks with a high level of consistency. Compared with the benchmark studies discussed earlier, the present study showed a stronger overall segmentation outcome in the current case-study setting, although direct comparison should still be interpreted carefully because Hou et al. (2021), Mayer

et al. (2021), Mayer et al. (2023), and Park et al. (2021) reported different evaluation metrics.

Table 2: Training-time diagnostic metrics of the proposed Mask R-CNN model

Metric	Value
Fast R-CNN classification accuracy	99.85%
Foreground classification accuracy	98.88%
Mask accuracy	98.99%
Fast R-CNN false negative	1.03%
Mask false negative	0.78%
Mask false positive	1.41%

Mayer et al. (2023) showed that stronger deep learning models can improve rooftop thermal detection performance, and Figure 5 in the present study further showed that the training process was stable throughout the learning stage. The loss curve decreased steadily to approximately 0.031 and became more stable in the later stage of training. This pattern showed that the model not only achieved strong final performance, but also learned rooftop thermal features in a consistent way over time. In this sense, Figure 5 provides additional support for the reliability of the proposed model by showing that the reported Table 2 results were achieved through stable convergence rather than unstable fluctuation.

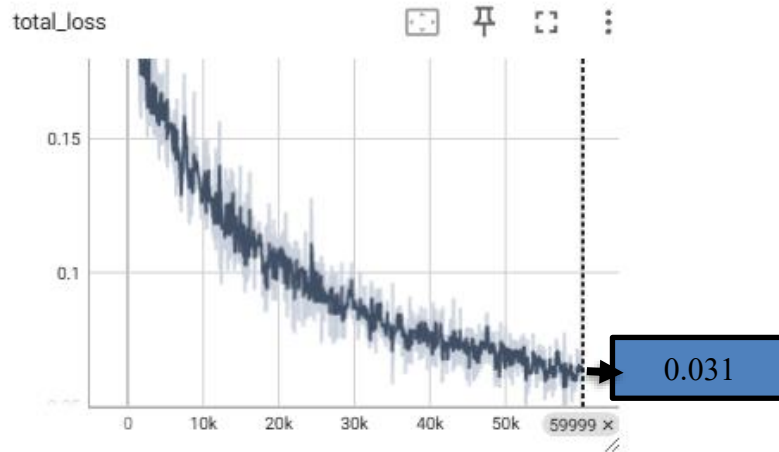


Figure 5: Total training loss of the proposed Mask R-CNN model

Visual segmentation performance

Although three rooftop classes were included, the qualitative discussion in Figure 6 focuses on the roof and solar panel because they are more visually prominent in the selected example. Krapf et al. (2022) showed that rooftop superstructure annotation can affect prediction quality, while Boccalatte and Chanussot (2025) showed that smaller rooftop superstructure-related classes tend to be less prominent and more difficult to interpret in segmentation outputs.

Figure 6 presents a representative visual result of the proposed Mask R-CNN model for rooftop thermal-image segmentation. As shown in Figure 6(c), the model was able to identify the main rooftop regions and generate segmentation masks that broadly followed the spatial extent of the rooftop scene. When read together with the original UAV thermal image and the gray-scale thermal image in Figure 6(a) and Figure 6(b), the segmentation output shows that the model not only produced a high numerical score, but also generated a visually meaningful object-level result. This supports the AP50 value of 95.1% reported in the present study and suggests that the model was effective in locating rooftop objects under the current case-study setting.

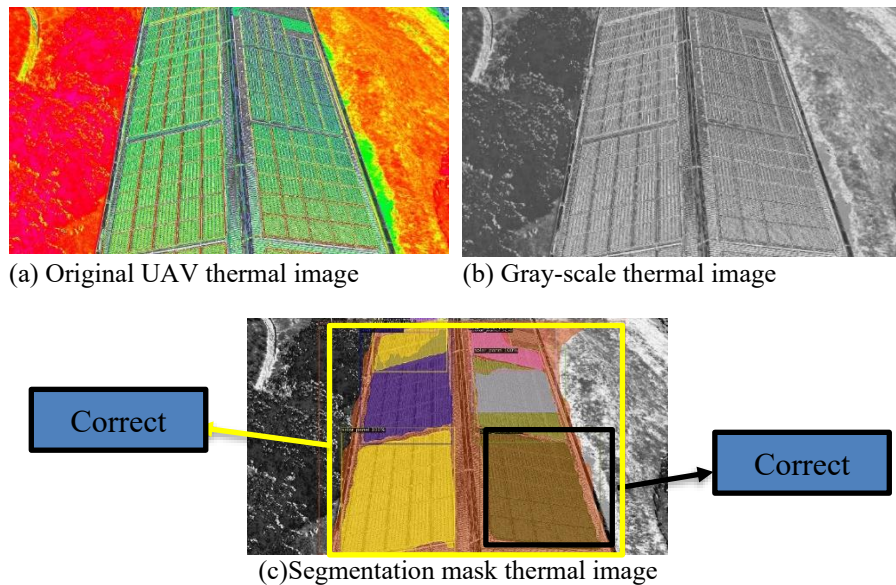


Figure 6: Representative visual result of rooftop thermal-image segmentation using the proposed Mask R-CNN model

This visual result supports earlier studies more directly and practically. Hou et al. (2021) showed that aerial image data can be used to segment building-related components, and the present study shows a similar result in rooftop

thermal imagery. Mayer et al. (2021) reported lower recall in early rooftop thermal bridge detection, while Mayer et al. (2023) later showed that stronger deep learning models could improve rooftop thermal detection performance. The present study is consistent with this general direction, because the segmented output in Figure 6 shows that rooftop elements can be identified clearly in the current dataset and case-study setting. However, these comparisons should still be interpreted carefully because the earlier studies focused on thermal bridge detection, whereas the present study focuses on roof, solar panel, and roof equipment segmentation.

The visual result is also meaningful in a building-related context. Park et al. (2021) showed that segmentation can improve the practical usefulness of thermal analysis in building-envelope studies. In the present study, the segmented rooftop image makes it easier to understand which rooftop part is being observed and where it is located. This gives the thermal image a clearer structure and makes it more useful for inspection review, condition assessment, and digital building workflows. In this sense, Figure 6 not only supports the numerical results reported earlier but also shows that rooftop thermal imagery can be translated into more structured rooftop information through segmentation.

These segmented outputs also support maintenance decisions: because the roof, solar panels, and roof equipment are separated into distinct regions, abnormal thermal patterns can be attributed to a specific component and prioritised for inspection, each element gains a stable spatial reference for tracking condition over repeated inspections, and roof and component extents can be measured directly for maintenance planning and documentation.

CONCLUSION

This study evaluated the performance of a Mask R-CNN model for rooftop thermal-image segmentation using UAV thermal images. The model was used to identify the roof, solar panel, and roof equipment in the case-study dataset. The results showed strong segmentation performance, with a mAP of 77.2%, an AP50 of 95.1%, and an AP75 of 90.2%, supported by high accuracy values and stable training behaviour. This indicates that the model was able to identify rooftop elements effectively from the thermal images. The study also shows that rooftop thermal imagery can be transformed into more structured rooftop information through segmentation. Instead of remaining as a raw thermal image, the rooftop scene can be separated into clearer object-level regions that are easier to interpret. This is important for building-related applications because clearer rooftop information is more useful for inspection, condition review, and maintenance-related observation. Overall, the research objective was achieved. The proposed model was able to identify rooftop elements and provide measurable segmentation performance from UAV thermal images. This study has several

limitations. It is based on a single building (E49, USM) and 100 thermal images, so the results indicate feasibility in this setting rather than general accuracy. The reported metrics were obtained on the validation subset; an independent test-set evaluation and per-class analysis are recommended for future work. Future work should expand the dataset across buildings and seasons and add a same-data baseline.

REFERENCES

- Acorsi, M. G., Gimenez, L. M., & Martello, M. (2020). Assessing the Performance of a Low-Cost Thermal Camera in Proximal and Aerial Conditions. *Remote Sensing*. <https://doi.org/10.3390/rs12213591>
- Balaras, C. A., & Argiriou, A. A. (2002). Infrared thermography for building diagnostics. *Energy and Buildings*. [https://doi.org/10.1016/s0378-7788\(01\)00105-0](https://doi.org/10.1016/s0378-7788(01)00105-0)
- Batchuluun, G., Lee, Y. W., & Park, K. R. (2020). Thermal Image Reconstruction Using Deep Learning. *IEEE Access*. <https://doi.org/10.1109/access.2020.3007896>
- Boccalatte, A., & Chanussot, J. (2025). Quantifying urban solar potential losses from rooftop superstructures via aerial imagery and convolutional neural networks. *Renewable Energy*, 249, 123088. <https://doi.org/10.1016/j.renene.2025.123088>
- Che Haron, R., & Muhd Zafir, N. B. (2025). Digital twin application in construction cost management. *Planning Malaysia*, 23(35). <https://doi.org/10.21837/pm.v23i35.1665>
- Chicherin, S., Zhuikov, A., & Junussova, L. (2023). District heating for poorly insulated residential buildings—Comparing results of visual study, thermography, and modeling. *Sustainability*, 15(20), 14908. <https://doi.org/10.3390/su152014908>
- Choi, J.-S., Kim, C., Jang, H., & Kim, E.-J. (2022). Dynamic thermal bridge evaluation of window-wall joints using a model-based thermography method. *Case Studies in Thermal Engineering*, 35, 102117. <https://doi.org/10.1016/j.csite.2022.102117>
- Chu, F., Cao, J., Song, Z., Shao, Z., Pang, Y., & Li, X. (2024). Toward generalizable multispectral pedestrian detection. *IEEE Transactions on Intelligent Transportation Systems*, 25(5), 3739–3750. <https://doi.org/10.1109/TITS.2023.3330155>
- Dai, X., Yuan, X., & Wei, X. (2020). TIRNet: Object detection in thermal infrared images for autonomous driving. *Applied Intelligence*, 51(3), 1244–1261. <https://doi.org/10.1007/s10489-020-01882-2>
- Deng, R., Yang, Z., Chen, X., You, G., & Yan, K. (2020). Object-based thermal image segmentation for fault diagnosis of reciprocating compressors. *Sensors*, 20(12), 3436. <https://doi.org/10.3390/s20123436>
- Dobbie, S., Strafford, H., Pickrell, O. W., Fonferko-Shadrach, B., Jones, C., Akbari, A., ... Lacey, A. (2021). Markup: A Web-Based Annotation Tool Powered by Active Learning. *Frontiers in Digital Health*, 3. <https://doi.org/10.3389/fdgth.2021.598916>
- Ejaz, N., Tariq, T. B., & Baik, S. W. (2012). Adaptive key frame extraction for video summarization using an aggregation mechanism. *Journal of Visual Communication and Image Representation*, 23(7), 1031–1040. <https://doi.org/10.1016/j.jvcir.2012.06.013>

- Garrido, I., Erazo-Aux, J., Lagüela, S., Sfarra, S., Ibarra-Castanedo, C., Pivarčiová, E., Gargiulo, G., & Maldague, X. P. V. (2021). Introduction of deep learning in thermographic monitoring of cultural heritage and improvement by automatic thermogram pre-processing algorithms. *Sensors*, 21(3), 750. <https://doi.org/10.3390/s21030750>
- Gou, Y., Li, Q., Yao, R., Chen, J., & Bai, T. (2023). Algorithm analysis of defect depth for PVC plate with flat-bottomed holes in active pulsed infrared thermography. *Measurement Science and Technology*, 34, 085404. <https://doi.org/10.1088/1361-6501/acd39e>
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 2961–2969). <https://doi.org/10.1109/ICCV.2017.322>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). <https://doi.org/10.1109/CVPR.2016.90>
- Hernandez, G., Navarro, M., & Perez-Cisneros, M. (2020). Failure Detection on Electronic Systems Using Thermal Images and Metaheuristic Algorithms. *IEEE Latin America Transactions*. <https://doi.org/10.1109/tla.2020.9111672>
- Hou, Y., Chen, M., Volk, R., & Soibelman, L. (2021). An approach to semantically segmenting building components and outdoor scenes based on multichannel aerial imagery datasets. *Remote Sensing*, 13(21), 4357. <https://doi.org/10.3390/rs13214357>
- Khan, N. A., & Bhattacharjee, B. (2021). Thermal and noise insulation performance interaction of building envelope during building simulation optimization in tropical climates. *Building and Environment*, 200, 107948. <https://doi.org/10.1016/j.buildenv.2021.107948>
- Kim, H., Lamichhane, N., Kim, C., & Shrestha, R. (2023). Innovations in building diagnostics and condition monitoring: A comprehensive review of infrared thermography applications. *Buildings*, 13(11), 2829. <https://doi.org/10.3390/buildings13112829>
- Krapf, S., Brabec, J., & Moser, M. (2022). RID—Roof information dataset for computer vision-based photovoltaic potential assessment. *Remote Sensing*, 14(10), 2299. <https://doi.org/10.3390/rs14102299>
- Mastrucci, A., van Ruijven, B., Byers, E., Poblete-Cazenave, M., & Pachauri, S. (2021). Global scenarios of residential heating and cooling energy demand and CO2 emissions. *Climatic Change*, 168(3–4). <https://doi.org/10.1007/s10584-021-03229-3>
- Mayer, Z., Kahn, J., Hou, Y., & Volk, R. (2021). AI-based thermal bridge detection of building rooftops on district scale using aerial images. In *Proceedings of the EG-ICE 2021 Workshop on Intelligent Computing in Engineering* (pp. 497–507).
- Mayer, Z., Kahn, J., Hou, Y., Götz, M., Volk, R., & Schultmann, F. (2023). Deep learning approaches to building rooftop thermal bridge detection from aerial images. *Automation in Construction*, 146, 104690. <https://doi.org/10.1016/j.autcon.2022.104690>

- Mohd Nazri, S. F. A., Hashim, K. S. H.-Y., & Mohd Suhaimi, M. S. N. (2025). BUILDING CONDITION ASSESSMENT FOR SCHOOL BUILDING. *Planning Malaysia*, 23(35). <https://doi.org/10.21837/pm.v23i35.1692>
- Mondal, S. A., Bhattacharaya, T., Rai, A., Sodhi, G. S., Bansal, R., Mondal, A., & Gupta, A. (2025). AutoTag: A framework for generating training data. *Progress in Artificial Intelligence*, 14(2), 191–210. <https://doi.org/10.1007/s13748-024-00360-x>
- Naiman, A., Farber, E., & Stein, Y. (2022). Edge detection and characterization of digitized images. *Pattern Analysis and Applications*, 26(1), 61–72. <https://doi.org/10.1007/s10044-022-01097-w>
- Nardi, I., Paoletti, D., & Sfarra, S. (2016). U-value assessment by infrared thermography: A comparison of different calculation methods in a Guarded Hot Box. *Energy and Buildings*. <https://doi.org/10.1016/j.enbuild.2016.04.017>
- Park, G., Lee, M., Jang, H., & Kim, C. (2021). Thermal anomaly detection in walls via CNN-based segmentation. *Automation in Construction*, 125, 103627. <https://doi.org/10.1016/j.autcon.2021.103627>
- Pedranzini, F., Alloni, E., Ficco, G., & Frattolillo, A. (2024). Proposal of a new method for the characterization and operational air leakages assessment in HVAC systems. *Energy and Buildings*, 305, 113881. <https://doi.org/10.1016/j.enbuild.2023.113881>
- Peng, Y., Zhao, S., & Liu, J. (2021). Segmentation of Overlapping Grape Clusters Based on the Depth Region Growing Method. *Electronics*. <https://doi.org/10.3390/electronics10222813>
- Rosti, B., Jankovic, A., Goia, F., Mathisen, H. M., & Cao, G. (2025). Design and performance analysis of a low-cost monitoring system for advanced building envelopes. *Building and Environment*, 269, 112344. <https://doi.org/10.1016/j.buildenv.2024.112344>
- Xu, Y., Chau, V., Wu, C., Zhang, Y., Zissimopoulos, V., & Zou, Y. (2023). A Semi Brute-Force Search Approach for (Balanced) Clustering. *Algorithmica*, 86(1), 130–146. <https://doi.org/10.1007/s00453-023-01158-4>
- Zhang, J., Liang, X., Wang, M., Yang, L., & Zhuo, L. (2020). Coarse-to-fine object detection in unmanned aerial vehicle imagery using lightweight convolutional neural network and deep motion saliency. *Neurocomputing*, 398, 555–565. <https://doi.org/10.1016/j.neucom.2019.03.102>

Received: 15th May 2026. Accepted: 6th June 2026