Towards Urban Heat Loss Modeling Using Building Digital Twin

Colin Richardson [1*], Anna T. Sullivan [1], Florence B. Berube [1], Shabnam Jabari [1], Travis Moore [2]

[1] University of New Brunswick, Department of Geodesy and Geomatics Engineering, Fredericton, NB, Canada
(colin.richardson, anna.sullivan, florence.berube, sh.jabari)@unb.ca

[2] National Research Council Canada, Fredericton, NB, Canada
(Travis.Moore@nrc-cnrc.gc.ca)

KEY WORDS: Thermal Mapping, Urban Digital Twins (UDT), Urban Heat Loss, Building Efficiency Modeling

Abstract

Seventy percent of global greenhouse gas (GHG) emissions originate from urban areas, with urban heat loss contributing significantly to energy consumption (UNEP, 2020). Digital twins offer a potential solution and insight into the problem and its causes. This is a study started as an undergraduate Engineering Capstone Project with a collaborative effort between the University of New Brunswick and the National Research Council of Canada to develop a workflow to aid thermal efficiency modeling using Digital twins. This project uses the University of New Brunswick (UNB) Fredericton campus as a case study to capture UAV, nadir perspective LiDAR, Panchromatic imagery and long wave infrared (LWIR) thermal imagery. The workflow includes 4 major steps following the preprocessing: (1) creating point clouds from the LiDAR and Panchromatic sources, (2) merging point clouds using grid-based segmentation and iterative closest point algorithm (ICP), (3) classifying the point cloud using Point CNN networks aided by manual refinement, and (4) overlaying thermal data. The resulting digital twin achieved a high level of spatial alignment accuracy, with 95% of points falling on building surfaces falling within an 11 cm tolerance as assessed by quadric cloud-to-cloud distance. Semantic classification performed using Point CNN and faster R-CNN object detection identified façade features such as windows and doors with a precision of 91.8% and an F1 score of 83%. Thermal data was successfully integrated and converted to approximate temperature values, enabling further analysis of surface heat behavior and laying the groundwork for future energy modeling applications. This case study demonstrates a scalable framework for high-detail drone based digital twin development with practical relevance to urban thermal efficiency analysis.

1 Introduction

Buildings hold a large amount of information within their walls, whether it be the occupancy, walking paths and common areas or if it be the very information of the structure itself such as materials, age and condition. This is the dream of a digital twin, to completely replicate and utilize all this geographic and contextual information to the best of its capabilities. Occupancy studies, energy efficiency, emergency planning and basic mapping and navigation are all examples of an industry that relies on this information that lies in front of our day to day lives (Ham & Kim, 2020). The biggest hurdle to this reality is the inclusion of a real time capture or an up-to-date model. Developments in UAV technology aid this step in big ways, the quick capture time and large scale with high detail provide a new venue and perspective for building research (Xie et al., 2023). UAV can quickly capture low-cost data with flexible operation characteristics making them well-suited to a range of specific data acquisition scenarios.

Recent advances in remote sensing and Deep learning have accelerated the potential of UAV-based digital twins, particularly for energy efficiency and thermal analysis (Ni et al., 2023). Thermal efficiency data is highly valuable and can yield very important insights when effectively applied. For example, (Sadhukhan et al., 2020) demonstrated the use of aerial thermal imagery to estimate building surface temperatures and derive thermal transmittance (U-values). More recently, (Zhu et al., 2024) proposed a method for co-registering thermal point clouds with semantic 3D building models, leveraging segmented point clouds for advanced thermal analysis. These developments reflect the broader impact of deep learning integration, especially in semantic segmentation and deep learning-based feature extraction. Hybrid models that combine LiDAR and photogrammetry have proven effective in mitigating geometric distortions (Zhu et al., 2024), while deep learning-based tools such as Mask R-CNN are enabling dynamic, real-time interpretation of thermal data (Wilson et al., 2023). This convergence of technologies allows for more accurate detection of heat-loss patterns and contributes to sustainable urban design. However, despite these advancements, thermal efficiency modeling has yet to be fully integrated into a scalable, real-time digital twin framework (Arowoiya et al., 2024).

Construction of digital twins have taken many forms across literature, often shaped by the specific use case or data availability. Some studies rely on simulation or building information materials (BIM) derived models to form the basis of the digital twin, such as in (Lu et al., 2020), who developed a multiscale DT of the West

Cambridge campus using city planning models. In contrast, (Borrmann et al., 2024) propose an AI-assisted pipeline for multi-scale digital twin generation, focusing on automatic integration rather than field-based acquisition. At the building scale, (Sundby et al., 2021) employ geometric change detection to refine DT models, but their methodology is based on fixed-position sensors. In the context of applied field methods, (Roda-Sanchez et al., 2024) demonstrated a smart campus DT integrating real-time sensor data and analytics, but the study does not integrate airborne sensors or multispectral workflows. All these methods are tailored to their specific purpose, our method is configured to use UAV-captured RGB, thermal, and LiDAR data in a multisensory configuration, processed through a structured photogrammetric and point cloud workflow to build a detailed, aligned, structural digital twin model.

While existing studies explore digital twins for building energy efficiency typically through simulation environments or single-sensor workflows, few demonstrate a detailed, applied UAV-based multi-sensor workflow that integrates LiDAR, RGB, and thermal imagery with semantic segmentation and deep-learning-based façade feature detection in a real-world case study. To the best of our knowledge, no prior work has combined these elements into a unified and reproducible pipeline for digital twin construction at the building scale. Hence, this research presents a practical case study that demonstrates a full-spectrum UAV data capture and processing workflow from multi-modal point clouds to window/door classification and thermal overlay for a real urban building façade. This applied methodology aims to contribute both an operational prototype and a base for future research into the study of energy modeling.

2 Study area

This case study explores the application of remote sensing and processing techniques to produce a foundational digital twin model. We use the Fredericton Campus of the University of New Brunswick, NB, Canada as a case study. This area spans from the bottom of campus at Head Hall engineering building up to the McLaughlin nursing building. The total area accumulating to approximately 3.4 ha with an elevation gain of 34m.

This section of campus is comprised of multistory brick buildings and a few single-story brick buildings. It features narrow structural corridors as well as a mix of flat and sloped roof surfaces. A dense forested section in the area gives a view into the inclusion of dense vegetation in UAV data. While this area is not a perfect representation of an urban environment, it exhibits many of the same characteristics and challenges commonly encountered in urban settings. This area was selected because of the structural complexity of the area and available permissions with data capture over urban environments in Fredericton.

2.1 Data Collection

Drone imagery and LiDAR has two common perspectives in most captures. A nadir perspective refers to imagery captured with the camera oriented directly downward, perpendicular to the ground surface (i.e., at a 90° angle). In contrast, an oblique perspective involves the camera being angled away from vertical typically around 45°allowing for the better capture of building façades and vertical structures. Due to an equipment limitation in this case study all data was captured from a nadir perspective.
Given the main goal of developing a detailed digital twin and thermal model suitable for future energy analysis applications, a multi-sensor approach was employed to maximize both spatial coverage and accuracy. Two complementary datasets LiDAR and photogrammetry were collected, each selected for their unique strengths and for their ability to compensate for limitations in the other. The LiDAR data was treated as the reference geometry for georeferencing the multispectral and thermal data. Data collected from the Altum-PT sensor produced high-density, detailed point clouds, which were intended to be co-registered with the LiDAR dataset.

All flights were captured from a nadir perspective to mitigate this limitation, flight parameters were adjusted between sensors to suit their respective objectives. The Zenmuse L1 system was flown at 100 meters AGL with 80% overlap, focusing on high spatial resolution and broad coverage. In contrast, the Altum-PT sensor was flown at a lower altitude of 60 meters AGL with terrain-following enabled and the same 80% overlap. These parameters were selected to improve the visibility and detail of building facades.

Data	Panchromatic	Thermal	LiDAR
Focal Length (mm)	16.3	4.5	Na
Flight Height (m)	60	60	100
GSD (cm/pixel)	2.59	22.66	2.73
Image Size	4112 x 3008	320 x 256	Na
Point Density (points/ m^2)	10,000	10.47	414

Data Collection Dates	December 17 2025	December 17 2025	December 16 2025

Table 1. Sensor and Flight Parameters

As the Altum-PT also captured thermal imagery characterized by a relatively coarse Ground Sampling Distance (GSD), the reduced flight altitude helped mitigate this limitation. Furthermore, thermal flights were scheduled for late afternoon, as close to dusk as possible. This action was done in an attempt to minimize natural solar radiance and try to enhance the view of the natural thermal signature of the building.

2.2 Preprocessing

The photogrammetric and thermal preprocessing was done in Pix4D a terrestrial and drone photogrammetry mapping software (Pix4D, 2025). The Panchromatic imagery was processed in conjunction with the Long Wave Infrared (LWIR) imagery, by doing this it helps portray the lower resolution thermal data on a higher resolution base. The resulting products, such as point clouds and meshes are then more detailed while having the lower resolution data overlayed. Then the LiDAR was processed in DJI Terra due to the nature of the sensor being used in the data capture requiring the specific software.

Following the initial raw data processing, all point clouds underwent spatial down sampling and noise filtering to ensure consistency across datasets and prepare them for integration into the processing workflows. A voxel-based down sampling approach was applied, standardizing point density to 1 cm^3, meaning that one point represents every square centimeter of space. To further improve data quality, Statistical Outlier Removal (SOR) filtering was implemented. This evaluates the local neighborhood of each point and eliminates those whose mean distance to their neighbors falls outside a defined statistical threshold.

However, variations in data density became evident due to the nadir acquisition. Ground surfaces and rooftops were well-represented due to their orientation relative to the sensor, vertical façades exhibited significantly lower point density. This discrepancy led to unintended cluster removal with some façade’s data being removed during filtering since it was interpreted as noise. These limitations were considered when interpreting classification and thermal analysis results in later stages.

3 Methodology

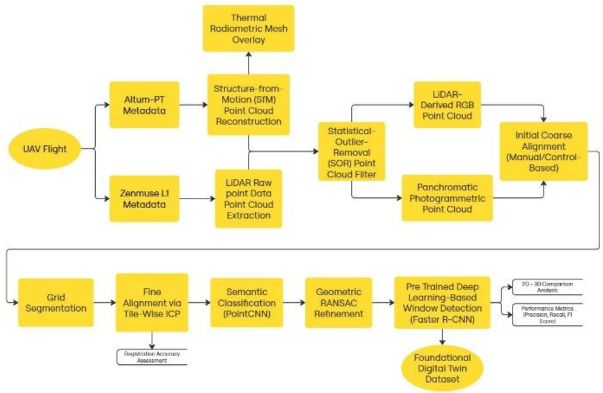

Figure 1. UAV based Workflow for Digital Twin Generation

Figure 1 shows the flowchart of the proposed work. The methodology outlined here shows a complete workflow used for generating a digital twin of the UNB campus study area. As shown in Figure 1, the process begins with data acquisition from two aerial sensor platforms and progresses through preprocessing, each preprocessing step is different for the Altum and the Zenmuse. With each input generating a point cloud, an initial coarse alignment is done. This is followed by a grid segmentation and then a tile wise ICP alignment. The aligned dataset is then subjected to semantic segmentation, followed by object classification focused on detecting windows and doors. This sequential process establishes the foundational dataset for subsequent analysis.

3.1 Multisource Point Cloud Alignment

Following the initial preprocessing steps, all datasets were converted into standardized point cloud format, subjected to noise reduction procedures, and prepared for subsequent integration. To prepare for segmentation two preparatory steps were done. First a initial manual course alignment was done to the LiDAR base to allow the later ICP alignment to easily match the geometry. Then a grid network is laid out of 50 m^3, grid squares, from a reference corner of the minimum bounding box of the cloud. The 50 m^3, grid squares were chosen to aid in the reduction of computational load on the processing and so that if there are any drifts or larger distortions from the photogrammetric point cloud then the alignment could be more accurate on a small local scale where it might struggle on a full project area. The photogrammetric i.e. Panchromatic point clouds of the project area were segmented from their original clouds on the grid square.

To reassemble the segmented point cloud grid tiles into a unified model, the LiDAR dataset was retained in its original georeferenced coordinate space, serving as the base reference. Each segmented chunk of photogrammetric data was then independently aligned back to the LiDAR using ICP. This was performed in CloudCompare by maintaining the LiDAR tile as the fixed (static) reference and transforming the segmented photogrammetric tiles using the "ICP” tools. This approach minimizes cumulative errors and maintains spatial integrity when working with multi-tile 3D point cloud datasets.

To evaluate this methodology, we do the same C2C distance computation but with a few changes. For this iteration, only building structures were retained from each point cloud, isolating them from transient or mobile elements such as vehicles and vegetation, which were subject to positional and structural changes between acquisitions. Given the study’s focus on structural energy modeling, buildings were chosen as the primary comparison target. The aligned panchromatic-derived point cloud was evaluated against the LiDAR-derived dataset, which served as a reference.

3.2 Semantic Point Cloud Classification

The usage of classified objects in 3D space allows for a wide array of in-depth analysis and detailed feature extraction. For this study the point cloud classification was done on two levels, first a semantic segmentation was carried out using Point CNN models then a refinement for windows and doors using Fast RCNN detection.

The initial phase of semantic classification was performed using Trimble Business Center (TBC). This was done to assign broad class labels to the now recombined point cloud. This step provided a foundational classification essential for downstream processing and was used to separate the dataset into nine primary classes, including buildings, ground, vegetation, power lines, and noise.

TBC's automated classification engine works with a combination of machine learning algorithms, specifically decision tree classifiers and random forest models, to distinguish between object types. These classifiers operate by analyzing multiple point attributes such as intensity, elevation, surface normal orientation, and local point density, comparing each point to its surrounding context. The underlying assumption is that features exhibiting similar geometric and radiometric properties likely belong to the same semantic category (Trimble, 2025).

Following the initial classification, a refinement was applied using TBC’s optional advanced classification module. This process uses RANSAC-based geometric fitting, where planar and linear models are fitted to point clusters to improve classification accuracy. For instance, vertical planar segments may be identified as building facades, while horizontal planes may represent ground or roof surfaces. However, in the TBC user’s interface there are no configurable settings or parameters for this step. This step plays a key role in reducing misclassification in complex urban scenes.

After establishing a preliminary automated classification of the point cloud, manual refinement was performed to correct significant misclassifications.. These adjustments primarily targeted structural elements that were incorrectly labeled or omitted.

Following the automated classification phases, additional effort was directed toward identifying and labeling specific thermally influential objects. This was a manual classification targeted at elements such as heat exchangers, HVAC vents, solar panels, and other rooftop or façade-mounted systems that may serve as thermal anomaly sources within the scene. Due to the limited frequency and non-uniform geometry of these features across the dataset, a consistent and repeatable training set was not available, rendering automated classification unfeasible. As a result, all relevant objects were manually extracted and classified using visual inspection of the coloured point cloud.

3.3 Deep Learning-Based Facade Feature Detection

With the base classification of the point cloud completed, specific façade and building features became the next focus of analysis. The detection was carried out using a pretrained faster R-CNN object detection model accessed via Esri’s Living Atlas and integrated within the ArcGIS Pro 3D Object Detection toolkit. Since the only user defined variable is the minimum confidence level no other parameters are able to be altered. The model operates on a textured 3D mesh in Arc GIS’s 3D space. This mesh was generated as an output from the Pix 4D software using the RGB imagery from the Zenmuse. This was done because the pretrained model performed best on a colourized mesh.

The ESRI pretrained faster R-CNN model generates 3D bounding boxes for detected windows and doors based on texture features, shape, and known facade configurations. In addition to this a preset confidence interval of 50% was set for detected objects, all objects below this confidence were discarded. The confidence interval was chosen due to certain resolution limitations with the mesh that will be discussed in the analysis. These bounding boxes are then projected onto the classified point cloud, where points falling within the defined spatial extent are reclassified accordingly.

3.4 Thermal Integration

Thermal overlays provide an invaluable layer of insight, enabling the visualization of temperature variations across a 3D model.

https://doi.org/10.5194/isprs-archives-XLVIII-M-8-2025-35-2025 | © Author(s) 2025. CC BY 4.0 License.

The result from the preprocessing of the thermal and panchromatic data is a uncalibrated LWIR band, after applying a formula provided in the official PIX4D documentation, the data is converted into temperature values in degrees Celsius (°C) (Pix4D, 2025).

$$T(C^{\circ}) = \left(\frac{\text{lwir}}{100}\right) - 273.5 \tag{1}$$

This formula is a temperature approximation, as it assumes that the surface being measured behaves as an ideal blackbody emitter with an emissivity of 1.0. In real-world conditions, materials exhibit varying emissivity values that are typically less than 1, depending on surface properties such as texture, color, and material composition. As a result, the derived temperature values may differ from true surface temperatures, especially when analyzing materials like glass, metal, or painted surfaces. However, the temperature measurements that are calculated from the raw imagery of the Altum are still representative and by extension the patterns and observations shown are still valid representations. After calibration Pix4D generates an orthomosaic representation of the LWIR band. This is then what is overlayed on a resulting mesh representation. Importantly, Pix4D does not embed temperature values directly into the mesh geometry; instead, it applies a texture-mapped overlay using a standard RGB image where each pixel color corresponds to a normalized temperature value on a 0–255 scale.

To extract quantitative temperature data from this visualization, the colours are baked into the mesh in mech lab then a scalar field is produced from the colour values. The following formula converts the colours back to the temperature values using the minimum temperature and maximum temperature that Pix 4D calculated.

$$Temp\ (°\text{C}) = \left(\frac{colour\ value}{255}\right) * (\max temp - \min temp) + \min temp \tag{2}$$

This scalar field allows the thermal information to be analyzed numerically in external software environments. While the thermal data can be used to compute zonal statistics, the primary objective in this case study was to generate a foundational framework that supports future research. This includes developing methods to enhance and fill energy model parameters and exploring the potential for object-level feature extraction from the point cloud using distinct thermal signatures.

4 Experiments

The concept of running the pretrained object detection model on raw imagery works on the same principles as the 3D viewshed but alternatively for batch processing was run in code. First, each input raster is evaluated to ensure it is in a RGB format. Inference was then carried out using the same pretrained Faster R-CNN architecture. Output detections were exported as vectorized feature classes for each input raster

This method yielded suboptimal results only achieving a Precision of 67.1%. While the pretrained object detection model was able to identify windows in the 2D imagery, the 2D-based detection demonstrated lower recall and a higher rate of false positives. This discrepancy may be attributed to two primary factors: first, the difference in viewing angle between the nadir-oriented orthophotos and the oblique perspective offered by the 3D mesh viewshed; and second, the inherent distortion and lack of façade visibility in nadir imagery, which hinders consistent detection.

Figure 1 shows one of the best cases from the 2D detection and shows window detection however, it exhibits double the amount of misclassification compared to the textured mesh-based detection for the same façade area.

Figure 2. Example Output of 2D Imagery-Based Window Detection Using Deep Learning

Despite geometric distortions in the mesh, the 3D model benefited from a more façade-centric perspective, which improved the contextual visibility of windows and doors.

5 Results

5.1 Multisource Point Cloud Alignment

To provide a baseline for alignment evaluation, an initial cloud-to-cloud (C2C) distance computation was performed using the raw positions derived from the onboard GNSS/RTK data associated with each sensor. This yielded a result showing an average distance of 3.4 m, this difference being attributed to the use of separate positioning systems for the Altum-PT and Zenmuse L1.

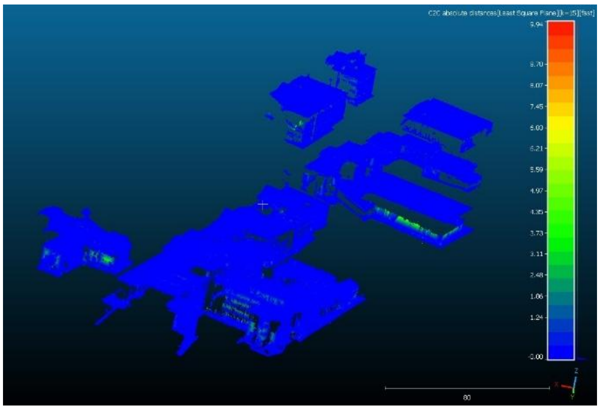

Figure 3. Building Comparison C2CResults

The results of the above image show that 95% of the alignment fell under 11cm distance in the cloud-to-cloud comparison. The notable exceptions to that being the highlighted green sections. These sections are sections that did not overlap in the reference LiDAR point cloud.

5.2 Façade Detection

The results of the object detection model, applied to textured mesh scenes within ArcGIS Pro's 3D environment, yielded a total of 246 window and door classifications. Out of that, there was a total of 295 ground truth windows.

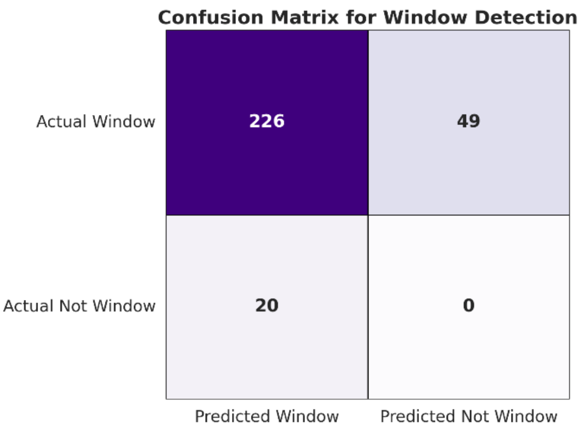

Figure 4. Facade Detection Confusion Matrix for Window Detection

The model achieved a precision of 91.8%, a recall of 75.9%, and an F1 score of 83%.

Figure 5. 3D Mesh-Based Window Detection Using Esri’s Pretrained Deep Learning Model

Despite the strong performance in precision, the model demonstrated a noticeable number of false negatives, with approximately 24% of identifiable windows not detected. This estimate is somewhat speculative, as it is based on visual inspection of windows that appeared sufficiently clear to be classified. Misclassifications primarily occurred where features such as vehicle roofs or regular square shapes mimicked the spatial or visual characteristics of windows. These issues were likely exacerbated by the geometric distortions introduced during mesh generation. In particular around window edges where surface warping and texture stretching were evident. Its is likely that with an improved mesh the resolution increase would lead to much better results than what is currently shown in this case study.

6 Discussion

This case study has demonstrated the practicality of using unmanned aerial vehicles (UAV) in the generation of digital twins of complex environments. But this has also left multiple areas to be grown further and expanded upon. The workflow presented establishes a foundational framework for integrating multisensory data, including thermal, multispectral, and LiDAR, toward advancing energy modeling applications. Specifically, it lays the groundwork for producing Level of Detail 3 (LOD3) digital twins, which are characterized by accurate representation of building facades and roof structures.

With all this said there are several keys’ areas with room for improvement. Adjustments to the data capture from both sources to help with resolution and complete coverage, is a key goal in the continuation of this research. As mentioned in section 3.4, this workflow is still very underdeveloped and leaves room for research in applications and how this can further aid energy modeling.

6.1 Data Capture Adjustments

Data capture is the foundation of any project and has profound implications on 3D model quality and subsequent analysis workflows. In this study, all datasets were acquired from a nadir perspective, using independent flight missions for each sensor system (Altum-PT and Zenmuse L1). This configuration introduced several limitations in facade detail capture, mesh fidelity, and thermal perspective, which consequently impacted classification accuracy and alignment quality. One of the main goals for the continuation of this research is the development of a gimbal for the Altum-pt camera so that oblique imagery can be done. Oblique perspectives are expected to significantly enhance the quality of reconstructed vertical surfaces, improve window and door detection accuracy, and reduce mesh distortions associated with steep viewing angles. Additionally, further improvements can be made by reducing flight altitude to increase ground resolution, and by incorporating higher-accuracy ground control targets to improve georeferencing consistency across flight sessions.

6.2 Prospective Thermal Applications

The addition of LWIR data allows for a valuable insight and snapshot of the thermal behaviour of buildings (Sun et al., 2023). While this is a major point to expand upon in this case study there are several avenues this data could be used for. One such example is the inclusion of thermal data to aid in point cloud classification, specifically in the addition of cross-attenuation models. This framework would look at the thermal signatures of objects to help classify them into specific classes.

From an energy modeling standpoint, there are two main methods energy modeling which is commonly employed. Physics based energy modeling aka building information modeling, which uses 2D floorplans to generate 3D geometry and material properties (LOD 2.5-3). These simulations rely on known thermal transmittance values (U-values) and are used to evaluate heat transfer under dynamic environmental conditions. While there are several versions of physics-based modeling, they all follow this general format (Malhotra et al., 2022).

The second method is thermal imaging-based estimation. This method relies on infrared imagery, although some models do incorporate a temporal time series aspect. These models extract surface temperatures and by using a known or estimated internal temperature derive the transmittance (U values). (Sadhukhan et al., 2020).

Both of these applications are of interest and consideration for future research. Energy modeling benefits greatly from the less intrusive faster modeling that remote sensing offers and the possibilities of maximizing LWIR imagery potential only furthers the main goals of this project.

7 Conclusion

This report outlines the development and execution of a UAV-based data integration workflow aimed at generating a digital twin for aiding energy modeling. By combining LiDAR, multispectral, and thermal imagery acquired from nadir perspectives, a structured methodology was developed for point cloud preprocessing, segmentation, alignment, semantic classification, and object detection.

The methodology demonstrated successful alignment of multisource point clouds, with 95% of building points falling within an 11 cm error threshold, as measured using a quadric cloud-to-cloud distance metric. The point cloud classification was carried out in three phases, beginning with automated general classification using Trimble Business Center, followed by deep learning-based detection of windows and doors via a faster R-CNN model, and finally manual classification of thermally influential rooftop features. The object detection model, when applied to textured 3D scenes, achieved a precision of 91.8% and an F1 score of 83%, although some misclassifications and missed detections were observed due to texture distortion and limited viewing angles. The limitations in data acquisition more specifically the exclusive use of nadir imagery, restricted the quality of façade reconstruction, mesh resolution, and ultimately classification accuracy. As such, future work should explore the use of oblique image capture.

Overall, this report demonstrates the feasibility of producing digital twins using UAV-acquired data and highlights the potential for extending these models into scalable, automated frameworks for urban thermal analysis and energy modeling. This research is the foundation which will be built on in future research.

8 REFERENCES

UNEP (2020): United Nations Environment Programme (UNEP), 2020: *Cities and Climate Change*. https://www.unep.org/publications-data

Zhu, J., Wysocki, O., Holst, C., & Kolbe, T. H. (2024). Enriching thermal point clouds of buildings using semantic 3D building models. *TUM School of Engineering and Design, Springer, 36(2)*, 150-180. 10.5194/isprs-annals-X-4-W5-2024-341-2024

Sun, Z., Zhang, Y., & Zhao, W. (2023): Digital Twin Technology for Thermal Comfort and Energy Efficiency in Buildings. *Monash Research Publications*, 10(2), 45-67. 10.1016/j.enbenv.2023.05.004

Arowoiya, V. A., Moehler, R. C., & Fang, Y. (2024). Digital twin technology for thermal comfort and energy efficiency in buildings: A state-of-the-art and future directions. *Energy and Built Environment,* 5, 641–656. https://doi.org/10.1016/j.enbenv.2023.05.004

Sadhukhan, D., Peri, S., Sugunaraj, N., Biswas, A., Selvaraj, D. F., Koiner, K., Rosener, A., Dunlevy, M., Goveas, N., Flynn, D., & Ranganathan, P. (2020). Estimating surface temperature from thermal imagery of buildings for accurate thermal transmittance (U-value): A machine learning perspective. *Journal of Building Engineering, 32*, 101637. https://doi.org/10.1016/j.jobe.2020.101637

Ham, Y., & Kim, J. (2020). Participatory Sensing and Digital Twin City: Updating Virtual City Models for Enhanced Risk-Informed Decision-Making. *Journal of Management in Engineering, 36*(3), 04020005. 10.1061/(ASCE)ME.1943-5479.0000748

Trimble. (2025). Trimble Business Center (Version 5.80) [Software]. Trimble Inc. https://www.trimble.com/Trimble-Business-Center

Esri. (2025). Living Atlas of the World [Geospatial data]. Esri. https://www.esri.com/en us/arcgis/products/living-atlas

Pix4D. (2025). *Processing thermal imagery – PIX4Dmapper documentation* [Technical documentation]. Pix4D. Retrieved from Pix4D Support.

Wilson, A. N., Gupta, K. A., Koduru, B. H., Kumar, A., Jha, A., & Cenkeramaddi, L. R. (2023). Recent advances in thermal imaging and its applications using machine learning: A review. IEEE Sensors Journal, 23(4), 3395–3407. https://doi.org/10.1109/JSEN.2023.3234335

Malhotra, A., Bischof, J., Nichersu, A., Häfele, K.-H., Exenberger, J., Sood, D., Allan, J., Frisch, J., van Treeck, C., O'Donnell, J., & Schweiger, G. (2022). Information modelling for urban building energy simulation—A taxonomic review. Building and Environment, 208, 108552. https://doi.org/10.1016/j.buildenv.2021.108552

Borrmann, A., Biswanath, M., Braun, A., Chen, Z., Cremers, D., Heeramaglore, M., Hoegner, L., Mehranfar, M., Kolbe, T. H., Petzold, F., Rueda, A., Solonets, S., & Zhu, X. X. (2024). Artificial Intelligence for the Automated Creation of Multi-scale Digital Twins of the Built World—AI4TWINNING. In T. H. Kolbe, A. Donaubauer, & C. Beil (Eds.), *Recent Advances in 3D Geoinformation Science* (pp. 233–247). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43699-4_14

Lu, Q., Parlikad, A. K., Woodall, P., Don Ranasinghe, G., Xie, X., Liang, Z., Konstantinou, E., Heaton, J., & Schooling, J. (2020). Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus. *Journal of Management in Engineering*, *36*(3), 05020004. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000763

Roda-Sanchez, L., Cirillo, F., Solmaz, G., Jacobs, T., Garrido-Hidalgo, C., Olivares, T., & Kovacs, E. (2024). Building a Smart Campus Digital Twin: System, Analytics, and Lessons Learned From a Real-World Project. *IEEE Internet of Things Journal*, *11*(3), 4614–4627. https://doi.org/10.1109/JIOT.2023.3300447

Sundby, T., Graham, J. M., Rasheed, A., Tabib, M., & San, O. (2021). Geometric Change Detection in Digital Twins. *Digital*, *1*(2), Article 2. https://doi.org/10.3390/digital1020009