Towards Roof Material Identification by Fusing Aerial and Street View Imagery

Faezeh Soleimani Vostikolaei [1], SeyedPooya Soofbaf[1], Travis Moore[2], Shabnam Jabari [1]

[1] Department of Geodesy & Geomatics Engineering, University of New Brunswick, Fredericton, NB, Canada- (fsoleima@unb.ca, sp.soofbaf@unb.ca, and sh.jabari@unb.ca)
[2] Construction Research Centre, National Research Council Canada, Travis.Moore@nrc-cnrc.gc.ca

Keywords: 3D building Modeling, Roof Material Detection, Roof Type Classification, 3D City Modeling, Digital Twins.

Abstract

Accurate roof material detection is essential for 3D building modeling, urban planning, climate change mitigation, and infrastructure maintenance. Roof materials directly influence key building properties, including thermal behavior and solar reflectance, and enable urban planners to prioritize climate-resilient infrastructure in areas exposed to extreme weather conditions, thereby helping to mitigate the urban heat island effect. Remote sensing and aerial imagery offer great potential for identifying roof materials. Recent advances in deep learning, particularly convolutional neural networks (CNNs), have significantly improved the performance of roof material detection using images.

Despite their success, approaches relying solely on aerial imagery remain inherently limited. From a top-down perspective, many roofing materials, such as asphalt and metal, exhibit similar spectral characteristics, making them difficult to distinguish. Furthermore, aerial imagery is often affected by shadows, occlusions, and limited texture visibility, particularly in dense urban environments or regions with complex illumination conditions.

In this study, we propose a novel multimodal framework for roof material classification that leverages both aerial orthoimages and ground-level imagery derived from GoPro video recordings. The proposed method utilizes spectral and textural features extracted from both nadir and off-nadir images to improve roof material detection. The framework employs a dual-branch architecture in which aerial and ground-level image features are extracted independently using CNN-based feature extractors. A decision-level late fusion strategy with learnable weights is designed to combine the predictions from each branch. The proposed method classifies four roof material classes, including asphalt, metal, membrane, and gravel, and achieves an overall accuracy of 91%.

1. Introduction

Accurate identification of roof materials is a critical component towards 3D building modeling, particularly for applications related to energy efficiency, climate resilience, urban heat modeling, and structural vulnerability assessment (Arzoumanidis et al., 2025; Law & Miura, 2025). Roof material choice significantly affects heat loss, snow load tolerance, wind resistance, and long-term maintenance requirements (Wang et al., 2024). As 3D city models evolve toward energy-aware digital twins, reliable roof details such as type and materials become a key step for estimating building thermal behaviour and photovoltaic suitability ((Li & Yang, 2025; Soleimani Vostikolaei & Jabari, 2023; Ujang et al., 2018; Vostikolaei & Jabari, 2025).

Remote sensing data, particularly high-resolution aerial and satellite imagery, provide an efficient means for large-scale roof material detection (Soleimani Vostikolaei & Jabari, 2025). Traditional approaches based on unsupervised clustering rely solely on spectral similarity and cannot separate materials with similar colors and textures. With the emergence of deep learning, especially convolutional neural networks (CNNs), significant improvements have been achieved in extracting multi-scale spectral and textural features from imagery, leading to more accurate roof material classification (Tavakoligargari et al., 2025). Kim et al. (2021) introduced a deep learning pipeline for roof material recognition from satellite imagery, where a heuristic detector first isolates roof regions before a custom 43-layer CNN performs material classification. Their architecture includes deeper convolutional stacks, optimized kernel designs, and enhanced skip connections.

Despite the advancement of CNN-based models for detecting roof materials, relying solely on nadir-view aerial imagery remains challenging. Many roofs, such as asphalt shingles and metal sheets, have spectral similarity among roofing materials, posing a significant challenge for accurate differentiation. Additionally, aerial imagery is covered by shadows, occlusions, or has limited texture visibility, particularly in dense urban environments or under complex illumination conditions. These factors reduce the availability of discriminative features and make it difficult to reliably extract roof material information from a single image.

To address these challenges, this study proposes a novel multimodal framework that integrates aerial orthoimages with ground-level imagery derived from GoPro video recordings. Unlike traditional approaches that rely on pre-existing street-view datasets, the proposed method introduces a flexible data acquisition strategy by extracting frames from field-survey videos. Roof regions are segmented from ground-level images to enable the model to focus on roof material features captured from off-nadir perspectives.

The proposed framework employs a dual-branch architecture in which aerial and ground-level images are processed independently using CNN-based feature extractors. To overcome the inherent spatial misalignment between modalities, a decision-level late fusion strategy is designed to combine predictions from both branches using learnable weights. This design allows the model to effectively leverage the complementary strengths of each modality, which provides global coverage from aerial imagery and rich texture information from ground-level views.

The main contribution of this work is a novel multimodal framework that integrates unpaired aerial orthoimages with GoPro-derived ground-level imagery, automatically segments roof regions from street-view data, and combines nadir and off-nadir features using a dual-branch CNN with decision-level late

fusion, without the need for alignment, to improve roof material classification accuracy.

2. Literature Review

Traditional approaches for rooftop and material detection are typically based on handcrafted spectral features and rule-based techniques, such as Support Vector Machines (SVMs) (Assouline et al., 2017) and Markov Random Fields (MRFs) (Katartzis & Sahli, 2008), which are often time-consuming, costly, and difficult to generalize across large geographic areas.

Recent advancements in CNN-based technologies have revolutionized the extraction of roof materials from high-quality, high-resolution street views and remote sensing images (Law & Miura, 2025; Mantas et al., 2025).

For example, Solovyev (2020) presented one of the first large-scale competition solutions for roof classification, employing multi-channel convolutional neural networks trained on Caribbean aerial imagery. The study introduced meta-features to improve prediction accuracy, achieving second place in the DrivenData "Open AI Caribbean Challenge". Similarly, Kim et al. (2021) proposed a two-stage pipeline combining heuristic building detection with a 43-layer convolutional network for roof-material recognition, demonstrating the advantages of architecture customization and preprocessing for large-scale satellite scenes.

In another work, Tingzon et al. (2023) applied instance-segmentation methods such as the Segment Anything Model (SAM) to delineate buildings from drone imagery in the Caribbean before classifying roof materials with convolutional networks. Their work emphasized the challenges of cross-country generalization, showing that models trained in one region can transfer reasonably well to another when roof geometry and color distributions are comparable. This study also underscored the importance of accurate footprint alignment and the integration of crowdsourced building inventories for climate-resilience analysis.

Satellite and airborne images are often shadowed or have poor contrast, making it challenging to extract roof material features from a single image. To improve the accuracy of the roof material detection network, some studies used multimodal datasets rather than relying on single-modal data. For example, Mantas et al. (2025) introduced RoofSense, a multimodal dataset integrating high-resolution RGB and airborne laser-scanning (ALS) data from several Dutch cities. The dataset provides pixel-level annotations for eight roof-material classes and supports semantic-segmentation research. Experiments with a ResNet-18-D and DeepLabv3+ architecture showed that pixel-wise RGB-ALS fusion improves accuracy, particularly for classes with subtle textural differences, although results remain sensitive to mis-registration between modalities. RoofSense highlights the value of dense labeling and multimodal integration for reliable roof-material mapping.

Most recently, Law & Miura, (2025) proposed RoofNet, a multimodal global benchmark dataset containing over 51,500 annotated roof material samples spanning 112 countries and covering 14 roof-material categories. The dataset utilizes a vision-language model (RemoteCLIP) with geographically aware prompts, demonstrating the potential of large-scale vision-language learning and human-in-the-loop verification to generate globally consistent roof-material datasets at unprecedented scale.

Srivastava et al. (2019) proposed a multimodal deep learning model that combines overhead imagery with Google Street View images for urban land-use classification. Their dual-stream CNN performs feature-level fusion through concatenation and achieves higher accuracy than single-modality baselines. They also introduce a CCA-based method to handle missing ground-level views at test time.

Hoffmann et al. (2019) evaluated multimodal fusion for building type classification using aerial and street-view images. Their work revealed that early-stage geometric fusion of features from two modalities often produces destructive effects due to spatial misalignment between aerial and street-level perspectives. Instead, they demonstrated that decision-level fusion, training independent CNN models on each image type and combining predictions at the output stage, significantly outperformed end-to-end fusion approaches, increasing precision from 68% to 76%. Their approach of maintaining model diversity through independent training on each modality provides a framework for leveraging complementary information from multiple viewpoints while avoiding the pitfalls of premature feature integration.

More recent studies have focused on improving accuracy, scalability, and integration with urban applications. For instance, Sun et al. (2025) introduces a dual-source deep learning framework that combines satellite imagery with Google Street View data to classify both roof and façade materials at the city scale. The study leverages geospatial data from OpenStreetMap and employs multiple CNN architectures with transfer learning to improve generalization across different cities, demonstrating strong scalability and applicability for urban sustainability and carbon assessment.

Overall, these studies show that while deep learning has significantly improved roof material detection, approaches relying on a single modality remain limited in their ability to capture material features in shadowed areas. In particular, aerial orthoimages alone do not provide sufficient spatial detail to distinguish between roofing materials, especially when materials such as asphalt shingles and metal sheet roofs exhibit similar spectral properties from nadir views. To overcome this limitation, this study incorporates an additional data source, the GoPro field-survey dataset, which provides oblique perspectives of facades, vegetation, and roofs.

3. Materials and methods

3.1 Data

This study focuses on New Brunswick, Canada, using two datasets. The first dataset consists of aerial orthoimages with a spatial resolution of 0.07 m provided by the Government of New Brunswick (GNB)[1], from which 360 manually labelled roof samples were extracted across four material classes: asphalt, metal, membrane, and gravel. The street-view dataset is extracted from GoPro video recording frames over 287 buildings in several urban areas, capturing oblique perspectives of roof edges and façade-visible materials from ground level. Roofs in both datasets were annotated using the same four-class and divided into training, validation, and test subsets following a consistent 70/15/15 split, respectively. The specifications of the datasets are shown in Table 1.

Data	Range of GSD	Date of acquisition
orthoimage	0.07 m	2015
Street view	1920 × 1080 px	2024

Table 1. Specification of the dataset.

[1] https://www.snb.ca/geonb1/e/dc/catalogue-E.asp

3.2 Methodology

This research employs a dual-branch deep learning architecture that integrates (1) aerial orthoimages extracted from aerial imagery and (2) GoPro field-survey imagery using an unpaired decision-level late-fusion strategy. The overall schema of the proposed methodology is shown in Figure 1.

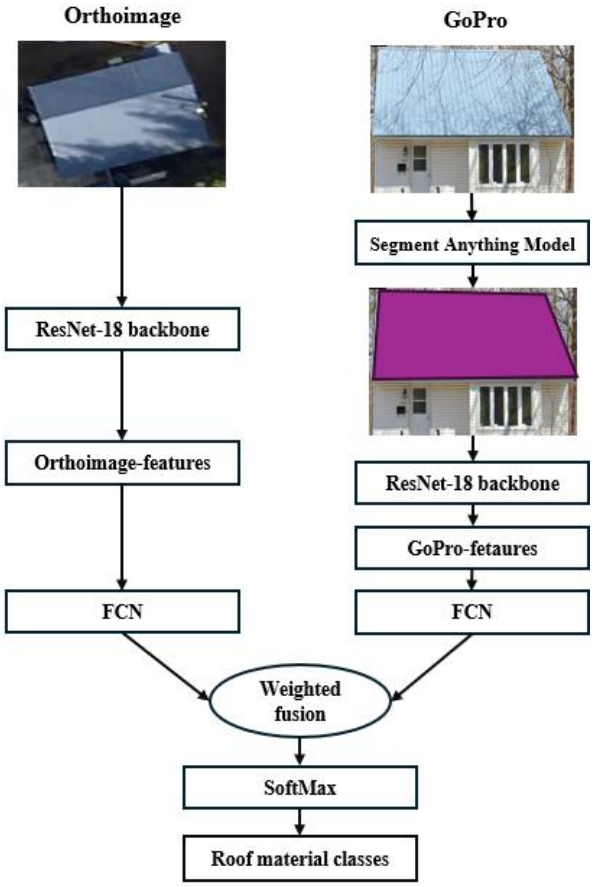

Figure 1. Overall schema of the proposed method

Aerial Orthoimage Branch

The aerial imagery baseline branch uses ResNet-18 as the backbone to extract the image features and classify the roof material. The model is initialized with ImageNet-pretrained weights and fine-tuned for four roof-material classes. During training, all layers of the network were unfrozen, which allowed both low-level and high-level features to adapt to the spectral, textural, and geometric characteristics of aerial imagery. The final fully connected layer was replaced with a new layer corresponding to the number of roof material classes and trained using cross-entropy loss.
Each image is resized to 224 × 224 pixels and normalized using ImageNet mean and standard deviation to ensure consistency with the pretrained model. After global average pooling, the network produces a 512-D feature vector, which is fed to a fully connected classifier that outputs class logits for the roof patch.

Street View Branch

The street view images capture facade-visible roof materials from ground-level perspectives. To collect these data, a GoPro camera, which is a compact wide-angle action camera commonly used for field surveys, was used to record videos around buildings. The recorded videos are then converted into individual frames at one-second intervals using an automated Python-based frame extraction process.
To isolate the roof region, the Segment Anything Model (SAM) (Kirillov et al., 2023) is applied to segment roofs from the other objects. The segmented roof crops are resized to 224 × 224 pixels. Data augmentation is applied to improve robustness, including:

- Random horizontal flipping
- Rotations
- Color jitter

A second ResNet-18 encoder, pretrained on ImageNet and fine-tuned on the street view dataset, outputs class logits for each sample; class probabilities are obtained by applying SoftMax during inference, using the same roof-material classes as the aerial branch.

Late Weighted Fusion

Because aerial orthoimage and street-view images are unpaired, and due to the lack of spatial correspondence between them, fusion is performed at the decision level, avoiding the need for pixel-wise or feature-level alignment.
For each class of roof material c, two learnable parameters $\alpha_a(c)$ and $\alpha_s(c)$ are introduced for aerial and street-view modalities, respectively. These are converted into normalized weights using a softmax function:

$$w_a(c) = \frac{e^{\alpha_a(c)}}{e^{\alpha_a(c)} + e^{\alpha_s(c)}}, w_s(c) = \frac{e^{\alpha_s(c)}}{e^{\alpha_a(c)} + e^{\alpha_s(c)}}$$

where, $w_a(c)$ and $w_s(c)$ are the weight assigned to the aerial and street view modality for class c, respectively. These learnable parameters allow the model to rely more on the modality that performs better for each class.
The fused probability vector is obtained by a weighted summation of the branch logits. This parameter for class c is computed as:

$$z_{\text{fused}}(c) = w_a(c) \cdot z_a(c) + w_s(c) \cdot z_s(c)$$

where, $z_a(c)$ and $z_s(c)$ are the output of fully connected layers in aerial and street view modalities, respectively.

A final SoftMax layer is applied to obtain class probabilities, and the predicted class is selected using the argmax function.

4. Experiments

We first designed two baseline networks for aerial and street view datasets based on the ResNet-18 module and initialized them with ImageNet-pretrained weights (Deng et al., 2009) and fine-tuned them independently on their respective datasets. The models are trained using cross-entropy loss and the Adam optimizer (X. Wang & Aitchison, 2025).
After training the two branches independently, their weights are frozen, and the fusion module, which is based on a customized ResNet-18, is trained separately. In this second stage, only the fusion parameters $\alpha_a(c)$ and $\alpha_s(c)$ are optimized during training, while the parameters of both backbone networks remain fixed. The fusion model operates on the logit outputs of each branch, which are the prediction scores output by the fully connected layer.
The fusion module learns the weights for each modality by optimizing the fusion parameters on unpaired data, in which the aerial and street-view samples are not spatially aligned. During training, modality-specific batches are processed independently,

and the model updates the fusion weights based on the classification loss, enabling it to estimate the relative reliability of each modality for each class without requiring paired observations.
To address the absence of paired samples, the fusion module is trained using an alternating-batch strategy, in which aerial and street-view data are processed independently. During aerial batches, only the aerial logits are provided to the fusion module, while the street-view logits are set to zero. Conversely, during street-view batches, only the street-view logits are used, and classes not present in the street-view dataset are excluded from the loss computation by masking. This training scheme enables the model to learn class-specific modality importance based on modality-specific supervision, without requiring spatially corresponding inputs across modalities.
The aerial and street-view branches were each trained for 200 epochs using Adam with an initial learning rate of 5e-5. After freezing both backbones, the fusion head was trained separately on unpaired batches for 20 epochs using Adam with an initial learning rate of 1e-3.

We evaluated the proposed framework using Precision, Recall, and F1-score metrics. The performance of the proposed method is evaluated using standard classification metrics, including precision, recall, F1-score, and accuracy, computed at the class level.
Precision measures the proportion of correctly predicted samples for a given class among all samples predicted as that class

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}$$

Recall measures the proportion of correctly predicted samples for a given class among all actual samples belonging to that class:

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}$$

The F1-score represents the harmonic mean of precision and recall, providing a balanced measure between them:

$$\text{F1} = \frac{2 \times Precision \times Recall}{\text{Precision+Recall}}$$

5. Results and Discussion

The performance of the proposed framework is evaluated under the aerial branch and multimodal aerial and street view fusion branch. The results are shown in Table 2.
The aerial orthoimages ResNet-18 baseline achieved an accuracy of 72.2%, an F1-score of 0.69, and a recall of 0.72, demonstrating strong performance given the limited dataset size and spatial resolution.
The street-view branch achieved 90% test accuracy with an F1-score of 0.84 and a recall of 0.858. The street-view branch achieved higher accuracy than the aerial orthoimage branch, primarily due to its rich texture information visible from oblique ground-level viewpoints. However, because street-view imagery cannot capture flat or low-slope roofs from the ground, only two roof material classes, asphalt and metal, were visible and annotated, and flat or low-slope roofs, such as membrane and gravel, are not observable. Despite these constraints, this branch provides valuable complementary information for pitched-roof materials and serves as an important component for subsequent bimodal fusions.
The proposed multimodal fusion model, which combines the strengths of aerial with street view modalities, achieves an overall accuracy of 91%, F1-score of 0.8, and a recall of 0.8, for four classes.

	Accuracy (%)	F1-score	Recall
Aerial baseline	72.2	0.69	0.72
Proposed method	91	0.84	0.80

Table 2. The overall accuracy of the proposed method

The results clearly show that multimodal fusion provides substantial improvements over the aerial branch. By combining these two modality features through decision-level fusion, the model reduces ambiguity between spectrally similar materials.

		Asphalt	Metal	Membrane	Gravel
Aerial	Train	102	84	39	27
	Validation	21	18	9	6
	Test	24	18	6	6
Street view	Train	147	54	0	0
	Validation	32	12	0	0
	Test	31	11	0	0

Table 3. Dataset distribution per class

Table 3 shows the distribution of samples across roof material classes for both aerial and street-view datasets. The aerial dataset contains four roof material classes (asphalt, metal, membrane, and gravel), which are relatively well distributed, although asphalt and metal dominate the dataset.
In contrast, the street-view dataset contains only two visible classes, asphalt and metal, due to the inherent limitations of ground-level observations, to capture the flat or low-slope roofs. Figure 2 shows some snapshots of aerial and GoPro datasets, which are shadowed.

Figure 2. Some snapshots of the aerial and street view (GoPro) datasets

6. Conclusion

This study presented a unpaired multimodal framework for roof material classification by integrating aerial orthoimages with ground-level imagery derived from GoPro video recordings. In

the proposed workflow, GoPro videos are first converted into individual frames, and roof regions are extracted using a segmentation model.

Unlike existing approaches that rely on spatially aligned datasets, the proposed method does not require paired aerial orthoimages and street-view data. Instead, it employs a dual-branch ResNet architecture combined with a decision-level late fusion strategy with class-specific weights, which allows the model to use information from both aerial and street view data.

The experimental results demonstrate that single-modality approaches are limited in finding all types of roof material precisely. The aerial orthoimage branch, although capable of capturing structural information, struggles to distinguish materials with similar spectral characteristics. In contrast, the street-view branch benefits from rich texture information captured from oblique viewpoints and achieves higher accuracy for visible classes; however, it lacks complete coverage of all roof types.

In particular, flat or low-slope roofs, such as membrane and gravel, are not visible from ground-level perspectives, as they are typically occluded by building façades and surrounding structures. As a result, these roof types are underrepresented or absent in the street-view dataset.

By combining these two modalities, the proposed fusion model significantly improves classification performance, achieving an overall accuracy of 91% and outperforming both individual branches. The learned class-specific fusion weights further confirm that the model adaptively assigns higher importance to the most informative modality for each roof material class.

References

Arzoumanidis, L., Nguyen, S. H., Johannsen, L., Rothaut, F., Li, W., & Dehbi, Y. (2025). Object Detection for the Enrichment of Semantic 3D City Models with Roofing Materials. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *X-4/W6-2025*, 9–16. https://doi.org/10.5194/isprs-annals-x-4-w6-2025-9-2025.

Assouline, D., Mohajeri, N., & Scartezzini, J.-L. (2017). *Building rooftop classification using random forests for large-scale PV deployment*. https://doi.org/10.1117/12.2277692.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. *2009 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848.

Hoffmann, E. J., Wang, Y., Werner, M., Kang, J., & Zhu, X. X. (2019). Model Fusion for Building Type Classification from Aerial and Street View Images. *Remote Sensing 2019, Vol. 11, Page 1259*, *11*(11), 1259. https://doi.org/10.3390/RS11111259.

Katartzis, A., & Sahli, H. (2008). A stochastic framework for the identification of building rooftops using a single remote sensing image. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(1), 259–271. https://doi.org/10.1109/TGRS.2007.904953.

Kim, J., Bae, H., Kang, H., Lee, S. G., Kim, J. ;, Bae, H. ;, Kang, H. ;, Kwan, C., & Lee, D. J. (2021). CNN Algorithm for Roof Detection and Material Classification in Satellite Images. *Electronics 2021, Vol. 10, Page 1592*, *10*(13), 1592. https://doi.org/10.3390/ELECTRONICS10131592.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W. Y., Dollár, P., & Girshick, R. (2023). Segment Anything. *Proceedings of the IEEE International Conference on Computer Vision*, 3992–4003. https://doi.org/10.1109/ICCV51070.2023.00371.

Law, N. T., & Miura, Y. (2025). *RoofNet: A Global Multimodal Dataset for Roof Material Classification*. https://arxiv.org/pdf/2505.19358v1

Li, Q., & Yang, X. (2025). *Identification of Roof Materials in Urban Hot Spots From Remote Sensing Imagery Using Deep Learning Methods*. 3659–3662. https://doi.org/10.1109/igarss55030.2025.11243545.

Mantas, D., Gao, W., & Ledoux, H. (2025). RoofSense: A Multimodal Semantic Segmentation Dataset for Roofing Material Classification. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *X-4-W6-2025*, 153–160. https://doi.org/10.5194/ISPRS-ANNALS-X-4-W6-2025-153-2025.

Soleimani Vostikolaei, F., & Jabari, S. (2023). Large-Scale LoD2 Building Modeling using Deep Multimodal Feature Fusion. *Canadian Journal of Remote Sensing*, *49*(1). https://doi.org/10.1080/07038992.2023.2236243.

Soleimani Vostikolaei, F., & Jabari, S. (2025). Roof Geometrical Component Extraction Using Bimodal Data and Graph Neural Network. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLVIII-G-2025*, 1353–1358. https://doi.org/10.5194/ISPRS-ARCHIVES-XLVIII-G-2025-1353-2025.

Solovyev, R. A. (2020). Roof material classification from aerial imagery. *Optical Memory and Neural Networks (Information Optics)*, *29*(3), 198–208. https://doi.org/10.3103/S1060992X20030133.

Srivastava, S., Vargas-Muñoz, J. E., & Tuia, D. (2019). Understanding urban landuse from the above and ground perspectives: A deep learning, multimodal solution. *Remote Sensing of Environment*, *228*, 129–143. https://doi.org/10.1016/J.RSE.2019.04.014.

Sun, K., Li, Q., Liu, Q., Song, J., Dai, M., Qian, X., Gummidi, S. R. B., Yu, B., Creutzig, F., & Liu, G. (2025). Urban fabric decoded: High-precision building material identification via deep learning and remote sensing. *Environmental Science and Ecotechnology* , *24*. https://doi.org/10.1016/j.ese.2025.100538.

Tavakoligargari, M., Ghasemzadeh, M., & Arefi, H. (2025, August). *Classification of Building Types Using High-Resolution Images and Deep Learning Models*.

Tingzon, I., Cowan, N. M., & Chrzanowski, P. (2023). Mapping Housing Stock Characteristics from Drone Images for Climate Resilience in the Caribbean. *Environmental Data Science*, *3*. https://doi.org/10.1017/eds.2024.46.

Ujang, U., Azri, S., Zahir, M., Abdul Rahman, A., & Choon, T. L. (2018). Urban heat island micro-mapping via 3D city model. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(4/W10),

201–207. https://doi.org/10.5194/ISPRS-ARCHIVES-XLII-4-W10-201-2018.

Vostikolaei, F. S., & Jabari, S. (2025). Automated LoD2 Building Reconstruction Using Bimodal Segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://dataverse.lib.unb.ca/dataset.

Wang, X., & Aitchison, L. (2025). *How to set AdamW's weight decay as you scale model and dataset size*. http://arxiv.org/abs/2405.13698.

Wang, Y., Wang, Z. H., Rahmatollahi, N., & Hou, H. (2024). The impact of roof systems on cooling and building energy efficiency. *Applied Energy*, *376*, 124339. https://doi.org/10.1016/J.APENERGY.2024.124339.