Semantic Segmentation of Structural Thermal Bridges Via DTEP-Enhanced UNET++: A Comparative Study on V6 And V7 Architectures

Oğuzhan Hasan Özdemir[1], Eylül Açmaz[2], Önder Halis Bettemir[3], Faruk Ergen[4]

[1] İnönü University, Dept. of Computer Engineering, 44280 Battalgazi Malatya, Türkiye - 02220201557@ogr.inonu.edu.tr
[2] İnönü University, Dept. of Computer Engineering, 44280 Battalgazi Malatya, Türkiye - 02215076021@ogr.inonu.edu.tr
[3] İnönü University, Dept. of Civil Engineering, 44280 Battalgazi Malatya, Türkiye – onder.bettemir@inonu.edu.tr
[4] İnönü University, Dept. of Civil Engineering, 44280 Battalgazi Malatya, Türkiye - 36223621010@ogr.inonu.edu.tr

Keywords: Semantic Segmentation, Deep Learning, Thermal Bridges, DTEP Enhancer, Unet++, Energy Efficiency.

Abstract

Heat insulation performances of the existing building stock are known rarely since the attributes of the utilized insulation material was not documented during the construction stage. Moreover, strong winds and earthquakes can dislodge building roofs and cause openings at the window frame-wall junctions, increasing heat loss from the building. For these reasons, it is necessary to regularly measure the thermal insulation performances of existing buildings. Considering the number of existing buildings, automating the thermal insulation performance assessment will provide significant savings in terms of time and money. In this study, a framework system has been developed that measures the thermal insulation performance of buildings with minimal human intervention. In the first stage, building facade is recorded in video format with a thermal camera. Image frames with the least blur is obtained from the using the Variance of Laplacian method. To ensure accurate classification, entropy-weighted contrast enhancement is applied. To make the system's training data available, areas of heat loss were manually marked on the image. The resulting image dataset was divided into training and test data, and the YOLOv8 Instance Segmentation model was trained using these data. Then, heat loss was determined on other thermal images using the trained model. Case studies have shown that the developed model can successfully identify heat loss regions.

1. Introduction

Energy efficiency of existing buildings is critical for structural health and urban sustainability. This is because the operational environmental impact of a residence is highly depends on its heating energy (Bettemir and Erzurum, 2025). Thermal bridges are areas in the building envelope caused by insulation discontinuities in structural elements and are responsible for up to 30% of total energy loss. Following earthquakes and strong winds, the gaps between window frames and walls can increase, leading to deterioration in the building's thermal insulation performance. Additionally, the deterioration of material quality due to environmental conditions also reduces thermal insulation properties. Therefore, regular measurement of thermal insulation performance of buildings is necessary. Regularly measuring the thermal insulation performance of buildings in city centres requires significant labour and expense. Existing thermal inspection methods generally rely on operator experience and struggle to accurately detect boundaries in noisy thermal images. This study proposes a unique deep learning pipeline to isolate thermal bridges within their structural boundaries in low-contrast thermal data. The main focus of the research is to stabilize the noise accumulation in the basic model (v6) using physical-based pre-processing and temporal filtering methods in the proposed methodology.

2. Literature Review

Wardlaw et al. (2010) created 3D thermal images of the building using the structure-from-motion algorithm with thermal images obtained from different angles and positions, and manually identified areas causing heat loss. Mayer et al. (2021a) obtained thermal images with an unmanned aerial vehicle (UAV) to determine the thermal insulation performance of buildings constructed in Germany between 1950 and 1969 and to identify areas causing heat loss. By analysing the obtained images with ThermCad software, they manually determined heat losses originating from window sills, balcony floors, lintels, roof windows, window frames, floor slabs, ridges, and attics. Mayer et al. (2021b) identified heat losses occurring in building roofs using an ANN fed with a database containing images and 6895 pieces of information. 717 training and 200 test images were used to train the ANN, which was created with the Mask R-CNN architecture and classification was performed with an error rate of approximately 15%. Hou et al. (2021a) detected heat losses in buildings from optical thermal images obtained by UAV using deep convolutional neural network tools. Optical and thermal images were combined with the Pyramid Scene Parsing Network algorithm to identify 5 objects: roof, building surface, skylights, and cars. Heat loss regions were determined using an ANN with DeepLabv3 and Mask R-CNN architecture using 4190 training and 1000 test images. It was stated that the most important problem in the application was identifying the vehicles on the street as heat loss points.

Hou et al. (2021b) converted optical images into thermal images using a Generative Adversarial Network-based method. Optical images were converted to their thermal equivalents using an ANN fed with a training set of 14,000 images. Arjoune et al. (2021) developed a method to estimate the surface temperature and cumulative U-value of a building using two classification methods: K-means and threshold-based classification. The ANN was trained with 42, 439 images obtained manually by UAVs and selected windows, glass, walls, and trees. The heat transfer coefficients of the elements forming the building facade were estimated using empirical formulas. However, the estimated values were approximately four times the actual value for walls and half for windows. Pavlović and Barbarić (2013) performed wall emissivity measurements by supporting the thermal image with temperature gauge data using Thermography Studio software. Heat losses resulting from moisture, cracks, and inadequate thermal insulation were identified manually. O'Grady (2018) investigated the interaction of multiple thermal bridges in windows using thermal imaging. The success rate of identifying thermal bridges with known locations and sizes was measured by analysing them using the finite element method. Kim et al. (2021) classified heat losses using the density-based spatial clustering method of noisy applications. 151 thermal bridges and 223 cases were manually identified from 134 thermal images, and the ANN was trained. Accuracy rates ranged from 89% at best to 76% at worst in the case studies. Despotovic et al. (2019) estimated the heating energy requirement of a building using an ANN trained with optical images of the building and architectural elements and the building's age. 2898 images were used for training, 387 for validation, and 580 for testing to feed the ANN. Accuracy rates of around 60% were obtained in the analyses. Bettemir (2020) classified thermal heat losses of building façades by binarization. Bettemir (2024) compared the performances of 12 edge detection kernels to detect heat loss portions of building façades.

Zheng et al. (2022) detected heat losses by establishing the 3D infrared model of the building. Lu et al. (2022) implemented Canny edge detector to identify building portions that cause heat loss due to hallow parts on the masonry walls. Kim et al. (2022) trained a neural network having 1024 nodes and 5 layers to detect heat losses from buildings. The trained model was used to predict the heat losses in a simulation model and the prediction deviated from the true values by 0.09% and 0.18% in the average. Mayer et al. (2023) used Thermal Bridges on Building Rooftops (TBBRv2) dataset which consists of 926 images. MMDetection library provided the most successful thermal bridges detection with an Average Recall of 50.7%. The classification algorithm uses a pretrained Swin-T Transformer model which enhances the thermal bridge detection. Waqas and Araji (2024) implemented advanced deep learning by YOLOv7 to detect heat loss from building façades. The classification is improved by temperature measurements and the developed model utilizing YOLOv7 provided P (0.805), R (0.851), F1 − score (0.827), and mAP@0.5 of 0.770 at the end of the classification process. Hassan et al. (2025) proposed a thermal field detector based on multimodal Neural Radiance Fields which merges red-green-blue and thermal images. The developed model presented results with an average mean absolute error of 1.13 °C for buildings. Zhang et al. (2026) acquired thermal images of the buildings via UAV and analysed the images by DeepLabV3+ based deep learning model. The developed system enriched the Intersection over Union (IoU) for wall regions by 15.33%. Challa et al. (2025) combined thermal images acquired by UAV and indoor thermal images. The analyses were conducted on YOLOv4 to detect thermal anomalies and 92% mean average precision rate were obtained.

3. Methodology

Thermal cameras record heat as a digital signal but are affected by ambient noise and low contrast. Classical segmentation models usually fail to distinguish this noise as illustrated in the literature review. The noise and the usual segmentation model cause reporting the leakage area as much larger and more ambiguous than it actually is. The focus of this research is to mathematically isolate only the structural source of the leakage.

In this study a new method is developed to autonomously detect the heat insulation spots of the building façades. The Autonomous Thermal Diagnosis Platform is built on a hybrid architecture that combines building physics and computer vision algorithms. The implementation of the knowledge on the building elements present advantages compared with the standard object detection problem. While thermal analysis studies in the current literature are generally limited to static thresholding or multiple-layer artificial neural networks. This study presents a heat spot detection framework which consists of five different algorithmic layers take into account the physical and geometric characteristics of the thermal data.

3.1. Image Frame Extractor

Thermal images are usually acquired in video format. The motion of the image acquisition system can cause blur in the video frames. In order to extract quality frames, blur check operation is done by the Variance of Laplacian. The convolution kernel is given in Eq. 1 (Russ, 1998).

$$K=\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{1}$$

The kernel given in Eq. 1 is applied to the whole image. Then the Laplacian values are recorded in matrix and the variance of the matrix is computed. The computation procedure is implemented for all of the video frames and Variance of Laplacian values are computed for all of the frames of the thermal video. 500 video frames with the highest Variance of Laplacian values were selected for the analysis.

3.2. Dynamic Thermal Edge Pre-processor (DTEP):

Thermal images may have low contrast and due to the characteristics of thermal images instead of linear contrast enhancement, entropy-weighted contrast enhancement can provide better results. Entropy-weighted contrast enhancement method analyses the local entropy of the image. The formulation of the applied entropy-based method is given in Equation 2.

$$E(x) = x +(\sigma(x) *\alpha*G(x)) \tag{2}$$

In Eq. 2, x is the input thermal image whose edges are uncertain with low contrast. The input x can be considered as raw image. σ(x) is the local standard deviation which is computed for each picture element within a predefined neighbourhood. In Eq. 2 α is a contrast multiplier parameter which is a constant that is used to strengthen the edge data of the image. In this study α is taken as 1.5. G(x) is the entropy gate that represents the structural lines. It acts as a gateway, determining where the "information" is located in the image. In areas with low information levels, such as plain walls, it suppresses noise by taking values close to zero. However, in areas where there is heat loss and high variation in the image, it opens the gateway by taking higher values, enhancing the details in the image. E(x) is the enhanced output with enhanced structural lines.

3.3. Architectural Framework with Unet++ and EfficientNet-B4:

The main body of the model uses Unet++ built on the EfficientNet-B4 backbone. This architecture prevents the loss of fine structural detail in low-resolution data thanks to its densely packed network of interconnections.

The developed system is based on an advanced deep learning architecture that focuses not only on heat loss but also on the geometry of the building's exterior. EfficientNet-B4 Encoder is the main feature engine that extracts the building's structural features such as columns, beams and walls with high accuracy.

Unlike the standard UNet architecture Unet++, Nested Skip Connections, prevents the loss of fine structural details in low-resolution thermal data by using densely nested connections. The developed framework learns thermal bridges not just as a brightness value but as a geometric object that overlaps with the building's column and beam lines. The considered engineering pre-knowledge is expected to improve the accuracy of the output of the model.

3.4. YOLOv8-Seg Based Instance Segmentation and Data Generalization by Mosaic Augmentation

Semantic Segmentation models applied by traditional artificial intelligence models such as U-Net may perceive all heat leaks in the image as a single-spot or class. This makes it difficult to isolate and examine heat loss occurring in multiple different regions in the same frame. In addition, standard bounding box-based determinations may cause large margins of error in area calculation because they do not present the actual geometry of the leak.

In order to overcome the specified bottleneck, YOLOv8 Instance Segmentation architecture, which classifies pixels and assigns a unique tracking ID to each leak, was preferred. Instead of enclosing the leak in a box, the system draws its polygonal boundaries with pixel precision. In order for the model to adapt to different environmental conditions in the field and to prevent overfitting by memorizing the training data, the Mosaic Augmentation algorithm, in which four different thermal images are randomly cropped and combined, was applied to the data set created within the scope of this study. In this way, the model can accurately generalize the geometric patterns of heat leaks even in backgrounds it has never seen before.

3.5. Temporal Masking

Inspections based on thermal cameras are generally performed using handheld cameras or UAVs (Bettemir, 2024). In video recordings affected by vibration during imaging, the tracking IDs assigned by artificial intelligence are constantly interrupted or displaced. This leads to the precipitation of time series analyses. Gimbal use is also not always cost-effective or accessible. To solve the problem with a purely algorithmic approach, the Lucas-Kanade Optical Flow algorithm has been integrated into the developed system. The embedded module calculates the motion vectors of specific corner pixels of a feature in consecutive video frames along the x and y axes. An Affine Transform matrix is created from these vectors, and the shaking in the frame is compensated in the opposite direction. Consequently, a software-based "Digital Gimbal" has been created to maximize the tracking stability of the artificial intelligence.

This module is an acknowledgment mechanism that eliminates momentary flickering errors caused by camera shake or sensor noise in the video stream. Decision criteria are formulated in Eq. 3.

$$M_{robust} = \begin{cases} 1 & if \sum_{i=t-12}^{t} M_i \geq 9 \\ 0 & otherwise \end{cases} \tag{3}$$

In Eq. 3, M_{robust} is the Robust Mask that enhances the image which becomes smooth and stable. M_i is the binary mask at the i^{th} frame. This is an independent estimate that the proposed model produces for each frame. t represents the current time which is the analysed frame at the corresponding time. In Eq. 3, 12 is the buffer size which is decided to be half a second for 24 fps thermal video records. The developed system remembers 12 frames for each analysis. The constant value 9 is a threshold value that should be spotted at least 9 times inside the regions which are classified as real thermal leak zone within the most recent 12 frames.

3.6. Surgical Threshold

Surgical thresholding is a critical engineering decision that facilitates finding the source of heat loss rather than simply determining its spread. The formula of surgical threshold is given in Eq 4.

$$T_{surgical} = \max\left(\max(P) \cdot 0.95, 0.92\right) \tag{4}$$

In Eq. 4, $T_{surgical}$ is a dynamic threshold value, minimum acceptance criteria, which is computed particularly for each frame. P is the probability map which is a confidence map that assigns a probability value for each pixel for being a thermal energy leak. max(P) provides the peak confidence which is the highest confidence score of the analysed frame. The constant value 0.95 is an adaptive factor which takes into account the 95% of the peak score which leads to the selection of only dense thermal leak points. The constant 0.92 is a hard limit value which prevents false detection of thermal leak spots within weak leakage zones.

Deep learning models produce a Confidence Score for the objects they detect. In this approach, the model builds a statistical model on the pixel values and calculates the likelihood of heat loss based on the model it creates. However, the calculated statistical ratio does not express how severe the leakage is. The desired inference after classification is to convert the mask formed by the pixels into a ratio that expresses the extent to which it carries a structural or operational risk. Traditional approaches statically control pixel temperatures. However, static thresholds collapse when the ambient temperature changes. To address this deficiency, the Thermodynamic Heat Loss Index (THLI) is proposed in this study. THLI simultaneously takes into account the spatial size of the applied mask (pixel area) and the grayscale intensity (heat intensity) of the pixels under this mask. The THLI equation is presented in Eq. 5.

$$THLI = MA * \frac{AVD}{255} \lambda \tag{5}$$

In Eq. 5, MA is Mask Area, AVD is average pixel density, λ is taken as the calibration coefficient 0.1. The coefficient 255 normalizes the brightness values of the thermal image pixels. THLI not only labels the detected heat loss but also autonomously categorizes it into "Weak", "Medium" or "Critical" risk levels. The categorization can also be used for the determination of intervention priority.

3.7. Time Series Field Derivative:

Thermodynamic processes can be time-dependent (dt) if they are occurring because of a malfunctioning of the utility system of the building. By analysing a single photograph or a static video frame, it is not possible to determine whether a heat spot on a wall is due to an insulation fault that has been there for years or a recent plumbing failure. While current systems perform Descriptive detection, this study aims to develop a decision support system that can also perform Diagnostic and Predictive detection.

A memory queue of length 8 frames is defined for each unique leak ID in the system. The system saves the THLI mask field in this queue for each frame. The time-dependent field change rate is calculated by taking the difference between the current field and the oldest field in memory as presented in equation 6:

$$\frac{\Delta A}{\Delta t} = A_{current} - A_{t-8} \tag{6}$$

If the detected expansion exceeds the determined threshold value by showing a critical amount of pixel expansion in a short time, the developed decision support system removes the heat loss from the static state and defines it as a dynamic state. The threshold value includes the growth of the mask area by 5% or more in the image containing 8 frames. If the threshold value is exceeded, the mask, which is defined as heat loss, is labelled as on-going spread. The aim of the proposed approach is to enable the autonomous detection of utility failures.

4. Case Study

The proposed framework method is tested by the thermal images taken by FLIR thermal camera inside the campus of İnönü University, Malatya, Türkiye. Initially thermal images were acquired and they were classified as test and train images. The samples of test images were given in Figure 1.

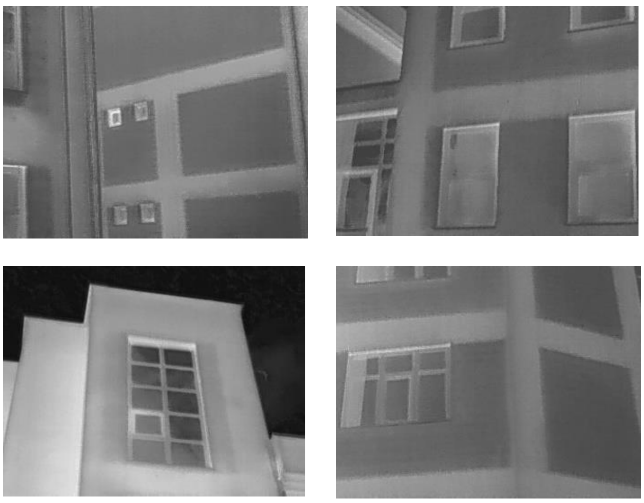


Figure 1.Sample images randomly taken from the collected 203 train images.

Samples of train images are presented in Figure 2.

Figure 2. Sample images randomly taken from the collected 51 train images.

The training process of the proposed model is conducted and the convergence curves are presented in Figure 3. In machine learning, and object detection models, the F1 confidence curve which is a graph showing how the F1 score of a model changes as the confidence threshold changes, is presented in Figure 3.a. Examining the curve, it can be seen that the F1 value is highest in the confidence range of 0.1 to 0.4. Besides the given interval the F1 score does not drop significantly for the confidence values less than 0.5.

Figure 3b presents the precision confidence curve. The Precision-Confidence Curve is a graphical tool used in machine learning and object detection models to show how the precision score of a model changes as the confidence threshold changes. It is generally known that confidence increases as the confidence level increases. However, increasing the confidence threshold too much will cause the model to only provide predictions that it is "very confident" in, thus avoiding false alarms. The price of this improvement is that areas of heat loss will be overlooked to avoid making incorrect predictions. As can be seen in Figure 3.a, it is a good choice not to set the confidence level higher than 0.5 in order to keep the F1 value high.

Figure 3.c presents the Precision-Recall Curve. This curve shows the trade-off between precision and recall values in a classification or object recognition model at different confidence thresholds. Figure 3.c demonstrates that when the recall rate is small the loss of precision is negligible. However, the rate of precision loss increases significantly when the recall rate exceeds 0.7. The graph illustrates that the benefit obtained by increasing the recall after 0.7 causes important losses in precision. If the recall is increased further it will detect every heat loss portion with many false alarms. The model suggests taking the recall rate as 0.781.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-L-4-W3-2026-123-2026

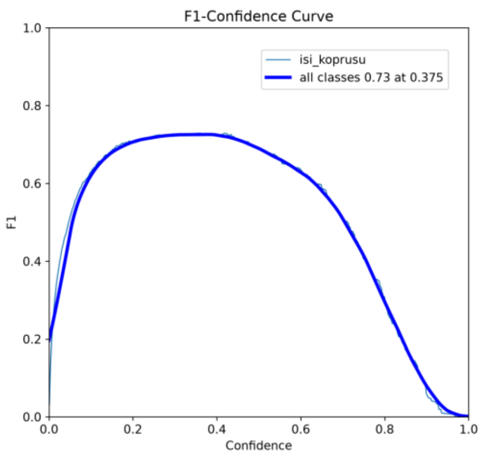


Figure 3a. F1-confidence curve.

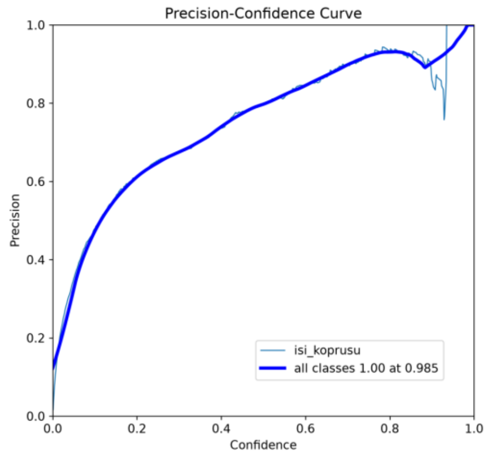


Figure 3.b.Precision versus confidence curve.

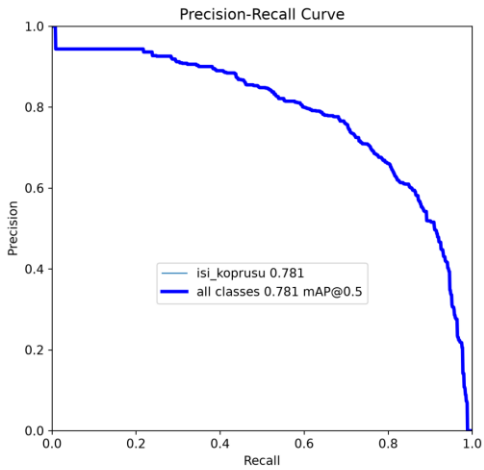


Figure 3.c. Precision-Recall versus recall curve.

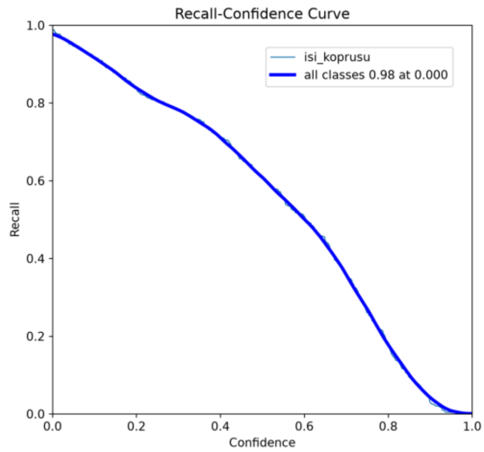


Figure 3.d. Recall versus Confidence Curve.

In Figure 3.d, Recall-Confidence Curve is presented. This curve shows how the recall score of a machine learning model changes as the confidence threshold changes. The curve demonstrates that if the confidence is increased the recall rate decreases. The curve is almost linear with its maximum gradient around 0.6 confidence value. This curve represents the trade-off between the false detection of an object as a heat loss portion and not detecting a heat loss portion. When all of the graphs are considered the confidence rate can be taken as 0.5.

Figure 4 presents the Normalized Confusion Matrix. This matrix is a performance evaluation table which shows the prediction results of a classification model in terms of ratios between 0 and 1. The confusion matrix demonstrates that all of the classified heat loss portions are correct. However, some of the heat loss portions were missed by the model as they were classified as background.

Similarly, none of the background object was detected as heat loss portion by the model as Figure 4 illustrates. Some of the building portions with heat loss were classified as background object. The normalized confusion matrix provides important information about the accuracy of the trained model.

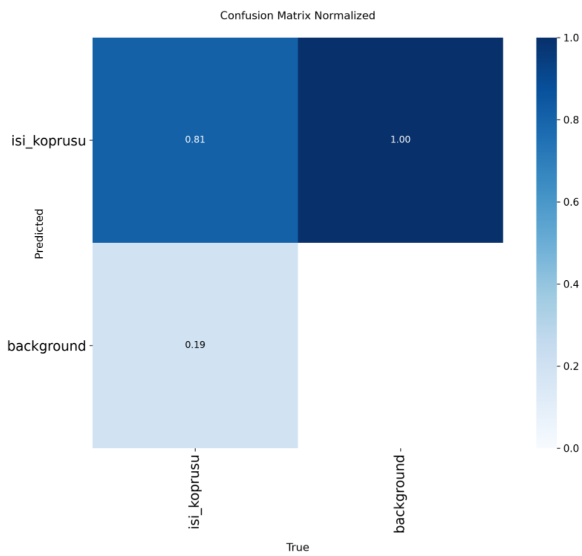
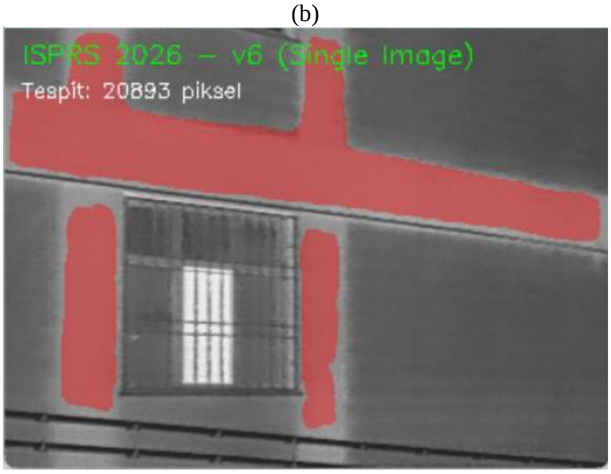


Figure 4. Normalized confusion matrix

The trained system is used for the detection of the thermal losses. A sample image thermal image and the detected heat-loss portions are shown in Figure 5a-c.

(a)

(b)



(c)

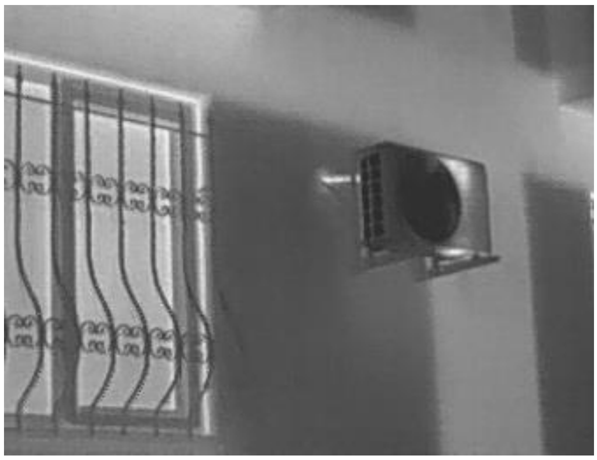
Figure 5. Thermal image and the resultant analysed images.

In Figure 5.a a thermal image extracted from the video captured by FLIR thermal camera of a building at the campus of İnönü University is shown. The poorly insulated reinforced concrete columns and beams are clearly shown in the image with bright pixels. In Figure 5.b the classification conducted by conventional method is shown. It is seen that the beams and columns were failed to be classified as heat loss building portion. In Figure 5.c the classification executed by the proposed algorithm is shown. It is seen that the beams and columns were successfully classified as heat loss portions.

Detecting windows with wrought iron fronts as a whole in thermal imaging is a significant challenge. In this study, such a thermal image is presented in Figure 6. The developed algorithm is shown to be more successful in thermal classification.

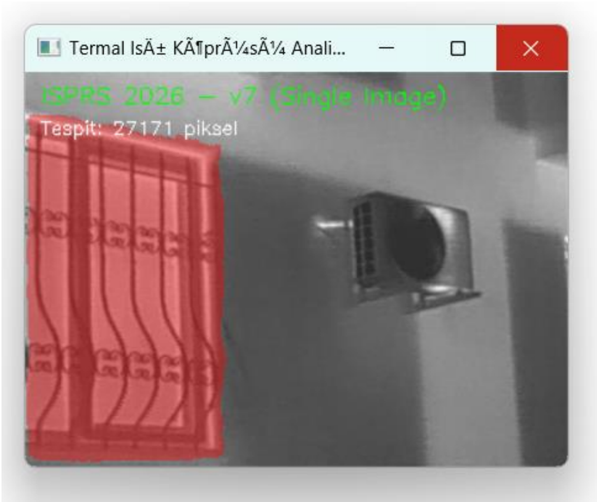
Figure 6.a illustrates the thermal image extracted from the thermal video. The thermal loss classification conducted by conventional algorithm is shown in Figure 9.6. The classification of the window by the proposed algorithm conducted by the Unet++ architecture is shown in Figure 6.c. The developed algorithm separated the edge information of window and railing boundaries from the thermal leakage. With a threshold of 0.92, the model only reported the core line, which is the source of the leakage.

The borders of the heat loss region classified by the conventional method is fluctuating which can be seen from Figure 6.b. The proposed algorithm successfully detected the heat loss region with sharper borders. Moreover, the original thermal image covers the ground floor of the building and for security reasons the windows were covered by perforated steel bars. The image was taken at night and the steel bars were cold enough to have dark pixel values while the window is bright as the heat dissipated from the glass and the window frame. The proposed classification algorithm successfully eliminated the curtaining perforate bars and provided an unfragmented heat loss classification.



(a)

(b)



(c)

Figure 9. Classification of heat loss regions from another thermal image.

5. Discussion of Results

The developed framework successfully detected thermal losses that occur at the façades of buildings. The developed system has several advantages when its facilitation of image acquisition and data analysis were concerned.

The developed system extracted the video frames with the least blur which is detected by Variance of Laplacian. Then the acquired thermal images are manually examined and the heat loss regions were marked. The marked images were binarized and they are randomly classified as train and test images. After the training process another set of thermal images were analysed. The developed system provided successful results.

The frame extraction from video offers conveniences to the users as the user can walk around the building or utilize a UAV and the developed system acquires the most quality video frames for the analysis. This shortens the data acquisition process which is conducted under severe winter conditions.

The test and train image set consists of 254 thermal images which is a relatively small set. Moreover, the set consists of only the buildings inside the İnönü University campus. The number of images and the variety of images should be increased to ensure the robustness of the detection of the thermal leakages.

The conducted case study demonstrated that the proposed algorithm successfully detects the heat bridges. However, when Figure 8.c is examined it is seen that the classified columns and beams are not continuous and some heat loss region were classified as non-heat loss region. This defect of the proposed algorithm can be removed by robust morphological operators that can fill the missing portions with a geometry based decision algorithms as a future study

6. Conclusion

The proposed heat loss detection algorithm is promising with its overall high accuracy and low computational demand. Besides it provides convenience during the data acquisition. Moreover, it does not need any training task that should be conducted with hundreds of thermal images. The proposed algorithm has the potential of reducing the required endeavour for the building inspection process.

The developed algorithm should be improved by enlarging the dataset in terms of number of images and the variety of the images. Complete detection of heat loss building portions can be warranted by inserting morphologic operators to the framework.

Acknowledgement

This study is supported by the Coordination of the Scientific Research Projects of İnönü University with the Grant number FDK-2025-4224.

References

Arjoune, Y., Peri, S., Sugunaraj, N., Biswas, A., Sadhukhan, D., and Ranganathan, P., 2021. An Instance Segmentation and Clustering Model for Energy Audit Assessments in Built Environments: A Multi-Stage Approach, *Sensors*, 21(13), 4375.

Bettemir, O. (2020). Bazi Yerel Benirizasyon Yöntemleri Ile Binalarin Isı Kaybina Yol Açan Kisimlarin Belirlenmesi. Computer Science, 5(1), 22-30.

Bettemir, Ö. H. (2024). Kızıl Ötesi Görüntülerden Binalardaki Isı Köprüsünün Belirlenmesi. Politeknik Dergisi, 27(3), 887-920.

Bettemir, Ö. H., & Erzurum, T. (2025). Comparative life cycle assessment of post-disaster housing according to ground type-outdoor temperature-lifespan relationship: Carbon, energy and cost. *Journal of Building Engineering*, 111, 113388.

Challa, K., AlHmoud, I. W., Islam, A. K., Graves, C. A., & Gokaraju, B. (2025). Comprehensive energy efficiency analysis in buildings using drone thermal imagery, real-time indoor monitoring, and deep learning techniques. IEEE Access, 13, 65094-65104.

Despotovic, M., Koch, D., Leiber, S., Doeller, M., Sakeena, M., and Zeppelzauer, M., 2019. Prediction and analysis of heating energy demand for detached houses by computer vision. *Energy and Buildings*, 193, 29-35.

Hassan, M., Forest, F., Fink, O., & Mielle, M. (2025). ThermoNeRF: A multimodal Neural Radiance Field for joint RGB-thermal novel view synthesis of building facades. Advanced Engineering Informatics, 65, 103345.

Hou, Y., Chen, M., Volk, R., and Soibelman, L., 2021a. An approach to semantically segmenting building components and outdoor scenes based on multichannel aerial imagery datasets. *Remote Sensing*, 13(21), 4357.

Hou, Y., Volk, R., and Soibelman, L., 2021b. A novel building temperature simulation approach driven by expanding semantic segmentation training datasets with synthetic aerial thermal images. *Energies*, 14(2), 353.

Kim, C., Choi, J. S., Jang, H., and Kim, E. J., 2021. Automatic detection of linear thermal bridges from infrared thermal images using neural network. *Applied Sciences*, 11(3), 931.

Kim, D. J., Kim, S. I., & Kim, H. S. (2022). Thermal simulation trained deep neural networks for fast and accurate prediction of thermal distribution and heat losses of building structures. Applied Thermal Engineering, 202, 117908.

Lu, Y., Duanmu, L., Zhai, Z. J., & Wang, Z. (2022). Application and improvement of Canny edge-detection algorithm for exterior wall hollowing detection using infrared thermal images. Energy and Buildings, 274, 112421.

Mayer, Z., Heuer, J., Volk, R., and Schultmann, F., 2021a. Aerial thermographic image-based assessment of thermal bridges using representative classifications and calculations. *Energies*, 14(21), 7360.

Mayer, Z., Kahn, J., Hou, Y., ve Volk, R., 2021b. 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*, Berlin, Germany, Vol. 30.

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.

O'Grady, M., Lechowska, A.A., and Harte, A.M. 2018 Application of infrared thermography technique to the thermal assessment of multiple thermal bridges and Windows. *Energy and Buildings*, 168, 347-362.

Pavlović, A., and Barbarić, Ž., 2013. Application of G100/120 thermal imaging camera in energy efficiency measuring in building construction. *Serbian Journal of Electrical Engineering*, 10(1), 153-164.

Russ, J. C. (1998). The image processing handbook, 3rd Edition, CTC Press and IEEE Press .

Waqas, A., & Araji, M. T. (2024). Machine learning-aided thermography for autonomous heat loss detection in buildings. Energy Conversion and Management, 304, 118243.

Wardlaw J., Gryka M., Wanner F., Brostow G., and Kautz J. 2010. A new approach to thermal imaging visualisation. EngD Group Project, University College London.

Zhang, D., Mei, Y., Wang, X., Wang, Z., & Cheng, Y. (2026). Heating building external wall heat loss measurement method based on UAV-IRT. Journal of Building Engineering, 115272.

Zheng, H., Gao, G., Zhong, X., & Zhao, L. (2022). Monitoring and diagnostics of buildings' heat loss based on 3D IR model of multiple buildings. Energy and buildings, 259, 111889.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-L-4-W3-2026-123-2026