CETRA2026 27-30 May 2026, Bol, Island Brač, Croatia
9th International Conference on Road and Rail Infrastructure

A MULTI-DAY THERMAL SIGNATURE AND MACHINE LEARNING FRAMEWORK FOR PREDICTING DEFECT DETECTABILITY IN PASSIVE INFRARED THERMOGRAPHY OF REINFORCED CONCRETE SLABS

Aseel Abdulazez Abdulridha[1], Ninel Alver[1], Kazuma Shibano[2], Tetsuya Suzuki[3]
[1]Ege University, Department of Civil Engineering, Turkey
[2]Niigata University, Graduate School of Science and Technology, Japan
[3]Niigata University, Institute of Agriculture, Japan

Abstract

Passive infrared thermography is widely used for non-destructive evaluation of concrete structures; however, its reliability under real outdoor conditions is limited by environmental variability and the transient nature of defect visibility. As a result, image-based interpretation often leads to inconsistent assessments of subsurface damage. This study presents a time-resolved, weather-aware framework to quantify the detectability of defects in passive IRT using multi-day surface temperature monitoring. A full-scale concrete slab including artificial delaminations at depths of 50, 100, and 250 mm was tested outdoors over seven daytime periods with one-minute temporal resolution. A physics-informed Random Forest regressor was used to model and remove the Weather-driven baseline surface temperature to isolate residual thermal responses. An objective, sound-referenced detectability score was defined, enabling threshold-based detection without manual labeling. Persistence-based metrics, including duty cycle and detectability windows, and supported by Probability-of-Detection (PoD) analysis, were used to evaluate the detectability. Results demonstrate that passive IRT detectability is fundamentally time-dependent; all delaminations at 250 mm depth, as well as smaller delaminations at shallower depths, remained non-detectable and provided a stable intact-reference group. These findings highlight the necessity of temporal reliability metrics for defensible interpretation of passive thermographic monitoring.

Keywords: passive infrared thermography, concrete delamination detection, weather-aware detectability, probability of detection (PoD), outdoor non-destructive testing

1 Introduction

Infrared thermography (IRT) has obtained widespread interest for concrete monitoring due to its non-contact processing, rapid data acquisition, and ability to inspect large surface areas. The principle of IRT is that subsurface anomalies disturb internal heat transfer, producing surface temperature variations that can be captured using thermal imaging systems [1, 2]. Passive IRT depends exclusively on natural thermal excitation, primarily solar radiation and ambient temperature fluctuations, which make it attractive for in-situ inspection of large concrete elements where active heating is impractical [3]. Field studies have indicated that passive IRT can successfully detect the delaminations in specific depths under convenient environmental conditions, especially during strong solar loading [4, 5]. Despite these advantages, defect detectability in passive IRT is highly affected by environmental conditions.

DOI: https://doi.org/10.5592/CO/CETRA.2026.1957

Variations in solar irradiance, ambient air temperature, wind speed, and surface moisture can significantly change thermal contrast, resulting in inconsistent defect visibility across inspection times and days [6]. Thus, shallow and relatively large defects are generally detectable, while deeper or smaller defects often remain invisible even when present [7, 8]. Furthermore, long-term outdoor monitoring demonstrates that defect-related thermal signatures may appear only within narrow temporal windows and disappear entirely under unfavorable conditions [9, 10]. These observations indicate that image-based thermographic inspection provides limited reliability for passive IRT. Recent studies have explored machine-learning techniques to improve the objectivity of thermographic interpretation, including deep learning–based image segmentation and feature-driven classifiers [11]. While promising, many of these approaches treat defect detectability as a static image classification problem and report reduced performance when environmental variability or increased defect depth is introduced [12, 13]. Feature-based machine learning using physically meaningful thermal descriptors has been shown to provide improved robustness and interpretability under noisy conditions [14], yet the temporal persistence of detectability remains largely unaddressed. Overall, existing studies indicate that passive IRT detectability is inherently time-dependent and environmentally conditioned, while most current approaches lack objective, persistence-based criteria for assessing inspection reliability. These gaps motivate the present study, which proposes a time-resolved, weather-aware framework for objectively quantifying subsurface defect detectability in concrete structures using passive infrared thermography under real outdoor conditions.

2 Experimental program

2.1 Test specimen and defect configuration

A full-scale reinforced concrete slab with dimensions of approximately 1500 × 1500 × 300mm was constructed using self-compacting concrete to ensure uniform material quality and minimize casting-induced variability. Nine artificial delamination-type defects were introduced using square Styrofoam sheets embedded at three depths measured from the exposed surface: 50 mm (P1–P3), 100 mm (P4–P6), and 250 mm (P7–P9). Defect sizes ranged from 100 × 100 mm to 300 × 300 mm. Figure 1 presents the geometric layout and cross-sectional configuration of the embedded defects. Based on experimental observations, defects P1 (50 mm, 200 × 200 mm), P2 (50 mm, 300 × 300 mm), and P4 (100 mm, 300 × 300 mm) exhibited detectable thermal behavior, with P4 detectable only intermittently. All remaining locations served as sound reference points and were never observed as detectable during the monitoring campaign.

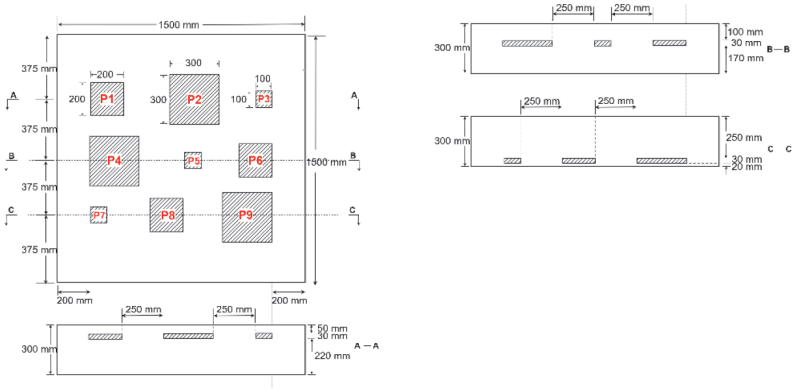

Figure 1 Geometric layout and cross-sectional views of embedded defects

2.2 Passive infrared thermography monitoring

Following curing, the slab was placed outdoors and monitored using passive infrared thermography under natural environmental excitation, including solar radiation, ambient temperature fluctuations, wind, and convective heat exchange. No artificial heating was used. A calibrated FLIR infrared camera was positioned normal to the specimen surface at a fixed distance to maintain consistent spatial resolution, as shown in Figure 2. Over seven separate days in November, thermal images were captured at one-minute intervals during a fixed daily monitoring window from 09:00 to 17:00. Radiometric thermal images were converted to surface temperature values (°C) and synchronized with meteorological data, including air temperature, relative humidity, dew point, wind speed, and categorical weather conditions.

Figure 2 Experimental setup and thermal data acquisition

3 Methodology

3.1 Data processing and time-series extraction

Orthogonal projection and geometric alignment were applied for raw thermal images to ensure spatial consistency across time. After that, to enable quantitative analysis, the radiometric grayscale values of thermal images were converted to surface temperature values (°C) using a camera-specific grayscale–temperature conversion function. Fixed surface locations corresponding to predefined inspection points (P1–P9) were then identified, and one-minute surface temperature time series were extracted for each point. Figure 3 shows the overall processing pipeline adopted in this study.

3.2 Physics-informed feature representation

Passive IRT under natural exposure exhibits strong diurnal behavior driven by solar heating and ambient temperature cycles. To preserve the periodic nature of time-of-day, cyclic sine and cosine encoding was used instead of linear time representation [15]. Lagged surface temperature values at 10, 30, and 60 minutes were used to represent the thermal inertia of the concrete surface, reflecting the transient heat conduction and storage effects in concrete [16]. Environmental forcing variables collectively capture the dominant boundary conditions, and these in turn govern surface heat transfer.

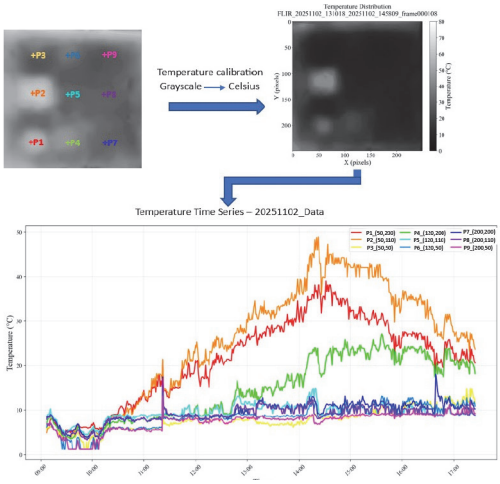

Figure 3 Thermal image calibration and time-series extraction pipeline

3.3 Weather-driven baseline modeling and residual response

To explicitly split global environmental forcing from localized defect effects, the measured surface temperature was decomposed into a baseline component and a residual anomaly component. A Random Forest regression model was trained to predict the expected baseline surface temperature using environmental variables, diurnal phase features, and thermal inertia terms [17]. To evaluate the model generalization, leave-one-day-out cross-validation was used to ensure that each test day was completely unseen during training. The defect-driven residual thermal response was calculated as the difference between the measured surface temperature and the predicted baseline temperature.

3.4 Objective detectability score

Residual signals at intact reference points were used to define a noise floor representative of non-defective behavior. At each time step, an intact-reference residual was computed as the average residual across all non-detectable points. A robust noise scale was estimated using the median absolute deviation (MAD) to mitigate the influence of outliers caused by transient environmental effects [18]. An objective detectability score was defined as the normalized absolute deviation of each point's residual from the intact reference. A conservative detection threshold was selected from the upper tail of the intact-score distribution, enabling binary detectability events to be identified without manual labeling.

3.5 Time-dependent detectability metrics

Detectability was treated as a time-dependent process rather than a single-image outcome. Contiguous intervals during which the detectability score exceeded the threshold were identified as detectability windows. For each defect and each day, the detectability duty cycle was computed as the fraction of the monitoring window during which the defect was detectable. These persistence-based metrics provide a quantitative measure of inspection reliability under passive IRT conditions.

4 Results

4.1 Baseline model performance

The weather-driven baseline model achieved high predictive accuracy across most monitoring days, with coefficients of determination exceeding 0.9 under dry conditions. Figure 4 presents the cross-day baseline regression performance evaluated using leave-one-day-out validation. Reduced performance was observed during rainy periods, reflecting physically plausible boundary-condition changes associated with surface wetting and evaporative cooling.

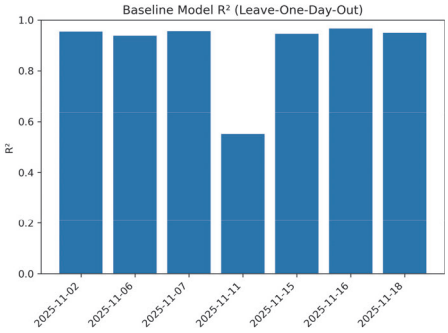

Figure 4 Cross-day generalization performance of the weather-driven baseline model evaluated using leave-one-day-out validation

4.2 Residual separation between defect and sound regions

Figure 5 compares the probability density distributions of baseline-removed residuals for sound reference points and detectable defect locations. Sound regions exhibit narrow, near-zero-centered residual distributions, whereas detectable defects exhibit broader distributions with larger amplitudes. This clear separation confirms that baseline removal effectively suppresses global environmental coupling while preserving localized defect-driven thermal behavior.

4.3 Time-dependent detectability and duty cycle

Figure 6 presents the detectability duty cycle for defects P1, P2, and P4 across the seven monitoring days. Shallow defects at 50 mm depth (P1 and P2) exhibit moderate to high detectability persistence, whereas the deeper defect at 100 mm depth (P4) exhibits substantially lower and more variable duty cycles. Sound reference points remain near the designed false-alarm level, confirming stable threshold behavior.

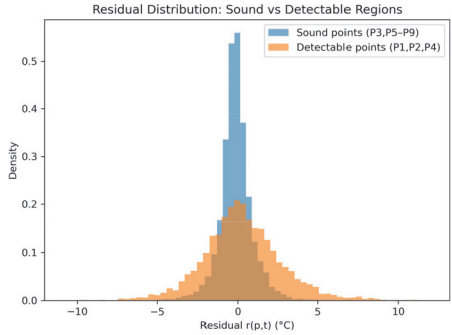

Figure 5 Probability density distributions of baseline-removed thermal residuals for sound reference points (P3, P5–P9) and detectable defect points (P1, P2, P4)

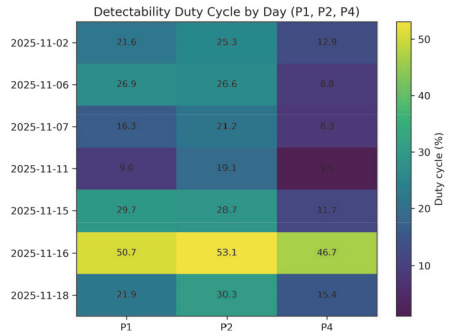

Figure 6 Day-wise detectability duty cycle (%) for defects P1, P2 (50 mm depth), and P4 (100 mm depth) during the daytime monitoring window (09:00–17:00)

4.4 Probability-of-detection analysis

Figure 7 presents Probability-of-Detection (PoD) curves as a function of the detectability threshold. Shallow defects maintain higher PoD across a wide threshold range, while the deeper defect exhibits reduced and more threshold-sensitive PoD. Sound regions remain near the target false-alarm rate, demonstrating the robustness of the detectability formulation.

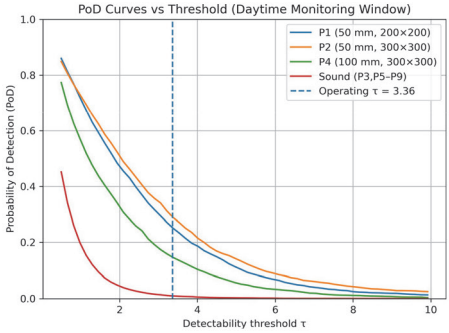

Figure 7 Probability-of-Detection (PoD) curves as a function of the detectability threshold for defects P1, P2 (50 mm depth), P4 (100 mm depth), and sound reference points

5 Conclusion

This study presented a time-dependent framework for assessing subsurface defect detectability in concrete using passive infrared thermography under real outdoor conditions. Based on seven days of high-resolution monitoring, the following conclusions are drawn:

- passive IRT detectability is fundamentally time-dependent rather than static
- weather-driven baseline removal is fundamental for separating defect-related thermal behavior
- persistence-based metrics provide meaningful measures of inspection reliability
- shallow delaminations exhibit consistent detectability, while deeper defects are intermittently observable and environmentally affected
- snapshot thermograms alone are insufficient for reliable defect inspection under passive conditions.

Acknowledgement

This study is part of a joint JSPS-TUBITAK project. The authors acknowledge the support of TUBITAK under grant number 223N061 and JSPS under grant number JPJSBP120249402.

References

[1] Vavilov, V.P., Burleigh, D.D.: Review of pulsed thermal NDT: Physical principles, theory and data processing, NDT & E International, 73 (2015), pp. 28–52

[2] Tomita, K., Chew, M.Y.L.: A review of infrared thermography for delamination detection on infrastructures and buildings, Sensors, 22 (2022) 2, 423

[3] Omar, T., Nehdi, M.L., Zayed, T.: Infrared thermography model for automated detection of delamination in RC bridge decks, Construction and Building Materials, 168 (2018), pp. 313–327

[4] Hiasa, S., Birgul, R., Catbas, F.N.: Effect of defect size on subsurface defect detectability and defect depth estimation for concrete structures by infrared thermography, Journal of Nondestructive Evaluation, 36 (2017) 3, pp. 1–21

[5] Pozzer, S., Omidi, Z., El Refai, A., López, F., Ibarra-Castanedo, C., Maldague, X.: Passive infrared thermography for subsurface delamination detection in concrete infrastructure: Capabilities, Construction and Building Materials, 419 (2024), 135542

[6] Rathod, H., Gupta, R.: Sub-surface simulated damage detection using non-destructive testing techniques in reinforced-concrete slabs, Construction and Building Materials, 215 (2019), pp. 754–764

[7] Al Gharawi, M., Adu-Gyamfi, Y., Washer, G.: A framework for automated time-lapse thermography data processing, Construction and Building Materials, 227 (2019), 116507

[8] Cheng, C., Na, R., Shen, Z.: Thermographic laplacian-pyramid filtering to enhance delamination detection in concrete structure, Infrared Physics & Technology, 97 (2019), pp. 162–176

[9] Gu, J., Unjoh, S.: Image processing methodology for detecting delaminations using infrared thermography in CFRP-jacketed concrete members, Composite Structures, 270 (2021), 114040

[10] Pozzer, S., Vieira De Souza, M.P., Hena, B., Hesam, S., Rezayiye, R.K., Rezazadeh Azar, E., López, F., Maldague, X.: Effect of different imaging modalities on the performance of a CNN: An experimental study on damage segmentation in infrared, visible, and fused images of concrete structures, NDT & E International, 132 (2022), 102709

[11] Abdulridha, A.A., Alzubaidi, L., Alver, N.: Integration of artificial intelligence techniques with infrared thermography for defect detection in concrete structures: A systematic review, Measurement, 267 (2026), 120491

[12] Pozzer, S., Ramos, G., Rezazadeh Azar, E., Osman, A., El Refai, A., López, F., Ibarra-Castanedo, C., Maldague, X.: Enhancing concrete defect segmentation using multimodal data and Siamese neural networks, Automation in Construction, 166 (2024), 105594

[13] Ta, Q. T., Huh, J., Yim, H. J., Lee, G., et al.: Automatic detection of subsurface defects in concrete structures using state-of-the-art deep learning-based object detectors on the infrared dataset, Engineering Structures, 329 (2025), 119829

[14] Shibano, K., Morozova, N., Shimamoto, Y., Alver, N., Suzuki, T.: Improvement of crack detectivity for noisy concrete surface by machine learning methods and infrared images, Case Studies in Construction Materials, 20 (2024), e02984

[15] Mahajan, T., Singh, G., Bruns, G.: An experimental assessment of treatments for cyclical data, Computer Science Conference for CSU Undergraduates, virtual, 2021.

[16] Bergman, T. L.: Fundamentals of heat and mass transfer, John Wiley & Sons, 2011.

[17] Breiman, L.: Random forests, Machine Learning, 45 (2001) 1, pp. 5–32

[18] Leys, C., Ley, C., Klein, O., Bernard, P., Licata, L.: Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, Journal of Experimental Social Psychology, 49 (2013) 4, pp. 764–766