

University of Nebraska - Lincoln

## DigitalCommons@University of Nebraska - Lincoln

---

Faculty Publications in Construction  
Engineering & Management

Durham School of Architectural Engineering  
and Construction

---

11-19-2019

### The application of gray-scale level-set method in segmentation of concrete deck delamination using infrared images

chongsheng cheng  
cheng.chongsheng@huskers.unl.edu

Zhigang Shen  
University of Nebraska - Lincoln, shen@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/constructionmgmt>

Part of the [Construction Engineering and Management Commons](#), [Data Science Commons](#), [Structural Engineering Commons](#), [Structural Materials Commons](#), and the [Transportation Engineering Commons](#) Network

---

Logo  
cheng, chongsheng and Shen, Zhigang, "The application of gray-scale level-set method in segmentation of concrete deck delamination using infrared images" (2019). *Faculty Publications in Construction Engineering & Management*. 20.  
<https://digitalcommons.unl.edu/constructionmgmt/20>

This Article is brought to you for free and open access by the Durham School of Architectural Engineering and Construction at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications in Construction Engineering & Management by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# The application of gray-scale level-set method in segmentation of concrete deck delamination using infrared images

Chongsheng Cheng<sup>1</sup>, and Zhigang Shen<sup>2</sup>

<sup>1</sup> Durham School of Architectural Engineering and Construction, University of Nebraska-Lincoln, 122 NH, Lincoln, NE 68588; e-mail: cheng.chongsheng@huskers.unl.edu

<sup>2</sup> Durham School of Architectural Engineering and Construction, University of Nebraska-Lincoln, 113 NH, Lincoln, NE 68588; e-mail: shen@unl.edu

## Abstract:

Conventional nondestructive delamination detection of concrete pavements through thermography is often based on temperature contrasts between delaminated and sound areas. Non-uniform backgrounds caused by the environmental conditions are often challenging for contrast-based methods to robustly differentiate the delaminated areas from the sound areas. Instead of focusing on the temperature contrast, this study proposes a temperature gradient-based level set method (LSM) to detect boundaries for delamination segmentations. A modified edge indicator function is developed to represent the normalized temperature gradient of a thermal image. The experimental study was conducted to evaluate its applicability and stability for boundary detection in terms of different burial depths and observation time windows. Then the LSM generated edges LSM were used for delamination segmentations. The results demonstrated significant improvements compared to the well-accepted k-mean method in terms of accuracy and stability of the detections. A field implementation of this method using an in-service bridge generated satisfactory outcomes when compared to the ground truth data. Major limitations and future works are also discussed in terms of parameter selections.

Keywords: Concrete Delamination; Thermography; Nondestructive Evaluation; Temperature Gradient Map; Edge-Based Level Set Function

## 1 Introduction

Accurately locating and quantifying the concrete pavement delamination is essential for decision-making of repair/replacement of bridge decks. Infrared thermography (IRT) as one of the nondestructive detection techniques (NDT) has been widely recognized as a fast and efficient method for shallow delamination detection [1–5]. Previous studies made efforts on handling the thermal image from the aspects of temperature histogram for the threshold method [6], temperature density for the k-mean clustering method [7], and temperature spatial relationship for region-growth method [8,9]. So far, these reported methods to distinguish the delaminated area from the sound area were primarily based on evaluating the temperature contrast at the global level of the inspected decks, which often led to over or under estimations of the delaminated areas. The level set method (LSM) has been successfully used in segmenting tissues or organs for medical image processing [10–13]. LSM has significant advantage over global temperature contrast methods in detecting local gray-scale edges, which makes it an ideal tool for identifying edges of delaminated areas. However, there is a lack of studies on using LSM in segmenting delaminated concrete pavement.

Based on previous studies, variations in the shape and depth of the delamination [14], the observation time window [15,16], and surface textures [17] played important roles for the detectability of delamination edges from thermography. Since previous studies focused on the developed

temperature contrast, the effect on temperature gradient at the boundary and non-boundary locations needs to be evaluated. Therefore, this paper will focus on investigating the applicability of the edge detection of delaminated pavement areas based on the temperature gradients in terms of different buried depths and time windows of observation in the experimental setup. To utilize the gradient information and shape characteristic of the delamination, an edge-based level set method (LSM) is introduced to profile the delamination along the boundary. The performance is compared with the conventional method and the detectability is also evaluated. Finally, a field implementation is conducted for further evaluation of its applicability in practice.

## 2 Background

### 2.1 The Active Contour Model (ACM) and the Level Set Method (LSM) for Curve Evolution

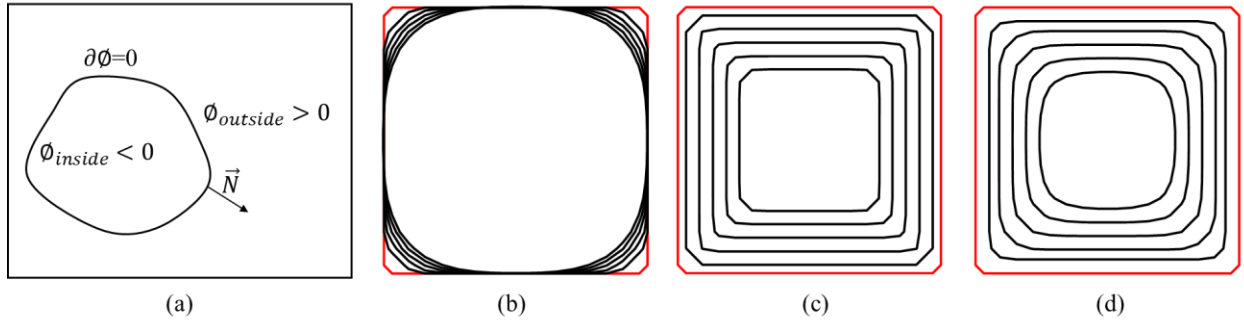
The ACM has been widely used in image segmentation due to its advantage on the deformable nature of models and the user-defined functional for constraints, which allowed the implementation of prior knowledge for object segmentation [18]. The basic idea is to define a curve (called active contour) in a computational space, moving it to the object's boundary from an initial position driven by the defined force and constraint over time until stable [19]. LSM was later used to better handle the ACM by introducing the implicit function, and thus replaced the parametric curve by the level set function for evolution [20]. There were two basic types of forces to be applied based on different constraints used: the edge-based and region-based models [21]. Edge-based active contour was forced by the image gradient and trapped the contour at the strongest edge location of the object such, as the geodesic ACM [22]. The region-based force was driven by the energy functional that derived from image intensity difference, such as C-V model [19]. Most relevant works [23,24] could be found in that a multilayer region-based model was used for thermal image segmentation for nondestructive evaluation of the composite structure and hot mix asphalt. The general form of the LSM for curve evolution could be written in a partial differential equation (PDE) form in **Eq.1-3** [20]. It defines a level set function  $\phi(x, t)$  to make a curve  $C(t) = \{x | \phi(C(t), t) = 0\}$  that always exists, any points  $x$  on the curve at any time  $t$  equals zero, and also satisfies  $\phi_{inside}(x, t) < 0$  and  $\phi_{outside}(x, t) > 0$  (**Fig. 1a**). Thus, the evolution of the curve is the function that updates by the time derivative in **Eq.2** where  $\phi_t$  refers to  $\partial\phi/\partial t$  meaning the derivative of level set function by time. **Eq.1** defines how the curve moves based on the force  $F$  on each point in terms of the curve gradient magnitude ( $|\nabla\phi|$ ) in the normal direction  $\vec{N}$  (**Fig. 1a**), and the **Eq.3** gives the definition of initial curve function.

$$\phi_t = (K + F)|\nabla\phi| \quad (1)$$

$$\phi = \phi + \phi_t \quad (2)$$

$$\phi(x, t = 0) = \phi_0(x) \quad (3)$$

Where  $K$  is the curvature of the level set function that is defined as  $K = div(\nabla\phi/|\nabla\phi|)$ . The  $div(\cdot)$  refers to the divergence of two vectors and the  $\nabla\phi/|\nabla\phi|$  is the normal of the curve at a point and could be calculated by the gradient vector field (GVF) [25]. Thus, the  $K$  constrains how ‘‘bent’’ a curve should be, and a large value indicates a very sharp corner at the curve. The force term  $F$  could be a constant value singularly or piece-wisely, which plays the major role to control the curve evolution. **Fig.1b** illustrates the curve evolution with the only  $K$  during 5 iterations, which shows the curve becoming smoother. **Fig.1c** shows the curve evolution with only a constant force  $F$  that evenly shrinks the curve during each iteration. **Fig.1d** shows both terms added for curve evolution and thus the curve shrinks and smooths during the iteration.



**Figure 1.** Definition of level set function (a), and the curve evolution with the same initial curve  $\Phi_0$  (red curve) during 5 iterations (black curves): (b)  $\Phi_t = K|\nabla\Phi|$ ; (c)  $\Phi_t = F|\nabla\Phi|$  with a constant  $F$ ; and (d)  $\Phi_t = (K + F)|\nabla\Phi|$  with a constant  $F$ .

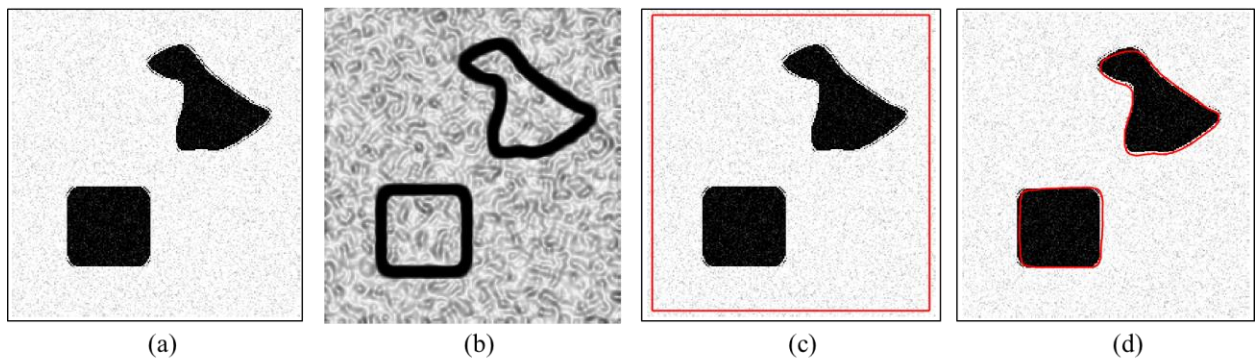
## 2.2 Edge-based Level Set Method for Image Segmentation

The most influential edge-based model for image segmentation could be the geodesic active contour (GAC) model [22]. It defined the force term and stopping criterion by introducing an edge indicator. Its level set formulation is shown in Eq.4. It is similar to Eq.1 with two additional terms ( $g$  and  $\nabla g \cdot \nabla\Phi$ ) introduced. The edge indicator  $g$  in Eq.5 describes a force that is driven by the gradient of the input image ( $\nabla I$ ). This function calculates the strongest edge closing to 0 and weakest edge closing to 1 and thus could drive the curve function  $\Phi_t$  to be trapped at the strongest edge location. Fig 2 shows the general procedure of the GAC model to segment objects with different shapes. Firstly, the gradient map was generated by the edge indicator function in Fig. 2b, and then an initial curve function  $\Phi$  was assigned (Fig. 2c). Fig.2d shows the result after 200 iterations: the final curve falls on the boundary of the objects. It is noticeable that the edge-based model is sensitive to the noise, the strong edge, and the initial contour, which brings challenges to directly implement the model on thermal images [26].

$$\frac{\partial\Phi}{\partial t} = g \cdot |\nabla\Phi| \cdot \left( \text{div} \left( \frac{\nabla\Phi}{|\nabla\Phi|} \right) + F \right) + \nabla g \cdot \nabla\Phi \quad (4)$$

$$g(|\nabla I|) = \frac{1}{1 + |\nabla G_\sigma * I|^2} \quad (5)$$

Where  $|\nabla\Phi|$  is the gradient magnitude of level set function  $\Phi$ ,  $\text{div} \left( \frac{\nabla\Phi}{|\nabla\Phi|} \right)$  is the curvature term,  $F$  is the balloon force term (often a constant), and  $\nabla g \cdot \nabla\Phi$  is the regularization term. In the edge indicator function (Eq.5),  $|\nabla I|$  refers to the gradient magnitude of the input image ( $I$ ), and  $\nabla G_\sigma * I$  is the convolution of input image with a Gaussian kernel.



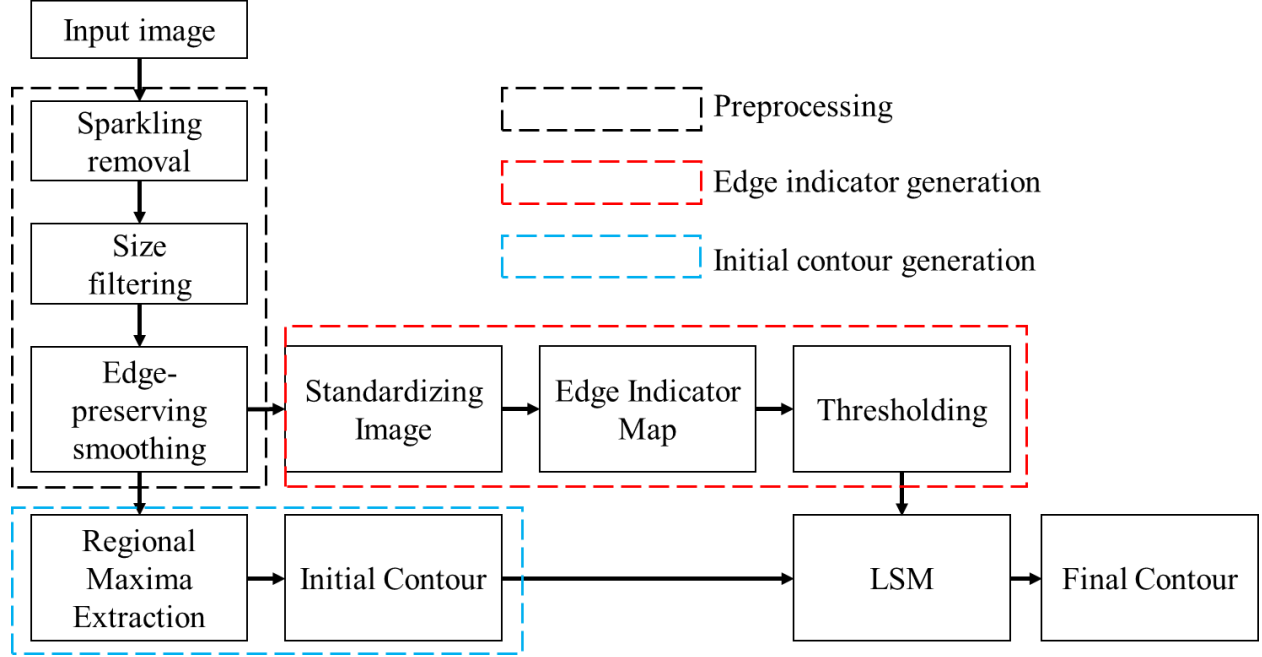
**Figure 2.** The procedure of the GAC model for image segmentation: (a) synthetic input image; (b) edge indicator function  $g$ ; (c) initial contour in red; (d) final contour in red.

## 3 Methodology

### 3.1 The General Framework

To evaluate the temperature gradient and appropriately implement the LSM for delamination segmentation, a general framework is proposed including preprocessing, modified edge indicator

function, and generation of initial contour. **Fig.3** shows the diagram of the proposed framework with three modules playing different roles of data processing. The thermal image first goes through the preprocessing module for outlier/sparkling removal, size filtering, and edge-preserving smoothing. Then the processed image is sent to the edge indicator generation module to get the normalized gradient map and the initial contour generation module to have the initial level set function ( $\Phi_0$ ). A modified level set function then takes the edge indicator function as the driving force to evolve the initial contour following the PDE-based GAC model [20].



**Figure 3. The diagram of the proposed framework**

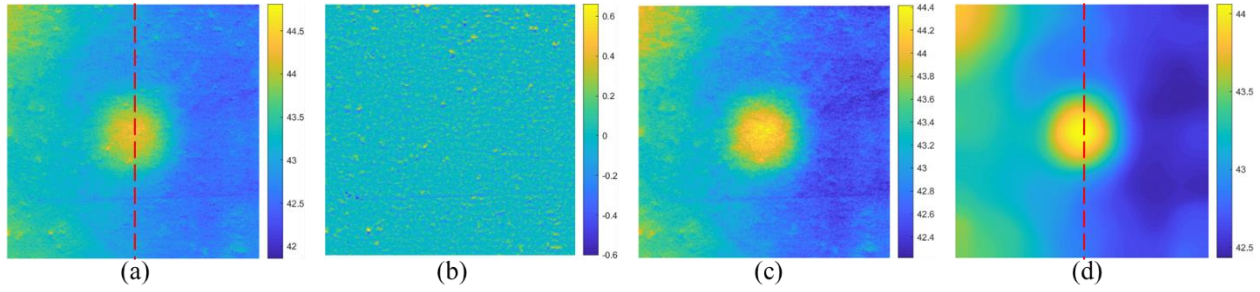
In order to generate a well-behaved edge indicator function for LSM, the raw thermal image needs to be preprocessed. The sparkling areas caused by the surface texture reflection from the solar radiation can be revealed by a difference image ( $I_d$ ), which is calculated based on **Eq.6**. They can be treated as outliers in the difference image (**Fig.4b**) and detected based on the median absolute deviation [27]. Then, the detected outliers are replaced by their nearest neighbor mapped from the original image. In **Fig.4**, a thermal image of the concrete slab with artificial delamination buried at 4.4cm depth is used to illustrate the intermediate results (see **Section 4.1** for the detailed description of the experimental setup). **Fig.4a** shows the raw thermal image taken at 5 p.m. outdoors in summer 2018. The warm color in the image refers to a high temperature and the cold color for low temperature, where the delamination often occurs as the hot region during the most daytime (middle area of **Fig. 4a**). **Fig.4b** shows the difference image which is related to the surface textures, and **Fig.4c** shows the raw image with sparkling points replaced. After the outlier is replaced, the image is further processed by a mean filter to smooth for certain-sized ( $\delta_{size}$ ) textures (**Eq.7**), which is achieved through the convolution of the image with the kernel of ones. Finally, the weighted median filter with the Gaussian Kernel introduced in Zhang et al. [28] is used to smooth the image with the edge-preserving consideration. The final image is shown in **Fig. 4d**. its comparison to the raw image is shown in **Fig. 5a**, where, in the cross-section view, the oscillations due to sparkling and texture are eliminated, and a smooth boundary is presented with subtle interruption to the edge shape.

$$\text{Sparkling removal: } I_d = I - \text{median}(I, \text{size}) \quad (6)$$

$$\text{Size filtering: } I_s = I_d * \text{ones}(\text{size}) \quad (7)$$

$I_d$  refers to the difference image, and  $\text{median}(\cdot)$  is the median filter with a pre-defined window size. In **Eq.7**, the  $I_s$  is the size filtered image,  $(*)$  denotes the convolution operation, and  $\text{ones}(\text{size})$

refers to the two-dimensional array of 1 with the pre-defined size ( $\delta_{size}$ ). The value of size in both **Eq.6** and **Eq.7** is determined based on task specification. We calculated it to smooth out the texture smaller than 2.5 cm<sup>2</sup> (1 square inch) in the paper which results in an 9-by-9-pixel window size.



**Figure 4. The diagram of intermediate results of preprocessing module (unit: Celsius degree): (a) raw thermal image from a sample slab of experimental study; (b) the difference image; (c) raw image with sparkling effects removed; (d) smoothed image after size filtering and edge-preserving smoothing.**

### 3.2 Modified Edge Indicator and Initial Contour Generation

Once the image is preprocessed, the modified edge indicator function is used to generate the edge map. First, the smoothed image is standardized according to its mean and standard deviation, and then the temperature gradient ( $\mathbf{G}$ ) is calculated using the ‘‘Sobel’’ kernel based on Vincent and Folorunso [29]. Then, the edge indicator function is calculated by **Eq.8** combining the anisotropic flux function [30] and modified edge indicator function [31]. Lastly, the function ( $\mathbf{g}_m$ ) is rescaled to [0,1] in which 1 means the weak edge and 0 means the strong edge. In general, the edge indicator function could be used to measure the relative strength of the temperature gradient across the thermal image.

$$g_m = e^{-[G/(1+G^2)]^2} \quad (8)$$

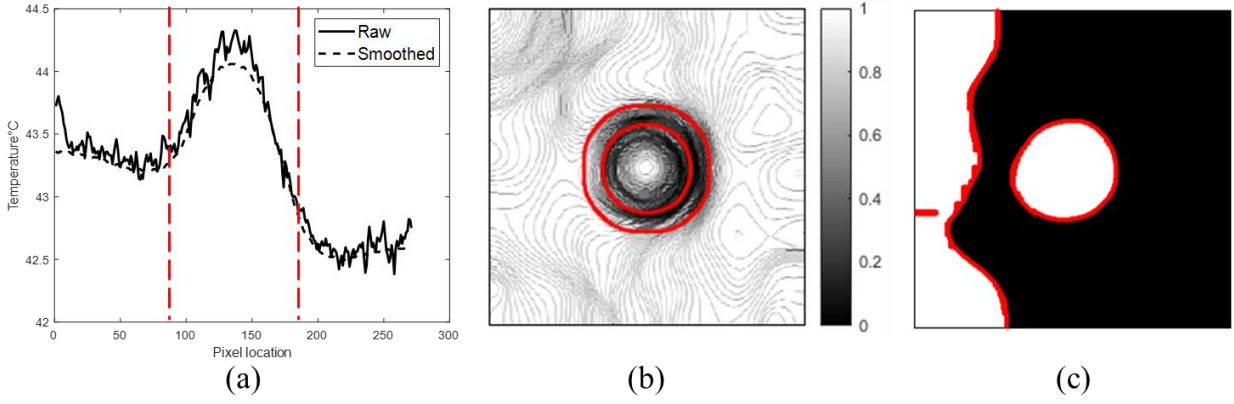
$e$  denotes the exponential function, and  $G$  the temperature gradient of the thermal image. Due to the diffusion nature of heat conduction within a medium, the temperature gradient (edge strength) of the boundary shown in the thermal image exhibits a strong blurring effect (**Fig. 5b**). To distinguish the potential boundary of delamination from the non-delaminated area, the threshold method could be used. In addition, variations in different depths of delamination buried, and different observation time could potentially dominate this relative strength for detectability. Thus, the experimental study is carried out to evaluate the applicability of the proposed scheme.

Another import factor affecting the edge-based LSM is the generation of initial contour, which could be handled by author’s previous study through utilizing the regional maxima of the thermal image [32]. The basic idea is to extract regional maxima as the initial contour for curve evolution. The rationale behind the idea is based on the finding that delamination in a thermal image is the subset of regional maxima defined in the mathematical morphology. Using the regional maxima as the starting contours, the contour is expected somewhere around the real boundary of delamination and thus will be more easily led by the force of the edge indicator function. **Eq. 9** shows the generation of initial contour  $\emptyset_0$ :

$$\emptyset_0(x) = \begin{cases} -1, & x \in R_{max} \\ 1, & x \in R_{back} \end{cases} \quad (9)$$

$R_{max}$  refers to the regional maxima detected by previous work shown as the white region in **Fig. 5c**, and  $R_{back}$  refers to the background area (the black region in **Fig. 5c**). The initial curve is the zero-level contour shown as the red line in **Fig. 5c**.





**Figure 5.** (a) section temperature profile after preprocessing module (red dash line indicates the real boundary location of mimicked delamination); (b) edge indicator function generated by Eq.8; (c) regional maxima of Fig. 4d with the initial contour in red.

### 3.3 Modified Level Set Function

The level set function used in this study is based on the classic GAC model in the PDE form with a few modifications. The regularization term ( $\nabla g \cdot \nabla \phi$ ) has been removed and is replaced by the reaction diffusion (RD) term with two-step splitting method (TSSM) according to Zhang et al [20]. These modifications eliminate the major drawbacks of LSM which often require the re-initialization of level set function during evolution and improve the weak boundary anti-leakage and numerical stability. **Eq. 10** shows the modified level set function:

$$\frac{\partial \phi}{\partial t} = \alpha \cdot |\nabla \phi| \cdot K + \beta \cdot g_m \cdot |\nabla \phi| \quad (10)$$

$$\phi^{n+1/2} = \phi^n - \Delta_{t1} \cdot \mathcal{L}(\phi) \quad (11)$$

$$\phi^{n+1} = \phi^{n+1/2} + \Delta_{t2} \cdot (\partial \phi / \partial t) \quad (12)$$

In **Eq.10**, the  $|\nabla \phi|$  is the gradient magnitude calculated by GVF of level set function,  $K$  refers to the curvature term defined in **Eq.1**, and the  $g_m$  is the edge indicator function in **Eq.8**. **Eq.11** and **Eq. 12** are the two-step splitting method used to add reaction diffusion term  $\mathcal{L}(\phi)$  and replace original function updates in **Eq.2**. The  $\mathcal{L}(\phi)$  refers to the Laplacian of the level set function. The parameter  $\alpha$  and  $\beta$  are constants that control the balance between the curvature constraint of the contour and the driving force to move the contour.  $\Delta_{t1}$  denotes the time step of the diffusion term and  $\Delta_{t2}$  refers to the time step for function updates. These parameters need to be tuned depending on implementation.

### 3.4 Performance Evaluation Metrics

Two metrics are used to evaluate the segmentation performance because the direct comparison of boundary similarity between segmentation and ground truth is not a feasible approach. Instead, we use the intersection over union (IoU) to evaluate the segmentation accuracy. IoU is defined as the overlapped area between prediction and ground truth over the union of them (**Eq.13**). It is a well-accepted metric to measure similarity between two sets and is often used for the performance evaluation in image segmentation task [33]. This metric returns 0 if there is no overlapping and 1 if it is perfectly matched. Area percentage ( $R_{area}$ ) of detection defined in **Eq.14** is another metric that calculates the detected area over the entire valid surveyed area.  $R_{area}$  is often used as a reference for maintenance decisions in practice and is expected to be as close as possible to the ground truth. To implement the result from level set function for performance evaluation, a signed function is used to transfer the region inside the contour as foreground (value of 1) and outside the contour as background (value of 0) in **Eq.15**.

$$IoU = \frac{|A_p \cap A_g|}{|A_p \cup A_g|} \quad (13)$$

$$R_{area} = \frac{A_p}{A_{total}} \quad (14)$$

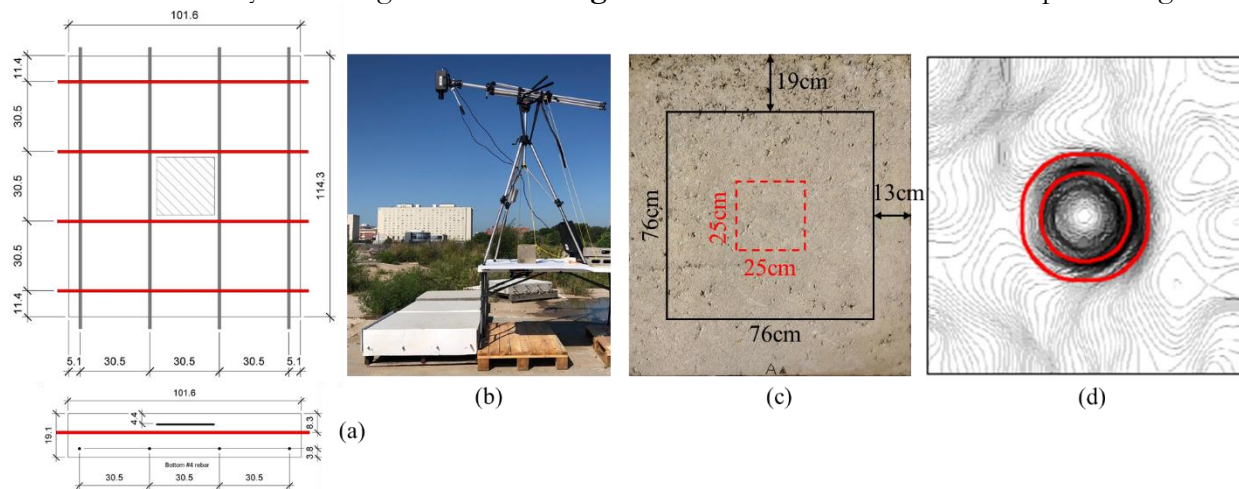
$$\phi_b = \text{sign}(\phi) = \begin{cases} 1, \phi \leq 0 \\ 0, \phi > 0 \end{cases} \quad (15)$$

In **Eq.13**,  $A_p$  is the predicted area by the algorithm,  $A_g$  is the ground truth area;  $A_p \cap A_g$  calculates the intersection between prediction and ground truth, and  $A_p \cup A_g$  returns the union area among prediction and ground truth. In **Eq.14**,  $A_{total}$  refers to the total surveyed area. In **Eq.15**,  $\phi_b$  refers to the binarized level set function.

## 4 Experimental Study of Temperature Gradient at Delamination Boundary

### 4.1 Experiment Setup and Data Collection

Experimental studies were conducted by using the mimicked delamination in a reinforced concrete slab outdoors in sunny weather. Three concrete slabs were cast and assigned mimicked delamination in three depths: shallow (4.4 cm), median (7 cm), and deep (9.5 cm). The slab design was based on the empirical deck design procedure from *Bridge Office Policies and Procedure* [34] by the Nebraska Department of Transportation. The layout of the slab is illustrated in **Fig.6**. The Styrofoam, with the size of 25 cm by 25 cm, was buried at a depth of 4.4 cm (shallow), 7 cm (median), and 9.5 cm (deep) from the top surface of each slab (**Fig.6a**). The thickness of foam is about 0.4 cm with the thermal conductivity of 0.03 W/(m·K) in order to stimulate the delamination in the real bridge deck, which is identified as a thin layer of air with a thermal conductivity around 0.02 W/(m·K). The slabs were cured indoor over 28 days and then moved outside for the data collection in August and September 2018. A thermal camera (**Table 1**) used to collect the surface temperature of three slabs (**Fig. 6b**) was set right on the center top of the slab. The supporting frame was set on the north side of each slab with the adjusted height so that there was no shadow effect from the camera and frame arms on the thermal image during the data collection. Also, the camera view was facing perpendicularly downward to the surface of the slab to ensure the least image distortion. This angle of setup also aimed to fit the configuration for the drone implementation on-site in the later section (**Section 5**). To reduce the side effect caused by the horizontal heat diffusion under the solar radiation, we first added a thin foam layer to insulate at the specimen's sides (**Fig. 6b**). In addition, the thermal image to be analyzed has been cut into a 76 by 76 cm region shown in **Fig.6c** as the black box for successive processing.



**Figure 6.** Experimental design, setup, and data collection and processing: (a) design of concrete slab and the layout of mimicked delamination; (b) data collection setup; (c) the optical image of slab surface (black box indicates the cropped region for thermal image and red box indicate the location of delamination); (d) edge indicator function of temperature gradient (red circles located 2.5cm from the real boundary of delamination inward and outward).

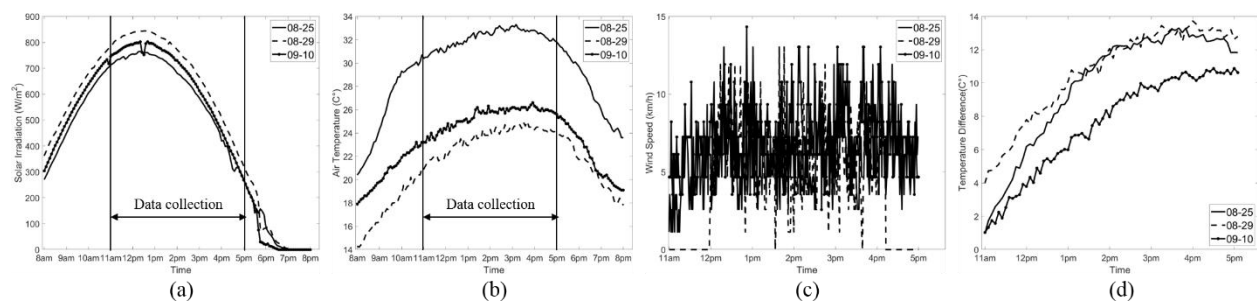


**Table 1. Thermal camera specification**

Thermal Camera	Spectral Range	Spatial Resolution	Thermal Sensitivity at 25 °C	Temperature Sensing Range	Max Frame Rate at Full Resolution
FLIR A8300	3~5 $\mu\text{m}$	1280x720	20 mk	-20°C~ 350°C	60 Hz

#### 4.2 Weather Conditions

Since the experiment was conducted in three different days, the weather conditions need to be evaluated. Weather information (solar irradiation, ambient temperature) was obtained from quality-controlled datasets ([www.ncdc.noaa.gov/crn/qcdatasets.html](http://www.ncdc.noaa.gov/crn/qcdatasets.html)) [35]. The chosen station is 5 km away from the site of experiment. Wind speed was recorded on site by a local weather station. **Fig.7(a)** shows the solar irradiance from 8 a.m. to 8 p.m. which illustrates all three experiments were under sunny condition with similar solar loading pattern. The solar loading has been recognized the most dominate factor contributing to the detectability of delamination in different burry depths [15,36]. With the similar loading pattern, the heating source for the three experiments could be treated as similar. **Fig.7(c)** shows the wind speed for all three days were under 15 km/h which was less than the 24 km/h recommended by ASTM (D4788) for accurate data collection. **Fig.7(b)** shows the ambient temperature variations. The ambient temperature showed a different absolute value for each experiment, but a similar climbing and declining pattern could be observed across three experiments. The effect of ambient temperature on the detectability of delamination was well studied in [37], which revealed that during certain time windows of a day, the delamination became undetectable when the ambient temperature was close to the concrete surface temperature. These time windows often occurred in the morning during sunrise and evening during sunset, where the delaminated area was experiencing a transition from cold to hot region and vice versa. Thus, the temperature difference between the slab surface and ambient played more important role than the absolute air temperature itself. **Fig.7(d)** plots the temperature difference for the three experiments. It shows that during the data collection time (11a.m. to 3:30 p.m.), the temperature contrast was continuously rising, which indicated the dominance of solar loading over the air temperature. Thus, this period could be treated at heating stage for all three experiments. During the time between 3:30 p.m. and 5 p.m., the temperature contrast became flat and then decreasing. This indicated the start of transition period, which the detectability might start to decrease. Although the weather conditions were not exactly same for the three experiments, with the evaluation on the leading factor (solar loading) and affective factors (ambient temperature and wind speed), the three experiment could be treated as a similar weather condition as “summer sunny condition”.

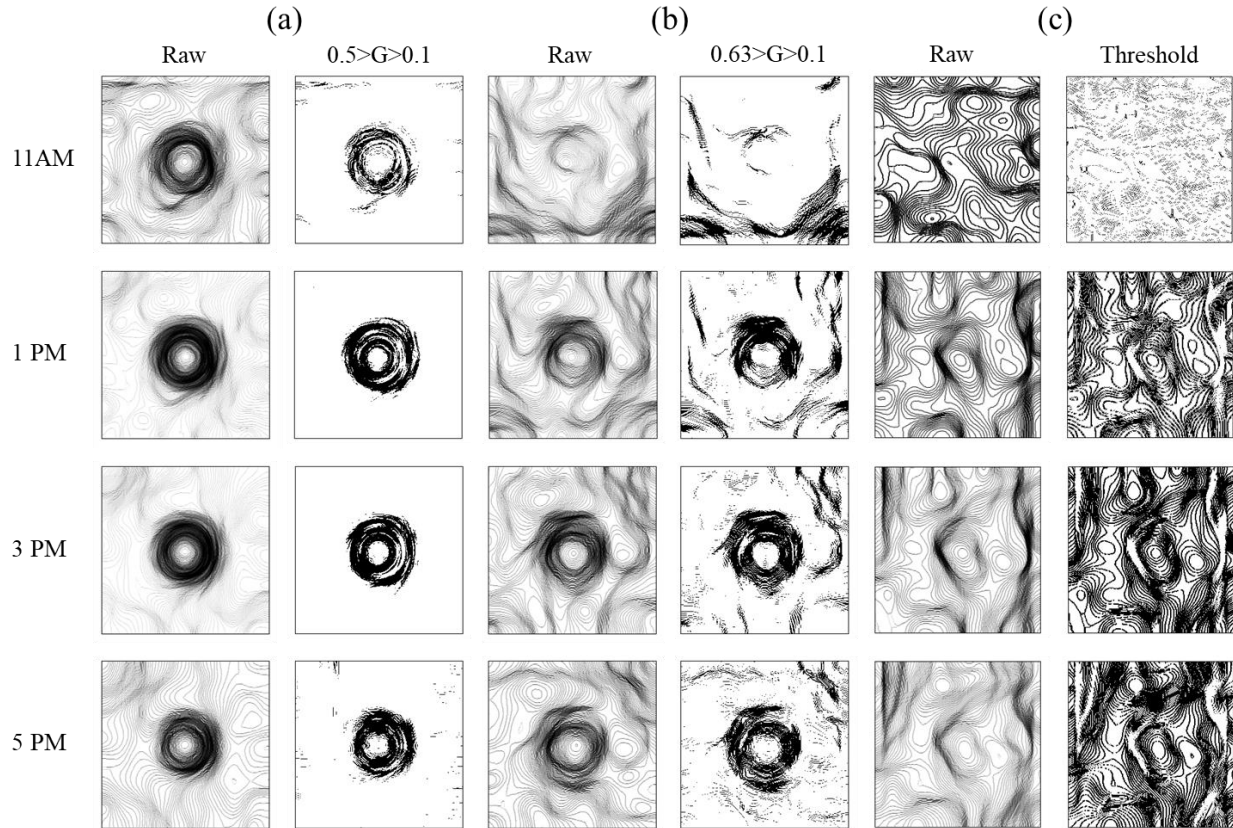


**Figure 7. Weather condition: (a) solar irradiation; (b) ambient temperature; (c) wind speed; and (d) temperature difference between slab surface and ambient. Data collected on 08-25-2019 for shallow delamination, 08-29-2019 for median delamination, and 09-10-201 for deep delamination.**

#### 4.3 Implementation and result comparison

**Fig.8** shows the temperature gradient map by implementing the threshold criterion (**Section 6.1**) on three slabs at four different time windows. For the slab with shallow delamination (**Fig.8a**), the threshold value (0.1 to 0.5) was applied on all four gradient images. It shows this threshold range was able to separate the boundary (black) and non-boundary (white) areas steadily during the experimentation. For the slab with median deep delamination in **Fig.8b**, the threshold could not separate boundary and the non-boundary regions at 11 a.m. At 1 p.m., the distinguishability increased but was not stable enough and caused a faulty classification (black lines that not belonged to the

boundary). After 2 p.m., the stability increased, and the fault classification decreased (**Fig.8b** for 3 pm and 5 pm). For the slab with deep delamination in **Fig.8c**, the low distinguishing power (see **Section 6.1, Fig.14c**) resulted in a boundary that cannot be easily distinguished from the non-boundary area.



**Figure 8.** Temperature gradient map calculated by the edge indicator function and its threshold result based on one-standard deviation at different time windows: (a) slab with 4.4 cm-deep delamination with raw result (left) and the threshold result (right); (b) slab with 7 cm-deep delamination with raw result (left) and the threshold result (right); slab with 4.4-cm deep delamination with raw result (left) and the threshold result (right) based on Fig.14c top (thresholds used at 11 a.m.:0.63~0.8, at 1 p.m.:0.95~0.4, at 3 p.m.:0.95~0.4, at 5 p.m.:0.95~0.4 ).

With the temperature gradient calculated and thresholds determined, the boundary of the delamination could be sketched by the LSM with customized initial contour. The processing was applied to the three slabs at the same time window mentioned in **section 4.2**. **Table 2** shows the parameters used for the initial contour ( $T_i, T_{step}$ ), the size filter ( $\delta_{size}$ ), and the level set function ( $\alpha, \beta, \Delta_{t1}, \Delta_{t2}$ ). **Fig.9** shows the segmentation results of the three slabs at four different time windows in the date described in **Fig.7**. **Fig.9a left** illustrates a smoothed thermal image for shallow delamination (4.4 cm deep) after the preprocessing module with the initial contour (**Eq.9**) is overlaid on the left side. **Fig.9a middle** shows the results for the proposed method (red), k-mean method with  $k=2$  (blue), and ground truth (green). **Fig.9a right** compares the results of the proposed method to k-mean method ( $k=3$ ) and ground truth. The k-mean clustering method was a well-accepted method introduced in Omar and Nehdi [7] and in Omar et al. [39] for delamination segmentation in the thermal image. It divided the image space into multiple groups based on the Euclidean distance of each pixel to the centroid  $k$ . Based on the proposed criterion [39], the  $k$  could be set as 1, 2, 3 and 4 depending on the age of deck, the thermal contrast of deck surface, and the size and depth of spalling occurred in the bridge deck. When  $k$  equals 2, the processed image is divided into monitoring and sound regions, where the monitoring region indicates the delamination is under developing and the sound region indicates no delamination. When  $k$  equals 3, a warning region is then added to indicate the acute delamination. Although the literature defined a practical procedure to determine the number of  $k$  for field implementation, the criterion for experimental study was not well established. Thus, we compared the result of k-mean method using both  $k=2$  and  $k=3$  to the proposed method for the analysis of delamination boundary profiling in **Fig.9** and the delamination segmentation in **Fig.10**.

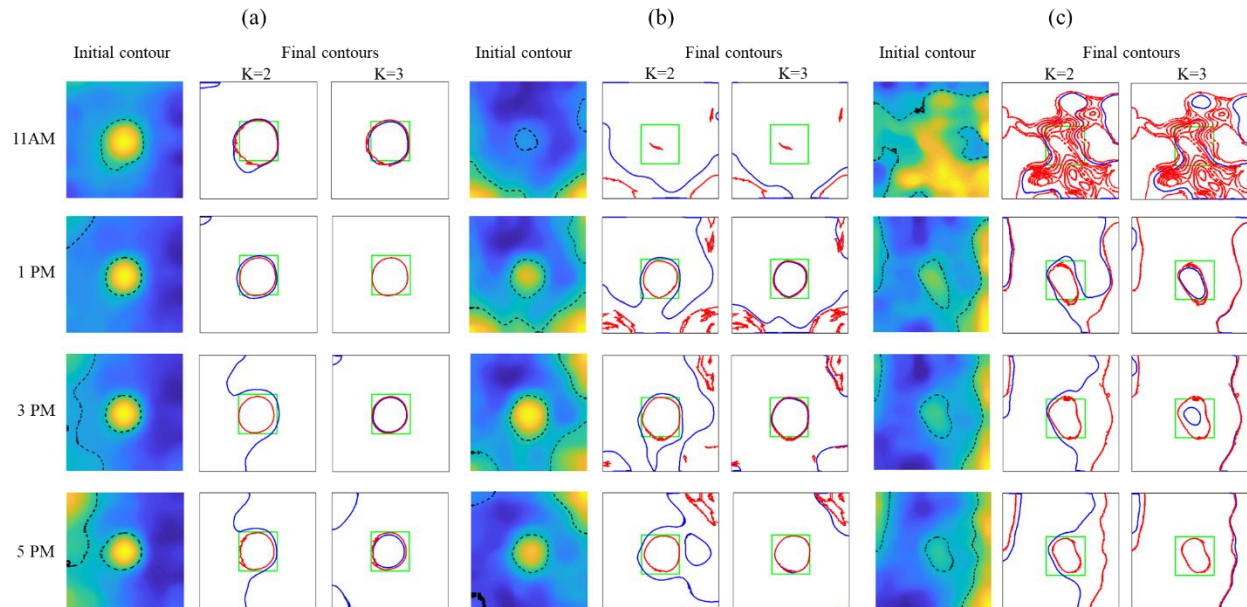
**Table 2. Parameter setting for experimental and field implementation**

Case	$T_i$	$T_{step}$	$\delta_{size}$	$\alpha$	$\beta$	$\Delta_{t1}$	$\Delta_{t2}$	Iteration
Experiment	0.1*1	0.1	9	0.001	1	0.0001	10	100*2
Field	0.5	0.2	5	0.001	1	0.0001	10	100

Note \*1: 0.5 used for the case with 4.4 cm deep delamination

\*2: 400 iterations used for the experimental case with 9.5 cm deep delamination at 11 am for convergence

When comparing the proposed method to the k-mean method (k=2) for the boundary profiling, the proposed method exhibits overall lower fault alarms and higher stability across different time windows (**Fig.9**). In **Fig.9a middle**, it is shown that the current method (red lines) sketches closer to the real boundary of the delamination (green lines) than the k-mean method (blue lines) does across all time windows. Especially in the afternoon (3 p.m. and 5 p.m.), the k-mean method intends to sketch larger area as the delamination boundary due to the non-uniform background. For the median delamination (7 cm deep) in **Fig.9b middle**, neither the proposed method nor k-mean work well at 11 a.m. due to the non-distinguishability based on the gradient analysis (**Section 6.1**). Later in the day (1, 3, and 5 p.m.), the proposed method successfully profiles the boundary of the delamination with some faulty detections exist. These faulty detections are affected by the unstable threshold used, and it is observed that their tracks are spatially related to the gradient strength map in **Fig.8b** correspondingly. The k-mean method failed to profile the shape of the delamination at 11 a.m. and over-profiled the boundary at other time windows in this case. For the slab with deepest delamination (9.5 cm deep) in **Fig.9c middle**, the current method fails to profile the boundary of the delamination at 11 a.m. and partially sketch its boundary at other time windows but has more fault detections. Because the distinguishing power is too weak to use, shown in **Fig.14c**, the processed result should be carefully handled for practical implementation. In terms of the k-mean method, it fails to sketch the boundary at all time windows for this case.



**Figure 9. The results of delamination profiling for comparison: initial contour (black dashed line), current method (red solid line), k-mean method (blue solid line for k=2 in the middle and k=3 in the right), and ground truth (green line): (a) slab with 4.4 cm deep delamination; (b) slab with 7 cm deep delamination; (c) slab with 9.5 cm deep delamination.**

When setting the k=3 for the k-mean method, improved performance is observed, and more agreement is found compared to the proposed method for the boundary profiling. Since the k-mean method with k=3 divided the image into three regions (**Fig.10**), the region with the highest mean value was selected [39] to indicate the most likely indication of delamination. In **Fig.9a right**, both the current method and k-mean method show similar outcome of sketching the delamination boundary. Some fault detections by the k-mean method occurred at 3 p.m. and 5 p.m. Similar results could be observed in **Fig.9b** at 1, 3, and 5 p.m. and in **Fig 9c** at 1 p.m. In these cases, the current method and

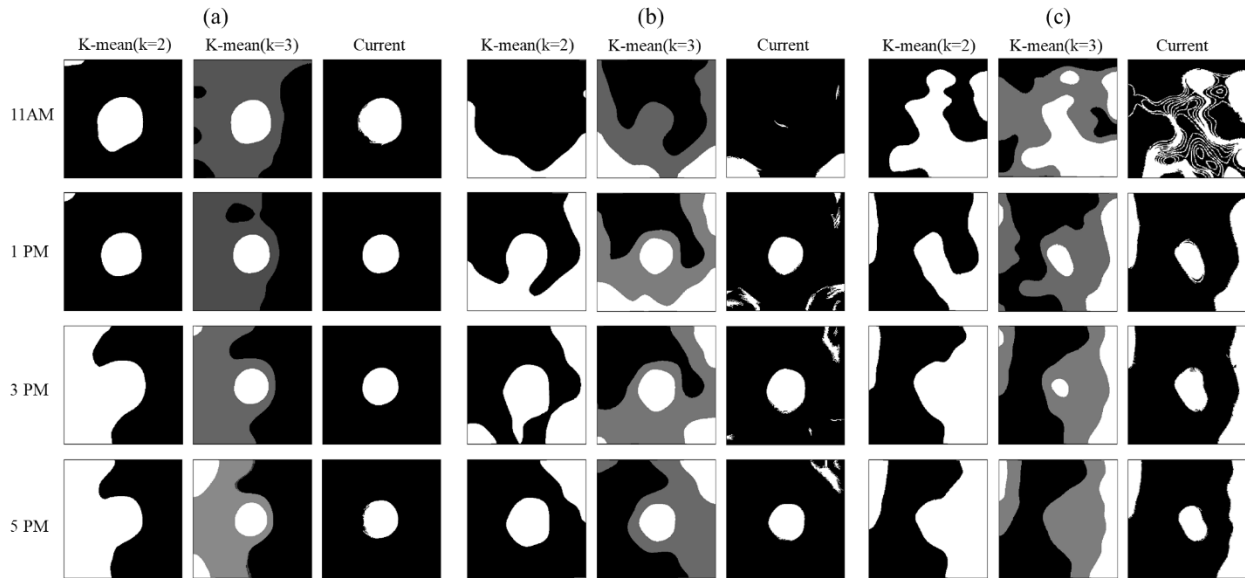
k-mean method both have fault detection of boundaries. Especially for the deep delamination (**Fig.9c**), the k-mean method starts to fail to detect the boundary at 3 p.m. and miss the detection at 5 p.m., while the proposed method shows constant results after 1 p.m. In summary, both the proposed and k-mean methods could sketch the delamination boundary close to the ground truth profile, while the proposed method exhibits more stability and fewer faulty detections. When using the k-mean method with the number of k larger than 2, the interpretation of data needs to be carefully handled.

To quantitatively compare the segmentation accuracy, we binarized the final level function (**Eq.15**) and evaluated it through the IoU and area ratio. **Fig.10** shows the segmentation results for the three slabs between the k-mean and current method at four different time windows. **Table 4** shows the quantitative results in terms of the IoU and area ratio error. For the slab with shallow delamination (**Fig.10a**), the current method outperformed the k-mean method at 11 a.m., 3 p.m. and 5 p.m. with around 68% to 79% accuracy to the ground truth. For the case at 1 p.m., the best performance is achieved by the k-mean method when k is set to 2. It returns 76% accuracy while the current method is evaluated as 72%. The area ratio error indicates the current method keeps the difference rate within 4% in this case. For the slab with median depth delamination (**Fig.10b**), the current method did not perform well at 11 a.m. in terms of accuracy. The accuracy increased to 49% and above after 1 p.m. The area ratio error was kept below 3.2%. On the other hand, the k-mean method exhibited lower accuracy and higher area ratio error for this case for both k=2 and k=3. For the slab with 9.5 cm deep delamination (**Fig.10c**), the current method shows some detectability after 1 pm but presents more fault alarms than the previous two cases. Therefore, the current method returns low accuracy (ranging from 17% to 27%). In terms of k-mean method, the segmentation accuracy stays around 21% across time when setting k=2 (**Fig.10c**), while the accuracy varies from 0% to 21% when k=3. In general, it is found that the proposed method shows a stable segmentation accuracy and area ratio across all time windows for the shallow delamination and the median delamination after 1 p.m. The performance of the k-mean method varies across time for both k=2 and k=3 for the shallow and the median delamination. For the deep delamination, both the proposed and k-mean methods do not perform well. It also found that the selection of k=3 often returns an improved performance compared to k=2 in segmentation performance. These findings share the agreements to the findings in the boundary profiling which supports the usefulness of the temperature gradient.

**Table 4. Results comparison with K-mean Clustering**

Time Window		11 AM			1 PM			3 PM			5 PM		
Depth (cm)	Metric	Current	K=2	K=3	Current	K=2	K=3	Current	K=2	K=3	Current	K=2	K=3
4.4	IoU	<b>0.7917</b>	0.6801	0.7790	0.7168	<b>0.7550</b>	0.7155	<b>0.6845</b>	0.2227	0.5936	<b>0.6868</b>	0.2066	0.3311
	$D_{AREA}$	<b>0.05</b>	+3.39	-0.91	-3.05	<b>-0.45</b>	-3.07	<b>-3.46</b>	+37.13	-3.70	-3.45	+36.69	<b>+2.65</b>
7	IoU	<b>0.0067</b>	0	0	<b>0.4923</b>	0.2119	0.2111	<b>0.6808</b>	0.2620	0.3902	<b>0.5738</b>	0.4254	0.4283
	$D_{AREA}$	<b>-3.21</b>	+12.06	+3.54	<b>-0.03</b>	+35.01	+13.93	<b>+0.31</b>	+28.86	+8.00	<b>-1.32</b>	+11.58	+3.46
9.5	IoU	0.1111	0.2087	<b>0.2092</b>	0.1859	<b>0.2158</b>	0.1669	<b>0.2713</b>	0.2071	0.0693	0.1687	<b>0.1865</b>	0
	$D_{AREA}$	<b>21.88</b>	+31.10	+22.33	+12.18	+22.49	<b>+6.01</b>	+10.24	+35.37	<b>+3.96</b>	+13.50	+36.54	<b>+4.84</b>

**Note:**  $D_{AREA}$  is shown as the difference between the calculated area ratio of each method and the ground truth based on **Eq.14**. The ground truth area ratio is 11.03% for 4.4 cm-deep delamination, 10.9% for 7 cm-deep delamination, and 11.03% for 9.5 cm-deep delamination. For IoU, the higher value the better performance, while for area ratio difference, the lower value the better performance.



**Figure 10.** Segmentation results comparison between the current method and k-mean method: the white color indicates the detection of delamination, gray color indicates the region having delamination under developing, and black color indicates the sound region. (a) slab with 4.4 cm deep delamination; (b) slab with 7 cm deep delamination; (c) slab with 9.5 cm deep delamination.

## 5 Implementation on the In-Service Bridge

### 5.1 Data Collection and Image processing

A field test was conducted on an in-service concrete bridge deck with concrete overlay at the northbound of US 77 over Rock Island and Pacific Railroad (RI&P), Nebraska. The bridge is 54.9 m in length and 12 m in width heading from south to north. The data was collected through a UAV equipped with optical and infrared cameras in October 2017 (**Fig.11**). The path of the UAV was designed to follow the center of the bridge at a fixed height and constant speed. The thermal image was recorded at a 1 Hz sampling rate and then post-processed through the customized MATLAB algorithms for perspective correction, rotation, and stitching. Then, the accuracy of the stitched thermal image was checked and calibrated through manual comparison with the visible image for this case. **Fig.12a** shows the regular image of the surveyed bridge deck surface superimposed with hammer sounding result provided by the Nebraska Department of Transportation (NDOT). Hammer sounding is a conventional method for shallow delamination detection for concrete bridge deck, based on ASTM 4850M-12. The hammer sounding results were verified by 12 coring samples at different locations in authors' previous study [32], where the defected regions were actually the deboned concrete overlays (very similar to delamination). The boundaries of these regions were directly marked by a chalk line on the surface of the deck during the test. **Fig.12b** shows the smoothed thermal image of the bridge processed by the preprocess module, and the black dash line indicates the initial contour generated by the author's previous work for regional maxima extraction [32]. The parameters for smoothing and level set function are presented in **Table 2**. The gradient threshold was set from 0.8 to 0.1 for this field case. This value is based on the recommended threshold by the gradient analysis in **section 4.2**, and it is adjusted to account for the strong gradient generated by a circumstance other than the delamination, such as a patch or road striping paint (**Fig.12c**). **Fig 12c** shows the gradient map calculated by the edge indicator function. Then, the level set function evolved based on this map after the threshold is applied.





Figure 11. The set-up of UAV (DJI M600 Pro) with the thermal camera (FLIR A8300)

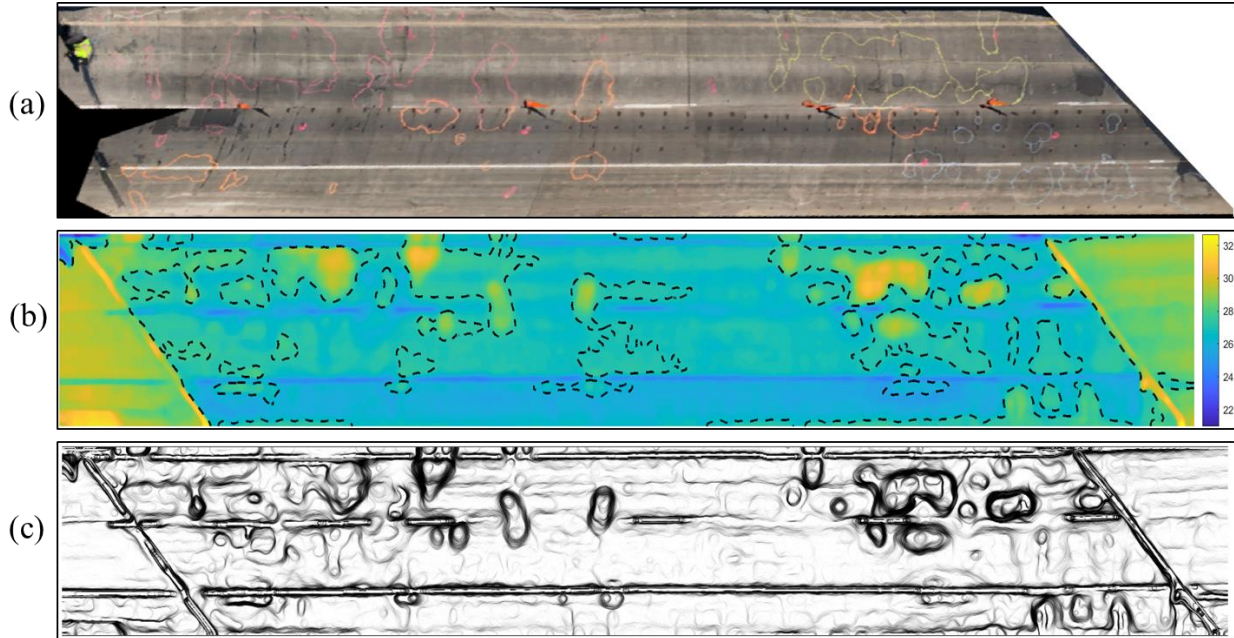


Figure 12. Field implementation: (a) bridge surface marked by hammer sounding result; (b) thermal image of the bridge deck marked by initial contour; (c) temperature gradient map generated by edge indicator function.

## 5.2 Result Comparison With K-Mean Method

The results compared the current method and the k-mean method to the hammer sounding test (Fig.13). During the core sampling, the hammer sounding result was confirmed as effective in this bridge [32], and thus it is used as the reference for the case comparison (green lines in Fig.13). It is found that the current method (red color) most closely resembles the hammer sounding result (Fig.13a), and thus returns highest IoU rate (Table 4). Fig.13b shows that the k-mean method ( $k=2$ ) missed some delaminated regions on the shoulder area, where the shoulder shows lower temperature than the lane area (Fig.12b). This is known as the effect of non-uniform background [40]. When  $k$  sets to 3, the area of the warning region (blue color) decreases and a large area of monitoring region (gray color) occurs, as shown in Fig.13c. This monitoring region represents the mediate temperature group between high and low temperature groups (shown in both experimental study in Fig.10 and field implementation in Fig.13c). Therefore, it could be affected by the non-uniformity of the temperature distribution which may lead to the inaccurate interpretation of result. In terms of the ratio of the delaminated area over the deck area, the proposed method shows the closest estimation of area over the k-mean method, which has the area ratio error of 1.5%. Although the current method has 9% higher accuracy than the k-mean method, the agreement to the hammer sounding test was not high. One of the potential issues was the geometric distortion in the provided hammer sounding test that caused the pattern shift when we extracted it for result comparison (Fig.13). In terms the k-mean method adopted for the bridge implementation with overlays, the selection of number of  $k$  and interpretation of data need to be adjusted.

Table 4. Quantitative evaluation on in-field implementation

Case	IoU			$D_{AREA}$ (%)		
	Current	K=2	K=3	Current	K=2	K=3
<b>Bridge</b>	<b>0.4462</b>	0.3505	0.3097	<b>-1.5</b>	-8.04	-10.42

Note:  $D_{AREA}$  is shown as the difference between calculated area ratio of each method and the reference based on Eq.14. The area ratio tested by hammer sounding is 18.2%. For IoU, the higher value the better performance while for area ratio difference, the lower value the better performance.

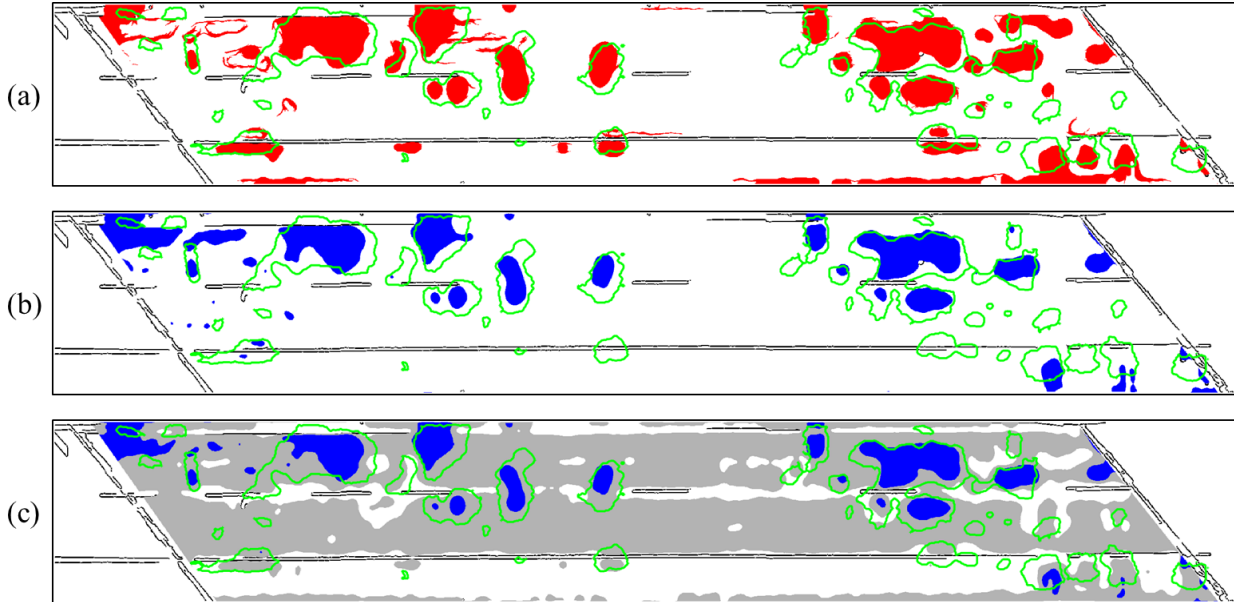


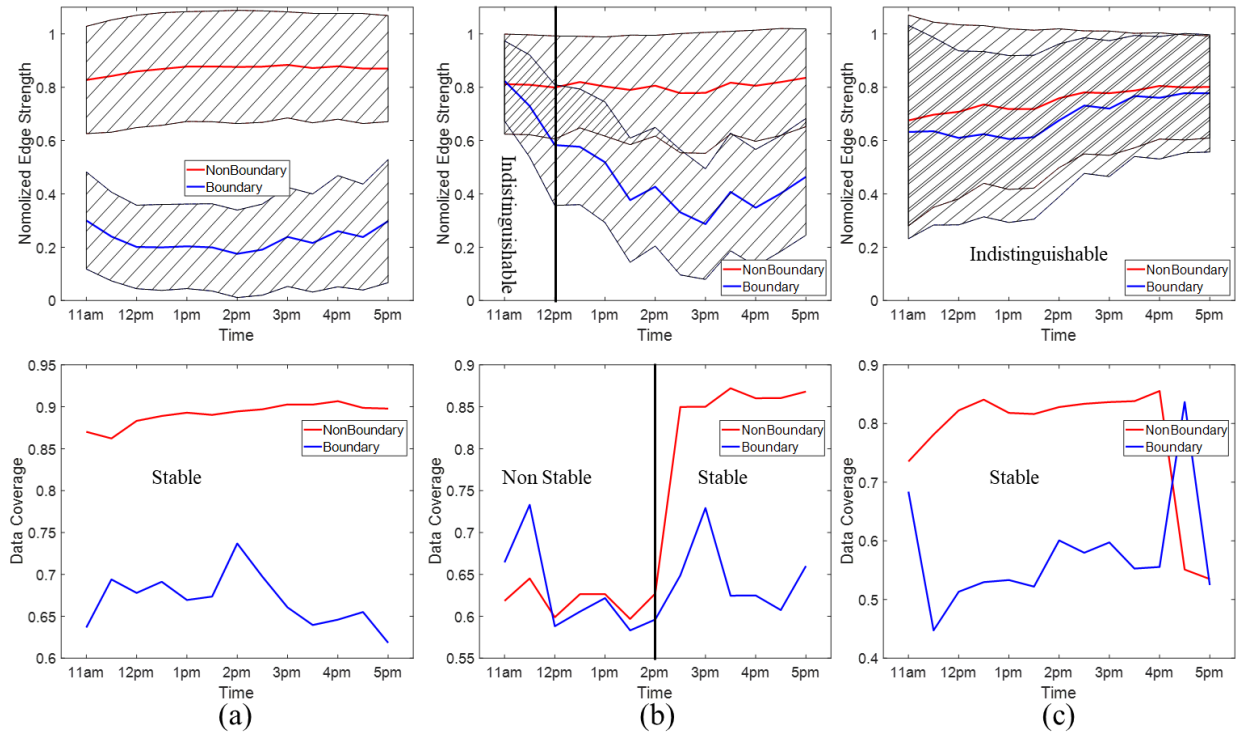
Figure 13. Field implement results (green lines show the result of hammer sounding test; white region indicates no delamination): (a) the proposed method (red region indicates delamination); (b) the k-mean method with k=2 (blue region indicates potential delamination); (c) the k-mean method with k=3 (blue and gray indicate potential delamination).

## 6 Discussion

### 6.1 Gradient Analysis

Since the current method relies on the temperature gradient map (Eq.8) as the leading factor, the separability and stability of the temperature gradient at delamination boundary need to be evaluated for the three experiment. To investigate these variations, the normalized gradient strength is calculated (Eq.8) and divided into the boundary and non-boundary areas for each experiment. Fig.15a shows a sample image of temperature gradient calculated by the edge indicator function, where the red round corner square is located 2.5 cm from the exact boundary of the delamination inward and outward (the effect of selecting different boundary width is discussed in section 6.2). Thus, the region between the two lines is used to represent the gradient of boundary and its complement region for the non-boundary area. The reason to choose a very rounded corner square instead of the sharp corner square for boundary gradient extraction is based on the observation that the sharp edge of the delamination is naturally smoothed out at corners due to heat diffusive propagation (in Fig.4a and Fig.5b, the pattern of delamination is a more rounded shape). The effect of choosing a square shape on the analytic result is compared in section 6.2. To evaluate how separable the two distributions are from boundary and non-boundary areas, their mean with standard deviation is used. Considering the mean as the criterion to separate two distributions, the mean needs to be as far as possible from the distributions. Fig.14 top shows the mean gradient strength of boundary and non-boundary area for three slabs over a day. It is noticed that in the slab with shallow delamination (Fig.14a top), the gradient strength around the delamination boundary (blue line) is constantly between 0.2 to 0.3, and the mean gradient of non-boundary area (red line) is around 0.8 to 0.9 from 11 a.m. to 5 p.m. For the slab with median delamination (Fig.14b top), the mean gradient difference was less than 0.2 before 12 p.m. and increased with as time passed. It reached around 0.4 for the boundary area from 1 p.m. to 5 p.m.; in the meantime, the non-boundary area was constantly around 0.8 from 11 a.m. to 5 p.m.

**Fig.14c top** shows that the mean gradients between boundary and non-boundary area of the slab with deep delamination were indistinguishable due to the difference less than 0.2 from 11 a.m. to 5 p.m.



**Figure 14. Gradient analysis of slabs. Top: mean gradient strength for boundary area (blue) and non-boundary region (red) and corresponding standard deviations (hatch pattern) for (a) slab with 4.4 cm deep delamination, (b) slab with 7 cm deep delamination, and (c) slab with 9.5 cm deep delamination. Bottom: data coverage when using one-standard deviation for boundary area (blue) and non-boundary region (red).**

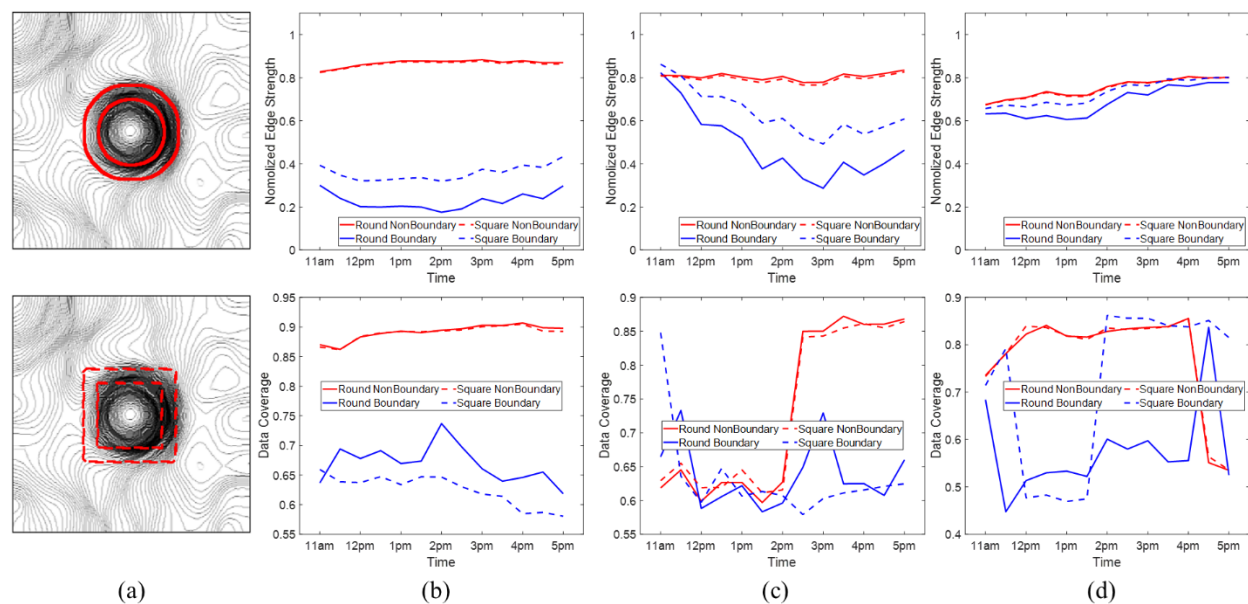
Adding one standard deviation above and below the mean value gives a range of threshold that could be used to group the gradient distribution for boundary and non-boundary areas from the gradient map of the thermal image. This interval is expected to cover as much data as possible that belongs to the target distribution but the least possible data belonging to the counter distribution. Graphically, the less intersection between them, the better separability (**Fig.14 top** hatch patterns). **Fig.14 top** shows that the shallow delamination has better separability than the deep delamination, which is non-distinguishable due to the mix of distribution. On the other hand, under the one standard deviation interval, the coverage of data needs to be investigated for its representative power in terms of stability. **Fig.14 bottom** shows the data coverage under one standard deviation from the corresponding mean for three slabs. **Fig.14a bottom** shows that for shallow delamination, around 85% of the data was covered when using one standard deviation from the mean to represent the non-boundary gradient. For the boundary gradient, this interval became around 65% to 75% of coverage in **Fig.14a**. It is expected to become a stable choice when both coverages for the two distributions are high (e.g. 80% to 90% coverage) or at least one is high. Increasing the range of interval would improve the coverage, but this tends to intersect more between distributions. The optimal decision could be achieved through the receiver operating characteristics (ROC) analysis [38]. However, that analysis is beyond the scope of this paper. Instead, the author adopted the one standard deviation as an empirical selection of threshold. Based on this choice, stability for data representation of the three slabs could be observed: the slabs with the shallow (**Fig.14a**) and the deep (**Fig.14c**) delamination present acceptable stability, while the slab with median delamination (**Fig.14b**) exhibits instability before 2 p.m. Although the analytic result exhibits an observation that the temperature gradient at shallower delamination has better separability and stability than the deeper delamination in general, the thorough conclusion on the relationship between burying depth and temperature gradient at boundary requires further experimentations under the identical weather condition. However, the empirical threshold ranges could be recommended based on current analysis. For the shallow delamination, the value



ranging from 0.1 to 0.5 could be used and the value ranging from 0.1 to 0.63 could be used for the median delamination under a similar weather condition, when the slab is under the heating stage on a summer sunny day. For deep delamination, no feasible threshold value could be completely distinguishable under current setup.

## 6.2 Effect of Boundary Corner Shape on Gradient Analysis

Since we used a rounded corner square (**Fig.15a top**) rather than the sharp corner square (**Fig.15a bottom**) for the gradient analysis, the effects on the distinguishability and stability need to be discussed. We compared the normalized gradient strength (**Fig.15 top**) and data coverage percentage (**Fig.15 bottom**) for the two shapes with boundary width of 2.5 cm as an example. It is found that there is minimal effect both on gradient strength and data coverage for the non-boundary area (red solid and dashed lines) in terms of corner sharpness. When changing the round corner to the sharp one, the gradient strength of boundary area increases in all slabs across all time windows (compare the blue dashed line to the blue solid line), which makes a smaller difference of gradient strength between boundary and non-boundary areas. As mentioned in **section 4.1**, the larger difference of gradient strength between boundary and non-boundary areas, the higher the distinguish power is expected to be. Thus, using the sharp cornered ribbon for gradient analysis would decrease the distinguish power for threshold selection. Similar decreasing of coverage is observed in the case of the slab with shallow delamination (**Fig.15b bottom**). For the case of median deep delamination (**Fig.15c bottom**), the coverage shows no significant difference when changing the shape, and no clear trending of this changing could be easily observed. For the case of deep delamination (**Fig.15d bottom**), the coverage increases to 80% after 2 p.m. when changing round to sharp. The possible reason is that there may not be significant difference in the gradient strength either (**Fig.15d top**). Overall, the rounded corner shape is preferred; how rounded the corner should be is case-dependent, which requires future studies.

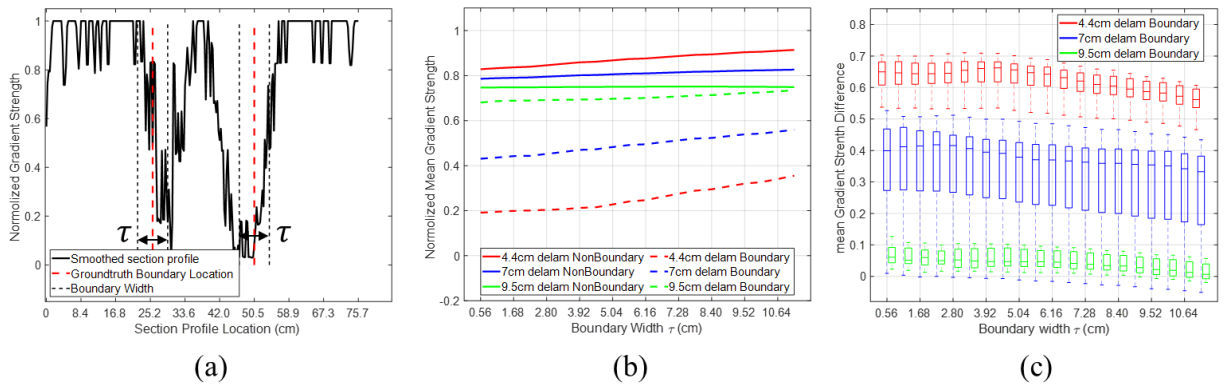


**Figure 15. Comparison of corner shapes for gradient analysis: (a) rounded corner on the top and square corner on the bottom; (b) the case of the slab with shallow delamination; (c) the case of the slab with median deep delamination; (d) the case of the slab with deep delamination.**

## 6.3 Effect of Boundary Width on Gradient Analysis

Another factor affecting the gradient strength is the width to be chosen to represent boundary gradient. To compare this effect, we average out the gradient strength across time for each depth delamination and evaluate them along different values of width in **Fig 16**. First, the section view of gradient strength from **Fig.5a** is shown in **Fig.16a**, where red dashed lines represent the exact locations of the delamination boundary. It is noticeable that the boundary is not laid on the strongest edge location (lowest peak values in **Fig.16a**), which motivates the authors to use a band to represent the gradient strength instead of a single value. The boundary width  $\tau$  ranges from 0.56 cm to 10.64 cm for the

analytical comparison in **Fig.16b**. It shows that the mean gradient strength increases as the width increases for both boundary and non-boundary areas. In addition, the boxplot of mean gradient difference (defined as non-boundary minus boundary gradient) is shown in **Fig.16c** for optimal width selection. For the shallow delamination (red in **Fig.16c**), there is an elbow point after 5.04 cm where the difference of gradient strength starts to drop, which indicates the width less than 5 cm could return a good threshold evaluation. For the case of median deep delamination (blue color in **Fig.16c**), a large variation of mean difference is observed (longer range of 25<sup>th</sup> and 75<sup>th</sup> percentile box), which indicates the mean gradient difference varies at different time windows. This suggests that a wider range of thresholds should be used to represent the boundary gradient strength. In addition to the width determination, there is no significant elbow point of dropping to be used, and thus the recommended width could follow the case of shallow delamination. For the case of deep delamination (green in **Fig.16c**), the mean gradient difference remains 0 to 0.1, which is too weak to exhibit the distinguish power. In this study, we used the 2.5 cm-width inward and outward the boundary, which sums 5 cm in width across all cases for simplified illustration. For the general usage for quantitative evaluation of gradient on delamination boundary, a similar analysis needs to be conducted.



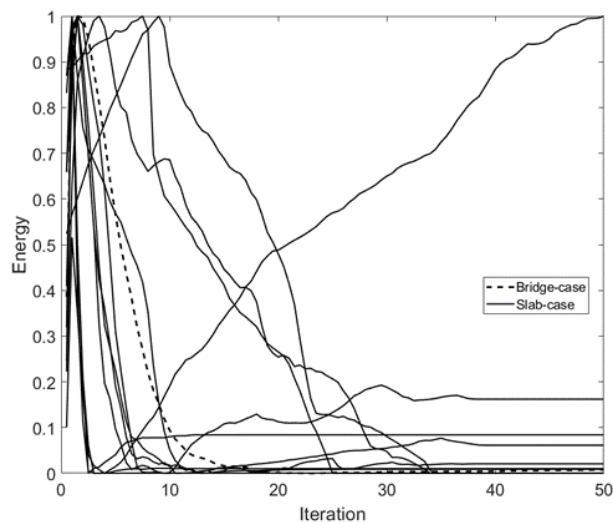
**Figure 16. Results of boundary effect on gradient analysis: (a) section view of gradient strength for boundary and non-boundary position; (b) mean gradient strength for different boundary widths; (c) mean gradient strength difference for different boundary widths.**

#### 6.4 Discussion of LSM Performance and Limitations

Because the level set approach has been used as the segmentation method for delamination boundary detection through thermography, the factors affecting its performance needs to be discussed. Because the edge based GAC model is used for curve evolution and constraints, the edge indicator function plays a significant role. Although the experimental study established a stable range of using thresholds, the present study only covered the condition in limited climate regions in the summer season. The variations above would affect the thermal dynamic that the heat transition of the delaminated area could respond to, which means further experimental evaluations are needed. Another factor affecting the gradient strength could be the size of delamination. Because the current study only evaluated the single sized delamination experimentally, the size variation needs to be further studied. The study conducted in Hiasa et al. [14,37] found that larger size of delamination would return higher temperature contrasts, and the contrast became stable when the size was larger than 40 cm by 40 cm. For the temperature gradient map, the strongest gradient occurred around the boundary instead of in the center of the delamination, and thus we could observe a ring-like pattern in the gradient map (**Fig.8**). The small gradient in the center of the delamination means slow temperature changes, which made the temperature contrast between the center of delamination and sound area constant. Thus, the delamination is too small to be detectable for both the contrast-based and the gradient-based method. Lastly, **Fig.17** shows the normalized total energy used by the level function during iteration for 12 slab cases and the bridge case. The energy is defined by the squared sum of  $\partial\phi/\partial t$  in **Eq.10**. All cases reached the convergence of energy which confirmed the stability of the proposed method. All cases converged rapidly within 50 iterations except the slab case with 9.5 cm deep delamination at 11 am, which took 380 iterations to converge (shown as the diagonally ascending line in **Fig.17**).



Several further improvements are needed in future studies to enhance this method's field performance in terms of robustness and accuracy. A thorough and complete experimental study is required to fully investigate and understand the effect of buried depth on the variation of temperature gradient. Due to the lack of large number of field samples the performance of the developed method in different type of bridge decks under different environmental conditions is yet to be further validated in further studies.



**Figure 17. The convergence of LSM for slab cases and bridge case: solid lines represent the total normalized energy used by the level set function for 12 slab cases and the dashed line represents the bridge case.**

## 7 Conclusions

This paper presented an edge based LSM to segment the delamination areas of the concrete pavement through thermography. A normalized temperature gradient map was generated based on the modified edge indicator function to represent the delamination boundary. The experimental studies showed that the delamination boundary could be profiled by the proposed gradient map through a simple one-standard deviation threshold. It recommends that under summer sunny weather condition, the thresholds ranging from 0.1 to 0.63 could be used to distinguish the temperate gradient at delamination boundary (shallow delamination of 4.4 cm deep and median delamination of 7 cm deep) for experimental studies during the heating stage. This value could be further adjusted when applying to the real bridge deck with concrete overlay due to the consideration of the strong gradient caused by non-delamination objects such as patches and road strips. The experimental study and field implementation results supported the applicability of the proposed method with improved performance compared to the well-accepted k-mean method. The current study is limited by the single-sized delamination evaluated in the experimental setup under the summer season as well as by the instability caused by the simple threshold used. Further work will also focus on the optimal threshold determination in terms of different area-to-depth ratios of delamination and deck layouts and evaluating the applicability on extended experimental studies and field implementations.

## ACKNOWLEDGMENT

The authors would like to thank the Nebraska Department of Transportation for their efforts in facilitating the data collection and sharing their non-destructive evaluation results.

## 8 References

- [1] S. Abu Dabous, S. Yaghi, S. Alkass, O. Moselhi, Concrete bridge deck condition assessment using IR Thermography and Ground Penetrating Radar technologies, *Autom. Constr.* (2017). doi:10.1016/j.autcon.2017.04.006.

- [2] S.H. Kee, T. Oh, J.S. Popovics, R.W. Arndt, J. Zhu, Nondestructive bridge deck testing with air-coupled impact-echo and infrared thermography, *J. Bridg. Eng.* (2012). doi:10.1061/(ASCE)BE.1943-5592.0000350.
- [3] S. Lin, D. Meng, H. Choi, S. Shams, H. Azari, Laboratory assessment of nine methods for nondestructive evaluation of concrete bridge decks with overlays, *Constr. Build. Mater.* (2018). doi:10.1016/j.conbuildmat.2018.08.127.
- [4] C. Maierhofer, R. Arndt, M. Röllig, R. Helmerich, A. Walther, B. Hillemeier, C. Rieck, Quantification of Voids and Delaminations in Real Concrete and Masonry Structures with Active Thermography: Case Studies, in: 2006. doi:10.21611/qirt.2006.049.
- [5] G. Washer, R. Fenwick, S. Nelson, R. Rumbayan, Guidelines for thermographic inspection of concrete bridge components in shaded conditions, *Transp. Res. Rec.* (2013). doi:10.3141/2360-02.
- [6] T. Oh, S.H. Kee, R.W. Arndt, J.S. Popovics, J. Zhu, Comparison of NDT methods for assessment of a concrete bridge deck, *J. Eng. Mech.* (2013). doi:10.1061/(ASCE)EM.1943-7889.0000441.
- [7] T. Omar, M.L. Nehdi, Remote sensing of concrete bridge decks using unmanned aerial vehicle infrared thermography, *Autom. Constr.* (2017). doi:10.1016/j.autcon.2017.06.024.
- [8] I. Abdel-Qader, S. Yohali, O. Abudayyeh, S. Yehia, Segmentation of thermal images for non-destructive evaluation of bridge decks, *NDT E Int.* (2008). doi:10.1016/j.ndteint.2007.12.003.
- [9] A. Ellenberg, A. Kontsos, F. Moon, I. Bartoli, Bridge deck delamination identification from unmanned aerial vehicle infrared imagery, *Autom. Constr.* (2016). doi:10.1016/j.autcon.2016.08.024.
- [10] S.S. Suganthi, S. Ramakrishnan, Anisotropic diffusion filter based edge enhancement for segmentation of breast thermogram using level sets, *Biomed. Signal Process. Control.* (2014). doi:10.1016/j.bspc.2014.01.008.
- [11] S. Zhou, J. Wang, S. Zhang, Y. Liang, Y. Gong, Active contour model based on local and global intensity information for medical image segmentation, *Neurocomputing.* (2016). doi:10.1016/j.neucom.2015.12.073.
- [12] L. Wang, C. Li, Q. Sun, D. Xia, C.Y. Kao, Active contours driven by local and global intensity fitting energy with application to brain MR image segmentation, *Comput. Med. Imaging Graph.* (2009). doi:10.1016/j.compmedimag.2009.04.010.
- [13] C. Li, R. Huang, Z. Ding, J.C. Gatenby, D.N. Metaxas, J.C. Gore, A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI, *IEEE Trans. Image Process.* (2011). doi:10.1109/TIP.2011.2146190.
- [14] S. Hiasa, R. Birgul, F.N. Catbas, Effect of Defect Size on Subsurface Defect Detectability and Defect Depth Estimation for Concrete Structures by Infrared Thermography, *J. Nondestruct. Eval.* (2017). doi:10.1007/s10921-017-0435-3.
- [15] G. Washer, R. Fenwick, N. Bolleni, J. Harper, Effects of Environmental Variables on Infrared Imaging of Subsurface Features of Concrete Bridges, *Transp. Res. Rec. J. Transp. Res. Board.* (2009). doi:10.3141/2108-12.
- [16] A. Watase, R. Birgul, S. Hiasa, M. Matsumoto, K. Mitani, F.N. Catbas, Practical identification of favorable time windows for infrared thermography for concrete bridge evaluation, *Constr. Build. Mater.* (2015). doi:10.1016/j.conbuildmat.2015.10.156.
- [17] C. Cheng, R. Na, Z. Shen, Thermographic Laplacian-pyramid filtering to enhance delamination detection in concrete structure, *Infrared Phys. Technol.* 97 (2019). doi:10.1016/j.infrared.2018.12.039.
- [18] X. Han, C. Xu, J.L. Prince, A topology preserving level set method for geometric deformable models, *IEEE Trans. Pattern Anal. Mach. Intell.* (2003). doi:10.1109/TPAMI.2003.1201824.
- [19] T.F. Chan, L.A. Vese, Active contours without edges, *IEEE Trans. Image Process.* (2001). doi:10.1109/83.902291.
- [20] K. Zhang, L. Zhang, H. Song, D. Zhang, Reinitialization-free level set evolution via reaction diffusion, *IEEE Trans. Image Process.* (2013). doi:10.1109/TIP.2012.2214046.
- [21] M. Wan, G. Gu, J. Sun, W. Qian, K. Ren, Q. Chen, X. Maldague, A Level Set Method for Infrared Image Segmentation Using Global and Local Information, *Remote Sens.* 10 (2018) 1039.
- [22] V. Caselles, R. Kimmel, G. Sapiro, Geodesic Active Contours, *Int. J. Comput. Vis.* (1997). doi:10.1023/A:1007979827043.
- [23] Y. Huang, C. Te Chiu, M.G. Lee, S.Y. Lin, Applying the multilayer level set approach to explore the thermal surface characteristics of hot mix asphalt, *Constr. Build. Mater.* (2014). doi:10.1016/j.conbuildmat.2013.10.097.
- [24] Y. Huang, J.W. Wu, Infrared thermal image segmentations employing the multilayer level set method

- for non-destructive evaluation of layered structures, *NDT E Int.* (2010). doi:10.1016/j.ndteint.2009.08.001.
- [25] C. Xu, J.L. Prince, Snakes, shapes, and gradient vector flow, *IEEE Trans. Image Process.* (1998). doi:10.1109/83.661186.
- [26] G. Huang, H. Ji, W. Zhang, A fast level set method for inhomogeneous image segmentation with adaptive scale parameter, *Magn. Reson. Imaging.* (2018). doi:10.1016/j.mri.2018.05.011.
- [27] C. Leys, C. Ley, O. Klein, P. Bernard, L. Licata, Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, *J. Exp. Soc. Psychol.* (2013). doi:10.1016/j.jesp.2013.03.013.
- [28] Q. Zhang, L. Xu, J. Jia, 100+ times faster weighted median filter (WMF), in: *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2014. doi:10.1109/CVPR.2014.362.
- [29] O. Vincent, O. Folorunso, A Descriptive Algorithm for Sobel Image Edge Detection, in: *Proc. 2009 InSITE Conf.*, 2017. doi:10.28945/3351.
- [30] J. Weickert, *Anisotropic Diffusion in Image Processing*, 1998.
- [31] S. Anam, E. Uchino, N. Suetake, Image boundary detection using the modified level set method and a diffusion filter, in: *Procedia Comput. Sci.*, 2013. doi:10.1016/j.procs.2013.09.095.
- [32] C. Cheng, Z. Shang, Z. Shen, Bridge deck delamination segmentation based on aerial thermography through regularized grayscale morphological reconstruction and gradient statistics, *Infrared Phys. Technol.* (2019). doi:10.1016/j.infrared.2019.03.018.
- [33] H. Yu, Z. Yang, L. Tan, Y. Wang, W. Sun, M. Sun, Y. Tang, Methods and datasets on semantic segmentation: A review, *Neurocomputing.* (2018). doi:10.1016/j.neucom.2018.03.037.
- [34] *Bridge Office Policies and Procedures (BOPP) Manual*, Nebraska Dep. Transp. (2016).
- [35] H.J. Diamond, T.R. Karl, M.A. Palecki, C.B. Baker, J.E. Bell, R.D. Leeper, D.R. Easterling, J.H. Lawrimore, T.P. Meyers, M.R. Helfert, G. Goodge, P.W. Thorne, U.S. climate reference network after one decade of operations status and assessment, *Bull. Am. Meteorol. Soc.* (2013). doi:10.1175/BAMS-D-12-00170.1.
- [36] G. Washer, R. Fenwick, N. Bolleni, Effects of solar loading on infrared imaging of subsurface features in concrete, *J. Bridg. Eng.* (2010). doi:10.1061/(ASCE)BE.1943-5592.0000117.
- [37] S. Hiasa, R. Birgul, F.N. Catbas, Investigation of effective utilization of infrared thermography (IRT) through advanced finite element modeling, *Constr. Build. Mater.* (2017). doi:10.1016/j.conbuildmat.2017.05.175.
- [38] T. Fawcett, An introduction to ROC analysis, *Pattern Recognit. Lett.* (2006). doi:10.1016/j.patrec.2005.10.010.
- [39] T. Omar, M.L. Nehdi, T. Zayed, Infrared thermography model for automated detection of delamination in RC bridge decks, *Constr. Build. Mater.* (2018). doi:10.1016/j.conbuildmat.2018.02.126.
- [40] A.A. Sultan, G. Washer, A pixel-by-pixel reliability analysis of infrared thermography (IRT) for the detection of subsurface delamination, *NDT E Int.* (2017). doi:10.1016/j.ndteint.2017.08.009.