Journal of Future Artificial Intelligence and Technologies

E-ISSN: 3048-3719

Research Article

A Lightweight Distance-Aware Loss for Thin Crack Segmentation in Building Facades under Limited-Data Conditions

Edona Krasniqi * and **Visar Shehu**

Faculty of Contemporary Sciences and Technologies, South East European University, 1200 Tetovo, North Macedonia; e-mail: ek08040@seeu.edu.mk; v.shehu@seeu.edu.mk

* Corresponding Author: Edona Krasniqi

Abstract: Thin crack segmentation in building facades is particularly challenging under limited data conditions due to the extremely narrow geometry and weak contrast of crack structures. Conventional overlap-based loss functions, such as Binary Cross-Entropy (BCE) and Dice loss, optimize pixel-wise agreement but do not explicitly account for spatial boundary relevance, often resulting in fragmented predictions and geometric misalignment. This study introduces a lightweight distance-aware weighting extension of Dice loss designed to improve boundary alignment without modifying the network architecture. The proposed approach integrates Euclidean distance information derived from ground-truth masks to assign higher importance to prediction errors near crack boundaries while reducing the influence of distant background regions. The method is evaluated on a real-world dataset of 108 facade images (87 training and 21 validation) using a standard U-Net architecture under identical training conditions. Experimental results demonstrate a consistent reduction in geometric boundary error. The 95th percentile Hausdorff Distance (HD95) decreases from 230.08 px with BCE and 217.55 px with Dice loss to 148.28 px with the proposed distance-aware formulation, corresponding to reductions of approximately 81.8 px and 69.3 px, respectively. In addition, the proposed loss improves overlap-based metrics, achieving IoU@0.1 = 0.2844 and Dice = 0.4087 on the validation set. These results indicate that incorporating spatial distance information into the optimization objective improves geometric alignment and structural continuity of thin crack predictions. The findings suggest that integrating lightweight distance-aware weighting into conventional loss formulations can improve segmentation quality for thin structures in constrained-data scenarios while maintaining computational simplicity.

Keywords: Boundary-aware loss; Crack detection; Distance-aware loss; Hausdorff distance; Infrastructure inspection; Low-data learning; Thin structure segmentation; U-net segmentation.

Received: February, 13th 2026
Revised: March, 7th 2026
Accepted: March, 9th 2026
Published: March, 15th 2026
Curr. Ver.: March, 15th 2026

1. Introduction

Thin crack segmentation in building facades is a challenging task in semantic image analysis, particularly when only a limited amount of labeled data is available. Thin cracks are narrow, elongated structures that often exhibit weak contrast against the surrounding surface. In many cases, they occupy only a very small fraction of the image area, which makes accurate localization difficult. As a result, even small prediction errors can break crack continuity or cause important segments to be missed. Segmenting thin cracks is therefore significantly more demanding than segmenting large and compact objects commonly addressed in conventional semantic segmentation tasks [1]–[3].

From an engineering perspective, accurate crack segmentation is important for structural inspection and maintenance planning. Crack patterns may reveal information about material degradation, stress concentration, and potential safety risks in civil infrastructure. Automated crack segmentation can therefore support faster and more objective inspection processes in routine maintenance and post-disaster assessments [4], [5]. Compared with simple image

classification or bounding-box detection, semantic segmentation provides pixel-level localization that enables detailed analysis of crack length, connectivity, and spatial distribution. For thin cracks, preserving geometric alignment and structural continuity is often more critical than maximizing pixel-wise overlap.

Deep learning models, particularly encoder–decoder architectures such as U-Net, have become a standard approach for semantic segmentation tasks [6]. U-Net combines global contextual information from deeper layers with spatial details from shallower layers through skip connections, enabling the network to capture both semantic and fine-grained structural features [7]. Variants of U-Net and related convolutional neural network architectures have demonstrated strong performance in applications including medical imaging, road extraction, and crack detection [1], [8], [9]. Nevertheless, extremely thin structures remain difficult to segment accurately. Downsampling operations in deeper layers may blur fine details, while shallow layers can be sensitive to noise and complex surface textures. Consequently, thin cracks may appear fragmented or partially suppressed in predicted segmentation masks.

In addition to architectural design, the choice of loss function plays a critical role in segmentation performance. Most segmentation networks are trained using overlap-based losses such as Binary Cross-Entropy (BCE) and Dice loss [10], [11]. BCE treats segmentation as a pixel-wise classification task and penalizes misclassified pixels independently, whereas Dice loss directly optimizes the overlap between predicted and ground-truth regions and partially mitigates class imbalance. While these loss functions are effective for large and compact objects, they treat all pixel errors equally regardless of their spatial location [12].

For thin crack segmentation, this uniform treatment of pixel errors can be problematic. Small boundary shifts or slight misalignments may not significantly reduce Dice or Intersection over Union (IoU) scores, yet they can produce visible discontinuities in crack structures. Conversely, predictions that are geometrically close to the ground truth may still be penalized strongly even when the spatial deviation is small. As a result, optimizing purely overlap-based metrics does not necessarily guarantee good geometric alignment for thin structures.

To address these limitations, several boundary-aware and distance-based loss formulations have been proposed. Boundary loss functions explicitly penalize contour misalignment between predicted and ground-truth masks [13], while Hausdorff-based losses incorporate distance-based evaluation metrics into the optimization process to reduce extreme boundary deviations [14]. More generally, distance-transform-based approaches weight pixel errors according to their spatial proximity to object boundaries, allowing models to focus more strongly on structurally relevant regions [15]. Although these approaches introduce geometric information into the optimization process and have demonstrated improved boundary accuracy in segmentation tasks, they often increase computational complexity and require careful tuning of multiple hyperparameters. In limited-data scenarios, complex loss formulations may also lead to unstable training behavior.

The challenges become more pronounced when only a small dataset is available. Pixel-level annotation of cracks is time-consuming and requires careful manual labeling. Cracks may also fade into textured or painted backgrounds, making annotation uncertain in some regions. In practical inspection scenarios, datasets are typically small and domain-specific. Under such conditions, increasing network depth or introducing complex regularization strategies may lead to overfitting rather than improved generalization [16], [17]. These constraints motivate the exploration of lightweight strategies that enhance spatial sensitivity within existing architectures without substantially increasing model complexity.

From a geometric perspective, thin crack segmentation differs from traditional object segmentation tasks. The goal is not to recover a large connected region but to accurately trace narrow crack trajectories across the facade surface. Maintaining continuity along these trajectories is essential for structural interpretation. Distance-based evaluation metrics such as the 95th percentile Hausdorff Distance (HD95) provide a more informative measure of boundary deviation than overlap-based metrics [18]. HD95 emphasizes extreme geometric differences between predicted and ground-truth boundaries while reducing sensitivity to outliers. In crack analysis, this metric captures spatial errors that may be meaningful in practice even when IoU or Dice scores remain relatively stable.

These observations raise an important question: can simple spatial weighting improve geometric alignment without introducing a complex loss formulation? Instead of designing highly sophisticated boundary-aware objectives, it may be sufficient to incorporate minimal

distance-based information into an existing loss function. Such an approach would preserve computational simplicity while reducing the risk of unstable optimization in small datasets.

In this work, we propose a lightweight distance-aware extension of Dice loss designed specifically for thin crack segmentation under limited-data conditions. The proposed method incorporates a Euclidean distance transform derived from the ground-truth mask to weight pixel-wise errors according to their spatial proximity to crack boundaries. Pixels located near crack structures receive higher weights during loss computation, whereas distant background pixels contribute less. This weighting mechanism increases the influence of geometrically critical regions during training while preserving the simplicity of the underlying loss formulation.

A key design principle of this study is controlled experimentation. To isolate the effect of the loss function, we employ a standard U-Net architecture without architectural modification. The network structure, optimizer, preprocessing pipeline, and training protocol remain identical across all experiments. This controlled setup ensures that observed performance differences arise solely from the choice of loss function rather than from architectural changes. The proposed method is evaluated on a real-world dataset consisting of 108 facade images containing predominantly thin hairline cracks and several moderately thicker cracks. All images are manually annotated at the pixel level. The dataset is divided into 87 training images and 21 validation images, reflecting a constrained-data scenario typical of practical inspection environments. Performance is evaluated using both overlap-based metrics, including IoU, and distance-based metrics, specifically the HD95.

Experimental results demonstrate that distance-aware weighting consistently reduces extreme boundary deviations compared with BCE and standard Dice loss. Although improvements in overlap-based metrics are modest, the reduction in HD95 indicates improved geometric alignment and better preservation of crack continuity. The contributions of this study are threefold.

- First, we introduce a lightweight distance-aware extension of dice loss tailored to thin crack segmentation under limited-data conditions.
- Second, we conduct a controlled experimental analysis that isolates the impact of the loss formulation while maintaining a fixed network architecture.
- Third, we provide quantitative and qualitative evidence demonstrating improved geometric alignment of thin crack predictions, as measured by the HD95 metric.

Together, these contributions highlight the importance of aligning the training objective with the geometric characteristics of thin structures, particularly in data-constrained structural inspection applications.

2. Related Work

2.1. Thin Structure Segmentation

Segmenting thin structures such as cracks, blood vessels, and road networks has long been recognized as a challenging problem in computer vision due to their narrow width, elongated geometry, and sensitivity to noise and low contrast [3], [6]. Unlike compact objects, thin structures occupy only a small fraction of the image area and are often visually similar to background textures. As a result, they are particularly vulnerable to fragmentation and false negatives during segmentation. Even minor prediction errors may disrupt structural continuity, producing broken or incomplete representations that are unsuitable for practical applications.

Traditional segmentation models often struggle to preserve the continuity of thin structures, resulting in fragmented predictions or missing segments, particularly in complex and textured backgrounds [1]. These difficulties are further exacerbated by variations in illumination, surface material, and viewpoint, which commonly occur in real-world imaging scenarios such as facade inspection. Consequently, segmentation quality for thin structures cannot always be reliably assessed using region-based overlap alone, since disconnected predictions may still achieve moderate overlap scores.

Encoder–decoder architectures, most notably U-Net, have become a standard choice for thin structure segmentation tasks due to their ability to retain spatial detail through skip connections [19]. By combining low-level spatial features with high-level semantic information, U-Net mitigates the loss of fine-grained details caused by repeated downsampling. U-Net and its variants have been successfully applied to crack detection [4], [5], vessel

segmentation [20], and road extraction [8], demonstrating their versatility across domains where thin structures are prominent.

Despite these advances, the performance of such models remains strongly dependent on the availability of large and well-annotated datasets [16]. Thin structure segmentation typically requires pixel-level annotations, which are expensive and time-consuming to produce. Consequently, many real-world applications rely on relatively small datasets that exhibit considerable variability and annotation noise. Under such conditions, even well-established architectures may struggle to generalize effectively, producing incomplete or noisy segmentation outputs.

Low-data scenarios are particularly common in practical structural inspection and post-disaster assessment settings. In such environments, segmentation models often exhibit limited generalization capability, leading to fragmented or inconsistent predictions [17]. This limitation highlights the need for strategies that improve learning efficiency and robustness without increasing model complexity. Rather than relying solely on architectural modifications, alternative approaches that directly address the characteristics of thin structures during optimization may provide a more practical solution.

2.2. Loss Functions for Semantic Segmentation

Loss functions play a dDice scores while still producing geometrically inaccurate or fragmented predictions [21]. This discrepancy highlights a fundamental limitation of overlap-based optimization when applied to thin structure segmentation.

Several studies have therefore emphasized that overlap-based metrics and losses are insufficient for capturing geometric fidelity in thin and elongated objects [18]. Instead, segmentation quality should be assessed not only by pixel-wise agreement but also by how closely the predicted structure aligns geometrically with the ground truth. This observation motivates the development of loss functions that incorporate spatial and geometric information directly into the optimization process.

2.3. Distance-Based and Boundary-Aware Approaches

To overcome the limitations of traditional loss functions, recent research has introduced boundary-aware, shape-aware, and distance-based loss formulations. Boundary loss functions explicitly penalize misalignment between predicted and ground-truth boundaries, improving contour accuracy and sensitivity to object outlines [13]. Shape-aware losses incorporate global geometric constraints to preserve object morphology and prevent unrealistic predictions [22].

Distance transform–based approaches leverage spatial distance maps derived from ground-truth masks to weight pixel-wise errors according to their proximity to object boundaries [14], [15]. In these methods, errors near the target structure contribute more strongly to the loss, while distant background pixels are down-weighted. Such formulations have demonstrated improved boundary precision in medical imaging and remote sensing applications [23], where accurate delineation of anatomical or geographic structures is essential.

Despite their promise, many distance-based approaches have been evaluated primarily on datasets with abundant training data or on objects with relatively thick structures [24]. Their effectiveness for extremely thin structures under limited-data conditions remains insufficiently explored. In such regimes, distance weighting may interact differently with optimization dynamics, and its potential benefits are not yet fully understood.

Several recent studies have emphasized the importance of preserving structural continuity and topology when segmenting thin and elongated objects, particularly in scenarios where pixel-wise accuracy alone fails to reflect practical segmentation quality [25], [26]. These works highlight that segmentation outputs must maintain connectivity and shape consistency in order to remain useful in downstream analysis and decision-making tasks. In addition, comprehensive reviews of segmentation loss functions indicate that geometry-aware, boundary-aware, and distance-based optimization strategies consistently outperform traditional overlap-based formulations when applied to highly imbalanced or structurally complex targets [27]–[29]. These findings reinforce the notion that loss function design plays a critical role in determining segmentation quality, especially for challenging object classes.

Related work in remote sensing has also demonstrated that segmentation of linear and elongated structures, such as roads, benefits from architectures and optimization strategies that explicitly account for object geometry rather than relying solely on region overlap [30].

These insights are directly applicable to crack segmentation, where the target structures exhibit similar geometric properties. To address data scarcity, transfer learning and training with sparse annotations have been widely adopted, enabling robust performance even when only limited labeled samples are available [19], [31], [32]. Such strategies improve feature reuse and convergence but do not fundamentally address the mismatch between overlap-based objectives and geometric accuracy. Finally, advances in network design, including residual learning, have contributed to more stable optimization and improved convergence in deep segmentation models, indirectly supporting more reliable learning of fine-grained structures [32]. However, architectural improvements alone cannot fully compensate for loss functions that are misaligned with task-specific geometric requirements.

A summary of representative approaches for thin structure segmentation is presented in Table 1, highlighting their optimization strategies, limitations, and relation to the proposed method. In this work, we build upon these ideas by proposing a simple and computationally efficient distance-aware loss function explicitly designed for thin structure segmentation in low-data scenarios. Unlike prior approaches, the proposed method focuses on improving structural continuity and reducing extreme boundary deviations, as quantified by distance-based evaluation metrics such as the HD95. By isolating the effect of the loss function in a controlled experimental setting, this study aims to clarify the role of distance-aware optimization in achieving geometrically meaningful segmentation of thin structures under limited-data conditions.

Table 1. Representative approaches for thin structure segmentation.

Category	Representative Works	Main Idea	Limitations	Relation to This Work
Thin structure segmentation	[1], [3], [6]	Early CNN-based methods for crack and thin object segmentation	Fragmented predictions and weak boundary preservation	Motivates the need for geometry-aware optimization
Encoder–decoder architectures (U-Net)	[4], [5], [8], [19], [20]	Preserve spatial detail through skip connections	Strong dependence on large annotated datasets	We adopt U-Net but do not modify the architecture
Low-data segmentation	[16], [17]	Data augmentation and transfer learning to mitigate data scarcity	Does not address geometric loss mismatch	Our method targets loss design rather than increasing data volume
Overlap-based loss functions (BCE, Dice)	[10], [11], [21], [33]	Optimize pixel-wise accuracy and region overlap	Ignore spatial and boundary relevance	Baseline methods in our experiments
Limitations of overlap-based metrics	[18]	IoU and Dice are insufficient for thin structures	Poor reflection of geometric errors	Justifies the use of HD95 in evaluation
Boundary-aware loss functions	[13], [22]	Explicit boundary penalization and shape constraints	Additional complexity and parameter tuning required	Our loss formulation is simpler and computationally efficient
Distance-based loss functions	[14], [15]	Use distance transform to weight pixel errors	Mostly tested on thick structures or large datasets	We validate distance-aware loss in an extreme low-data regime
Applications in medical imaging and remote sensing	[23], [24], [30]	Geometry-aware optimization improves boundary precision	Domain-specific assumptions	Confirms the general relevance of distance-based losses
Reviews and surveys	[27]–[29]	Geometry-aware losses outperform overlap-based ones	Limited focus on low-data thin structures	Our work addresses this gap
Transfer learning and sparse supervision	[19], [23], [31]	Improve convergence with limited labeled data	Loss functions remain overlap-driven	Complementary but orthogonal to our approach

3. Proposed Distance-Aware Loss Function

This section describes the proposed distance-aware loss formulation designed to improve segmentation of thin structures such as cracks. The formulation aims to introduce spatial sensitivity into the training objective while preserving the simplicity of conventional loss functions. The key idea is to assign higher importance to prediction errors occurring near thin structures and lower importance to distant background pixels.

3.1. Problem Formulation

Let

$$X \in \mathbb{R}^{H \times W} \tag{1}$$

denote an input image with spatial resolution $H \times W$, where H and W represent image height and width, respectively.

Let

$$Y \in \{0,1\}^{H \times W} \tag{2}$$

denote the corresponding ground-truth segmentation mask. In this mask, pixels labeled with value 1 represent foreground structures (e.g., cracks), while pixels labeled with 0 represent the background.

Given a segmentation model $f_\theta(\cdot)$ parameterized by θ, the predicted probability map is defined as

$$P = f_\theta(X), \quad P \in [0,1]^{H \times W} \tag{3}$$

where each element p_i represents the predicted probability that pixel i belongs to the crack class.

The goal of model training is to learn parameters θ such that the predicted segmentation map aligns with the ground truth not only at the pixel level but also in terms of geometric structure. For thin structures, this means preserving spatial continuity and minimizing boundary deviations.

3.2. Limitations of Standard Loss Functions

Most segmentation models are trained using overlap-based loss functions such as BCE or Dice loss. These losses measure prediction accuracy by comparing predicted probabilities with ground-truth labels at each pixel location [10]. For example, the BCE loss is defined as

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)] \tag{4}$$

where N denotes the total number of pixels, y_i is the ground-truth label of pixel i, and p_i is the predicted probability.

Although such formulations work well for large and compact objects, they treat all pixel errors equally regardless of spatial location. This assumption becomes problematic when segmenting thin structures. For instance, small boundary deviations may cause visible breaks in crack continuity even when the overall overlap score remains high. Therefore, optimizing standard overlap-based losses does not necessarily guarantee geometrically accurate predictions for thin structures.

3.3. Distance Transform of the Ground Truth

To introduce spatial awareness into the training objective, we compute a distance transform based on the ground-truth mask [15]. Let $D(i)$ denote the Euclidean distance from pixel i to the nearest foreground pixel in the ground-truth mask. The distance transform is defined as

$$D(i) = \min_{j \in Y_{fg}} \|i - j\| \tag{5}$$

where Y_{fg} represents the set of foreground pixels (i.e., crack pixels) in the ground-truth mask.

In this formulation, pixels belonging to crack structures have distance value $D(i) = 0$. Background pixels have increasing distance values as their spatial distance from the crack increases. The resulting distance map encodes spatial information about the proximity of each pixel to the crack structure. An example of the distance transform derived from a crack mask is illustrated in Figure 1. The visualization shows how distance values gradually increase away from crack boundaries.

3.4. Distance-Based Weighting Function

Using the distance map $D(i)$, we define a spatial weighting function that determines how strongly each pixel contributes to the training loss [14], [15]. The weighting function is defined as

$$w(i) = \exp\big(-\alpha D(i)\big) \tag{6}$$

where $\alpha > 0$ is a scaling parameter that controls the rate of exponential decay.

This formulation assigns higher weights to pixels located close to crack structures, lower weights to pixels located farther away in the background. As a result, prediction errors near crack boundaries contribute more strongly to the optimization objective. This encourages the model to focus on geometrically critical regions where segmentation accuracy is most important.

For clarity, Figure 1 illustrates the entire weighting mechanism, including the input image, the ground-truth mask, the computed distance map, and the resulting weight map. Pixels located near crack boundaries receive higher weights, while distant background pixels gradually receive smaller weights.

3.5. Distance-Aware Loss Formulation

Using the spatial weights defined above, the proposed distance-aware loss function is constructed by weighting a standard pixel-wise loss [12], [25]. The distance-aware loss is defined as

$$L_{DA} = \frac{1}{N}\sum_{i=1}^{N} w(i) \cdot L_{base}(y_i, p_i) \tag{7}$$

where $w(i)$ is the spatial weight derived from the distance transform, L_{base} represents a conventional loss function such as BCE or Dice loss, N denotes the total number of pixels.

This formulation maintains the simplicity of traditional loss functions while introducing spatial sensitivity to thin structures. Instead of redesigning a completely new objective, the proposed method modifies the contribution of each pixel according to its geometric relevance.

3.6. Optimization Perspective

During backpropagation, the gradient of the proposed loss with respect to the network parameters is scaled by the distance-based weights:

$$\frac{\partial L_{DA}}{\partial \theta} = w(i) \cdot \frac{\partial L_{base}}{\partial \theta} \tag{8}$$

This formulation indicates that gradients originating from pixels near crack structures are amplified, while gradients from distant background pixels are suppressed. Consequently, the model focuses more strongly on learning accurate representations of thin structures. From an optimization perspective, this mechanism acts as a form of spatially adaptive weighting that emphasizes geometrically important regions without altering the underlying network architecture.

3.7. Relation to Distance-Based Evaluation Metrics

The proposed loss formulation is closely related to distance-based evaluation metrics commonly used to assess segmentation quality. One widely used metric is the HD95, which measures extreme boundary deviations between predicted and ground-truth masks. Unlike overlap-based metrics such as IoU or Dice score, HD95 explicitly quantifies geometric misalignment between predicted and reference boundaries.

By emphasizing prediction errors near structural boundaries during training, the proposed loss encourages the model to reduce spatial deviations along thin structures. This

behavior aligns naturally with the objective measured by HD95 during evaluation. Consequently, the proposed distance-aware formulation is expected to improve geometric alignment and structural continuity of thin crack predictions, particularly in scenarios where only limited training data are available.

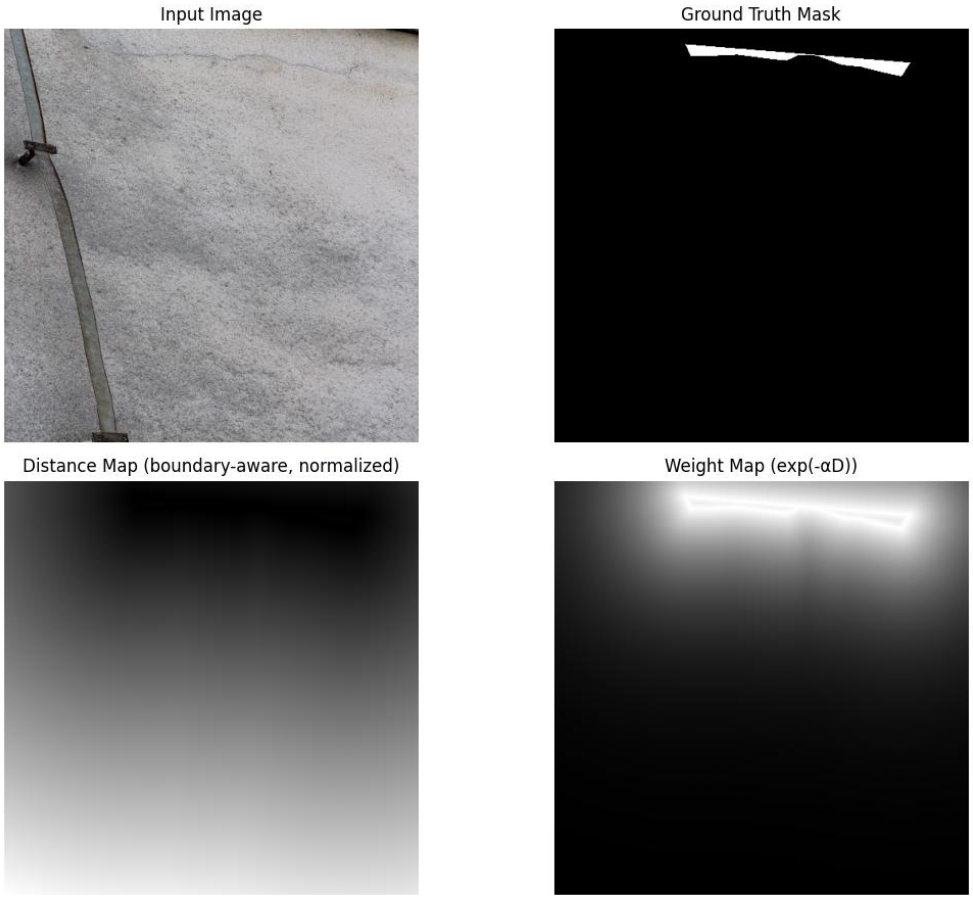


Figure 1. Visualization of the proposed distance-aware weighting mechanism. From left to right: input facade image, binary crack mask, boundary-aware normalized distance map, and the corresponding exponential weight map. Pixels closer to crack boundaries receive higher weights during training.

4. Experimental Setup

4.1. Dataset and Low-Data Scenario

The experimental evaluation was conducted on a privately collected dataset of real-world building facade images acquired in Ferizaj on January 12, 2026, using an Apple iPhone 17 Pro Max. Both the telephoto camera (100 mm) and the standard 24 mm camera were used depending on the distance from the facade. The dataset consists of 108 images captured at resolutions of 12 MP and 24 MP. All photographs were taken on the same day under natural daylight conditions. Although illumination conditions were generally consistent, minor variations occurred due to localized shadowing effects on the facade surface.

The facade material is largely consistent across the dataset, consisting primarily of painted exterior surfaces. The primary source of variability arises from crack morphology rather than surface material. The dataset contains predominantly thin hairline cracks, as well as several moderately thicker cracks. Crack structures vary in length, continuity, and local branching behavior. Due to the extremely small width of hairline cracks relative to the image resolution, crack pixels occupy only a minimal fraction of the total image area, resulting in severe class imbalance. This characteristic significantly increases the difficulty of segmentation, particularly for overlap-based optimization objectives.

All images were manually annotated at the pixel level to produce binary segmentation masks. The dataset was divided into 87 training images and 21 validation images, reflecting a constrained-data scenario typical of practical inspection settings. For computational consistency and fair comparison across loss functions, all images were resized to 512 × 512 pixels prior to training. The same preprocessing pipeline was applied to all experimental configurations. Although resizing changes the absolute pixel representation of crack thickness, it does not introduce relative bias between methods because preprocessing was strictly identical for all models. During training, standard data augmentation was applied, resulting in 260 training samples, while the number of unique real-world images remained 108. Representative examples from the dataset are shown in Figure 2, illustrating variations in crack thickness and the influence of local shadowing effects on the facade surface.

Figure 2. Representative facade images illustrating variations in crack thickness and local shadowing effects within the dataset.

4.2. Network Architecture

All experiments employ the standard U-Net architecture to ensure a fair and reproducible comparison between loss functions. The network follows the classical encoder–decoder design consisting of four encoder levels and a symmetric decoder with skip connections that preserve spatial information during upsampling. The encoder progressively increases the number of feature channels (64, 128, 256, and 512) while reducing spatial resolution through max-pooling operations. Each encoder level contains two 3×3 convolutional layers followed by ReLU activation functions, allowing the network to capture increasingly complex feature representations at deeper levels.

The decoder mirrors the encoder structure by progressively restoring spatial resolution through upsampling layers. At each decoding stage, feature maps from the corresponding encoder level are concatenated via skip connections. These skip connections help preserve fine structural information that might otherwise be lost during repeated downsampling operations. The final layer applies a 1×1 convolution followed by a sigmoid activation function to produce a pixel-wise probability map representing the predicted crack segmentation. No architectural modifications were introduced between experiments. Consequently, all observed performance differences arise solely from the choice of loss function rather than from changes in network design.

4.3. Training Configuration

All images were resized to a fixed spatial resolution before training. Models were trained using the Adam optimizer with a constant learning rate and a batch size adapted to the dataset size. Three loss functions were evaluated under identical training conditions, i.e., BCE, Dice loss, and proposed distance-aware loss. This experimental setup ensures that any observed performance differences can be attributed directly to the loss formulation rather than changes in architecture or training configuration. All experiments were implemented in PyTorch and executed on an NVIDIA T4 GPU to maintain consistent training conditions across all evaluated loss functions.

4.4. Evaluation Metrics

Model performance was evaluated using a combination of overlap-based and distance-based metrics dice coefficient, IoU, and HD95. Using multiple evaluation metrics allows the analysis of both pixel-level agreement and geometric boundary accuracy.

The Dice coefficient measures the degree of overlap between the predicted segmentation mask and the ground-truth mask. It is defined as the harmonic mean of precision and recall for the foreground class. Dice is particularly useful in segmentation tasks with class imbalance because it directly measures how well the predicted region overlaps with the true structure.

The IoU metric, also known as the Jaccard index, measures the ratio between the intersection and the union of the predicted and ground-truth regions. IoU provides a strict measure of region overlap and is commonly used as a standard evaluation metric in semantic segmentation benchmarks.

Although Dice and IoU capture overall region agreement, they are not always sufficient for evaluating thin structures. Small spatial shifts along crack boundaries may have only a minor impact on overlap scores while still producing visually significant geometric errors.

To address this limitation, the HD95 is used as a distance-based metric. HD95 measures the spatial deviation between predicted and ground-truth boundaries by computing the 95th percentile of pairwise boundary distances. Compared with the standard Hausdorff distance, HD95 reduces sensitivity to extreme outliers while still capturing significant boundary misalignments. In the context of thin crack segmentation, HD95 provides a more informative measure of geometric accuracy than overlap-based metrics alone. By combining Dice, IoU, and HD95, the evaluation captures both region-level agreement and boundary-level precision.

4.5. Hyperparameter Configuration (β and σ)

The proposed distance-aware loss introduces two hyperparameters: β, which controls the relative contribution of the distance-weighted term, and σ, which determines the spatial decay rate of the distance transform weighting. Parameter β regulates how strongly pixels near crack structures influence the loss computation. Larger values increase the contribution of boundary-adjacent pixels during optimization, whereas smaller values reduce the effect of distance-based weighting.

Parameter σ controls the spatial decay of the distance transform weighting function. Smaller values restrict the influence of the weighting to a narrow region surrounding crack boundaries, while larger values extend the influence over a broader spatial area. In the experiments, β was set to 2.0 and σ was set to 0.25. These values were selected empirically based on preliminary experiments to ensure stable optimization while preserving sufficient emphasis on boundary-adjacent pixels. To illustrate the parameter space explored during preliminary experiments, Figure 3 presents a visualization of the mean HD95 values obtained for different combinations of β and σ on the validation set after three training epochs. A detailed analysis of the influence of these parameters on segmentation performance is provided in Section 5.2.

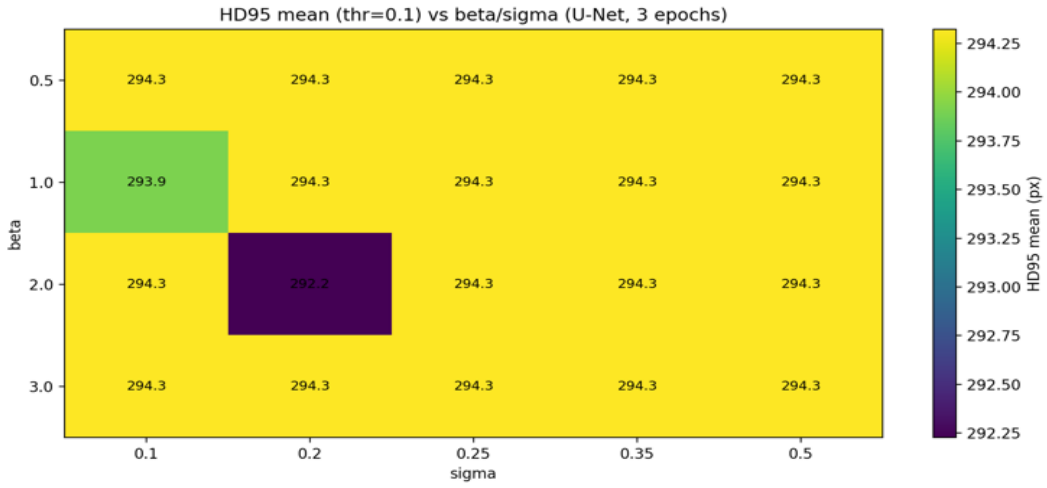


Figure 3. Mean HD95 on the validation set as a function of the distance-aware loss parameters β and σ. Lower values indicate improved boundary alignment.

5. Results and Discussion

5.1. Quantitative Results

Table 2 presents a quantitative comparison of the evaluated loss functions on the validation set.

Table 2. Performance comparison of loss functions

Loss function	IoU@0.1	Dice@0.1	HD95 (px)
BCE	~0.057	~0.105	230.08
Dice (lr = 1e-4 + clip)	~0.056	~0.103	217.55
Distance-Aware (β = 2.0, σ = 0.25 + Dice)	0.2844	0.4087	148.28

The proposed distance-aware loss demonstrates a clear improvement in thin crack segmentation compared with conventional loss formulations. Using BCE as a baseline result in an HD95 value of 230.08 pixels, indicating substantial boundary deviations between predicted and ground-truth crack structures. Replacing BCE with Dice loss slightly improves geometric alignment, reducing HD95 to 217.55 pixels. However, the proposed distance-aware formulation achieves a substantially lower HD95 value of 148.28 pixels. This corresponds to a reduction of approximately 82 pixels compared with BCE and 69 pixels compared with the Dice-based baseline.

In addition to improving boundary accuracy, the proposed loss also increases overlap-based metrics. The model achieves an IoU@0.1 of 0.2844 and a Dice coefficient of 0.4087 on the validation set, which are substantially higher than those obtained using BCE or Dice loss alone. These results indicate that incorporating spatial distance information into the optimization objective enables the network to better capture thin crack structures and align predictions more closely with the ground-truth geometry.

5.2. Discussion: Hyperparameter Sensitivity and Comparison to Boundary-Aware Losses

To further analyze the behavior of the proposed distance-aware formulation, we examine the influence of its hyperparameters β and σ. These parameters control the strength and spatial range of the distance-based weighting mechanism introduced in Section 4.5. Figure 3 illustrates the sensitivity of the proposed loss to different combinations of β and σ. The visualization shows the mean HD95 measured on the validation set after three training epochs. Lower values indicate improved boundary alignment. The results reveal a clear local minimum at $\beta = 2.0$ and $\sigma = 0.2$, which corresponds to the lowest observed HD95 value. This configuration places moderate emphasis on pixels located near crack boundaries while restricting the spatial influence of the distance weighting to a narrow region surrounding the crack structure. Such a balance appears to provide the most effective guidance for the segmentation model during early training.

When σ becomes too small or too large, the influence of the distance weighting diminishes. Extremely small values concentrate the weighting too narrowly around crack pixels, while larger values spread the weighting across broader background regions. In both cases, the resulting HD95 values approach those obtained by baseline loss formulations, indicating that excessive or insufficient spatial emphasis reduces the effectiveness of the distance-aware mechanism. It should be noted that the sensitivity analysis is performed after only three training epochs. The purpose of this analysis is to observe early optimization behavior rather than final convergence. Consequently, many parameter combinations produce similar HD95 values. Nevertheless, the presence of a consistent local minimum suggests that moderate boundary emphasis combined with a narrow spatial influence improves geometric alignment during early training.

Beyond the hyperparameter analysis, it is also useful to compare the proposed approach with specialized boundary-aware losses. Methods such as Boundary Loss and Hausdorff-based losses explicitly incorporate geometric constraints during training and can potentially improve boundary accuracy. However, these approaches often introduce additional computational overhead, such as repeated distance transform computations, and require careful balancing between multiple loss components to avoid unstable training. In contrast, the

proposed distance-aware formulation introduces spatial awareness through a simple weighting mechanism applied to an existing loss function. This design maintains computational simplicity while still encouraging the model to prioritize pixels located near crack boundaries. As a result, the proposed approach provides a lightweight alternative for improving geometric segmentation behavior without modifying the network architecture.

These observations are particularly relevant in constrained-data scenarios. Thin crack datasets typically contain only a limited number of annotated samples and exhibit strong class imbalance, where crack pixels occupy only a small portion of the image. By incorporating spatial distance information into the optimization process, the proposed loss encourages the model to focus on geometrically important regions even when training data are limited. These results suggest that improving the training objective can be an effective strategy for thin-structure segmentation under limited-data conditions. Rather than increasing network depth or architectural complexity, modifying the loss formulation provides a lightweight way to improve geometric alignment and structural continuity of segmentation outputs. However, further experiments across multiple datasets would be required to fully validate this behavior.

5.3. Boundary Accuracy and Error Reduction Analysis

Overlap-based metrics such as Dice and IoU provide a general measure of segmentation quality, but they are often insufficient for thin and elongated structures. In such cases, even small spatial deviations can cause severe structural fragmentation while producing only minor changes in overlap scores. To better evaluate geometric accuracy, boundary precision is therefore analyzed using the HD95, which measures extreme deviations between predicted and ground-truth crack boundaries.

Let $HD95_{BCE}, HD95_{Dice}, HD95_{DA}$ denote the HD95 values obtained using BCE, Dice loss, and the proposed distance-aware loss, respectively. Based on the validation results summarized in Table 2, the following values are obtained:

$$HD95_{BCE} = 230.08\text{px}, \qquad HD95_{Dice} = 217.55\text{px}, \qquad HD95_{DA} = 148.28\text{px}$$

These values exhibit a clear ordering: $HD95_{DA} < HD95_{Dice} < HD95_{BCE}$ indicating that increasing spatial awareness in the loss formulation leads to improved boundary alignment.

5.3.1. Reduction Relative to Binary Cross-Entropy

Compared with BCE, the proposed distance-aware loss reduces the mean HD95 by $\Delta_{BCE \to DA} = 230.08 - 148.28 = 81.80\text{px}$. This corresponds to a relative reduction of $\frac{81.80}{230.08} \times 100 = 35.55\%$.

The magnitude of this reduction indicates a substantial decrease in extreme boundary errors. Considering the fixed image resolution of 512×512 pixels, such a reduction represents a meaningful improvement in geometric alignment for extremely thin structures such as facade cracks.

5.3.2. Reduction Relative to Dice Loss

When compared with Dice loss, the distance-aware formulation achieves an additional reduction of $\Delta_{Dice \to DA} = 217.55 - 148.28 = 69.27\text{px}$, which corresponds to a relative improvement of $\frac{69.27}{217.55} \times 100 = 31.83\%$. Although smaller than the improvement observed relative to BCE, this result demonstrates that incorporating explicit distance information provides additional benefits beyond overlap-based optimization alone.

5.3.3. Interpretation and Consistency with the Loss Formulation

The observed reductions in HD95 are consistent with the mathematical design of the proposed loss function. By weighting pixel-wise errors according to their distance from crack structures, the loss increases the gradient contribution near thin boundaries during training. This mechanism encourages the network to prioritize structural continuity and boundary alignment, thereby reducing extreme spatial deviations.

The empirical relationship $HD95_{DA} < HD95_{Dice} < HD95_{BCE}$ therefore reflects the increasing degree of spatial awareness embedded in the respective loss functions. Importantly, these improvements are achieved without modifying the network architecture. All performance gains originate solely from the training objective, supporting the hypothesis that, in

constrained-data scenarios, improving the loss formulation can be an effective strategy for enhancing geometric segmentation performance.

5.4. Qualitative Results

Figure 4 presents qualitative segmentation results obtained using different loss functions. The columns correspond to representative validation examples, while the rows show the input façade image, the ground-truth crack annotation, and the predictions generated using BCE, Dice loss, and the proposed Distance-Aware loss under two probability thresholds (0.1 and 0.5). Visual comparison reveals clear differences in how each loss formulation captures thin crack structures. Predictions produced using BCE and Dice often miss portions of cracks or generate fragmented segments, particularly when cracks are extremely thin or appear in low-contrast regions of the façade surface. This behavior reflects the characteristics of overlap-based losses, which treat pixel errors uniformly and do not explicitly prioritize boundary alignment.

In contrast, the model trained with the proposed Distance-Aware loss produces more continuous crack predictions and better follows the natural trajectories of cracks. By incorporating spatial distance information during training, the loss emphasizes pixels located near crack boundaries, encouraging improved structural continuity and reduced boundary deviations. These qualitative observations are consistent with the quantitative results reported in Table 2, where the proposed loss achieves the lowest HD95 and improved overlap-based metrics compared with BCE and Dice loss.

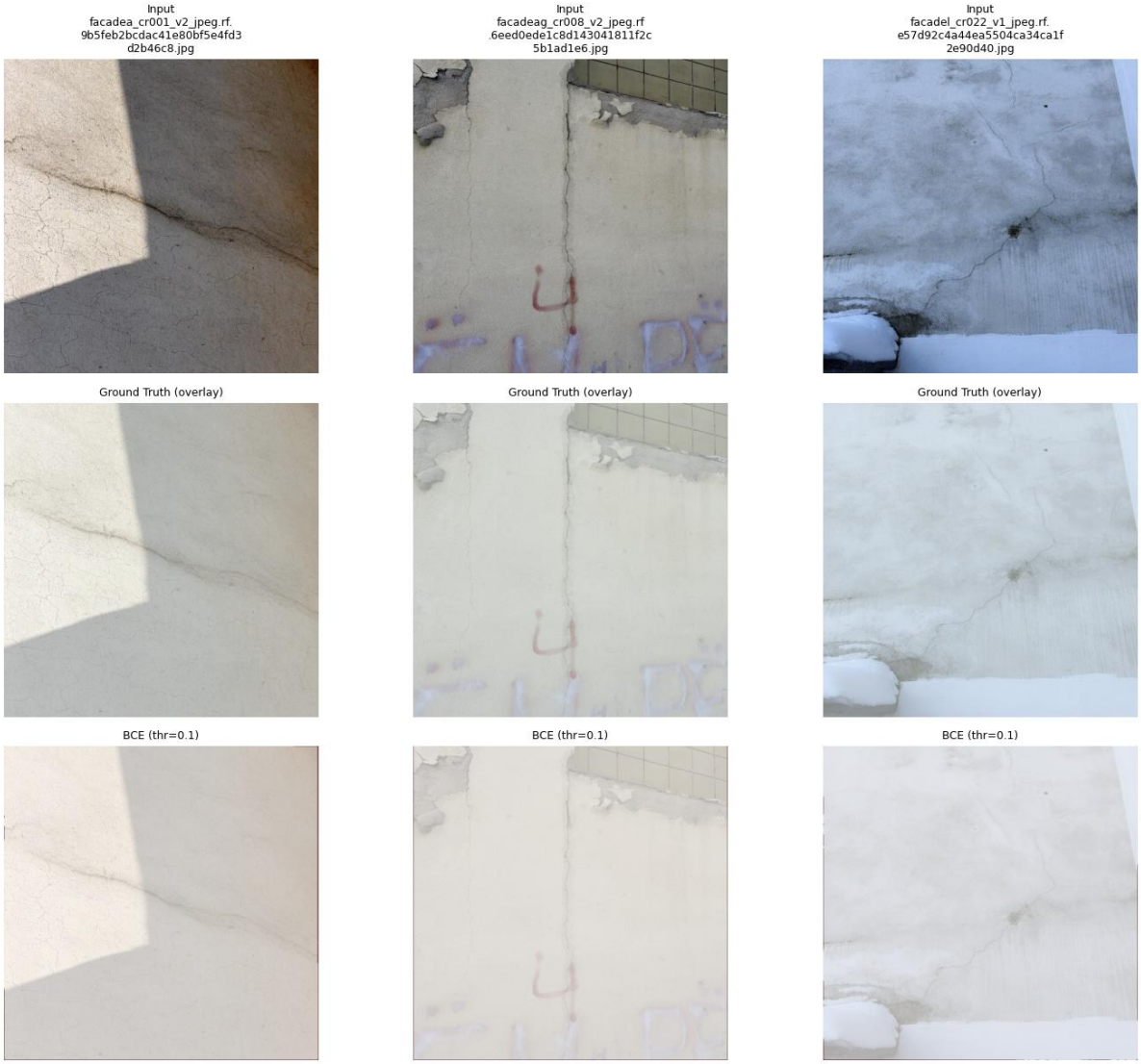
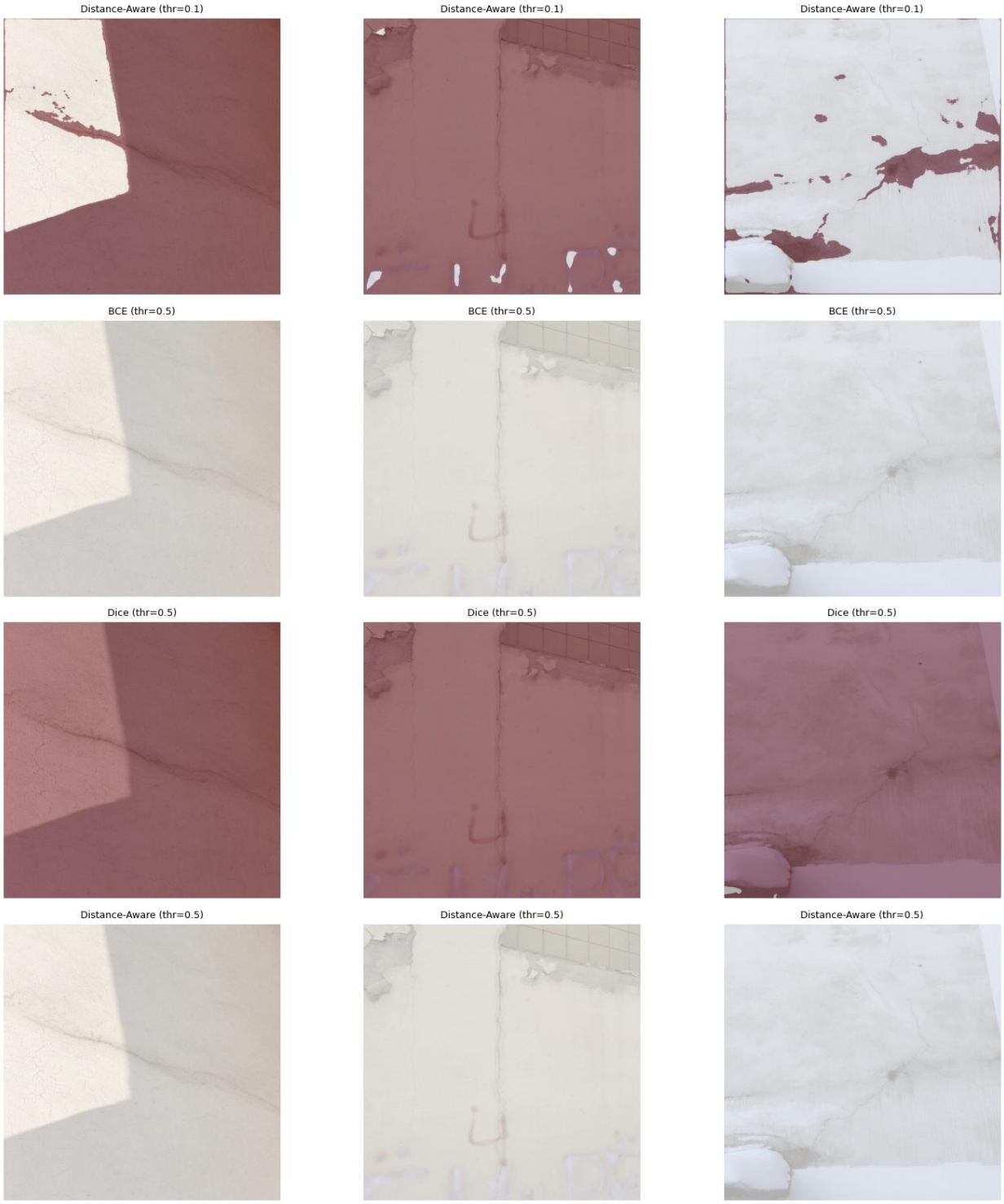


Figure 4. Qualitative comparison of segmentation results obtained using different loss functions.

Figure 4. (continued)



Columns correspond to representative validation examples. Rows show the input façade image, ground-truth crack mask, and segmentation predictions produced using BCE, Dice loss, and the proposed Distance-Aware loss with probability thresholds of 0.1 and 0.5. Ground-truth cracks are shown in green and predicted regions in red.

5.5 Discussion of Findings

The experimental results provide consistent evidence that incorporating spatial distance information into the training objective improves the segmentation of thin crack structures under limited data conditions. Across all experiments, the proposed distance-aware loss

achieved substantially lower HD95 values than both BCE and Dice loss while also improving overlap-based metrics. These results support the central hypothesis of this study: that introducing spatial awareness into the loss formulation can improve geometric alignment without modifying the network architecture. Thin cracks are highly sensitive to spatial deviations, and small positional errors may disrupt structural continuity even when overlap metrics remain relatively unchanged. By assigning greater importance to prediction errors near crack boundaries, the proposed loss encourages the model to preserve thin structural trajectories more accurately.

The improvements observed in HD95 are particularly significant for practical structural inspection tasks. In real-world applications, accurate localization of crack boundaries is often more important than maximizing global region overlap. A reduction in boundary deviation indicates that predicted crack paths more closely follow the true structural geometry, which is essential for assessing crack length, connectivity, and potential structural risks. Another important observation is that the proposed improvement is achieved without modifying the network architecture. All experiments were conducted using the same U-Net configuration and identical training settings. Consequently, the performance gains can be attributed directly to the loss formulation rather than architectural complexity. This finding suggests that optimizing the training objective can be an efficient strategy for improving segmentation performance in constrained-data environments.

The results also highlight the importance of selecting evaluation metrics that reflect the geometric characteristics of the target structures. While Dice and IoU remain useful indicators of overall segmentation performance, they are less sensitive to structural fragmentation in thin objects. In contrast, HD95 captures extreme boundary deviations and therefore provides a more informative measure of geometric accuracy in thin structure segmentation. The findings indicate that lightweight distance-aware optimization provides a practical mechanism for improving the geometric behavior of segmentation models when training data are limited.

6. Conclusions

This study investigated a lightweight distance-aware loss formulation for thin crack segmentation in building facades under limited data conditions. Thin crack detection is challenging because cracks are narrow and elongated structures that occupy only a small fraction of the image area, making segmentation highly sensitive to spatial deviations. The proposed method integrates Euclidean distance information derived from ground-truth masks into a Dice-based loss formulation. By increasing the influence of prediction errors near crack boundaries while reducing the contribution of distant background pixels, the method introduces spatial awareness into the optimization process without modifying the segmentation architecture.

Experimental evaluation on a real-world facade crack dataset demonstrates that the proposed loss consistently reduces boundary deviation compared with BCE and standard Dice loss. In particular, the results show a substantial reduction in HD95, indicating improved geometric alignment and better preservation of crack continuity. These improvements are achieved using a standard U-Net architecture and identical training conditions, confirming that the observed gains originate from the loss formulation rather than architectural modifications. Despite these encouraging results, several limitations remain. The experiments were conducted on a single facade crack dataset, and additional validation across multiple datasets would be necessary to assess the general applicability of the proposed approach. Furthermore, the distance-weighting mechanism relies on fixed hyperparameters that may depend on image resolution and crack morphology.

Future research may explore adaptive or learnable distance-weighting strategies and evaluate the method on additional thin-structure segmentation tasks, including applications in medical imaging and remote sensing. Integrating topology-aware or connectivity-preserving constraints may also further improve segmentation quality for extremely thin or fragmented structures.

Author Contributions: Conceptualization: E.K. and V.Sh.; Methodology: E.K.; Software: E.K.; Validation: E.K. and V.Sh.; Formal analysis: E.K.; Investigation: E.K.; Resources: E.K.; Data curation: E.K.; Writing—original draft preparation: E.K.; Writing, review and editing:

E.K. and V.Sh.; Visualization: E.K.; Supervision: V.Sh.; Project administration: E.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: The dataset used in this study is publicly available for research purposes. The façade crack segmentation dataset (FCD-108), including images and corresponding segmentation masks, can be accessed at: https://github.com/edonaa/facade-crack-dataset.git

Acknowledgments: The authors would like to acknowledge the academic support provided by South East European University during the doctoral research in which this study was conducted. The guidance and constructive feedback received throughout the doctoral research process are gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Y. Liu, J. Yao, X. Lu, R. Xie, and L. Li, "DeepCrack: A deep hierarchical feature learning architecture for crack segmentation," *Neurocomputing*, vol. 338, pp. 139–153, Apr. 2019, doi: 10.1016/j.neucom.2019.01.036.

[2] T. L. Mien, N. D. Tu, and N. Van Lam, "Deploying YOLOv8 for Real-Time Road Crack Detection on Smart Road Length Measurement Devices," *J. Futur. Artif. Intell. Technol.*, vol. 2, no. 1, pp. 135–144, May 2025, doi: 10.62411/faith.3048-3719-102.

[3] L. Zhang, F. Yang, Y. Daniel Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sep. 2016, pp. 3708–3712. doi: 10.1109/ICIP.2016.7533052.

[4] B. Oliveira Santos, J. Valença, J. P. Costeira, and E. Julio, "Domain adversarial training for classification of cracking in images of concrete surfaces," *AI Civ. Eng.*, vol. 1, no. 1, p. 8, Dec. 2022, doi: 10.1007/s43503-022-00008-6.

[5] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, "CrackTree: Automatic crack detection from pavement images," *Pattern Recognit. Lett.*, vol. 33, no. 3, pp. 227–238, Feb. 2012, doi: 10.1016/j.patrec.2011.11.004.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Lecture Notes in Computer Science*, Springer, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[7] H. Zunair and A. Ben Hamza, "Sharp U-Net: Depthwise convolutional network for biomedical image segmentation," *Comput. Biol. Med.*, vol. 136, p. 104699, Sep. 2021, doi: 10.1016/j.compbiomed.2021.104699.

[8] V. Mnih and G. E. Hinton, "Learning to Detect Roads in High-Resolution Aerial Images," in *Lecture Notes in Computer Science*, 2010, pp. 210–223. doi: 10.1007/978-3-642-15567-3_16.

[9] K. Pyar, "Segmentation Performance Analysis of Transfer Learning Models on X-Ray Pneumonia Images," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 1, pp. 64–74, Jun. 2024, doi: 10.62411/faith.2024-10.

[10] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, "Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations," in *Lecture Notes in Computer Science*, 2017, pp. 240–248. doi: 10.1007/978-3-319-67558-9_28.

[11] T.-Y. Lin *et al.*, "Focal Loss for Dense Object Detection," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, vol. 42, no. 2, pp. 2999–3007. doi: 10.1109/ICCV.2017.324.

[12] S. M. Hosseini, "Pixel-wise Modulated Dice Loss for Medical Image Segmentation," *arXiv*. Jun. 17, 2025. [Online]. Available: http://arxiv.org/abs/2506.15744

[13] H. Kervadec, J. Bouchtiba, C. Desrosiers, E. Granger, J. Dolz, and I. Ben Ayed, "Boundary loss for highly unbalanced segmentation," *Med. Image Anal.*, vol. 67, p. 101851, Jan. 2021, doi: 10.1016/j.media.2020.101851.

[14] D. Karimi and S. E. Salcudean, "Reducing the Hausdorff Distance in Medical Image Segmentation With Convolutional Neural Networks," *IEEE Trans. Med. Imaging*, vol. 39, no. 2, pp. 499–513, Feb. 2020, doi: 10.1109/TMI.2019.2930068.

[15] J. Ma *et al.*, "How Distance Transform Maps Boost Segmentation CNNs: An Empirical Study," in *Proceedings of the Third Conference on Medical Imaging with Deep Learning*, 2020, pp. 479–492. [Online]. Available: https://proceedings.mlr.press/v121/ma20b.html

[16] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[17] H. Cui, D. Wei, K. Ma, S. Gu, and Y. Zheng, "A Unified Framework for Generalized Low-Shot Medical Image Segmentation With Scarce Data," *IEEE Trans. Med. Imaging*, vol. 40, no. 10, pp. 2656–2671, Oct. 2021, doi: 10.1109/TMI.2020.3045775.

[18] D. Müller, I. Soto-Rey, and F. Kramer, "Towards a guideline for evaluation metrics in medical image segmentation," *BMC Res. Notes*, vol. 15, no. 1, p. 210, Dec. 2022, doi: 10.1186/s13104-022-06096-y.

[19] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," in *Lecture Notes in Computer Science*, 2016, pp. 424–432. doi: 10.1007/978-3-319-46723-8_49.

[20] H. Fu, Y. Xu, S. Lin, D. W. Kee Wong, and J. Liu, "DeepVessel: Retinal Vessel Segmentation via Deep Learning and Conditional Random Field," in *Lecture Notes in Computer Science*, 2016, pp. 132–139. doi: 10.1007/978-3-319-46723-8_16.

[21] L. Maier-Hein *et al.*, "Why rankings of biomedical image analysis competitions should be interpreted with care," *Nat. Commun.*, vol. 9, no. 1, p. 5217, Dec. 2018, doi: 10.1038/s41467-018-07619-7.

[22] O. Oktay *et al.*, "Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation," *IEEE Trans. Med. Imaging*, vol. 37, no. 2, pp. 384–395, Feb. 2018, doi: 10.1109/TMI.2017.2743464.

[23] C. You *et al.*, "Contour Knowledge-Aware Perception Learning for Semantic Segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 35, no. 5, pp. 4560–4575, May 2025, doi: 10.1109/TCSVT.2024.3515088.
[24] Y. Fu, J. Chen, and K. Li, "Structure-aware Loss Function for Ultrasound Image Segmentation," in *2021 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2021, pp. 1–4. doi: 10.1109/IUS52206.2021.9593773.
[25] A. Mosinska, P. Marquez-Neila, M. Kozinski, and P. Fua, "Beyond the Pixel-Wise Loss for Topology-Aware Delineation," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 3136–3145. doi: 10.1109/CVPR.2018.00331.
[26] S. Shit *et al.*, "clDice - a Novel Topology-Preserving Loss Function for Tubular Structure Segmentation," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2021, pp. 16555–16564. doi: 10.1109/CVPR46437.2021.01629.
[27] J. Ma *et al.*, "Loss odyssey in medical image segmentation," *Med. Image Anal.*, vol. 71, p. 102035, Jul. 2021, doi: 10.1016/j.media.2021.102035.
[28] Y. Yang, J. Zhuang, G. Sun, R. Wang, and J. Su, "Boundary-Guided Contrastive Learning for Semi-Supervised Medical Image Segmentation," *IEEE Trans. Med. Imaging*, vol. 44, no. 7, pp. 2973–2988, Jul. 2025, doi: 10.1109/TMI.2025.3556482.
[29] S. Jadon, "A survey of loss functions for semantic segmentation," in *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, Oct. 2020, pp. 1–7. doi: 10.1109/CIBCB48159.2020.9277638.
[30] Z. Zhang, Q. Liu, and Y. Wang, "Road Extraction by Deep Residual U-Net," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, May 2018, doi: 10.1109/LGRS.2018.2802944.
[31] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.
[32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, vol. 2016-Decem, pp. 770–778. doi: 10.1109/CVPR.2016.90.
[33] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, Oct. 2016, pp. 565–571. doi: 10.1109/3DV.2016.79.