Distance-based Loss Function for Machine Learning in Image Segmentation of Concrete Cracks

Baris Özcan and Fabian Alexander Schmiedecke
Geodetic Institute and Chair for Computing in Civil Engineering & Geo Information Systems, RWTH Aachen University, Mies-van-der-Rohe-Str. 1, 52074 Aachen, Germany
E-Mail: oezcan@gia.rwth-aachen.de

Abstract: Crack detection is essential for maintaining the integrity and safety of infrastructure such as bridges, roads, and buildings. Traditional manual inspection methods are labor-intensive and error-prone, which underscores the need for automated solutions. Advances in machine learning have shown promise in automating this task, but further improvements are necessary to address specific challenges, such as the thin and elongated nature of cracks, varying lighting conditions, and background noise. This paper presents a refined loss function designed to enhance the performance of machine learning models in detecting fine-structured concrete cracks. This loss function considers the spatial distances between predicted and actual crack locations, penalizing false positives that are far from true crack positions. In this way, our approach emphasizes the thin nature and continuity of cracks, ensuring that even the smallest features are accurately detected. Our initial experiments demonstrate that this strategy offers a more accurate and reliable solution to crack detection.
Keywords: Concrete Crack Detection, Computer Vision, Machine Learning, Image Segmentation, Loss Functions

Erschienen in Tagungsband 35. Forum Bauinformatik 2024, Hamburg, Deutschland, DOI: 10.15480/882.13533

1 Introduction

Crack detection is crucial in civil engineering and infrastructure maintenance, ensuring the safety and longevity of structures like bridges, roads, and buildings. Early detection and timely repair of cracks can prevent catastrophic failures, making crack detection a critical task. Traditional methods rely on manual inspection, which is labor-intensive, time-consuming, and prone to human error. Consequently, automated crack detection using advanced image processing and machine learning techniques has seen significant attention in recent years.

In automated crack detection, deep learning models, particularly convolutional neural networks (CNNs), have demonstrated remarkable performance improvements. These models can learn complex patterns and features from large image datasets, enabling highly accurate detection of cracks. Despite these advancements, there are several challenges. One primary challenge is that traditional loss functions and evaluation metrics often fail to adequately capture and consider the thin, elongated nature and continuity of cracks, as well as account for varying lighting conditions and background noise.

This research proposes a distance-based loss function customized for detecting fine structures, such as cracks, in images using machine learning. The proposed loss function enhances the performance of crack detection models by penalizing false positives that are far from the true crack locations, thereby improving the detection and continuity of cracks. By focusing on the spatial distance between predicted and actual crack locations, this loss function effectively addresses the challenges associated with the thin and elongated nature of cracks, leading to more complete, precise and reliable automated crack detection.

The remainder of this paper is organized as follows: Section 2 reviews crack detection techniques and loss functions in machine learning contexts. Section 3 outlines our methodology, including the data preparation and the proposed loss function. Section 4 describes the experimental setup, including model architecture and training protocol. Section 5 presents our preliminary results and Section 6 provides a discussion of the results. Finally, Section 7 concludes the paper and suggests directions for future research.

2 State-of-the-art

2.1 Crack Detection Techniques

Early automated methods for crack detection [1] focused on image processing techniques, such as edge detection algorithms (e.g., Canny, Sobel), thresholding methods, and morphological operations. While these techniques showed some success, they were often limited by their sensitivity to noise, lighting variations, and the complexity of the background.

The emergence of machine learning, particularly deep learning, has revolutionized crack detection. Convolutional neural networks (CNNs) have been widely adopted due to their ability to automatically learn and extract hierarchical features from images. Several studies demonstrated the efficiency of CNNs in detecting cracks in various materials and conditions

[2]. However, these models, however, still often face challenges in accurately detecting fine, elongated cracks and differentiating them from similar patterns in the background.

2.2 Loss Functions in Image Segmentation

The choice of loss function is crucial in training deep learning models, as it directly impacts the ability of the model to learn from the data. In the context of image segmentation, several loss functions have been employed.

Cross-Entropy Loss: Commonly used in classification tasks, cross-entropy loss has also been applied to crack detection. However, it may not be the most effective for segmenting thin and elongated cracks due to its focus on pixel-wise classification without considering spatial arrangement or continuity.

Dice Loss: Dice loss is particularly useful for segmentation tasks, as it maximizes the overlap between the predicted and ground truth masks. This loss function often shows better performance in crack detection tasks compared to cross-entropy loss [3], particularly because it is less sensitive to class imbalance and better suited for cases where the object of interest occupies a small area relative to the background, such as thin cracks.

Intersection over Union (IoU) Loss: IoU loss measures the overlap between predicted and ground truth masks relative to their union. While similar to Dice loss in its focus on overlap, IoU loss provides a different perspective by emphasizing the union of the predicted and ground truth areas, which can be beneficial when minimizing false positives is crucial.

However, while these loss functions are effective in many scenarios, they may still face challenges in fully capturing the thin and continuous nature of cracks, particularly when spatial continuity and elongation are critical. Without further customization, machine learning models might produce fragmented or incomplete detections, potentially reducing their effectiveness in practical applications. Therefore, tailored loss functions are often necessary to enhance segmentation performance and ensure reliable automated crack detection.

3 Methodology

This section outlines the methodology of our research, including the data preparation and the loss function tailored for crack detection.

3.1 Data Preparation

For this study, we utilized a publicly available dataset comprising images of various surfaces with annotated cracks, accessible at Kaggle [4]. This dataset is a compilation of 12 publicly available crack segmentation datasets, containing approximately 11,200 images. All images in the dataset has been resized to a resolution of 448 x 448 pixels. The dataset includes images from diverse environments, ensuring variability in crack appearance due to different lighting conditions, materials, and background noise. However, inaccurate labels in the dataset were identified and manually removed to maintain reliability in model training and evaluation. Some sample images of this dataset is shown in Figure 1.

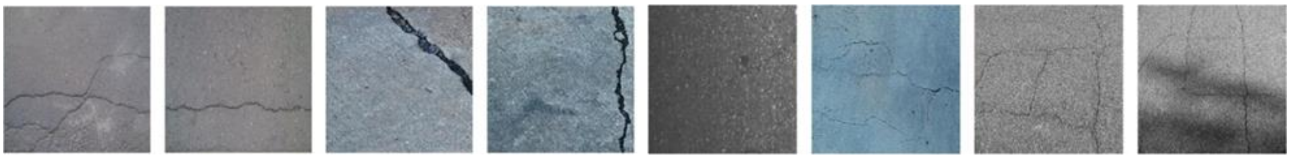

Figure 1: Sample images of the crack dataset.

To enhance the robustness and generalizability of the model, data augmentation techniques were applied. These techniques include rotation and flipping to introduce geometric variability, adjustments on brightness and contrast to simulate different lighting conditions and addition of Gaussian noise to make the model more resilient to real-world image noise.

Further preprocessing steps included resizing images to a uniform size of 128 x 128 Pixel with bicubic interpolation and normalizing pixel values.

3.2 Distance-based Loss

The Distance-based Loss function (DL) is designed to emphasize the continuity of cracks and prevent the prediction of background noise as separate cracks. This loss function integrates a distance map for each image, where every pixel is assigned a normalized value, depending on its distance to the nearest crack pixel priors to penalize FP predictions that are further away from the cracks in the predicted crack segments. The mathematical formulation is as follows:

$$\mathcal{L}_{DL} = \mathcal{L}_{Dice} \cdot (1 + \mathcal{L}_{Dist} \cdot (\alpha - 1)) \tag{1}$$

where $\mathcal{L}_{Dice}$ is the Dice loss ensuring overlap maximization, and $\mathcal{L}_{Dist}$ is a term that penalizes false positives based on their distance from the nearest crack pixel. The parameter α controls the balance between the two components.

$$\mathcal{L}_{Dice} = \frac{2 \cdot \mathrm{TP}}{2 \cdot TP + FP + FN} \tag{2}$$

$$\mathcal{L}_{Dist} = \frac{1}{N_{FP}} \sum_{i=1}^{N_{FP}} ND(FPI(i)) \tag{3}$$

with

N_{FP}: Number of False Positives

ND: Normalized Distance from Nearest Crack Pixel

FPI: False Positive Index

4 Experimental Setup

This chapter describes the experimental setup used to evaluate the proposed loss function for crack detection, including details about the model architecture and the training protocol.

4.1 Model Architecture

For our experiments, we used the U-Net architecture [5], a state-of-the-art CNN known for its effectiveness in image segmentation tasks. We specifically chose U-Net due to its proven performance in medical image segmentation, which shares similarities with crack detection in terms of requiring precise localization and boundary delineation. U-Net's encoder-decoder structure with skip connections allows for combining high-resolution features from the encoder with upsampled features from the decoder, capturing both fine details and contextual information. The encoder consists of convolutional and max-pooling layers, reducing spatial dimensions while increasing feature maps. The decoder includes upsampling and convolutional layers, restoring spatial dimensions and integrating features through skip connections, crucial for accurate crack detection.

4.2 Training Protocol

We split our dataset into 80% training, 10% validation, and 10% test sets to ensure comprehensive training and evaluation. The model was trained with the Adam optimizer (learning rate 5×10^{-4}, a batch size of 25, and for 50 epochs). We used the proposed Distance-based Loss function with the parameter $\alpha = 4.0$. To prevent overfitting and improve generalization, dropout layers with a rate of 0.3 were used in the decoder and encoder, and extensive data augmentation techniques such as rotation, flipping, brightness and contrast adjustment, and noise addition were employed.

5 Results

In this section, we present the results of our experiments for the proposed loss function for crack detection and compare it with standard loss functions.

5.1 Quantitative Results

We evaluated our model trained with the proposed loss function using five common metrics and compared it with two models trained with standard loss functions (Binary Cross-Entropy and Dice Loss). The results are summarized in Table 1.

Table 1: Performance comparison of our proposed loss function with two standard loss functions based on several metrics.

Metric	Binary Cross-Entropy Loss	Dice Loss	Distance-based Loss
Precision	0.4771	0.5202	0.5856
Recall	0.2336	0.7407	0.7707
F1 Score	0.2860	0.6001	0.6589
IoU	0.2147	0.3824	0.5027
Dice	0.3000	0.5609	0.6078

5.2 Qualitative Results

Additionally, the visual results for the predicted image masks of the cracks offer a more detailed evaluation of the model's performance. Figure 2 illustrates a comparison of crack detection outcomes using the three different loss functions, together with the original input images and the ground truth masks.

6 Discussion

Our initial experiments show that the model trained with the proposed Distance-based Loss function outperforms those using traditional loss functions. The proposed loss function led to better results for all evaluation metrics, indicating superior performance in the segmentation of cracks. Additionally, visual results show that the Distance-based loss function produces more accurate and continuous crack detections compared to traditional loss functions.

Original Image	Ground Truth	BCE Loss	Dice Loss	Distance-based Loss

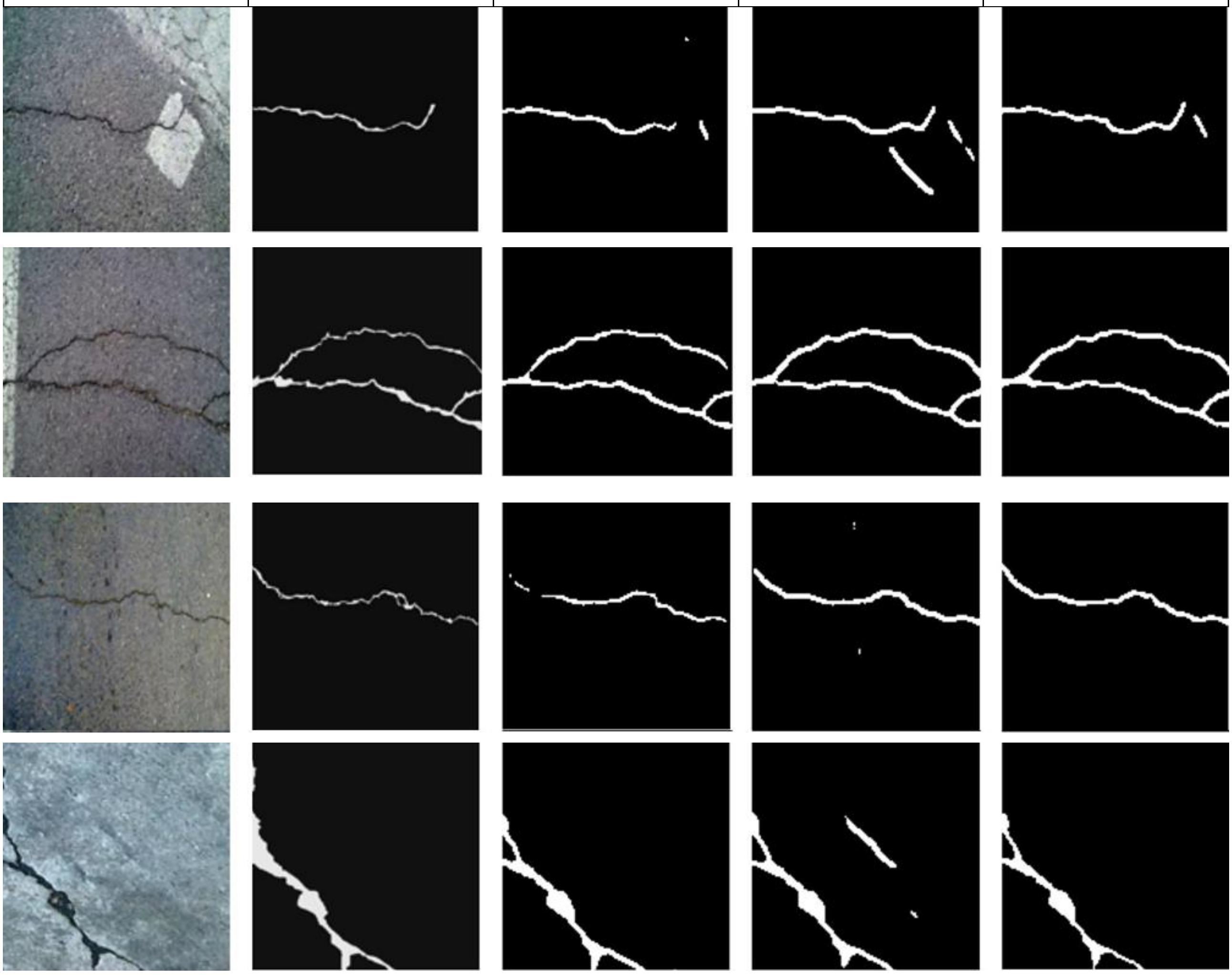

Figure 2: Crack detection results, from left to right: Original image, Ground Truth, Binary Cross-Entropy Loss, Dice Loss, and Distance Loss

The improved performance of the distance-based loss function comes from its ability to penalize false positives that are spatially distant from the true crack locations. This added spatial dimension reduces inaccuracies by focusing on the spatial relationship between the predicted and actual crack locations, unlike traditional loss functions that treat each pixel independently. Furthermore, the Distance-based loss function facilitates the model to detect cracks as continuous entities rather than seperated segments by emphasizing the spatial distance between connected components. This approach promotes spatial coherence, leading to smoother and more continuous predictions. Moreover, by focusing on the spatial consistency of cracks, the loss function helps the model differentiate between true crack signals and random noise, leading to more reliable crack detection in the presence of image noise. Additionally, handling class imbalance more effectively, the distance-based loss function prioritizes the accurate identification of crack pixels, improving detection performance in imbalanced datasets where non-crack pixels dominate. Also, the spatial relationships in the loss function helps the model learn more generalized features of cracks, making it more robust to different types of cracks and variations in new datasets. Lastly, the loss function aligns predicted cracks with human perception by ensuring that detected cracks are continuous and structured, which makes the results more intuitive.

7 Conclusion

In this paper, we proposed a Distance-based loss function customized for crack detection to address the limitations of traditional loss functions. This loss function penalizes false positives that are far from the true crack locations, improving the continuity and accuracy of detected cracks. Our initial experiments demonstrated that this approach outperforms traditional loss functions.

Future work will focus on integrating additional loss functions and evaluation metrics to further enhance crack detection models. Building on our proposed Distance-based loss function, we aim to develop and incorporate more such functions and evaluation metrics that better capture the structures and nuances of cracks. This approach will help refining models for crack dectection, making them more robust and effective in practical applications.

References

[1] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," Alexandria Engineering Journal, vol. 57, no. 2, pp. 787-798, June 2018.

[2] Y. Hamishebahar, H. Guan, S. So, and J. Jo, "A Comprehensive Review of Deep Learning-Based Crack Detection Approaches," Applied Sciences, vol. 12, no. 3, 2022.

[3] Q. D. Nguyen and H.-T. Thai, "Crack segmentation of imbalanced data: The role of loss functions," *Engineering Structures*, vol. 297, 2023.

[4] L. Middha, "Crack Segmentation Dataset," Kaggle, 2024. Accessed on: Aug. 23, 2024. [Online]. Available: https://www.kaggle.com/datasets/lakshaymiddha/crack-segmentation-dataset.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, 2015.