

Master's Thesis Presentation for the 148th Dissertation Evaluation

A Study on Unsupervised Anomaly Detection Based on Branch-Aware Multiple Memory Banks for Considering Crack Direction

균열 방향 고려를 위한 가지 반영형 다중 메모리 기반
비지도 이상 탐지에 관한 연구

Seowon Kim

ksw1912@cau.ac.kr

17th Sept, 2026



Contents

- 1. Introduction**
- 2. Related Work**
- 3. Motivation**
- 4. Proposed Method**
- 5. Experimental Settings**
- 6. Experimental Results**
- 7. Conclusion**

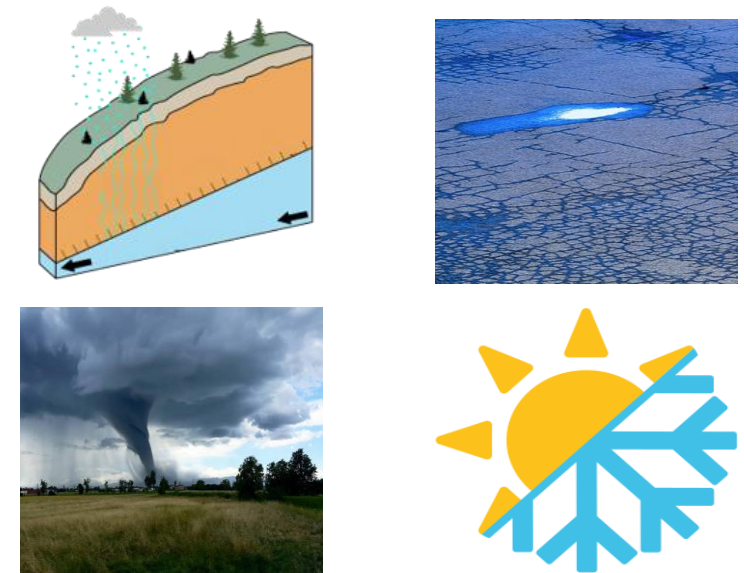
Introduction

What is Unsupervised Crack Detection & Why it is Important?

- Early crack detection is essential for assessing the condition and safety of civil infrastructure, as cracks indicate deterioration of structures losing their expected conditions due to long-term operation, internal defects, aggressive usage, and load changes [1, 2].
- Crack images are scarce and vary widely in shape and surface, so unsupervised crack detection learns only from normal images and flags a test image whose anomaly score deviates substantially from the normal distribution [1, 3].



[Figure 1] Comparison of normal and crack images



[Figure 2] Environmental factors

[1] Deng, et al., "Review on computer vision-based crack detection and quantification methodologies for civil structures," Construction and Building Materials, 2022

[2] Ai, Dihao, et al., "Computer vision framework for crack detection of civil infrastructure—A review," Engineering Applications of Artificial Intelligence, 2023

[3] Pang, et al., "Deep learning for anomaly detection: A review," ACM computing surveys (CSUR), 2021

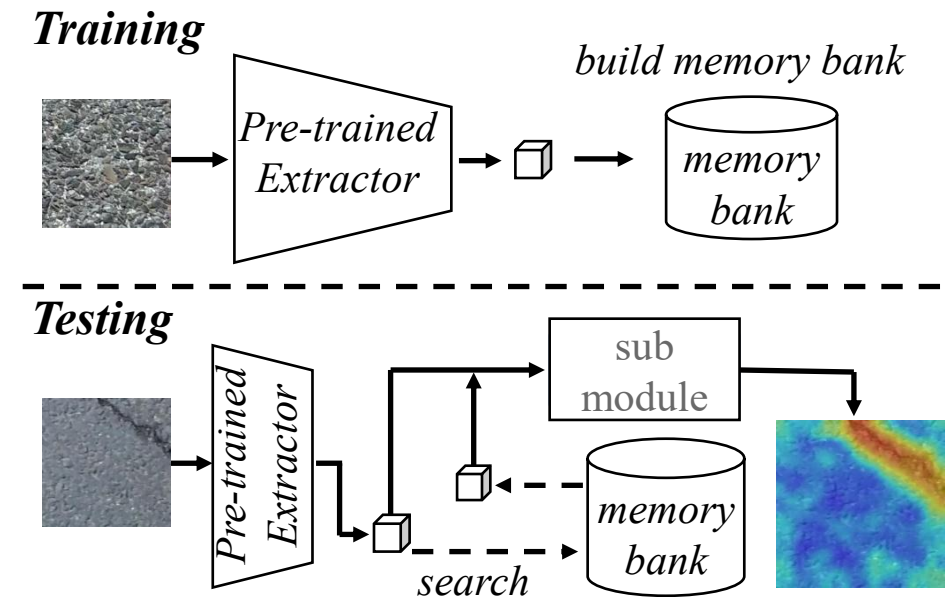
Introduction

Memory Bank methods

- Memory bank methods, one of the feature embedding-based strategies, store patch features of normal samples extracted by a pre-trained network and detect anomalies by comparing each test patch with its nearest stored normal feature in the memory bank [4].
- In this study, this approach is adopted because it requires neither additional network training nor simulated anomalies and enables fine-grained patch-level comparison, which is well suited for detecting thin and localized cracks across diverse surface backgrounds.

[Table 1] Comparison of Unsupervised Anomaly Detection Strategies

Strategy	Core Mechanism / Principle
Memory Bank-based	Store pre-trained patch features of normal samples in a memory bank and detect anomalies by the distance between each test feature and its nearest stored normal feature.
Knowledge Distillation & Reconstruction-based	Train a student or reconstruction network on normal samples and detect anomalies from its discrepancy with the teacher features or the input.
Distribution Estimation-based	Model the distribution of normal features with normalizing flows or Gaussians and detect anomalies as features with low likelihood under it.

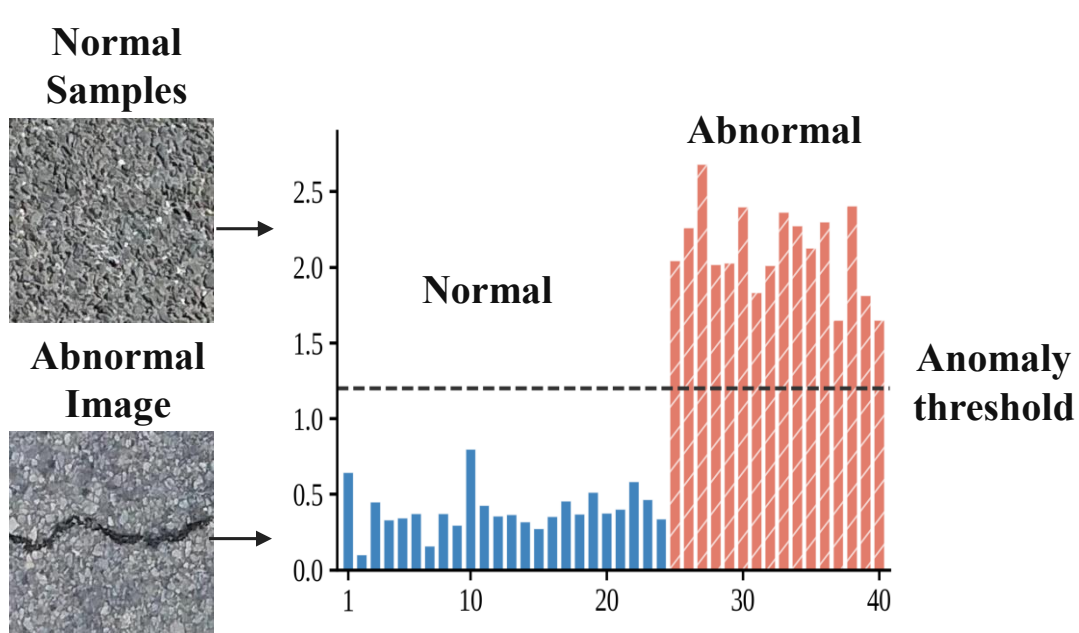


[Figure 3] Defect-free training using only normal samples

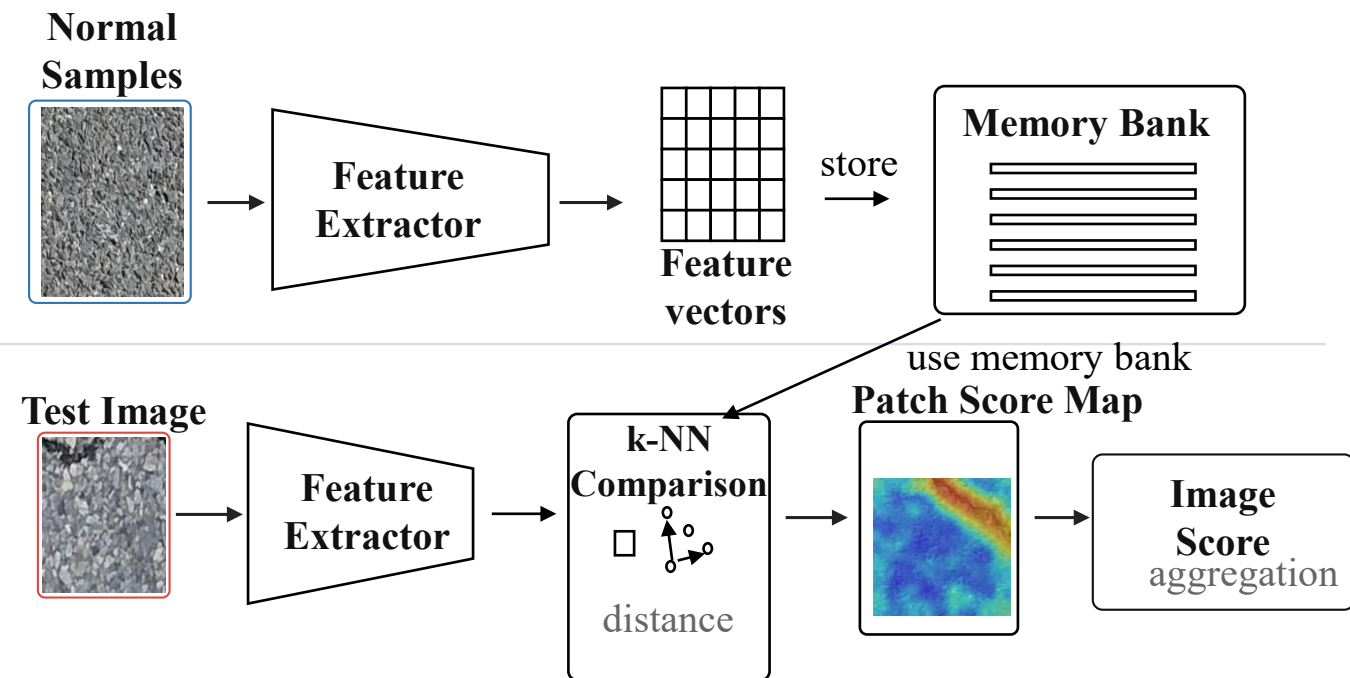
Introduction

I/O and General Procedure

- In the training phase, a pre-trained feature extractor extracts patch features from normal images to build a memory bank, and the input image is scored at test time by measuring the distances between its patch features and their nearest stored normal features [4].
- In the test phase, the input is either a normal or a crack test image, and the output is an image-level anomaly score reflecting how far the image deviates from the normal distribution, where a higher score indicates a higher likelihood that the image contains cracks [4].



[Figure 4] Input and output of unsupervised crack detection

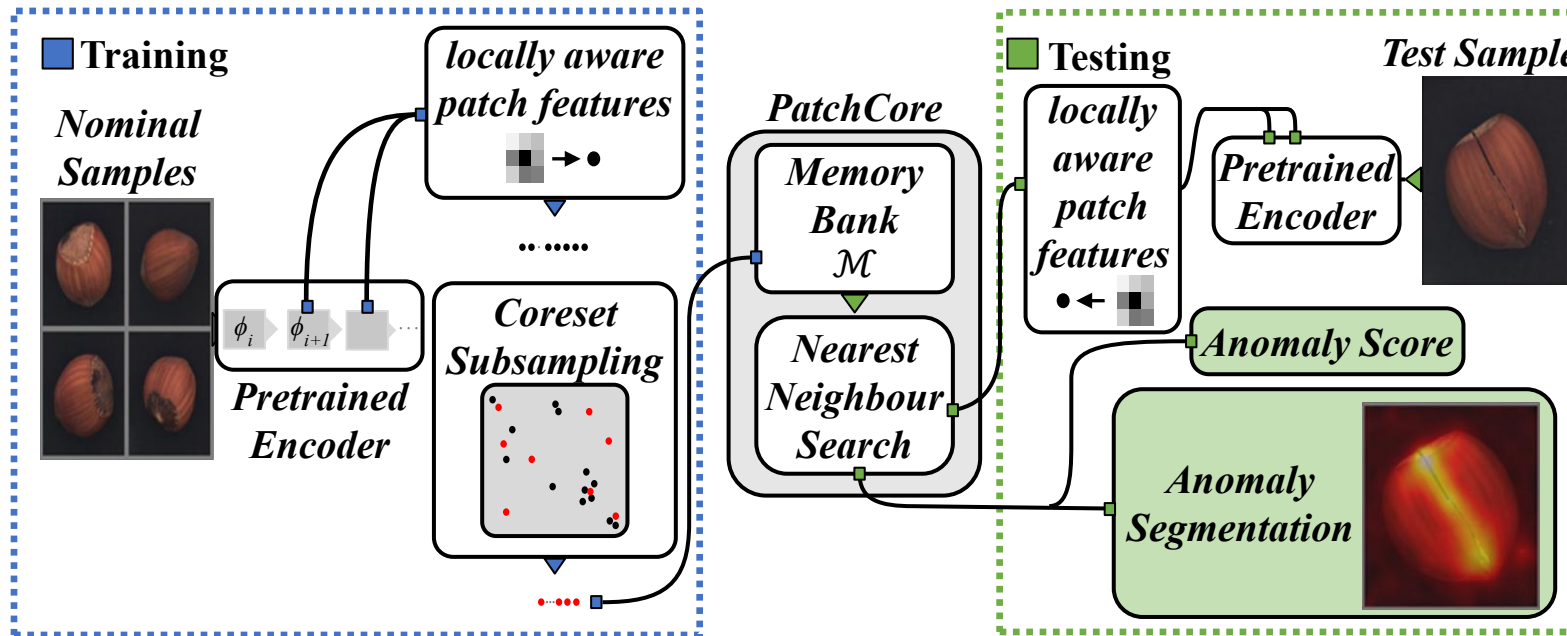


[Figure 5] General Procedure of Patch-Based Anomaly Detection

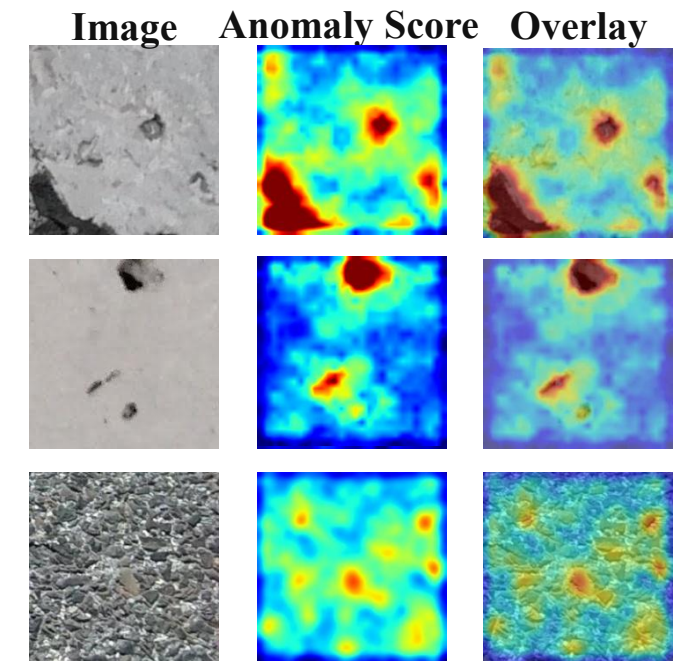
Related Work

Towards Total Recall in Industrial Anomaly Detection [5]

- To reduce the ImageNet bias of deep features, PatchCore stores locally aware mid-level patch features in a coreset-subsampled memory bank shared by all test patches, enabling anomaly detection without requiring aligned input images.
- Although PatchCore averages each 3×3 neighborhood into one feature, this averaging discards how the patch relates to each neighbor, so normal patches that differ from their neighbors lie far from the memory bank and become false positives.



[Figure 6] Overall framework of PatchCore

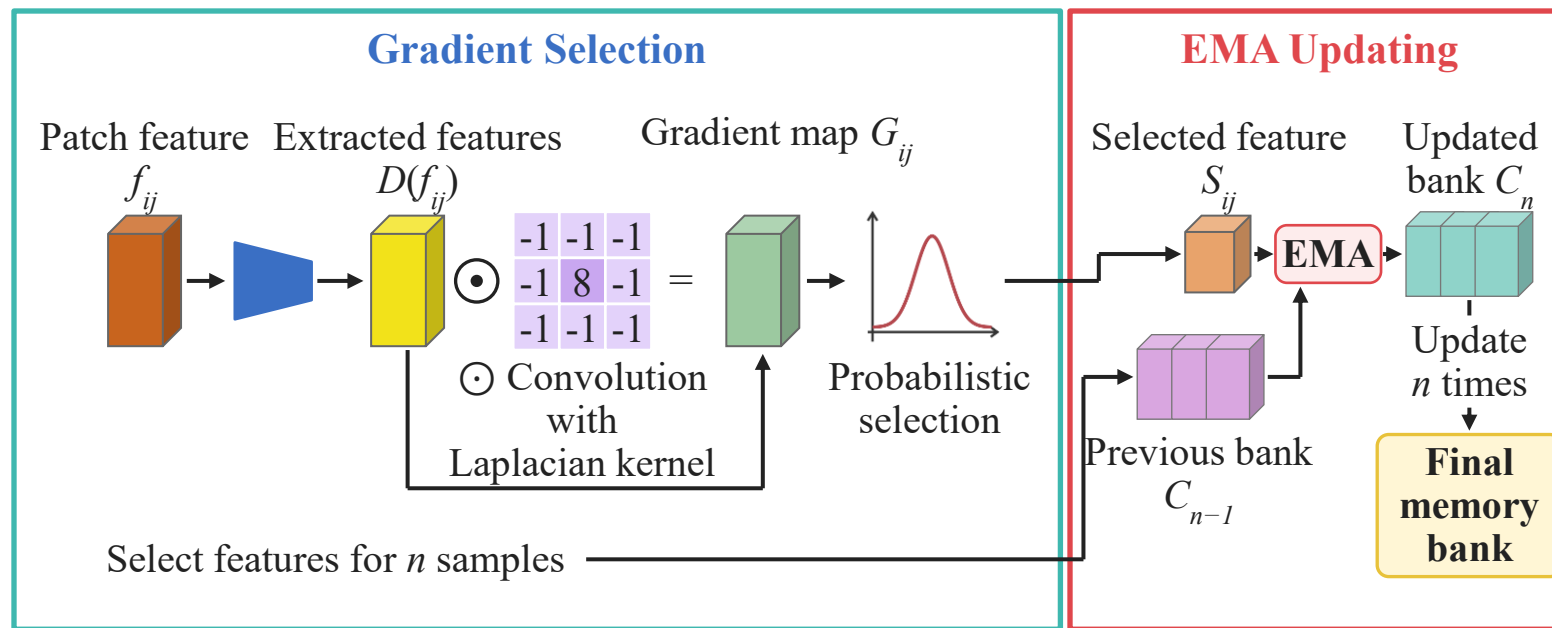


[Figure 7] Anomaly maps of PatchCore on crack images

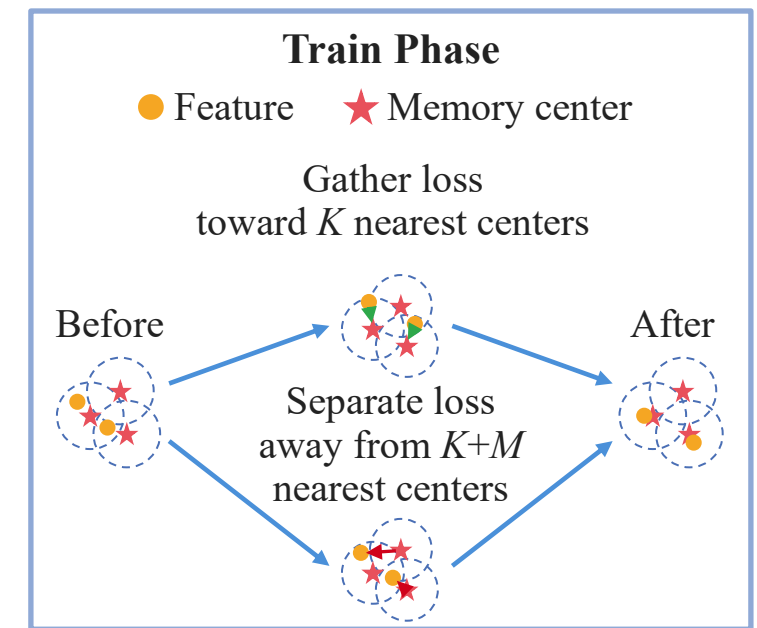
Related Work

MGSCA: Multiscale Gradient Selection-Based Coupled-Hypersphere Adaptation for Unsupervised Pavement Defect Detection [6]

- To handle pre-trained feature bias under complex pavement backgrounds and weak textures, MGSCA builds a representative gradient-selected memory bank and adapts the features with coupled hyperspheres, improving the defect detection accuracy.
- Because MGSCA scores each adapted feature by its distance to the memory centers independently of adjacent patches, normal patches that differ from their neighbors, such as holes and stains, remain far from every center and become false positives.



[Figure 8] Overall framework of MGSCA

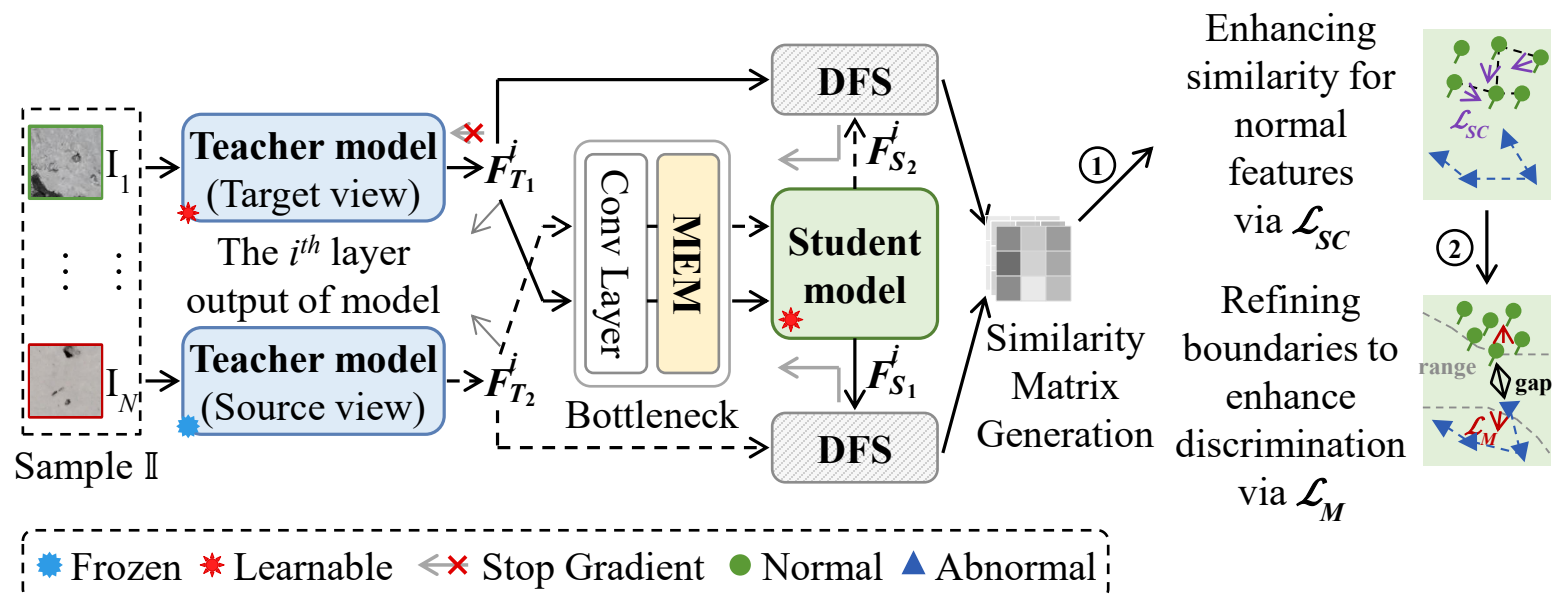


[Figure 9] Training of MGSCA with gather and separate losses

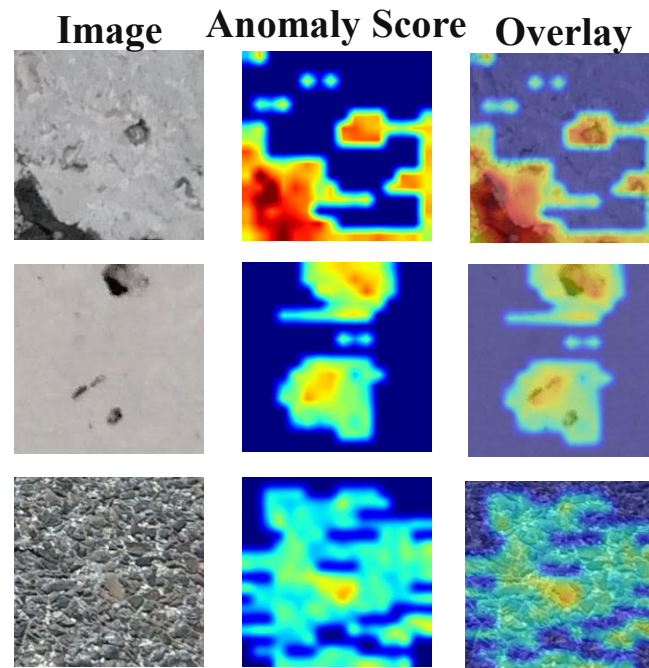
Related Work

UniNet: A Contrastive Learning-guided Unified Framework with Feature Selection for Anomaly Detection [7]

- UniNet introduced a teacher–student distillation framework with contrastive learning to prevent student networks from reproducing anomalous features, encouraging compact normal feature distributions and improving anomaly discrimination.
- Because the student is trained mainly on frequent normal patterns, it fails to reproduce teacher features on rare normal patterns such as holes and stains, and fusing multi-stage maps spreads these discrepancies into blob-like false positives.



[Figure 10] Overall framework of UniNet

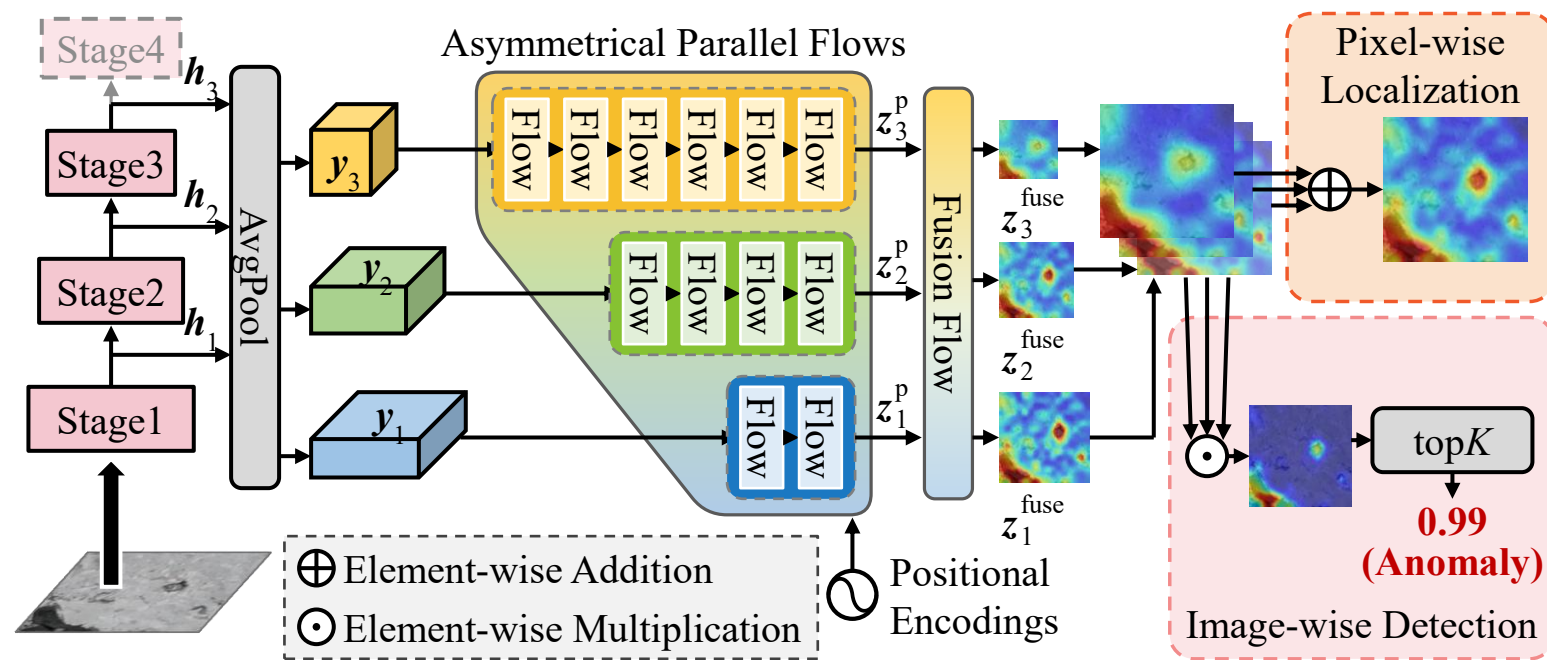


[Figure 11] Anomaly maps of UniNet on crack images

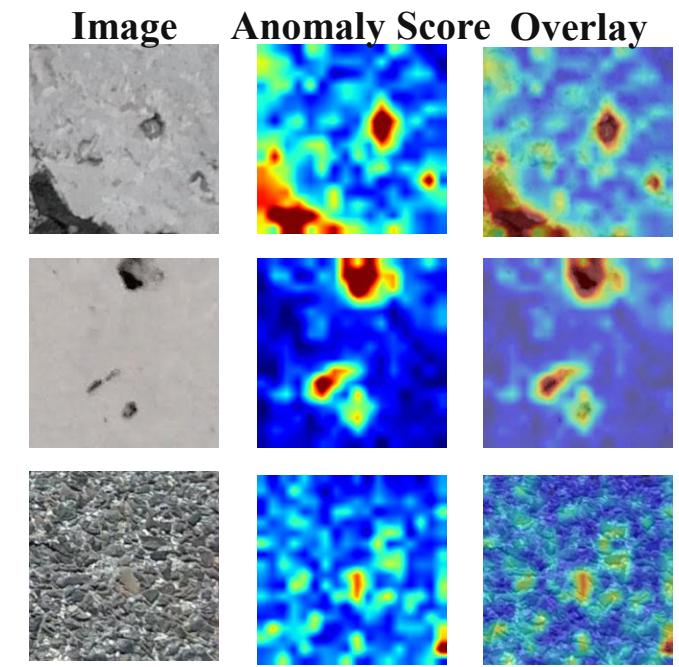
Related Work

MSFlow: Multi-Scale Flow-based Framework for Unsupervised Anomaly Detection [8]

- MSFlow proposed a multi-scale normalizing flow framework combining parallel flows and a fusion flow to address variations in anomaly sizes and spatial characteristics, improving detection and localization through multi-scale feature integration.
- Because the estimated distribution assigns low likelihood to rare normal textures on complex concrete surfaces, and coarse-scale responses are fused with fine ones, the anomaly maps become diffuse, causing false positives on crack-free surfaces.



[Figure 12] Overall framework of MSFlow

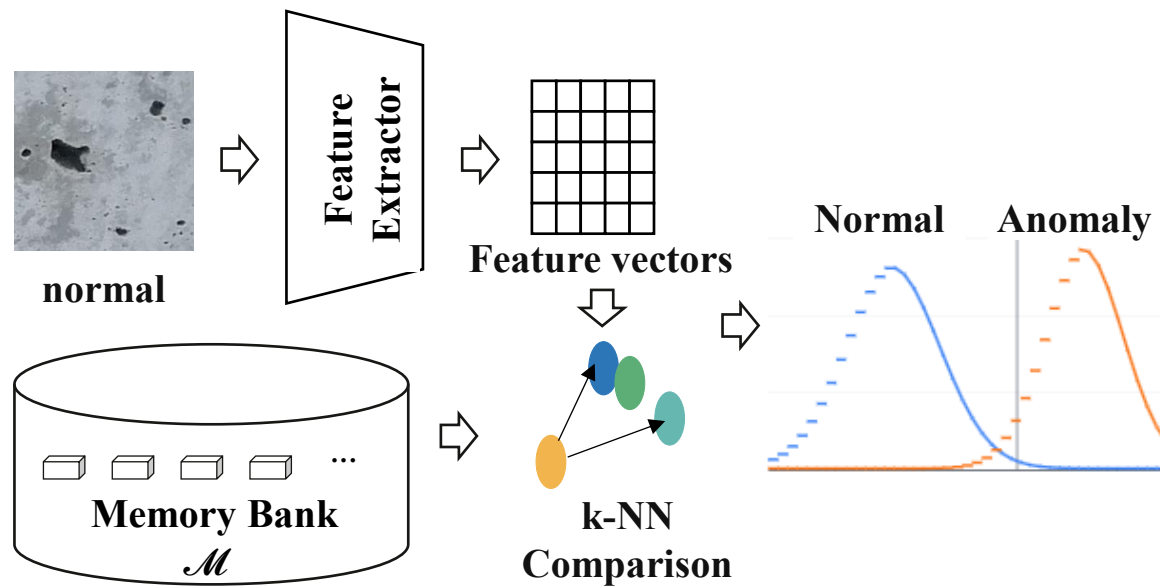


[Figure 13] Anomaly maps of MSFlow on crack images

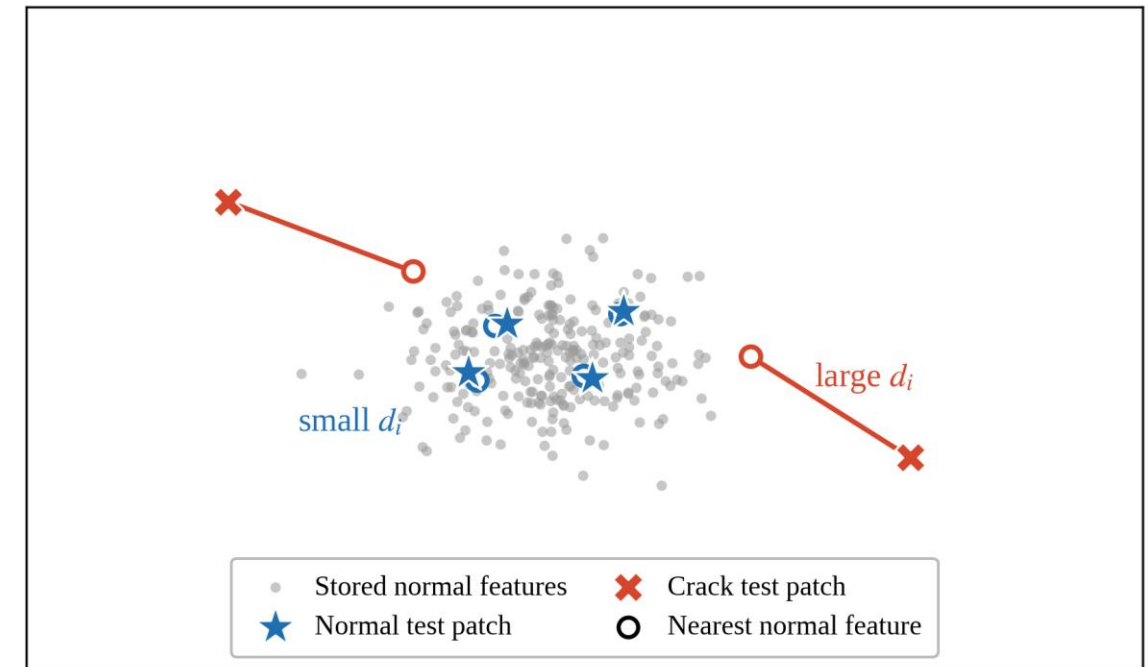
Motivation

Target Process

- In the target process, each test patch feature is matched to its single nearest normal reference feature stored in the memory bank, and the feature-space distance between the two is then directly used as the local anomaly score assigned to the corresponding patch.
- Ideally, every normal test patch is matched to a close normal feature and receives a small distance, while every crack patch remains far from all stored normal features and receives a clearly large distance, so that normal and crack patches are separated by one threshold.



[Figure 14] input and output for target process

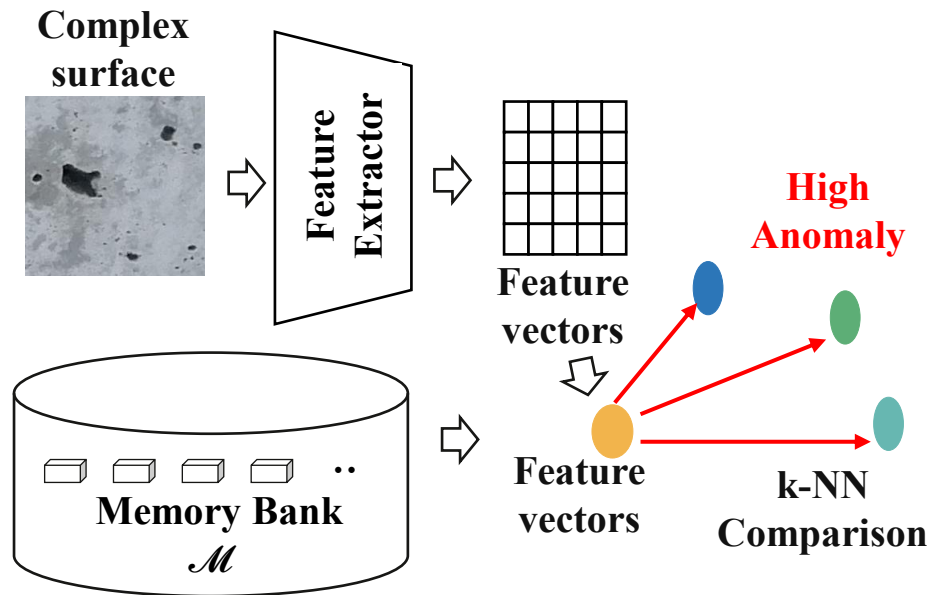


[Figure 15] ideal feature space

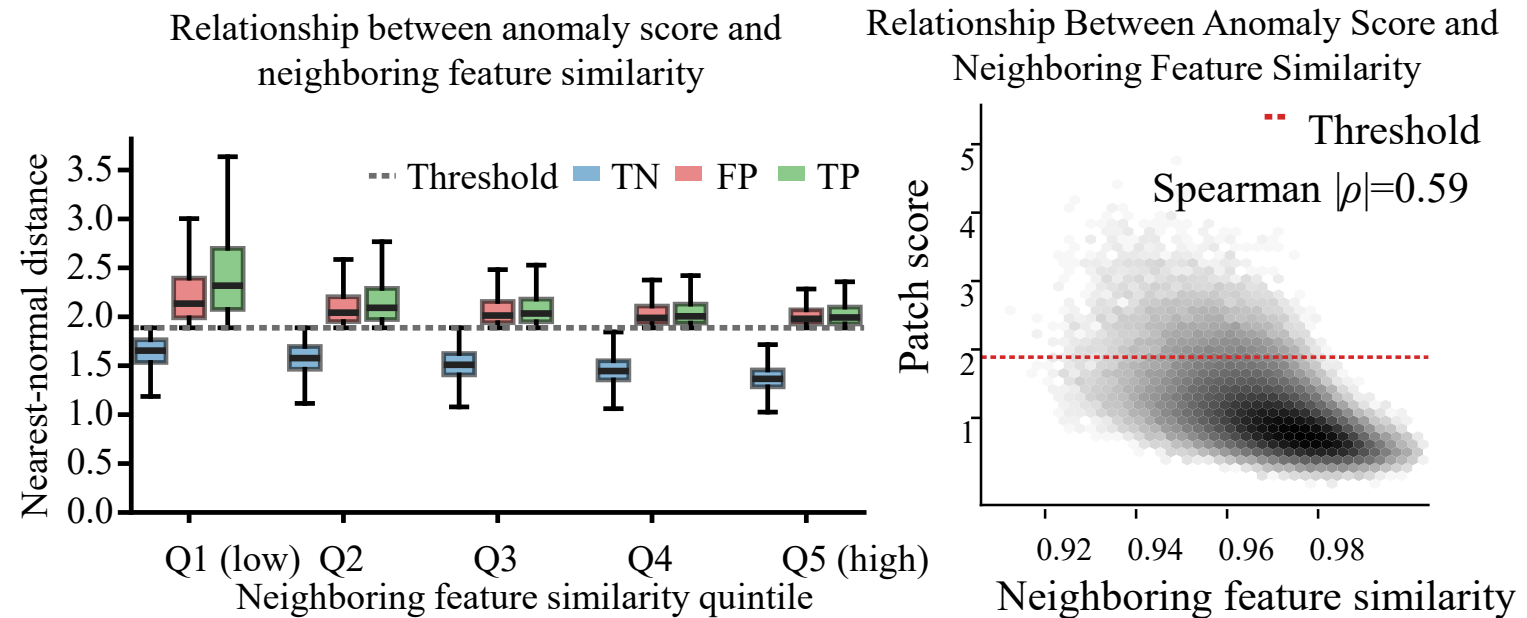
Motivation

Drawback of Existing Process

- When a crack-free patch looks different from its neighbors due to holes, stains, or exposed aggregates common on concrete and pavement surfaces, its representation may lie far from the stored normal features because such patterns are rare in the training data.
- Normal patches with low neighboring feature similarity are widely dispersed even within the same similarity group, so their distances to the nearest stored normal features can be comparable to those of crack patches, resulting in false positives in crack-free regions.



[Figure 16] Existing memory-bank matching on a complex normal surface

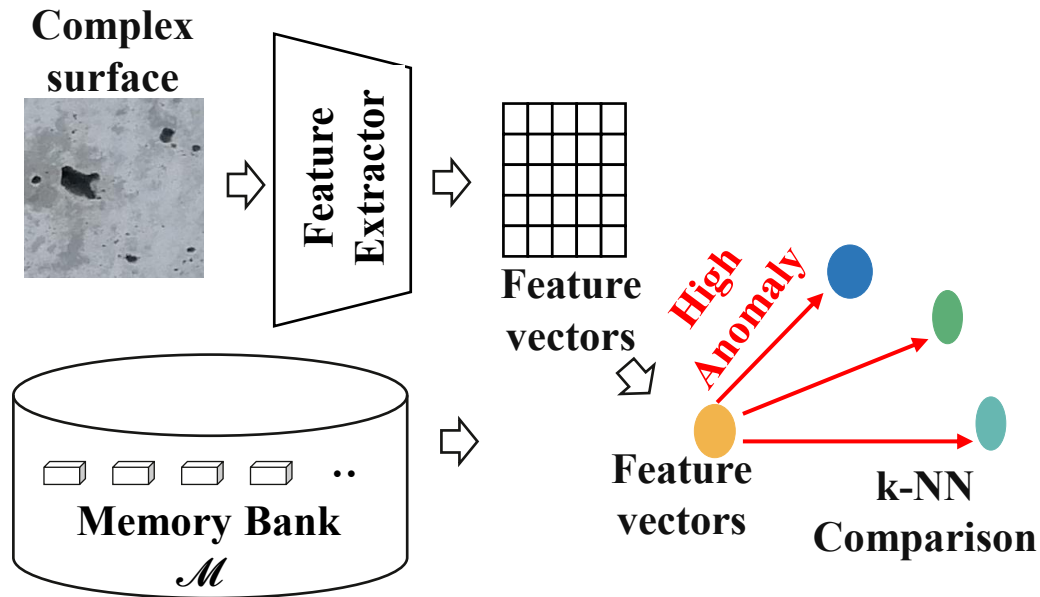


[Figure 17] Nearest-normal distance of TN, FP, and TP patches and its relationship with neighboring feature similarity

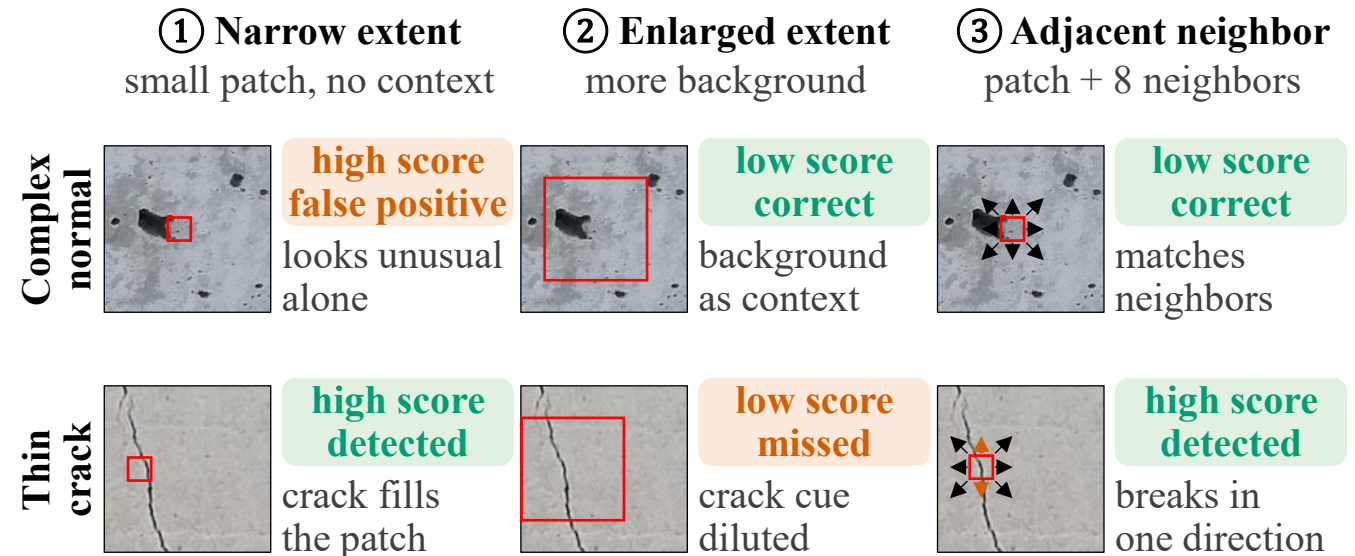
Motivation

How to Fix It?

- Because each patch is compared with the stored normal features independently, a narrow spatial extent cannot see the neighboring context and causes false positives on holes or stains, whereas an enlarged extent dilutes thin-crack cues and causes missed cracks.
- To resolve this, the small patch size should be kept to preserve thin-crack cues, while each patch is judged by its relations to the adjacent neighboring patches in each direction, so that normal variation is recognized from context without averaging it away.



[Figure 18] Existing memory-bank matching on a complex normal surface



[Figure 19] Spatial Extent of a Patch vs. Limited-Range Neighbor Context Along the Crack Direction

Proposed Method

Notation and Definition

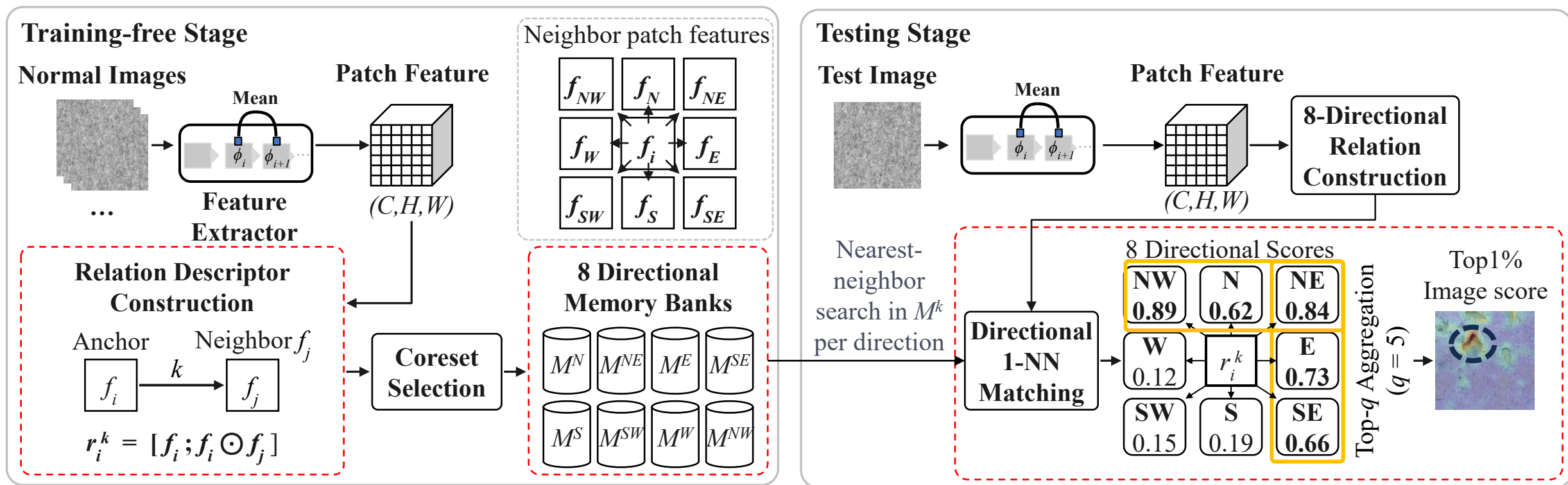
- The crack dataset is defined as $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$, where x_i denotes an input crack image, $y_i \in \{0, 1\}$ represents its image-level normal or crack label, and N denotes the total number of images included in the dataset for image-level anomaly detection tasks.
- The proposed process takes a reference patch feature f_i and its adjacent patch feature f_j as inputs, constructs a directional relation r_i^k , and compares it with the corresponding direction-specific normal memory bank $\mathcal{M}^k$ to obtain the directional distance d_i^k .

Notation	Name	Definition
$\mathcal{D}$	Dataset	Entire crack anomaly dataset consisting of $\{(x_i, y_i)\}$ pairs
x_i	Input Image	i-th input crack image
y_i	label	Image-level label, where $y_i \in \{0, 1\}$ denotes normal or crack
f_i	Patch Feature	Patch feature extracted at reference position i
k	Direction	Spatial direction from the reference patch to its neighboring patch
$j = i + \delta^k$	Neighbor Position	Neighbor of position i in direction k
r_i^k	Relation Feature	Relation between patches i and j, $[f_i; f_i \odot f_j]$
$\mathcal{M}^k$	Memory Bank	Set of normal relation features stored for direction k
d_i^k	Nearest-Neighbor Distance	Squared L2 distance from the standardized r_i^k to its nearest relation in $\mathcal{M}^k$
q	top-q	Number of largest d_i^k averaged ($q = 5$)
s_i	Patch Anomaly Score	Average of the q largest d_i^k over directions at position i
s^*	Image-level Anomaly Score	Image-level anomaly score obtained by averaging the top-p ratio of patch scores s_i

Proposed Method

Overall Architecture

- This study proposes an unsupervised crack anomaly detection method that compares directional relations between neighboring patches with branch-aware multiple memory banks, in order to reduce the false positives on normal patches that differ from their neighbors.
- In the training-free stage, the directional relation features extracted from normal images are stored in eight direction-specific memory banks, and at test time, each patch is scored by how closely its relations to its adjacent neighbors match the stored normal relations.



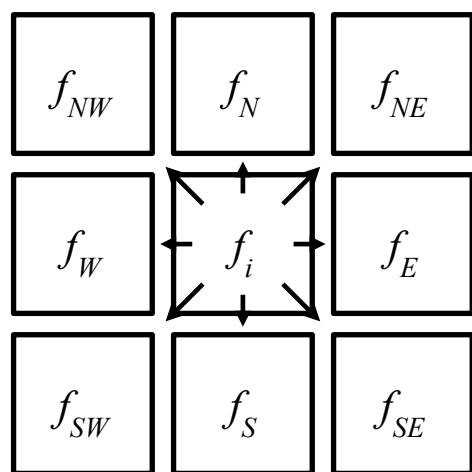
[Figure 20] Overview of the proposed model architecture

Proposed Method

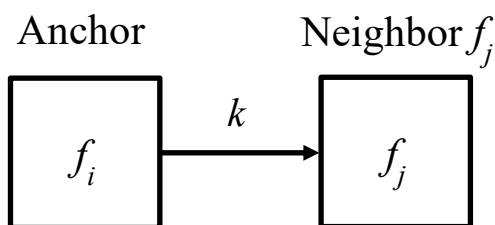
Directional Relation Memory Bank (DRMB)

- DRMB constructs a relation feature $r_i^k = [f_i; f_i \odot f_j]$ between each patch feature f_i and its neighboring patch feature f_j in each of the eight directions k , and then stores the resulting normal relation features in the corresponding direction-specific memory bank $\mathcal{M}^k$.
- At test time, each r_i^k is matched to its memory bank $\mathcal{M}^k$ to obtain the nearest-neighbor distance d_i^k , and the q largest of the eight distances are averaged into the patch score s_i , which reflects how a patch relates to its neighbors rather than its isolated appearance.

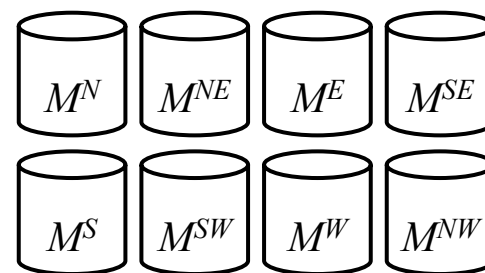
Relation Descriptor Construction



Neighbor patch features

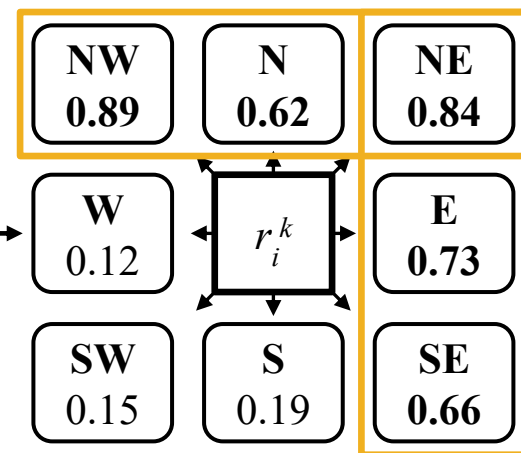


8 Directional Memory Banks



Directional
1-NN
Matching

8 Directional Scores



Top- q Aggregation
($q = 5$)

[Figure 21] Relation descriptor construction and directional scoring of DRMB

Proposed Method

Pseudocode

- The inference procedure takes a test image x , the eight direction-specific memory banks $\mathcal{M}^k$, the normal relation statistics (μ, σ) , the number of selected directions q , and the top-patch ratio p as inputs, and outputs the image-level anomaly score s^* of the test image.
- For each patch i and direction k , the relation feature r_i^k is compared with the memory bank $\mathcal{M}^k$ to obtain the distance d_i^k , the top- q distances are averaged into the patch score s_i , and the top- p fraction of patch scores is averaged to obtain the image score s^* .

Algorithm 2: Inference procedure of DRMB

Input: Test image x , $\{\mathcal{M}^k\}$, (μ, σ) , q , p

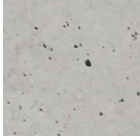
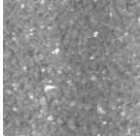
Output: Image score s^*

```
1: Extract  $\ell_2$ -normalized patch features  $\{f_i\}$  of  $x$ ;  
2: foreach patch position  $i$  do  
3:   foreach direction  $k \in \Omega_i$  do  
4:      $j \leftarrow i + \delta^k$ ;  
5:      $r_i^k \leftarrow [f_i; f_i \odot f_j]$ ;  
6:      $\bar{r}_i^k \leftarrow (r_i^k - \mu) \oslash \sigma$ ;  
7:      $d_i^k \leftarrow \min_{m \in \mathcal{M}^k} \|\bar{r}_i^k - m\|_2^2$ ;  
8:   end  
9:    $Q_i \leftarrow$  the  $q$  directions with the largest  $d_i^k$ ;  
10:   $s_i \leftarrow \frac{1}{|Q_i|} \sum_{k \in Q_i} d_i^k$ ;  
11: end  
12:  $s^* \leftarrow$  the average of the  $K$  largest patch scores  $s_i$ ;  
13: return  $s^*$ 
```

Line 1 $\rightarrow$ Extract ℓ_2 -normalized patch features f_i from x .
Line 2 $\rightarrow$ Iterate over each patch position i .
Line 3 $\rightarrow$ Iterate over each valid direction $k \in \Omega_i$.
Line 4 $\rightarrow$ Locate the neighbor position j in direction k .
Line 5 $\rightarrow$ Construct the relation feature $r_i^k = [f_i; f_i \odot f_j]$.
Line 6 $\rightarrow$ Standardize r_i^k with the normal statistics (μ, σ) .
Line 7 $\rightarrow$ Compute the directional distance d_i^k to $\mathcal{M}^k$.
Line 8 $\rightarrow$ Complete the directional scoring.
Line 9 $\rightarrow$ Select the q directions with the largest d_i^k .
Line 10 $\rightarrow$ Average the selected distances into the patch score s_i .
Line 11 $\rightarrow$ Complete the patch scoring.
Line 12 $\rightarrow$ Average the K largest s_i into the image score s^* .
Line 13 $\rightarrow$ Return the image-level anomaly score s^* .

Experimental Settings

Dataset

Dataset	Image	Train Image	Test Image	Label	Input Size
NCCD-PF		3,958	1,430	Image-level label	224×224
Crack500		700	300	Image-level label	224×224
IPAS		500	600	Image-level label	224×224
IRT-V		371	1,945	Image-level label	224×224

Cross Validation

- Data Split Ratio: Train : Test = 100% : 100%
- In the unsupervised setting, only normal images were used to build the memory banks, while both normal and crack images were used for testing.

Experimental Settings

Counterpart and Hyperparameter Setting

NN = Nearest Neighbor

Method	Batch Size	Epochs	Loss / Score	Optimizer	Learning Rate
PatchCore [5] (CVPR, 2022)	8 (feature extraction)	0 (memory build only)	k-NN (L2) distance; coreset + FAISS	None	(not used)
MGSCA [6] (IEEE TIM, 2024)	8 (feature extraction)	30	Coupled-hypersphere loss / multi-center distance	AdamW	1e-4
UniNet [7] (CVPR, 2025)	8 (feature extraction)	100	Contrastive + margin loss / cosine distance	AdamW	5e-3 (student), 1e-6 (teacher)
MSFlow [8] (IEEE TNNLS, 2025)	8 (feature extraction)	100	Negative log-likelihood	AdamW	1e-4

Evaluation Measures

- Image-level F1 Score: $F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$ represents the harmonic mean of precision and recall, serving as a complementary quantitative metric that balances the trade-off between false positives and false negatives for image-wise anomaly detection in crack inspection tasks.
- Image-level Average Precision (AP): $AP = \int_0^1 \text{Precision}(\text{Recall}) d\text{Recall}$ represents the area under the precision-recall curve, evaluating the trade-off between precision and recall of image-wise anomaly scores across all possible threshold settings and score distributions.
- Image-level AUROC: $AUROC = \int_0^1 \text{TPR} d(\text{FPR})$ measures how well normal and anomalous image scores are separated over all thresholds.
- Image-level FPR = $FP / (FP + TN)$: the proportion of normal images misjudged as anomalous at the F1-max threshold.

Experimental Results

Performance Comparison

- To verify that comparing directional relations improves crack detection on complex normal surfaces, the proposed method was compared with eleven existing unsupervised methods on four crack datasets using image-level AUROC and AP over all thresholds.
- The proposed method achieved the highest AUROC and AP on NCCD-PF, Crack500, and IPAS and the highest mean of 0.928 for both, exceeding PatchCore, which uses the same features, by 0.035 in AUROC and the best baseline by 0.025 in AP on NCCD-PF.

[Table 2] Results of Detection experiments on Image-level AUROC performance

	AUROC			
	Proposed	PatchCore	MSFlow	MGSCA
NCCD-PF	0.7706 ± 0.0010	0.7355 ± 0.0011▼	0.5593 ± 0.0301▼	0.6955 ± 0.0085▼
Crack500	0.9649 ± 0.0010	0.9525 ± 0.0006▼	0.9240 ± 0.0119▼	0.8702 ± 0.0139▼
IPAS	0.9986 ± 0.0000	0.9979 ± 0.0002▼	0.9541 ± 0.0085▼	0.9788 ± 0.0069▼
IRT-V	0.9774 ± 0.0002▼	0.9842 ± 0.0003	0.9742 ± 0.0068▼	0.9777 ± 0.0025▼

[Table 3] Results of Detection experiments on Image-level AP performance

	AP			
	Proposed	PatchCore	MSFlow	MGSCA
NCCD-PF	0.7547 ± 0.0010	0.7267 ± 0.0014▼	0.5408 ± 0.0251▼	0.7299 ± 0.0082▼
Crack500	0.9596 ± 0.0011	0.9473 ± 0.0009▼	0.8886 ± 0.0206▼	0.8949 ± 0.0102▼
IPAS	0.9996 ± 0.0000	0.9993 ± 0.0000▼	0.9739 ± 0.0060▼	0.9931 ± 0.0022▼
IRT-V	0.9975 ± 0.0000▼	0.9990 ± 0.0000	0.9977 ± 0.0007▼	0.9986 ± 0.0001▼

*Best results are marked in bold. A triangle (▼) indicates statistical significance at the 0.05 level based on paired t-tests against the proposed model.

Experimental Results

Performance Comparison

- At the F1-max threshold of each method, the proposed method achieved the highest F1-max on all four datasets and the lowest FPR on NCCD-PF, Crack500, and IPAS, so fewer crack-free images were misjudged as anomalous than with the existing baselines.
- On NCCD-PF, whose normal images contain many holes and stains, the FPR decreased from 0.5351 with PatchCore to 0.4716, whereas on IRT-V, whose normal surfaces are relatively uniform, PatchCore retained the lowest FPR of 0.1391 compared with 0.1739.

[Table 4] Results of Detection experiments on Image-level F1-Max performance

	F1-Max			
	Proposed	PatchCore	MSFlow	MGSCA
NCCD-PF	0.7305 ± 0.0014	0.7051 ± 0.0022▼	0.6703 ± 0.0027▼	0.6742 ± 0.0043▼
Crack500	0.9165 ± 0.0024	0.8908 ± 0.0024▼	0.8820 ± 0.0129▼	0.8151 ± 0.0099▼
IPAS	0.9955 ± 0.0000	0.9900 ± 0.0007▼	0.9611 ± 0.0043▼	0.9603 ± 0.0089▼
IRT-V	0.9941 ± 0.0000	0.9934 ± 0.0004▼	0.9940 ± 0.0008▼	0.9897 ± 0.0016▼

[Table 5] Results of Detection experiments on Image-level FPR performance

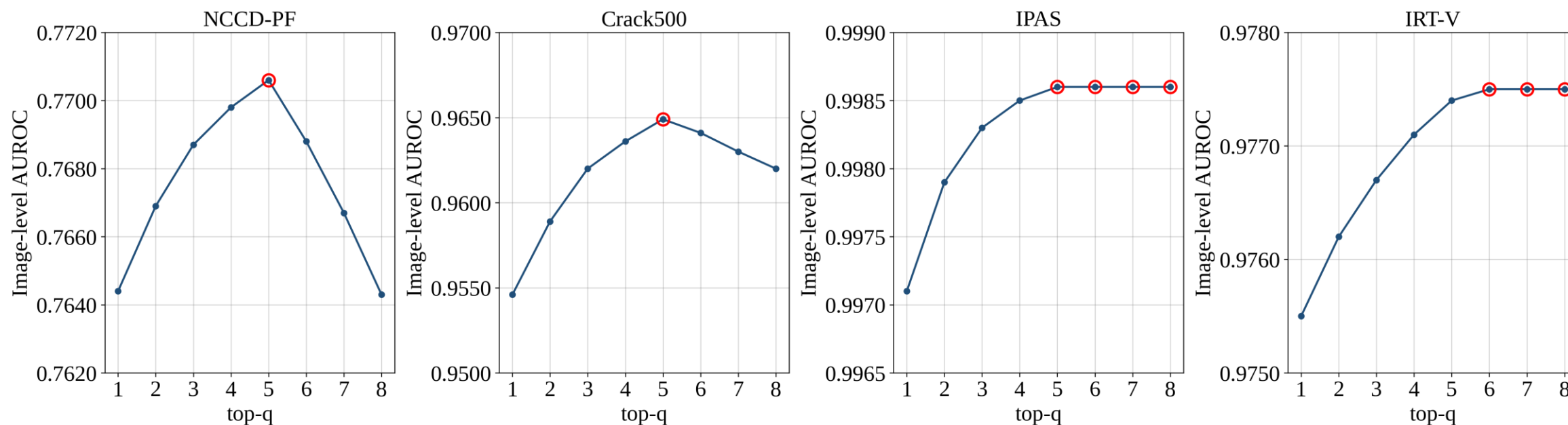
	FPR			
	Proposed	PatchCore	MSFlow	MGSCA
NCCD-PF	0.4716 ± 0.0070	0.5351 ± 0.0966▼	0.9561 ± 0.0414▼	0.7852 ± 0.0972▼
Crack500	0.1067 ± 0.0202	0.2053 ± 0.0186▼	0.1747 ± 0.0412▼	0.1840 ± 0.0369▼
IPAS	0.0067 ± 0.0000	0.0173 ± 0.0068▼	0.1333 ± 0.0283▼	0.1227 ± 0.0478▼
IRT-V	0.1739 ± 0.0194▼	0.1391 ± 0.0211	0.2000 ± 0.0499▼	0.2674 ± 0.0341▼

*Best results are marked in bold. A triangle (▼) indicates statistical significance at the 0.05 level based on paired t-tests against the proposed model.

Experimental Results

Comparison of the Number of Selected Directions (q)

- To examine how the number of selected directions affects detection, q is varied from 1 to 8 with other settings fixed and the image-level AUROC is measured on four crack datasets. Here, $q = 1$ keeps only the largest distance, and $q = 8$ averages all eight.
- Setting $q = 5$ yields the highest or near-highest AUROC on all four datasets, while using a single direction performs worst on Crack500, IPAS, and IRT-V, so $q = 5$ is adopted as a fixed default for all crack datasets without requiring dataset-specific tuning.

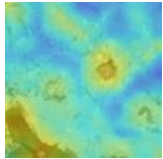
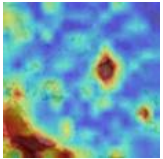
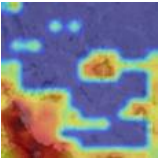
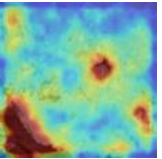
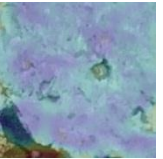
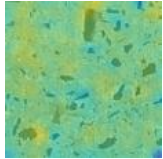
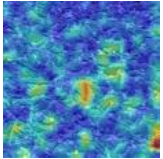
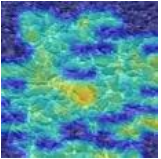
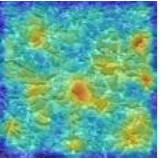
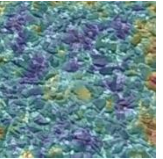
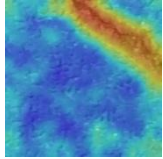
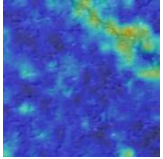
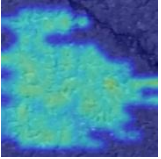
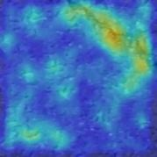
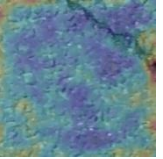
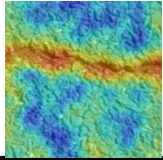
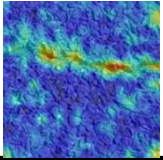
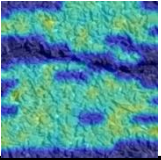
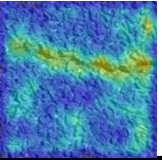
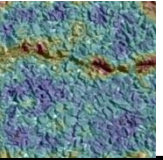


[Figure 22] Image-level AUROC according to the number of selected directions q on four crack datasets

Experimental Results

Qualitative Analysis

- Qualitative analysis examines how local directional relation modeling reduces false-positive responses in normal regions with low neighboring feature similarity by incorporating neighboring patch information within a limited spatial range for local comparison.
- Anomaly maps of the proposed method, MSFlow, UniNet, PatchCore, and MGSCA are compared on normal and crack images, focusing on false-positive responses in complex normal regions and on whether high responses are maintained along actual cracks.

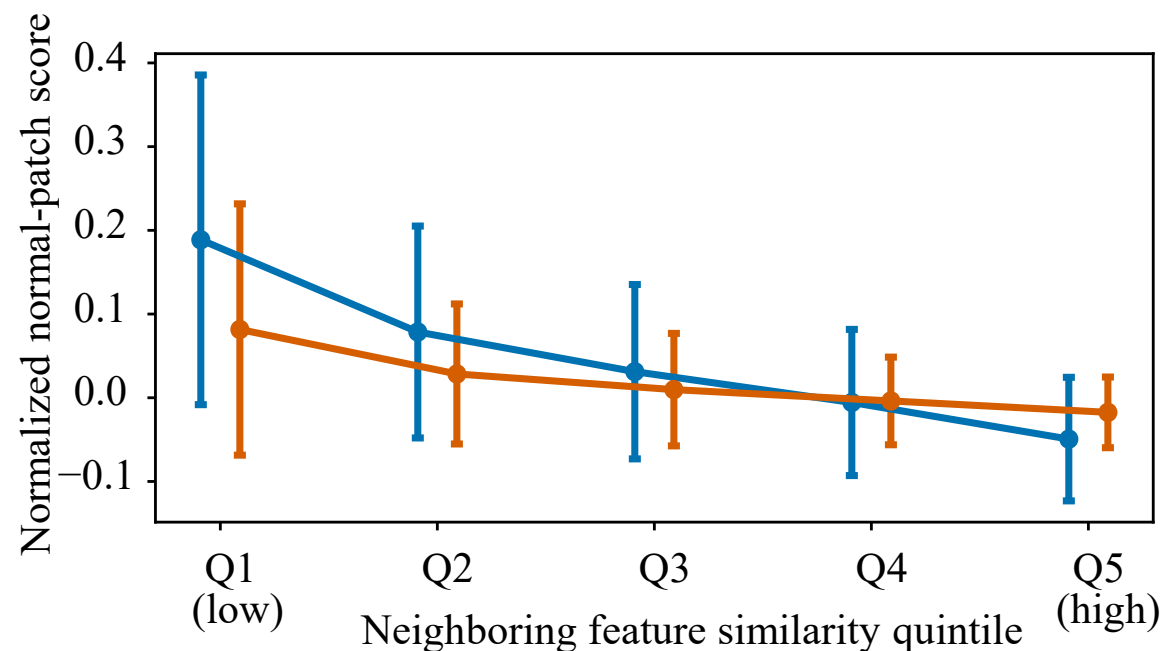
Images / Method	Input Images	Proposed	MSFlow	UniNet	PatchCore	MGSCA
Normal						
Normal						
Crack						
Thin crack						

[Figure 23] Qualitative comparison of anomaly maps from DRMB and existing UAD methods on normal and crack images

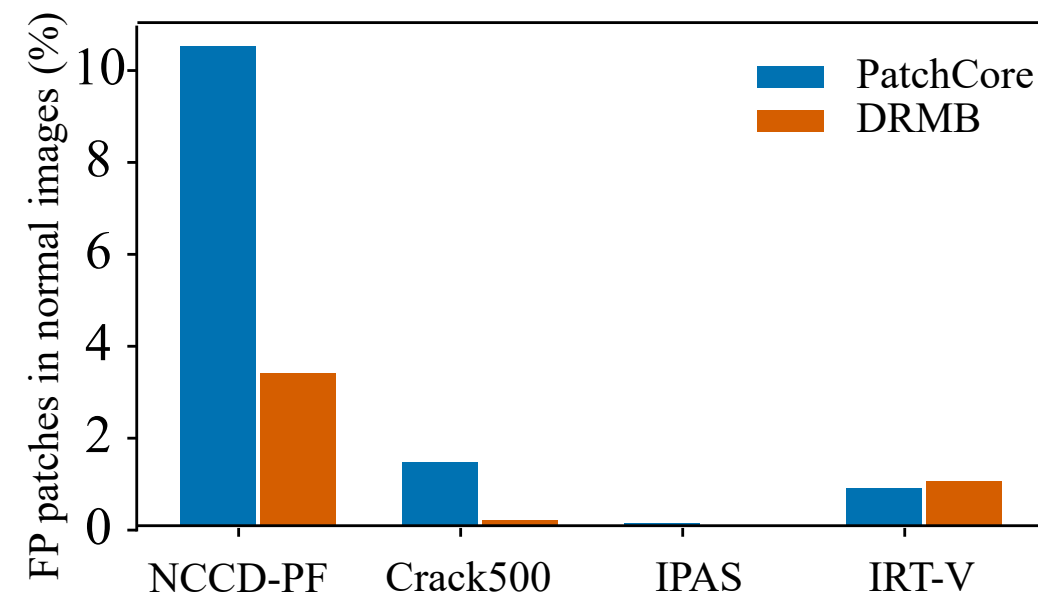
Experimental Results

In-depth Analysis of Normal Patches

- On NCCD-PF, PatchCore assigned higher scores to normal patches with lower neighboring feature similarity, whereas DRMB lowered both the scores and their spread in the lowest-similarity group by comparing each patch with normal neighbor relations.
- At the F1-max threshold, the false-positive patch ratio in normal images fell from 10.43% with PatchCore to 3.31% with DRMB on NCCD-PF and from 1.38% to 0.13% on Crack500, while on IRT-V both methods stayed below 1%, with DRMB slightly higher.



[Figure 24] Normal patches vs. neighboring feature similarity (NCCD-PF)

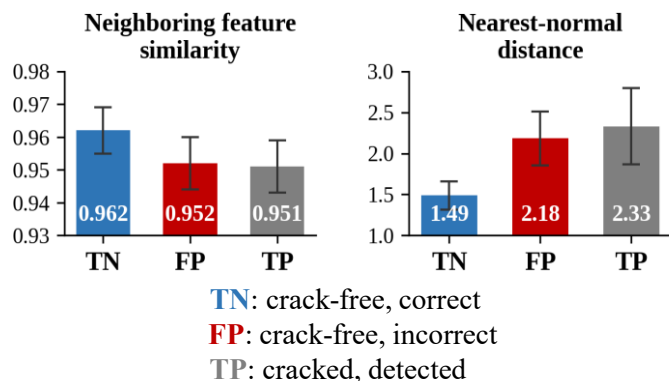


[Figure 25] False-positive patches at F1-max threshold

Conclusion

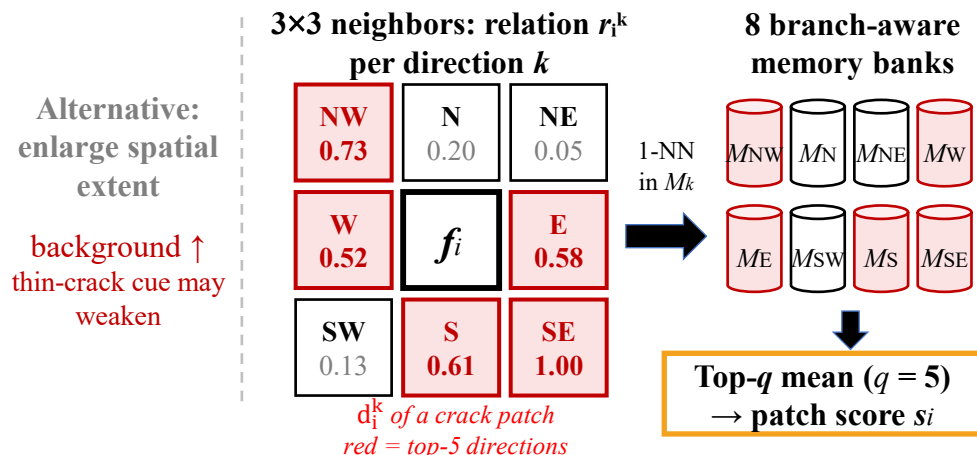
- This study aims to improve unsupervised crack anomaly detection by reducing false positives on normal patches with low neighboring feature similarity, such as holes and stains, and by exploiting local spatial context to distinguish complex normal patterns from actual cracks.
- Unlike existing studies, the proposed method models directional relations between neighboring patches and stores them in branch-aware multiple memory banks, enabling local crack continuity to be considered without simply enlarging the spatial extent of individual patches.
- Results show that DRMB improves detection on three of the four datasets by reducing false positives in normal regions with low neighboring feature similarity while preserving crack responses, indicating that directional relations help mainly on complex normal surfaces.

(a) Problem: single-patch matching



FP $\approx$ TP < TN in neighboring similarity
→ normal patches unlike their neighbors get large distances, inseparable from cracks

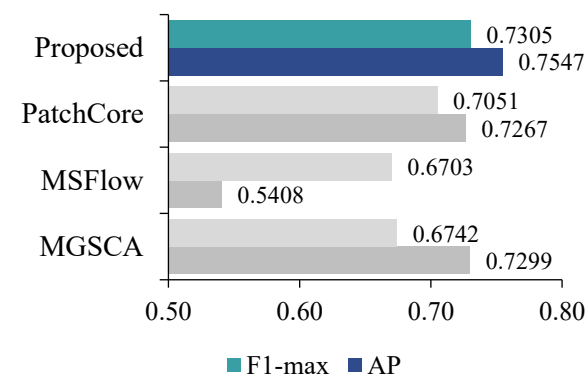
(b) Proposed: directional relations + branch-aware memory banks



$r_i^k = [f_i; f_i \odot f_j]$ (memory 5% = PatchCore 10% storage)
→ Local crack continuity without enlarging each patch

(c) Result: NCCD-PF

Image-level AP / F1-max (NCCD-PF)



AP + 0.025, F1 + 0.025
over the best baseline
(AP: MGSCA, F1: PatchCore)

[Figure 26] Summary of the Proposed Approach and Results.

Thank you

References

- [1] J. Deng, et al., “Review on computer vision-based crack detection and quantification methodologies for civil structures,” *Construction and Building Materials*, vol. 356, Art. no. 129238, 2022.
- [2] D. Ai, et al., “Computer vision framework for crack detection of civil infrastructure—A review,” *Engineering Applications of Artificial Intelligence*, vol. 117, Art. no. 105478, 2023.
- [3] G. Pang, et al., “Deep learning for anomaly detection: A review,” *ACM Computing Surveys*, vol. 54, no. 2, Art. no. 38, pp. 1–38, 2021.
- [4] Y. Lin, et al., “A survey on RGB, 3D, and multimodal approaches for unsupervised industrial image anomaly detection,” *Information Fusion*, vol. 121, Art. no. 103139, 2025.
- [5] K. Roth, et al., “Towards total recall in industrial anomaly detection,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, Jun. 18–24, 2022, pp. 14318–14328.
- [6] Y. Zhang, et al., “MGSCA: Multiscale gradient selection-based coupled-hypersphere adaptation for unsupervised pavement defect detection,” *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–13, 2024.
- [7] S. Wei, J. Jiang, and X. Xu, “UniNet: A contrastive learning-guided unified framework with feature selection for anomaly detection,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, Jun. 11–15, 2025, pp. 9994–10003.
- [8] Y. Zhou, X. Xu, J. Song, F. Shen, and H. T. Shen, “MSFlow: Multiscale flow-based framework for unsupervised anomaly detection,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 36, no. 2, pp. 2437–2450, 2025.
- [9] K. Tomaszekiewicz and T. Owerko, “A pre-failure narrow concrete cracks dataset for engineering structures damage classification and segmentation,” *Scientific Data*, vol. 10, Art. no. 925, 2023.
- [10] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, “Road crack detection using deep convolutional neural network,” in *Proc. IEEE ICIP*, Phoenix, AZ, USA, Sep. 25–28, 2016, pp. 3708–3712.
- [11] F. Yang, L. Zhang, S. Yu, D. Prokhorov, X. Mei, and H. Ling, “Feature pyramid and hierarchical boosting network for pavement crack detection,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 4, pp. 1525–1535, 2020.
- [12] Y. Wang, M. Niu, K. Song, P. Jiang, and Y. Yan, “Normal-knowledge-based pavement defect segmentation using relevance-aware and cross-reasoning mechanisms,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 4, pp. 4413–4427, 2023.
- [13] F. Liu, J. Liu, and L. Wang, “Deep learning and infrared thermography for asphalt pavement crack severity classification,” *Automation in Construction*, vol. 140, Art. no. 104383, 2022.