JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

FUDAN UNIVERSITY

Prototypical Residual Networks for Anomaly Detection and Localization

Hui Zhang [1,2] Zuxuan Wu [1,2] Zheng Wang [3] Zhineng Chen [1,2] * Yu-Gang Jiang [1,2]

[1] Shanghai Key Lab of Intell. Info. Processing, School of CS, Fudan University
[2] Shanghai Collaborative Innovation Center of Intelligent Visual Computing
[3] School of Computer Science, Zhejiang University of Technology

Poster: WED-PM-374

Prototypical Residual Networks for Anomaly Detection and Localization

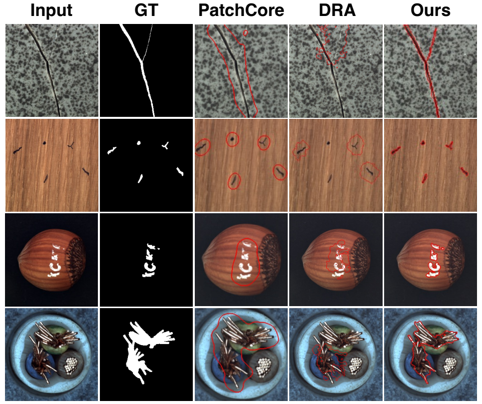
Hui Zhang [1,2] Zuxuan Wu [1,2] Zheng Wang [3] Zhineng Chen [1,2]* Yu-Gang Jiang [1,2]

[1] Shanghai Key Lab of Intell. Info. Processing, School of CS, Fudan University [2] Shanghai Collaborative Innovation Center of Intelligent Visual Computing

[3] School of Computer Science, Zhejiang University of Technology

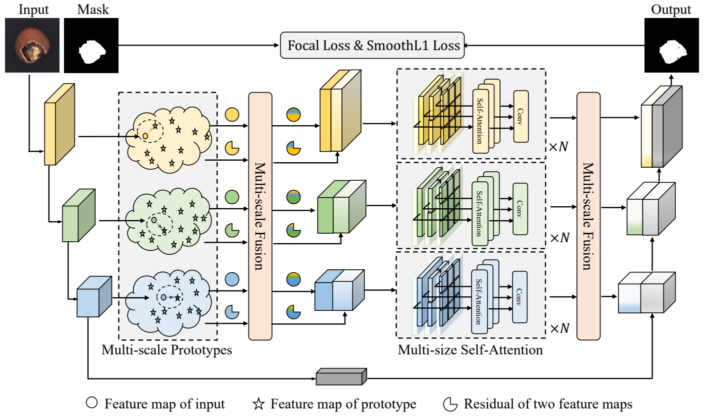
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Anomaly Detection

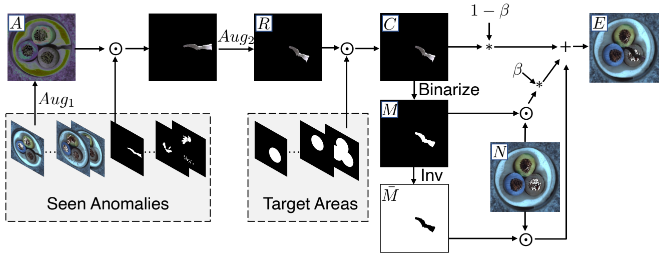


- We propose a novel network named PRN for anomaly detection.
- We propose a variety of anomaly generation strategies.
- PRN outperforms current SOTA on four datasets.

Overview of PRN



Anomaly Generation Strategies



Extended anomaly: $E = \bar{M} \odot N + (1-\beta)C + \beta(M \odot N)$

A - augmented anomaly

R - augmented anomaly region

M - the mask of anomaly

β - opacity parameter

Training

Multi-scale prototypes obtained by clustering:

$$\mathcal{P}_j \in \mathbb{R}^{K \times c^j \times h^j \times w^j}$$

Residual representation:

$$\mathcal{D}_{i,j} = D(\mathcal{F}_{i,j} - \mathcal{P}_j^*),$$

$$\text{s.t. } \mathcal{P}_j^* = \underset{\mathcal{P}_j^k \subset \mathcal{P}_j}{\arg\min} \|\mathcal{F}_{i,j} - \mathcal{P}_j^k\|_2$$

Multi-scale Fusion:

$$\mathcal{F}_{i,j}^* = f_{1j}(\mathcal{F}_{i,1}) + f_{2j}(\mathcal{F}_{i,2}) + f_{3j}(\mathcal{F}_{i,3})$$

Multi-size Self-Attention:

$$\mathcal{A}_{i,j}^s = \text{softmax}\left(\frac{\mathcal{Q}_{i,j}^s (\mathcal{K}_{i,j}^s)^T}{c^s}\right)\mathcal{V}_{i,j}^s \quad p_s \in \{h^j, h^j/2, h^j/4, h^j/8\}$$

Loss:

$$\mathcal{L}_{total} = \text{Smooth}_{\mathcal{L}1}(\mathcal{M}_o, \mathcal{M}) + \lambda \mathcal{L}_{focal}(\mathcal{M}_o, \mathcal{M})$$

Experiments

	MVTec				DAGM				BTAD				KolektorSDD2			
	I↑	P↑	O↑	A↑	I↑	P↑	O↑	A↑	I↑	P↑	O↑	A↑	I↑	P↑	O↑	A↑
DRAEM	97.6	96.7	91.3	68.1	91.1	83.4	70.5	35.6	89.0	87.1	61.6	19.2	81.1	85.6	67.9	39.1
CFLOW	97.5	97.7	93.4	59.6	91.2	95.1	87.6	45.2	90.5	96.1	71.6	**54.0**	95.2	97.4	93.8	46.0
SSPCAB	97.1	96.3	90.8	65.5	90.4	84.5	71.9	33.9	88.3	83.5	54.1	13.0	83.4	86.2	66.1	44.5
RD4AD	98.7	97.8	93.9	55.4	90.7	94.1	85.5	40.8	94.4	96.9	75.8	53.5	96.0	**97.6**	94.7	43.5
PatchCore	99.2	98.1	93.9	56.3	92.5	96.1	88.0	49.0	92.6	96.9	76.3	51.5	94.6	97.1	89.3	49.8
DRA	96.1	85.3	73.3	26.0	93.5	95.1	88.8	47.6	94.2	75.4	56.2	12.4	86.8	84.4	56.9	3.6
Ours	**99.4**	**99.0**	**96.1**	**78.6**	**98.2**	**96.6**	**93.8**	**49.4**	**94.7**	**97.1**	**78.0**	**54.0**	**96.4**	**97.6**	**94.9**	**72.5**

Table 6. PRN outperforms current SOTA on four datasets. "I", "P", "O", "A" and "T" respectively refer to the five metrics of image auroc, pixel auroc, pixel pro, pixel ap, and inference time per image.

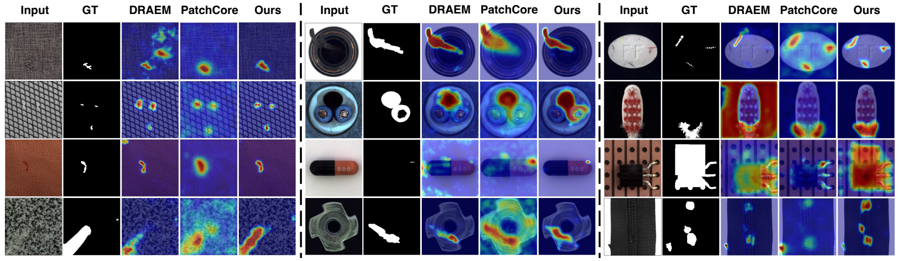


Figure 6. Qualitative examples on MVTec [4]. PRN achieves more accurate localization results for various types of anomalies.

Ablation Study

Module				Performance			
U-Net	MP	MSA	MF	I↑	P↑	O↑	A↑
✓				97.4	91.7	88.6	58.5
✓	✓		✓	98.9	98.5	95.3	77.0
✓		✓	✓	97.8	97.0	92.1	74.0
✓	✓	✓		98.7	98.5	95.4	78.1
✓	✓	✓	✓	**99.4**	**99.0**	**96.1**	**78.6**

Table 4. Ablations of different modules in PRN.

Anomaly Generation				Performance			
EA	HEA	HOA	TA	I↑	P↑	O↑	A↑
✓			✓	98.6	97.2	93.4	75.7
✓	✓		✓	99.1	98.4	95.4	77.4
✓		✓	✓	98.6	98.4	95.7	75.2
	✓	✓	✓	98.7	98.2	95.1	73.4
✓	✓	✓		98.4	98.4	94.9	77.6
✓	✓	✓	✓	**99.4**	**99.0**	**96.1**	**78.6**

Table 5. Ablations of anomaly generation strategies.

	I↑	P↑	O↑	A↑	T↓
$\text{PRN}_{5\%}$	99.2	98.6	95.4	78.1	**0.063**
$\text{PRN}_{10\%}$	**99.4**	**99.0**	**96.1**	**78.6**	0.064
$\text{PRN}_{20\%}$	99.2	98.8	95.7	77.3	0.066
$\text{PRN}_{100\%}$	86.2	91.4	75.4	49.9	0.074

Table 6. Ablations of the ratio of prototypes to total normal samples.

	DevNet [35]				DRA [13]				**PRN(Ours)**			
	I↑	P↑	O↑	A↑	I↑	P↑	O↑	A↑	I↑	P↑	O↑	A↑
1	79.6	75.3	51.0	16.5	88.9	78.8	58.2	19.1	98.8	98.3	95.4	74.7
5	86.7	83.7	66.9	22.7	93.5	82.8	68.6	21.9	99.2	98.6	95.6	76.4
10	92.2	85.3	71.4	24.4	96.1	85.3	73.3	26.0	**99.4**	**99.0**	**96.1**	**78.6**

Table 7. Impact of the number of seen anomalies used.

CONTENTS

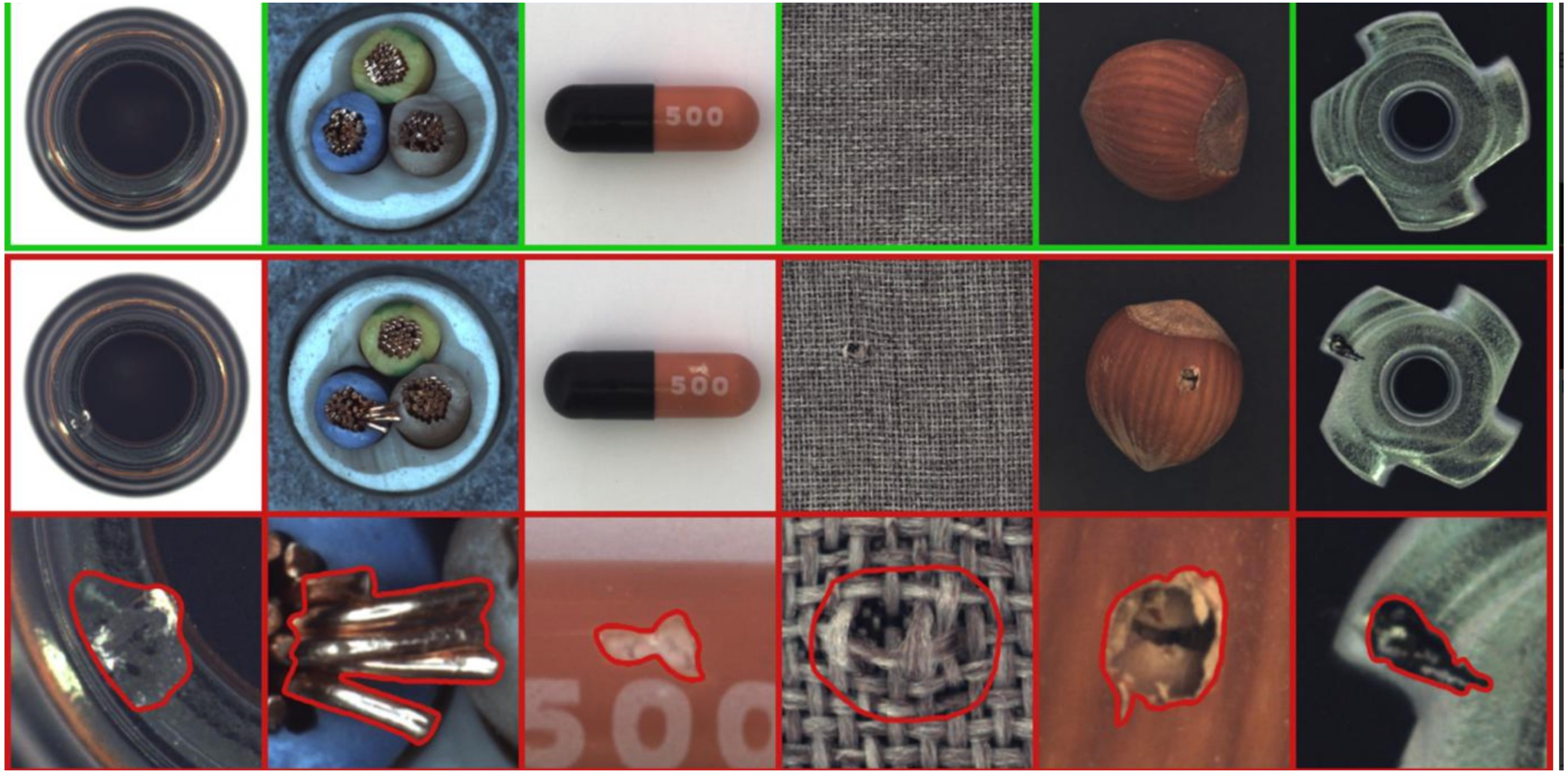
01 Introduction

Background

Anomaly detection and localization are widely used in industrial manufacturing for its efficiency and effectiveness.

Normal

Abnormal



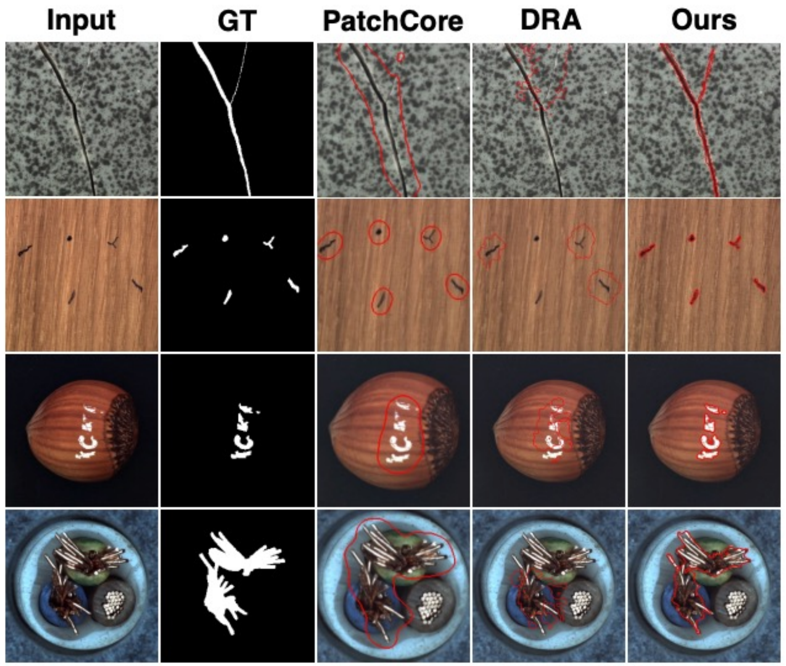
Motivation

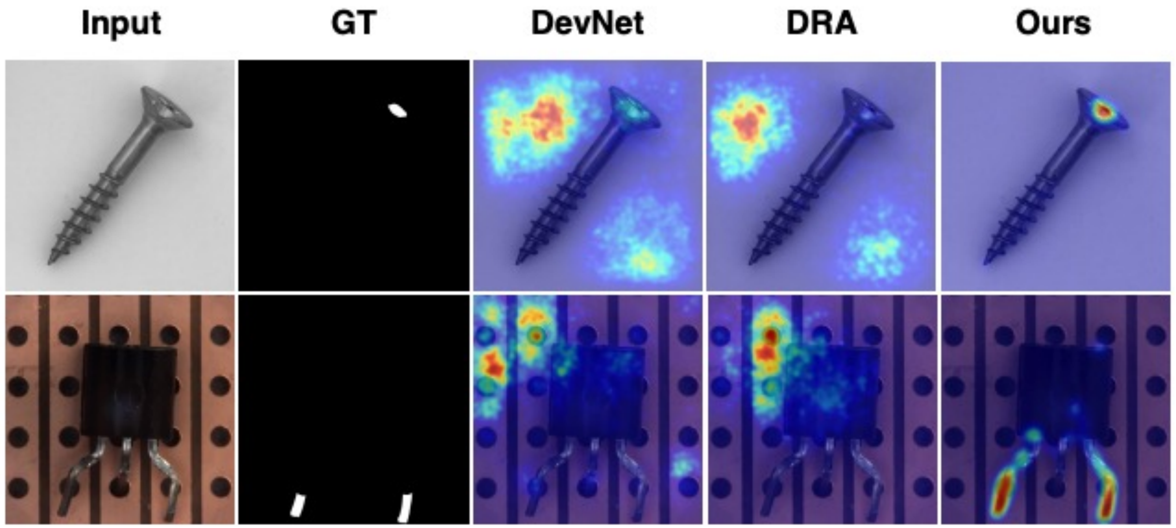
Difficulties & dominant paradigms

- ❖ Imbalanced learning
- ❖ Identifying abnormal regions
- ❖ Appearance variations

Unsupervised → Good performance | Implicit decisions

Supervised → Real-world | Label bias

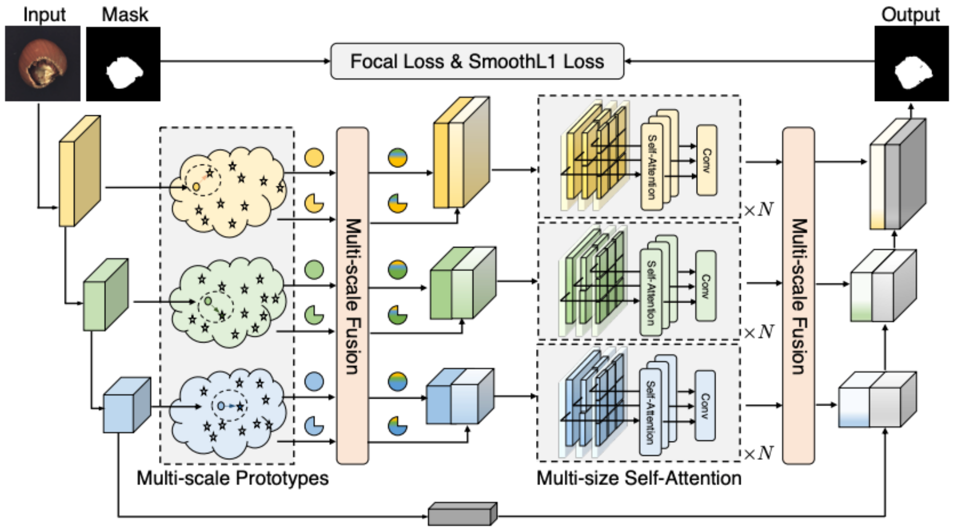


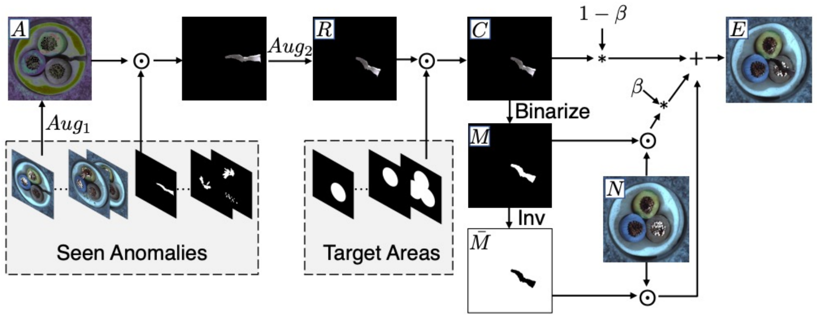


Contribution

We propose a framework called Prototypical Residual Network (PRN) as an effective remedy for aforesaid issues.

- PRN learns feature residuals of varying scales and sizes between anomalous and normal patterns, aiming to **address identifying abnormal regions and appearance variations**.
- We propose various anomaly-generation strategies to **address imbalanced learning.**
- PRN outperforms current SOTA on four datasets.



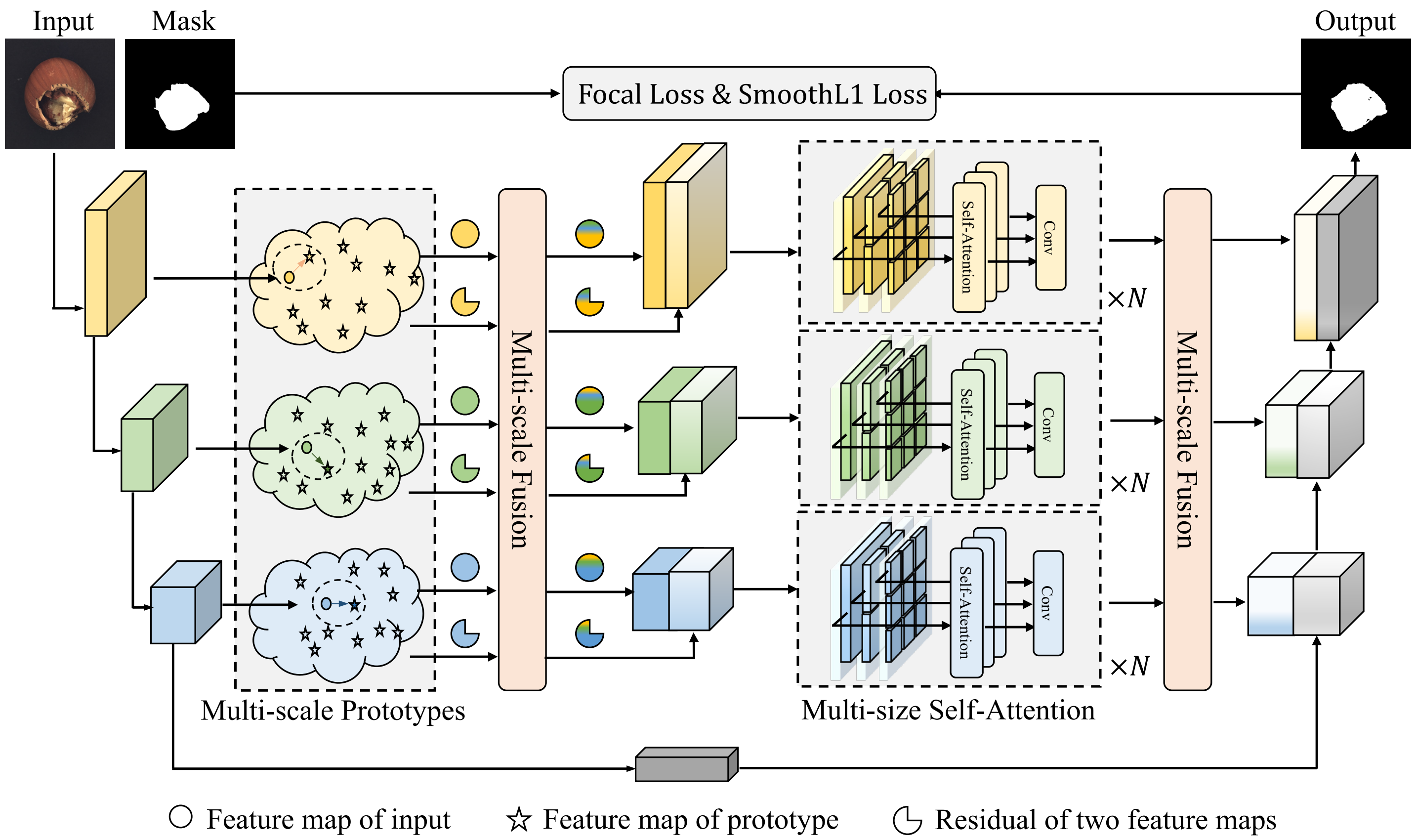


02

Method

Method

Overview



Training

Multi-scale prototypes : $\mathcal{P}_j \in \mathbb{R}^{K \times c^j \times h^j \times w^j}$

Residual representation: $\mathcal{D}_{i,j} = D(\mathcal{F}_{i,j} - \mathcal{P}_j^*)$, s.t. $\mathcal{P}_j^* = \arg\min_{\mathcal{P}_j^k \subset \mathcal{P}_j} \|\mathcal{F}_{i,j} - \mathcal{P}_j^k\|_2$

Multi-scale Fusion: $\mathcal{F}_{i,j}^* = f_{1j}(\mathcal{F}_{i,1}) + f_{2j}(\mathcal{F}_{i,2}) + f_{3j}(\mathcal{F}_{i,3})$

Multi-size Self-Attention: $\mathcal{A}_{i,j}^s = \text{softmax}\left(\frac{\mathcal{Q}_{i,j}^s (\mathcal{K}_{i,j}^s)^T}{c^s}\right)\mathcal{V}_{i,j}^s$

$p_s \in \{h^j, h^j/2, h^j/4, h^j/8\}$

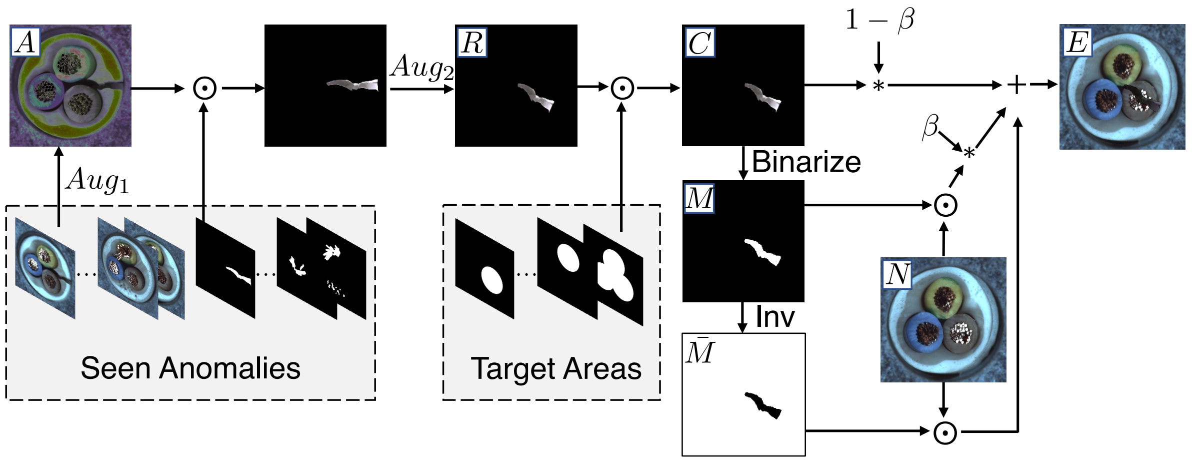
Loss: $\mathcal{L}_{total} = \text{Smooth}_{\mathcal{L}1}(\mathcal{M}_o, \mathcal{M}) + \lambda \mathcal{L}_{focal}(\mathcal{M}_o, \mathcal{M})$

Inference

Image-level anomaly score:
The average of the top-K anomalous pixels in the output.

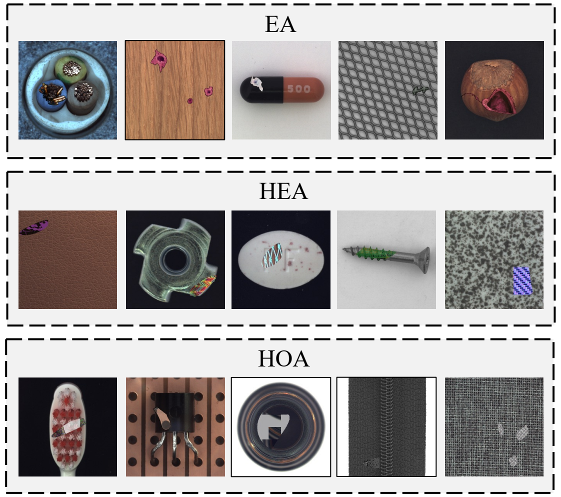
Anomaly Generation Strategies

Extended anomalies

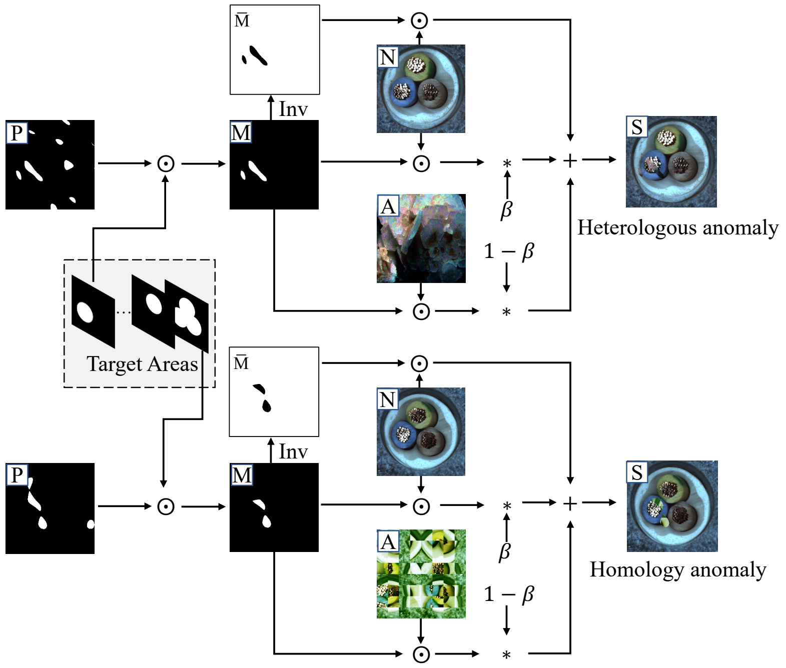


$$E = \bar{M} \odot N + (1-\beta)C + \beta\,(M \odot N)$$

More examples



Simulated Anomalies



$$S = \bar{M} \odot N + (1-\beta)(M \odot A) + \beta\,(M \odot N)$$

Notes

A - an augmented anomaly

R - augmented anomaly region

M - the mask of anomaly

β - opacity parameter

P - Threshold Perlin noise

03

Experiments

Numerical and visualization results

	MVTec				DAGM				BTAD				KolektorSDD2			
	I ↑	P ↑	O ↑	A ↑	I ↑	P ↑	O ↑	A ↑	I ↑	P ↑	O ↑	A ↑	I ↑	P ↑	O ↑	A ↑
DRAEM	97.6	96.7	91.3	68.1	91.1	83.4	70.5	35.6	89.0	87.1	61.6	19.2	81.1	85.6	67.9	39.1
CFLOW	97.5	97.7	93.4	59.6	91.2	95.1	87.6	45.2	90.5	96.1	71.6	**54.0**	95.2	97.4	93.8	46.0
SSPCAB	97.1	96.3	90.8	65.5	90.4	84.5	71.9	33.9	88.3	83.5	54.1	13.0	83.4	86.2	66.1	44.5
RD4AD	98.7	97.8	93.9	55.4	90.7	94.1	85.5	40.8	94.4	96.9	75.8	53.5	96.0	**97.6**	94.7	43.5
PatchCore	99.2	98.1	93.9	56.3	92.5	96.1	88.0	49.0	92.6	96.9	76.3	51.5	94.6	97.1	89.3	49.8
DRA	96.1	85.3	73.3	26.0	93.5	95.1	88.8	47.6	94.2	75.4	56.2	12.4	86.8	84.4	56.9	3.6
Ours	**99.4**	**99.0**	**96.1**	**78.6**	**98.2**	**96.6**	**93.8**	**49.4**	**94.7**	**97.1**	**78.0**	**54.0**	**96.4**	**97.6**	**94.9**	**72.5**

Table 6. PRN outperforms current SOTA on four datasets. “I”, “P”, “O”, “A” and “T” respectively refer to the five metrics of image auroc, pixel auroc, pixel pro, pixel ap, and inference time per image.

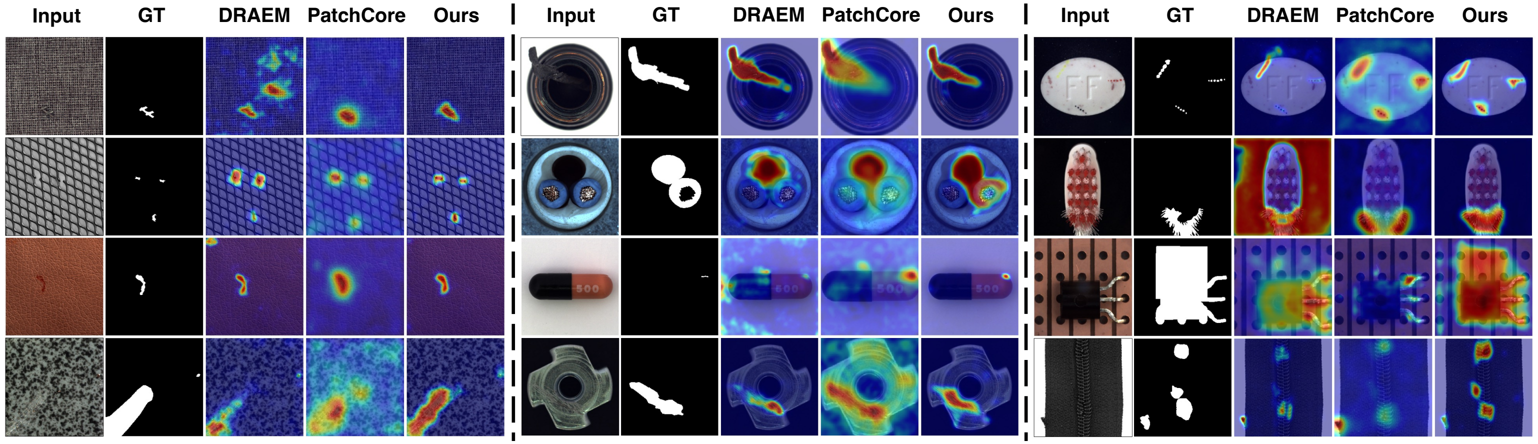


Figure 6. Qualitative examples on MVTec [4]. PRN achieves more accurate localization results for various types of anomalies.

Ablation Study

Module				Performance			
U-Net	MP	MSA	MF	I ↑	P ↑	O ↑	A ↑
✓				97.4	91.7	88.6	58.5
✓	✓		✓	98.9	98.5	95.3	77.0
✓		✓	✓	97.8	97.0	92.1	74.0
✓	✓	✓		98.7	98.5	95.4	78.1
✓	✓	✓	✓	**99.4**	**99.0**	**96.1**	**78.6**

Table 4. Ablations of different modules in PRN.

Anomaly Generation				Performance			
EA	HEA	HOA	TA	I ↑	P ↑	O ↑	A ↑
✓			✓	98.6	97.2	93.4	75.7
✓	✓		✓	99.1	98.4	95.4	77.4
✓		✓	✓	98.6	98.4	95.7	75.2
	✓	✓	✓	98.7	98.2	95.1	73.4
✓	✓	✓		98.4	98.4	94.9	77.6
✓	✓	✓	✓	**99.4**	**99.0**	**96.1**	**78.6**

Table 5. Ablations of anomaly generation strategies.

	I ↑	P ↑	O ↑	A ↑	T ↓
$PRN_{5\%}$	99.2	98.6	95.4	78.1	**0.063**
$PRN_{10\%}$	**99.4**	**99.0**	**96.1**	**78.6**	0.064
$PRN_{20\%}$	99.2	98.8	95.7	77.3	0.066
$PRN_{100\%}$	86.2	91.4	75.4	49.9	0.074

Table 6. Ablations of the ratio of prototypes to total normal samples.

	DevNet [35]				DRA [13]				**PRN(Ours)**			
	I ↑	P ↑	O ↑	A ↑	I ↑	P ↑	O ↑	A ↑	I ↑	P ↑	O ↑	A ↑
1	79.6	75.3	51.0	16.5	88.9	78.8	58.2	19.1	98.8	98.3	95.4	74.7
5	86.7	83.7	66.9	22.7	93.5	82.8	68.6	21.9	99.2	98.6	95.6	76.4
10	92.2	85.3	71.4	24.4	96.1	85.3	73.3	26.0	**99.4**	**99.0**	**96.1**	**78.6**

Table 7. Impact of the number of seen anomalies used.

04

Conclusion

Conclusion

Contributions & Limitations

Contributions

- We proposed a novel framework called Prototypical Residual Network (PRN) for anomaly detection and localization
- We proposed various anomaly generation strategies to expand and diversify the anomalies
- We conduct in-depth experiments on four popular datasets to confirm the effectiveness and generalizability of PRN

Limitations

- Our method requires ground truth masks of the seen anomaly samples
- Uniform image-level anomaly scores for anomalous images with different defect sizes do not favor small defects

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

FUDAN UNIVERSITY

Prototypical Residual Networks for Anomaly Detection and Localization

Thanks !

Poster: WED-PM-374