Explicit Boundary Guided Semi-Push-Pull Contrastive Learning for Supervised Anomaly Detection

CVPR 2023

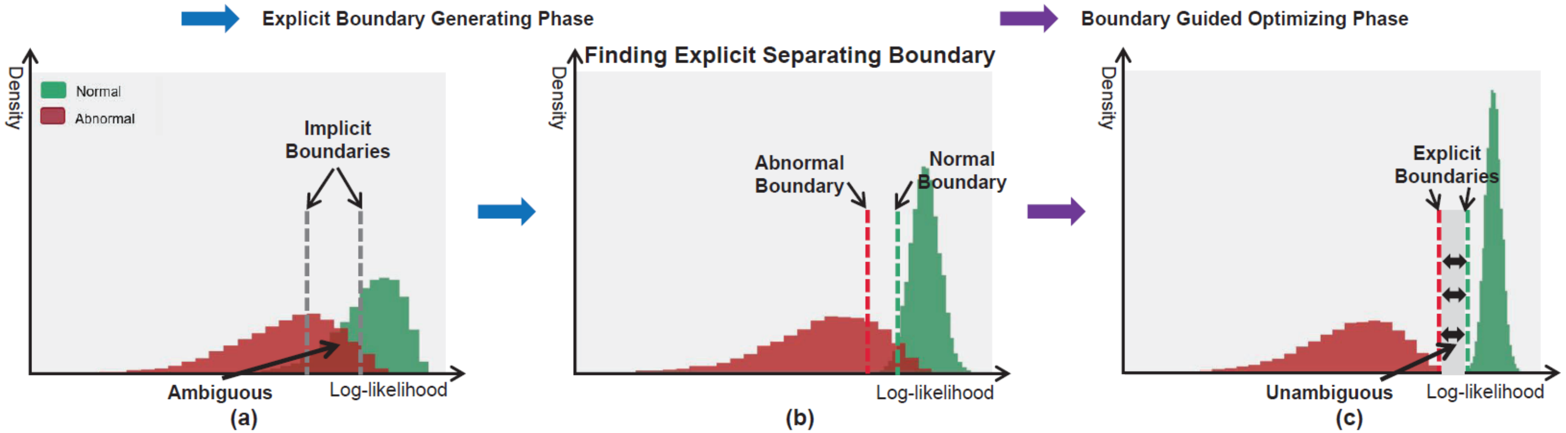
Xincheng Yao
Shanghai Jiao Tong University

THU-PM-372

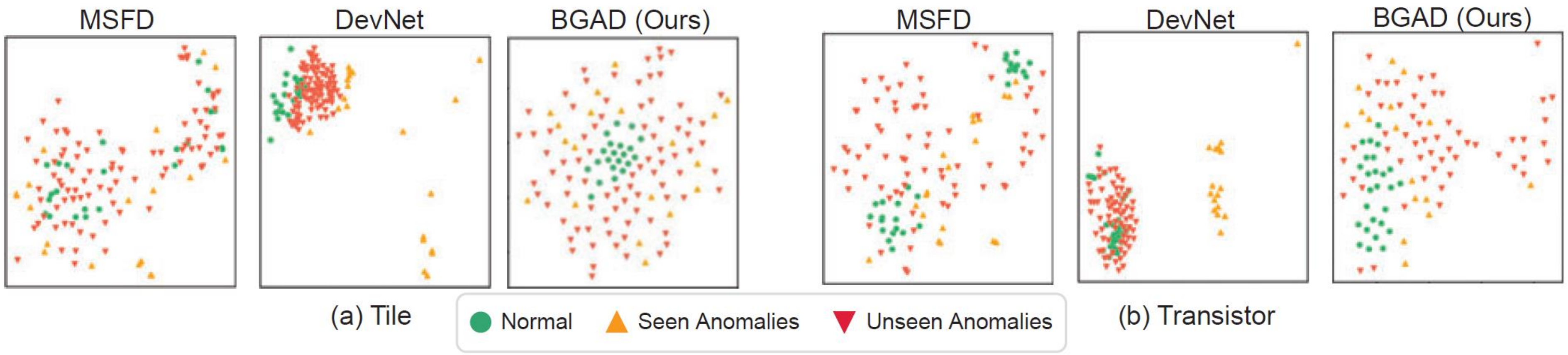
Coauthors:
Ruoqi Li, Jing Zhang, Jun Sun, Chongyang Zhang*

Preview

Supervised Anomaly Detection



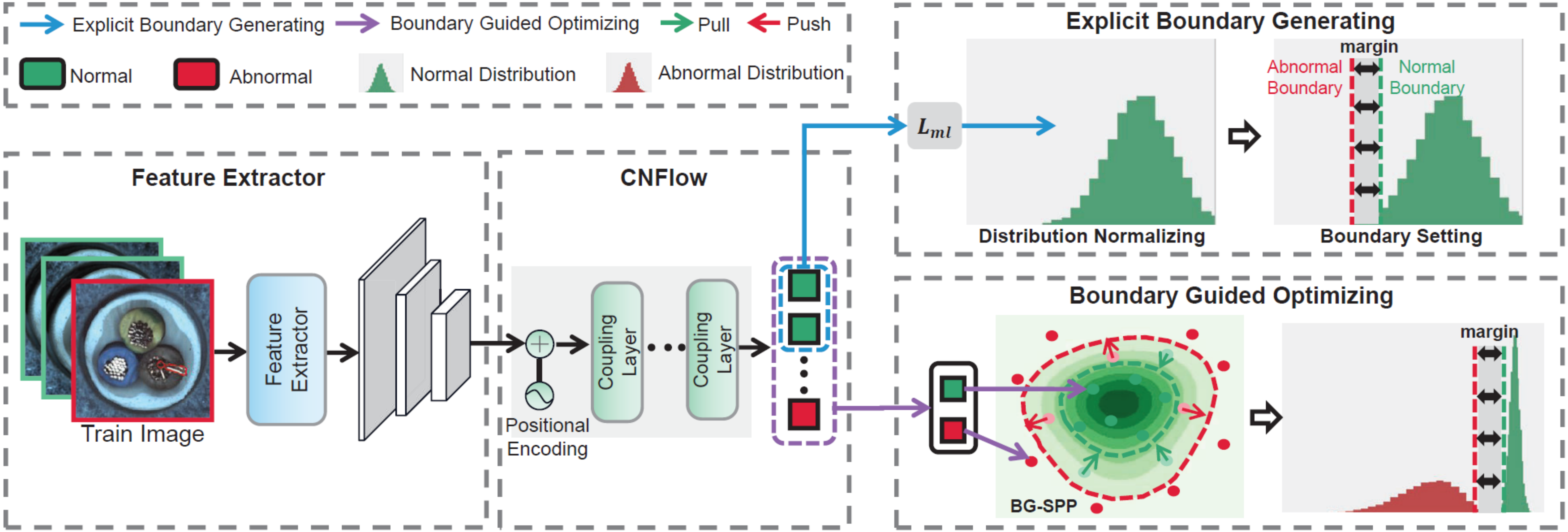
The **insufficient discriminability** issue is common in unsupervised anomaly detection.



Model needs to **generalize well to unseen anomalies!**

Preview

BGAD: Boundary Guided Anomaly Detection



Four parts: Feature Extractor, Conditional Normalizing Flow, Explicit Boundary Generating, Boundary Guided Optimizing.

Outline

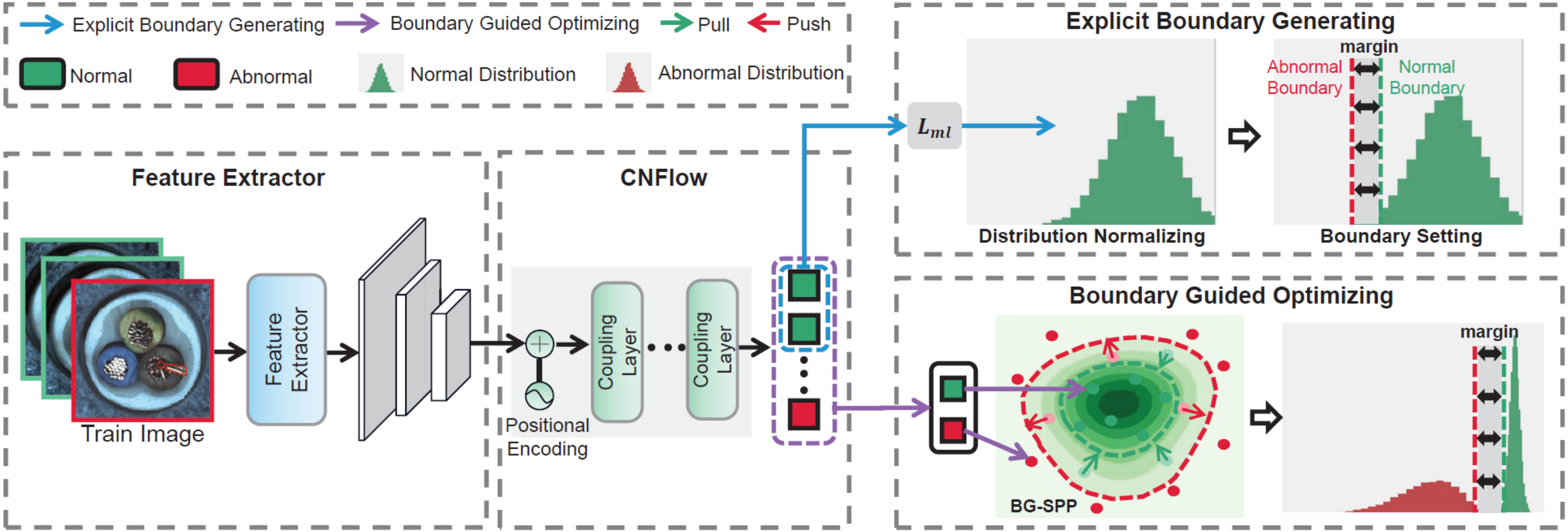
Supervised Anomaly Detection

A few known anomalies can be effectively exploited to train discriminative AD models with the objective to improve detection performance on the known anomalies and generalize well to unseen anomalies.

- Compared with unsupervised AD, supervised AD is more meaningful for real-world AD applications, because the detected anomalies can be used to further improve the discriminability and generalizability of the model.

Our Approach: BGAD

- **Boundary Guided Anomaly Detection, Model Overview:**



Four parts: Feature Extractor, Conditional Normalizing Flow, Explicit Boundary Generating, Boundary Guided Optimizing.

Our Approach: BGAD

- **Learning Normal Feature Distribution:**

- In order to find one anomaly-independent separating boundary, one simplified distribution of normal features should be learned firstly.
- Normalizing flow is employed to learn normal feature distribution.
- The log-likelihood of input feature x can be estimated by:

$$\log p_\theta(x) = \log p_{\mathcal{Z}}(\varphi_\theta(x)) + \sum_{l=1}^{L} \log\left|\det J_{\varphi_l}(y_{l-1})\right|$$

- The set of parameters is obtained by maximizing log-likelihoods:

$$\theta^* = \underset{\theta\in\Theta}{\operatorname{argmin}}\, \mathbb{E}_{x\sim p_{\mathcal{X}}}[-\log p_\theta(x)]$$

- The latent variable distribution can generally be assumed to obey $\mathcal{N}(0,\mathbf{I})$, the loss function is defined as:

$$\mathcal{L}_{ml} = \mathbb{E}_{x\in\mathcal{X}^n}\Big[\frac{d}{2}\log(2\pi) + \frac{1}{2}\varphi_\theta(x)^T\varphi_\theta(x) - \sum_{l=1}^{L} \log\left|\det J_{\varphi_l}(y_{l-1})\right|\Big]$$

Our Approach: BGAD

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

SHANGHAI JIAO TONG UNIVERSITY

- **Explicit Boundary Generating:**

- With the learned normal feature distribution, we can further find one explicit and compact separating boundary.
- Due to the high dimensional characteristics of the features, we therefore consider finding the boundary from the log-likelihood distribution.
- **Estimating log-likelihoods**:

$$\log p(x) = -\frac{d}{2}\log(2\pi) - \frac{1}{2}\varphi_\theta(x)^T\varphi_\theta(x) + \sum_{l=1}^{L}\log\left|\det J_{\varphi_l}(y_{l-1})\right|$$

- **Building normal log-likelihood distribution.** The $\mathcal{P}_n = \{\log p_i\}_{i=1}^{N}$ can be used to approximate the log-likelihood distribution of all normal features.
- **Finding explicit separating boundaries.** We select the β-th percentile of sorted normal log-likelihood distribution as the normal boundary b_n, and define an abnormal boundary $b_a = b_n - \tau$.

Our Approach: BGAD

- **Boundary Guided Optimizing:**

- With the explicit separating boundaries, we propose a **boundary guided semi-push-pull (BG-SPP) loss** for more discriminative feature learning.
- Our BG-SPP loss only pull together normal features whose log-likelihoods are smaller than b_n (semi-pull), while pushing abnormal features whose log-likelihoods are larger than b_a (semi-push):

$$\mathcal{L}_{bg-spp} = \sum_{i=1}^{N} |\min((\log p_i - b_n), 0)| + \sum_{j=1}^{M} |\max((\log p_j - b_n + \tau), 0)|$$

- The ℓ_1 norm based formulation can encourage the sparse log-likelihood distribution in the margin (b_a, b_n).
- We first only train with $\mathcal{L}_{ml}$; In the second training phase, the overall loss is $\mathcal{L} = \mathcal{L}_{ml} + \lambda \mathcal{L}_{bg-spp}$.

Our Approach: BGAD

- **Generalization Capability to Unseen Anomalies:**

- The obtained explicit separating boundary only relies on normal feature distribution, this means that the final decision boundary mainly depends on the normal distribution rather than being affected greatly by the anomalies.

- Our method still employs the normal distribution to determine anomalies, and can form a more compact normal feature distribution.

- The semi-push-pull mechanism does not enforce the anomalies to deviate from the normal distribution as far as possible, but only pushes the anomalies outside the margin region.

Experiments

- **Datasets:**
  - MVTecAD: 5534 high-resolution images, 15 categories, 73 anomaly types, and 1900 abnormal regions.
  - BTAD: This dataset contains 2830 real-world images of 3 industrial products.
- **Metrics:**
  - Area under the curve of the receiver operating characteristic (AUROC), image-level and pixel-level.
- **Settings:**
  - Multi-Class Setting: the known anomalies are randomly drawn from existing anomaly classes in the test set.
  - One-Class Setting: the known anomalies are randomly sampled only from one anomaly class, the others are used as unseen anomalies.

Experiments

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

SHANGHAI JIAO TONG UNIVERSITY

- **Results under the Multi-Class Setting:**

	Category	Unsupervised AD Methods							Supervised AD Method
		DRAEM* [52]	PaDiM* [10]	MSFD* [47]	PatchCore* [32]	CFA* [20]	NFAD‡	BGAD$^{w/o}$ (Ours)	BGAD (Ours)
Textures	Carpet	0.954/0.947	0.983/0.946	0.990/0.958	0.985/0.959	0.989/0.943	0.994/0.983	0.994/0.982	**0.996**±0.0002/**0.989**±0.0004
	Grid	0.997/0.984	0.963/0.894	0.986/0.937	0.974/0.891	0.977/0.932	0.993/0.980	0.994/0.980	**0.995**±0.0002/**0.986**±0.0001
	Leather	0.992/0.981	0.984/0.966	0.978/0.924	0.992/0.974	0.991/0.958	0.997/0.994	0.997/0.994	**0.998**±0.0001/**0.994**±0.0003
	Tile	0.994/0.949	0.958/0.884	0.952/0.841	0.960/0.939	0.960/0.860	0.969/0.929	0.968/0.927	**0.994**±0.0077/**0.978**±0.0021
	Wood	0.962/0.935	0.963/0.891	0.953/0.925	0.968/0.857	0.948/0.882	0.969/0.957	0.970/0.957	**0.982**±0.0053/**0.970**±0.0007
Objects	Bottle	0.993/0.955	0.978/0.936	0.985/0.940	0.986/0.956	0.987/0.944	0.988/0.965	0.989/0.964	**0.994**±0.0009/**0.971**±0.0011
	Cable	0.961/0.910	0.979/0.973	0.972/0.922	0.986/**0.980**	**0.987**/0.931	0.975/0.944	0.980/0.968	0.986±0.0010/0.977±0.0030
	Capsule	0.869/0.901	0.980/0.924	0.979/0.878	0.990/0.946	0.989/0.943	0.989/0.952	0.992/0.959	**0.992**±0.0021/**0.964**±0.0033
	Hazelnut	0.997/0.985	0.980/0.951	0.982/0.968	0.988/0.924	0.986/0.953	0.984/0.976	0.985/0.976	**0.995**±0.0040/**0.982**±0.0028
	Metal nut	0.992/0.935	0.979/0.929	0.972/0.985	0.986/0.935	0.987/0.918	0.971/0.942	0.976/0.948	**0.996**±0.0003/**0.970**±0.0012
	Pill	0.979/0.959	0.978/0.957	0.971/0.929	0.983/0.947	0.986/0.965	0.976/0.978	0.980/0.980	**0.996**±0.0002/**0.988**±0.0005
	Screw	0.992/0.965	0.974/0.923	0.983/0.924	0.984/0.928	0.985/0.944	0.988/0.945	0.992/0.960	**0.993**±0.0003/**0.968**±0.0010
	Toothbrush	0.970/0.940	0.980/0.894	0.986/0.877	0.987/0.939	0.989/0.894	0.983/0.904	0.986/0.938	**0.995**±0.0003/**0.961**±0.0026
	Transistor	0.970/0.935	0.983/0.967	0.886/0.781	0.964/0.967	**0.985**/0.960	0.923/0.788	0.940/0.830	0.983±0.0005/**0.972**±0.0015
	Zipper	0.984/0.966	0.978/0.948	0.981/0.935	0.986/0.963	0.988/0.944	0.986/0.957	0.987/0.957	**0.993**±0.0003/**0.977**±0.0002
	Mean	0.969/0.947	0.976/0.932	0.970/0.915	0.981/0.940	0.982/0.931	0.979/0.946	0.982/0.955	**0.992**±0.0007/**0.976**±0.0006
	Image-level Mean	0.978	0.975	0.964	0.988	0.989	0.968	0.974	**0.993**±0.0012

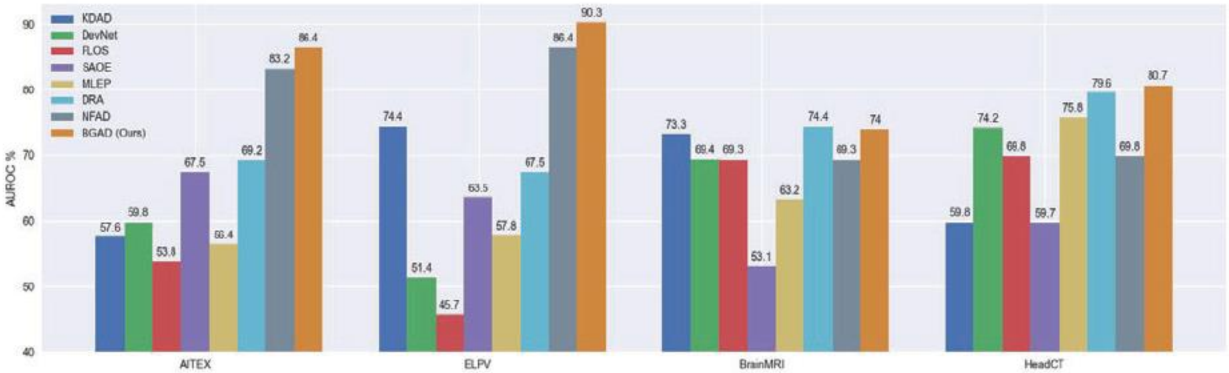
- Our BGAD reaches the best performance under all three evaluation metrics.
- We further surpass unsupervised baseline NFAD by 2.5% and 1.3% AUROC, 3.0% PRO.

Experiments

- **Results under the Multi-Class Setting:**

Category	Supervised AD Methods (Ten Abnormal Samples)			
	FCDD* [24]	DevNet* [27]	DRA* [12]	BGAD (Ours)
Carpet	0.981/0.952	-/-	-/-	**0.996**±0.0002/**0.989**±0.0004
Grid	0.949/0.897	-/-	-/-	**0.995**±0.0002/**0.986**±0.0001
Leather	0.984/0.973	-/-	-/-	**0.998**±0.0001/**0.994**±0.0003
Tile	0.977/0.938	-/-	-/-	**0.994**±0.0077/**0.978**±0.0021
Wood	0.950/0.901	-/-	-/-	**0.982**±0.0053/**0.970**±0.0007
Bottle	0.966/0.939	-/-	-/-	**0.994**±0.0009/**0.971**±0.0011
Cable	0.963/0.980	-/-	-/-	0.986±0.0010/0.977±0.0030
Capsule	0.970/0.922	-/-	-/-	**0.992**±0.0021/**0.964**±0.0033
Hazelnut	0.970/0.958	-/-	-/-	**0.995**±0.0040/**0.982**±0.0028
Metal nut	0.966/0.934	-/-	-/-	**0.996**±0.0003/**0.970**±0.0012
Pill	0.975/0.960	-/-	-/-	**0.996**±0.0002/**0.988**±0.0005
Screw	0.963/0.925	-/-	-/-	**0.993**±0.0003/**0.968**±0.0010
Toothbrush	0.967/0.907	-/-	-/-	**0.995**±0.0003/**0.961**±0.0026
Transistor	0.942/0.935	-/-	-/-	0.983±0.0005/**0.972**±0.0015
Zipper	0.968/0.948	-/-	-/-	**0.993**±0.0003/**0.977**±0.0002
Mean	0.966/0.938	-/-	-/-	**0.992**±0.0007/**0.976**±0.0006
Image-level Mean	0.965	0.948	0.961	**0.993**±0.0012

Categories	AE-MSE	AE-SSIM	VT-ADL	NFAD	BGAD (Ours)
1	0.490	0.530	**0.990**	0.972/0.767	0.982±0.0027/**0.830**±0.0318
2	0.920	0.960	0.940	0.967/0.578	**0.979**±0.0018/**0.648**±0.0173
3	0.950	0.890	0.770	0.996/0.988	**0.998**±0.0003/**0.993**±0.0005
Mean	0.780	0.790	0.900	0.978/0.778	**0.986**±0.0015/**0.824**±0.0163



- Compared with supervised AD methods, our method can also surpass these SOTA methods.
- The results on BTAD and other datasets also show the superiority of our method.

Experiments

- **Results under the One-Class Setting:**

	Known Class	Baseline	Ten Training Anomaly Samples					
		NFAD	DevNet	FLOS	SAOE	MLEP	DRA	BGAD (Ours)
Carpet	Color	0.998/**0.993**	0.767/-	0.760/-	0.467/-	0.689/-	0.886/-	**1.000/0.993**
	Cut	**0.998**/0.995	0.819/-	0.688/-	0.793/-	0.653/-	0.922/-	**0.998/0.996**
	Hole	0.997/0.993	0.814/-	0.733/-	0.831/-	0.674/-	0.922/-	**0.998/0.995**
	Metal	0.998/0.993	0.863/-	0.678/-	0.883/-	0.764/-	0.933/-	**1.000/0.994**
	Thread	**1.000**/0.995	0.972/-	0.946/-	0.834/-	0.967/-	0.989/-	**1.000/0.996**
	Mean	0.998/0.994	0.847/-	0.761/-	0.762/-	0.751/-	0.935/-	**0.999/0.995**
Metal_nut	Bent	0.977/0.959	0.904/-	0.827/-	0.901/-	0.956/-	0.990/-	**1.000/0.972**
	Color	0.977/0.963	0.978/-	0.9788/-	0.879/-	0.945/-	0.967/-	**0.999/0.973**
	Flip	0.976/0.977	0.987/-	0.942/-	0.795/-	0.805/-	0.913/-	**0.995/0.982**
	Scratch	**1.000**/0.965	0.991/-	0.943/-	0.845/-	0.805/-	0.911/-	**1.000/0.972**
	Mean	0.983/0.966	0.965/-	0.922/-	0.855/-	0.878/-	0.945/-	**0.998/0.975**

- The results show substantially better generalizability of our model in detecting unseen anomaly classes than the other supervised AD methods.
- Our model can outperform the baseline NFAD across all the datasets, which validates that better generalizability to unseen anomalies of our model.

Ablations

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

SHANGHAI JIAO TONG UNIVERSITY

- **Ablation study results:**

Metric \ Dataset	MVTecAD		Hard Subsets		Unseen Subsets	
	NFAD	BGAD	NFAD	BGAD	NFAD	BGAD
Image AUROC	0.968	0.992(+2.5%)	0.948	0.984(**+3.6%**)	0.948	0.971(**+2.3%**)
Pixel AUROC	0.979	0.992(+1.3%)	0.960	0.986(**+2.6%**)	0.960	0.982(**+2.2%**)
PRO	0.946	0.976(+3.0%)	0.863	0.949(**+8.6%**)	0.863	0.930(**+6.7%**)

- **Experiments on Hard Subsets.** This study demonstrates that our model is more beneficial for harder anomaly classes.
- **Generalization to Hard Subsets.** Even only trained with easy anomalies, our BGAD can generalize well to hard anomalies.

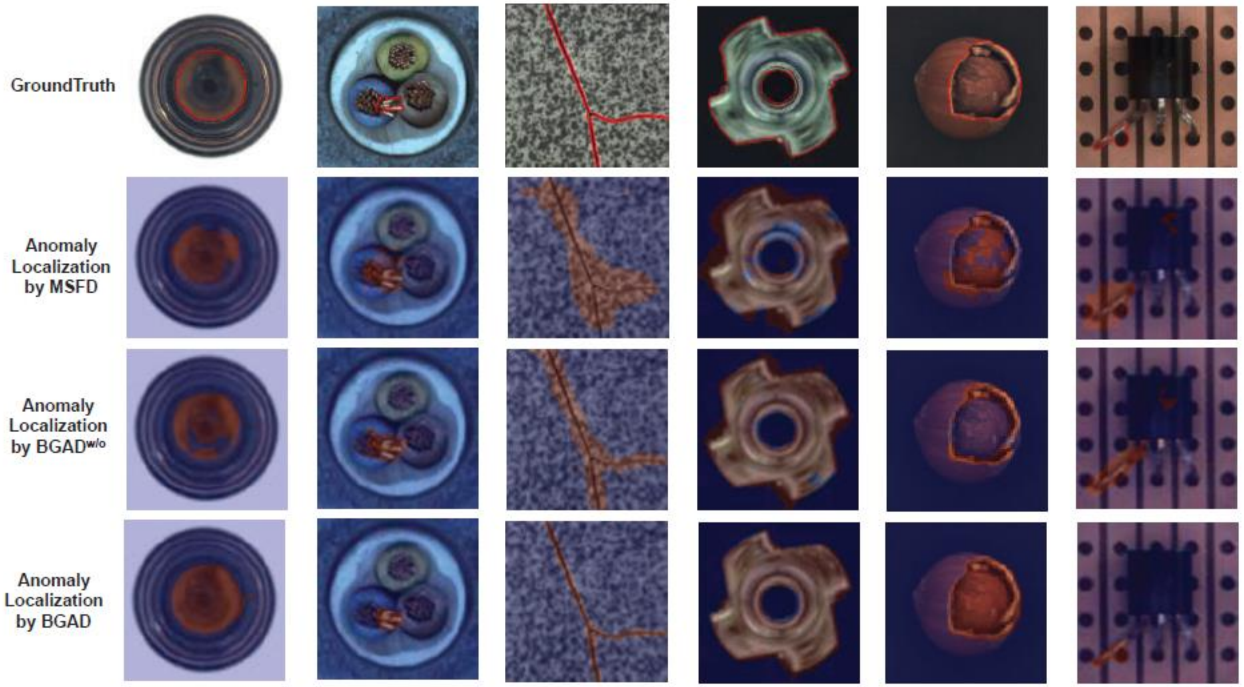
Method	Category				
	Carpet	Metal_nut	Capsule	Screw	Transistor
NFAD (baseline)	0.998/0.994	0.983/0.966	0.941/0.990	0.885/0.989	0.984/0.929
BGAD†	0.998/0.994	0.997/0.927	0.868/0.963	0.823/0.980	0.933/0.847
BGAD (Ours)	**0.999/0.995**	**0.998/0.975**	**0.988/0.991**	**0.947/0.991**	**0.994/0.942**

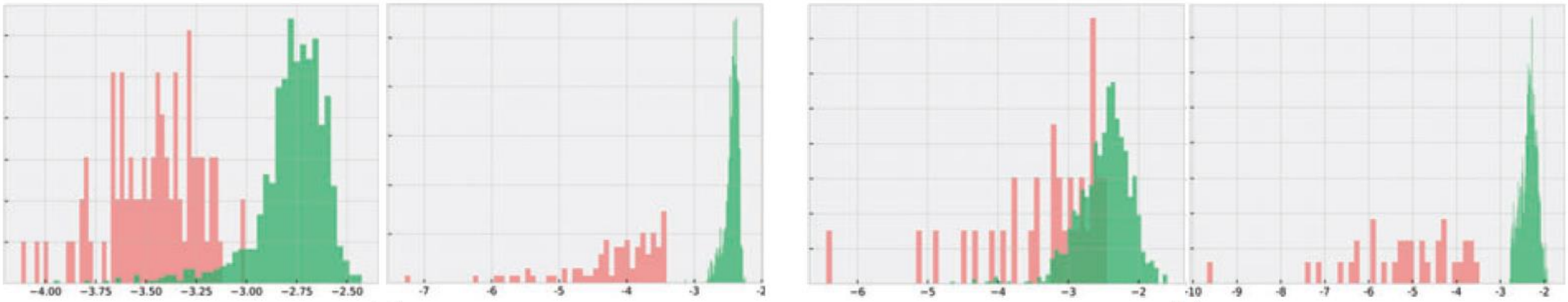
- **Effect of Semi-Push-Pull Mechanism.** The BGAD can outperform baseline NFAD and the variant BGAD† (employing the conventional contrastive loss).

Qualitative Results

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

SHANGHAI JIAO TONG UNIVERSITY





(a) (b)

Conclusions and Limitations

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

SHANGHAI JIAO TONG UNIVERSITY

- **Conclusions:**
  - We propose a novel AD model to tackle the **insufficient discriminability** issue and the **bias** issue simultaneously.
  - Compared with unsupervised AD models, our model can learn more discriminative features by exploiting a few anomalies effectively.
  - Compared with supervised AD methods, our method can mitigate the bias issue with the explicit separating boundary and the semi-push-pull mechanism.
- **Limitations:**
  - We employ normalizing flow to obtain the explicit separating boundary. However not all anomaly detection models can generate log-likelihoods.
  - Further improving our model’ s generalizability and theoretical analysis of model’ s generalizability are valuable future works.

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

SHANGHAI JIAO TONG UNIVERSITY

Thanks!

Contact Us:
sunny_zhang@sjtu.edu.cn