NeurIPS 2022

A Unified Model for Multi-class Anomaly Detection

Zhiyuan You[1] (***Presenter***) · Lei Cui[2] · Yujun Shen[3] · Kai Yang[4] · Xin Lu[4] · Yu Zheng[1] · Xinyi Le[1]

[1]Shanghai Jiao Tong University · [2]Tsinghua University · [3]CUHK · [4]SenseTime

Existing Separate Setting *V.S.* Our Unified Setting for Anomaly Detection

- **Existing Separate Setting**

  Train ***separate*** models for different classes of objects. (Fig. 1)

- **Drawbacks of the Separate Setting**

  1. ***Memory-consuming*** with a large number of classes.
  2. ***Uncongenial*** to the scenarios where the normal samples have some intra-class diversity.

- **Our Unified Setting**

  Train a ***unified*** model for all classes of objects. (Fig. 2)

- **Advantages of the Unified Setting**

  1. ***Memory-saving*** with a unified model for various classes.
  2. ***More practical*** since the industrial normal samples usually cover a range of categories.
  3. ***Easy*** to prepare the training data w/o the class labels.

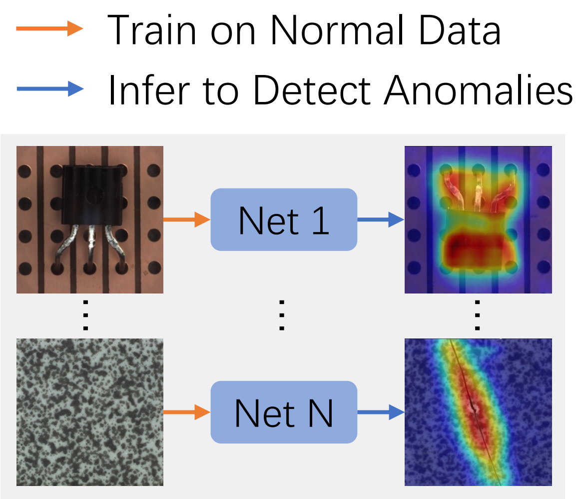

Fig. 1. Separate setting.

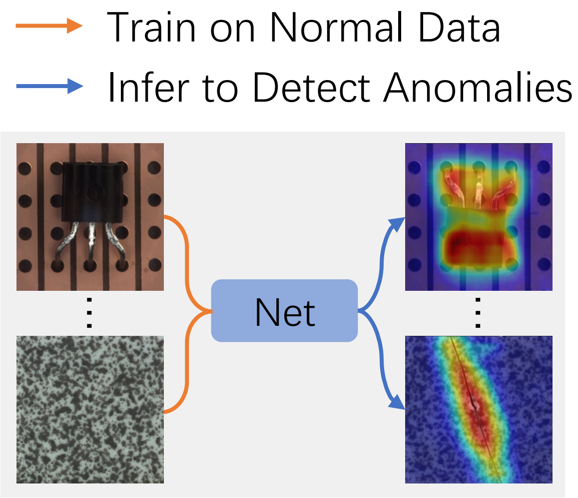

Fig. 2. Unified setting.

“Identical Shortcut” Problem in Reconstruction-based Methods

- **Analysis Method**

  Based on the feature reconstruction paradigm, we test 3 reconstruction nets (***MLP, CNN, & Transformer***).

- **Analysis Results**

  1. ***Observation***. During training, the loss becomes quite small (blue in Fig. 3a), but the performance (red for localization & green for detection in Fig. 3a) drops dramatically after reaching the peak.
  2. ***Reason***. The 3 models all suffer from the “***identical shortcut***” problem (visualized in Fig. 3b), which reconstructs both normal samples and anomalies well.

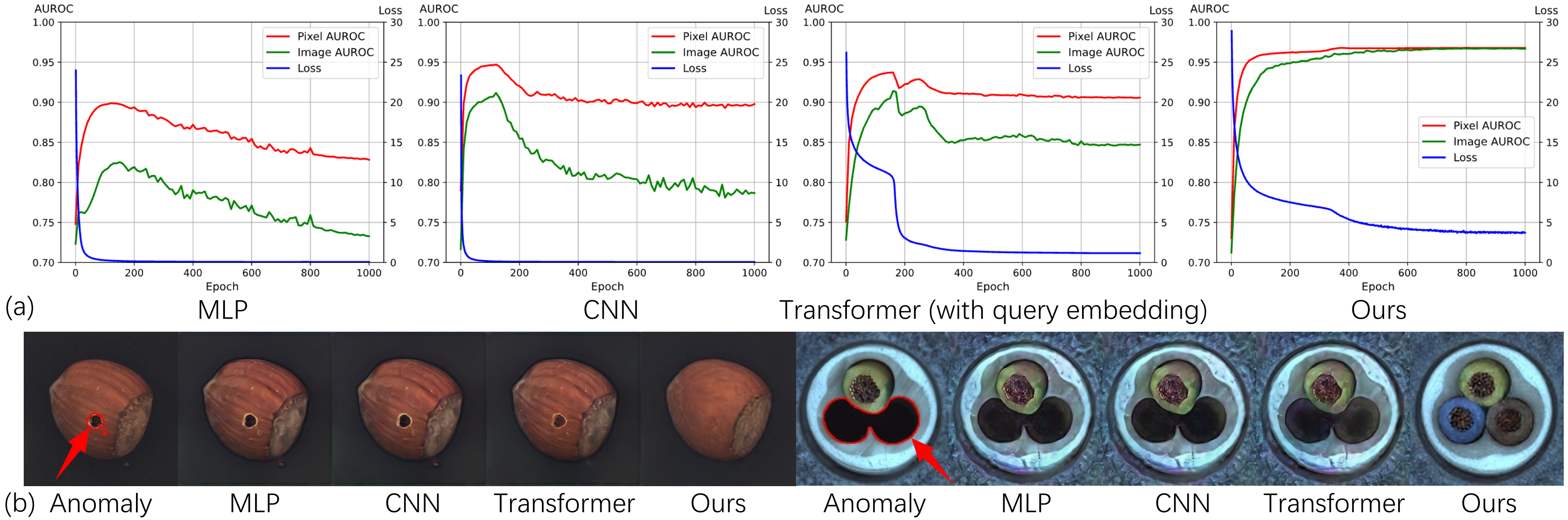

Fig. 3. Illustration of the “identical shortcut” problem.

Innovations of Proposed UniAD

1. *Layer-wise Query Embedding*. We add query embedding in ***every*** decoder layer to increase the ability to prevent the shortcut (Fig. 4).

2. *Neighbor Masked Attention*. We mask some neighbor tokens in the attention layer to prevent the information leak (Fig. 4).

3. *Feature Jittering*. We add noise to input features, leading the model to learn normal distribution by removing noise (Fig. 5).

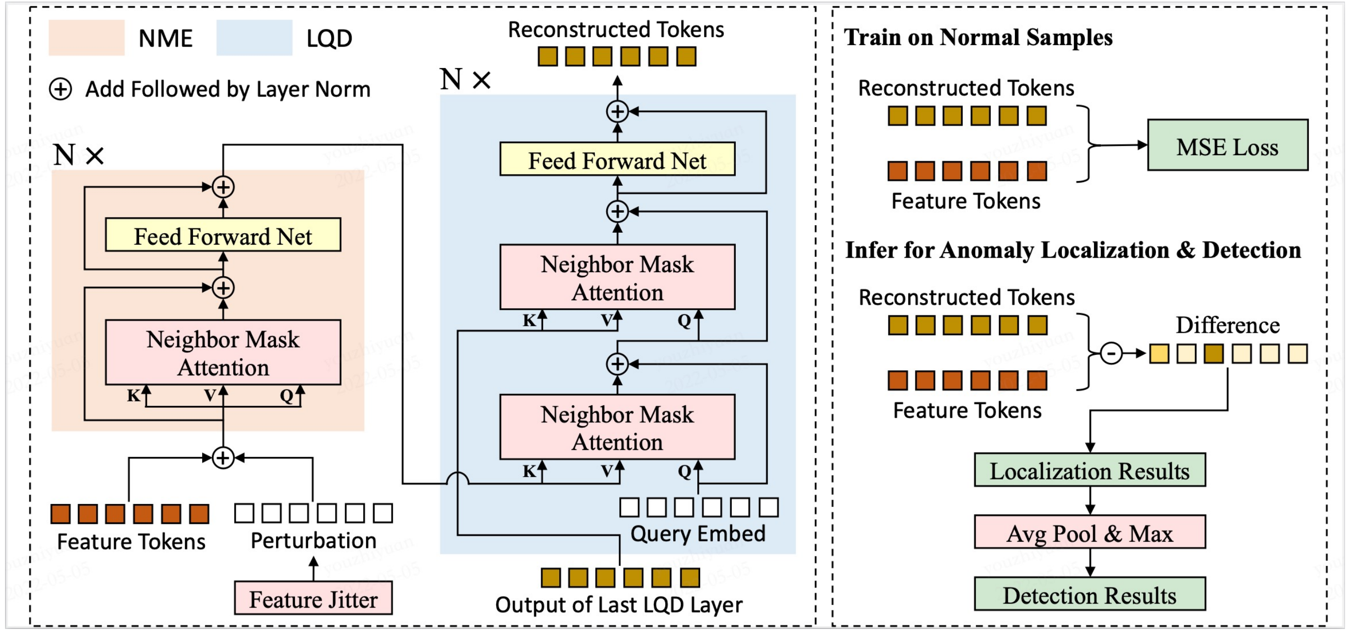

Fig. 4. Illustration of our model.

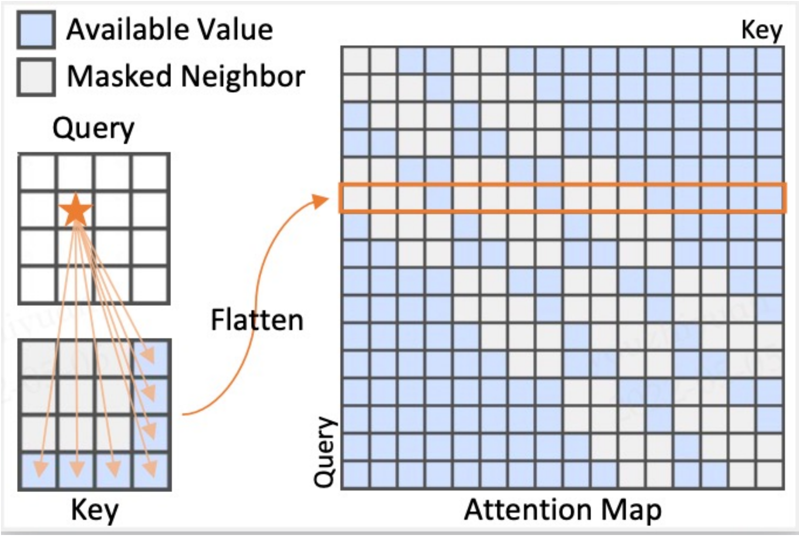

Fig. 5. Neighbor masked attention.

Results of Anomaly Detection and Localization on MVTec-AD

In the unified setting of MVTec-AD, we significantly outperform the best baseline by ***8.4% & 7.3%*** with anomaly detection and localization tasks, respectively.

Table 1: **Anomaly detection results with AUROC metric on MVTec-AD** [3]. All methods are evaluated under unified case / separate case. Our method is run with 5 random seeds.

	Category	US [5]	PSVDD [46]	PaDiM [8]	CutPaste [23]	MKD [36]	DRAEM [50]	Ours
Object	Bottle	84.0 / 99.0	85.5 / 98.6	97.9 / 99.9	67.9 / 98.2	98.7 / 99.4	97.5 / 99.2	**99.7** ± 0.04 / 100
	Cable	60.0 / 86.2	64.4 / 90.3	70.9 / 92.7	69.2 / 81.2	78.2 / 89.2	57.8 / 91.8	**95.2** ± 0.84 / 97.6
	Capsule	57.6 / 86.1	61.3 / 76.7	73.4 / 91.3	63.0 / 98.2	68.3 / 80.5	65.3 / 98.5	**86.9** ± 0.73 / 85.3
	Hazelnut	95.8 / 93.1	83.9 / 92.0	85.5 / 92.0	80.9 / 98.3	97.1 / 98.4	93.7 / 100	**99.8** ± 0.10 / 99.9
	Metal Nut	62.7 / 82.0	80.9 / 94.0	88.0 / 98.7	60.0 / 99.9	64.9 / 73.6	72.8 / 98.7	**99.2** ± 0.09 / 99.0
	Pill	56.1 / 87.9	89.4 / 86.1	68.8 / 93.3	71.4 / 94.9	79.7 / 82.7	82.2 / 98.9	**93.7** ± 0.65 / 88.3
	Screw	66.9 / 54.9	80.9 / 81.3	56.9 / 85.8	85.2 / 88.7	75.6 / 83.3	**92.0** / 93.9	87.5 ± 0.57 / 91.9
	Toothbrush	57.8 / 95.3	**99.4** / 100	95.3 / 96.1	63.9 / 99.4	75.3 / 92.2	90.6 / 100	94.2 ± 0.20 / 95.0
	Transistor	61.0 / 81.8	77.5 / 91.5	86.6 / 97.4	57.9 / 96.1	73.4 / 85.6	74.8 / 93.1	**99.8** ± 0.09 / 100
	Zipper	78.6 / 91.9	77.8 / 97.9	79.7 / 90.3	93.5 / 99.9	87.4 / 93.2	**98.8** / 100	95.8 ± 0.51 / 96.7
	Mean	68.1 / 85.8	80.1 / 90.8	80.3 / 93.8	71.3 / 95.5	79.8 / 87.8	82.6 / 97.4	**95.2** ± 0.11 / 95.4
Texture	Carpet	86.6 / 91.6	63.3 / 92.9	93.8 / 99.8	93.6 / 93.9	69.8 / 79.3	98.0 / 97.0	**99.8** ± 0.02 / 99.9
	Grid	69.2 / 81.0	66.0 / 94.6	73.9 / 96.7	93.2 / 100	83.8 / 78.0	**99.3** / 99.9	98.2 ± 0.26 / 98.5
	Leather	97.2 / 88.2	60.8 / 90.9	99.9 / 100	93.4 / 100	93.6 / 95.1	98.7 / 100	**100** ± 0.00 / 100
	Tile	93.7 / 99.1	88.3 / 97.8	93.3 / 98.1	88.6 / 94.6	89.5 / 91.6	**99.8** / 99.6	99.3 ± 0.14 / 99.0
	Wood	90.6 / 97.7	72.1 / 96.5	98.4 / 99.2	80.4 / 99.1	93.4 / 94.3	**99.8** / 99.1	98.6 ± 0.08 / 97.9
	Mean	87.4 / 91.5	70.1 / 94.5	91.9 / 98.8	89.8 / 97.5	86.0 / 87.7	99.1 / 99.1	**99.2** ± 0.07 / 99.1
	Mean	74.5 / 87.7	76.8 / 92.1	84.2 / 95.5	77.5 / 96.1	81.9 / 87.8	88.1 / 98.0	**96.5** ± 0.08 / 96.6

Table 2: **Anomaly localization results with AUROC metric on MVTec-AD** [3]. All methods are evaluated under unified case / separate case. Our method is run with 5 random seeds.

	Category	US [5]	PSVDD [46]	PaDiM [8]	FCDD [26]	MKD [36]	DRAEM [50]	Ours
Object	Bottle	67.9 / 97.8	86.7 / 98.1	96.1 / 98.2	56.0 / 97	91.8 / 96.3	87.6 / 99.1	**98.1** ± 0.04 / 98.1
	Cable	78.3 / 91.9	62.2 / 96.8	81.0 / 96.7	64.1 / 90	89.3 / 82.4	71.3 / 94.7	**97.3** ± 0.10 / 96.8
	Capsule	85.5 / 96.8	83.1 / 95.8	96.9 / 98.6	67.6 / 93	88.3 / 95.9	50.5 / 94.3	**98.5** ± 0.01 / 97.9
	Hazelnut	93.7 / 98.2	97.4 / 97.5	96.3 / 98.1	79.3 / 95	91.2 / 94.6	96.9 / 99.7	**98.1** ± 0.10 / 98.8
	Metal Nut	76.6 / 97.2	**96.0** / 98.0	84.8 / 97.3	57.5 / 94	64.2 / 86.4	62.2 / 99.5	94.8 ± 0.09 / 95.7
	Pill	80.3 / 96.5	**96.5** / 95.1	87.7 / 95.7	65.9 / 81	69.7 / 89.6	94.4 / 97.6	95.0 ± 0.16 / 95.1
	Screw	90.8 / 97.4	74.3 / 95.7	94.1 / 98.4	67.2 / 86	92.1 / 96.0	95.5 / 97.6	**98.3** ± 0.08 / 97.4
	Toothbrush	86.9 / 97.9	98.0 / 98.1	95.6 / 98.8	60.8 / 94	88.9 / 96.1	97.7 / 98.1	**98.4** ± 0.03 / 97.8
	Transistor	68.3 / 73.7	78.5 / 97.0	92.3 / 97.6	54.2 / 88	71.7 / 76.5	64.5 / 90.9	**97.9** ± 0.19 / 98.7
	Zipper	84.2 / 95.6	95.1 / 95.1	94.8 / 98.4	63.0 / 92	86.1 / 93.9	**98.3** / 98.8	96.8 ± 0.24 / 96.0
	Mean	81.2 / 94.3	86.8 / 96.7	92.0 / 97.8	63.6 / 91	83.3 / 90.8	81.9 / 97.0	**97.3** ± 0.02 / 97.2
Texture	Carpet	88.7 / 93.5	78.6 / 92.6	97.6 / 99.0	68.6 / 96	95.5 / 95.6	**98.6** / 95.5	98.5 ± 0.01 / 98.0
	Grid	64.5 / 89.9	70.8 / 96.2	71.0 / 97.1	65.8 / 91	82.3 / 91.8	**98.7** / 99.7	96.5 ± 0.04 / 94.6
	Leather	95.4 / 97.8	93.5 / 97.4	84.8 / 99.0	66.3 / 98	96.7 / 98.1	97.3 / 98.6	**98.8** ± 0.03 / 98.3
	Tile	82.7 / 92.5	92.1 / 91.4	80.5 / 94.1	59.3 / 91	85.3 / 82.8	**98.0** / 99.2	91.8 ± 0.10 / 91.8
	Wood	83.3 / 92.1	80.7 / 90.8	89.1 / 94.1	53.3 / 88	80.5 / 84.8	**96.0** / 96.4	93.2 ± 0.08 / 93.4
	Mean	82.9 / 93.2	83.1 / 93.7	84.6 / 96.7	62.7 / 93	88.0 / 90.6	**97.7** / 97.9	95.8 ± 0.04 / 95.3
	Mean	81.8 / 93.9	85.6 / 95.7	89.5 / 97.4	63.3 / 92	84.9 / 90.7	87.2 / 97.3	**96.8** ± 0.02 / 96.6