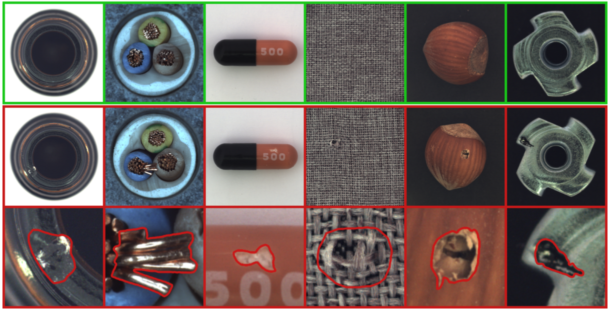
Evaluating State-of-the-art Methods for Industrial Anomaly Detection

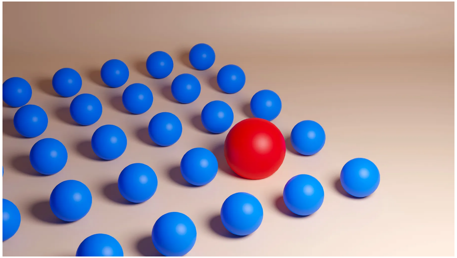
Final Presentation

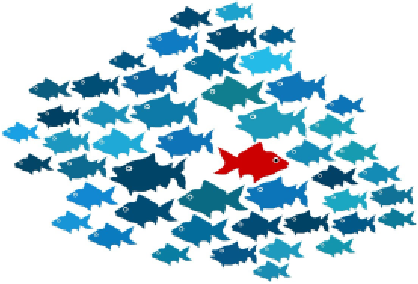
Table of Contents



What is an anomaly

- something different, abnormal compared to the majority of data points
- deviate from the common patterns
- unbounded, unknown events
- ...





Anomaly Detection (AD)

TECHNISCHE UNIVERSITÄT DARMSTADT

Anomaly detection, or outlier detection, is the identification of observations, events, or data points that deviate from what is usual, standard, or expected, making them inconsistent with the rest of a data set [1].

Real-world application domains

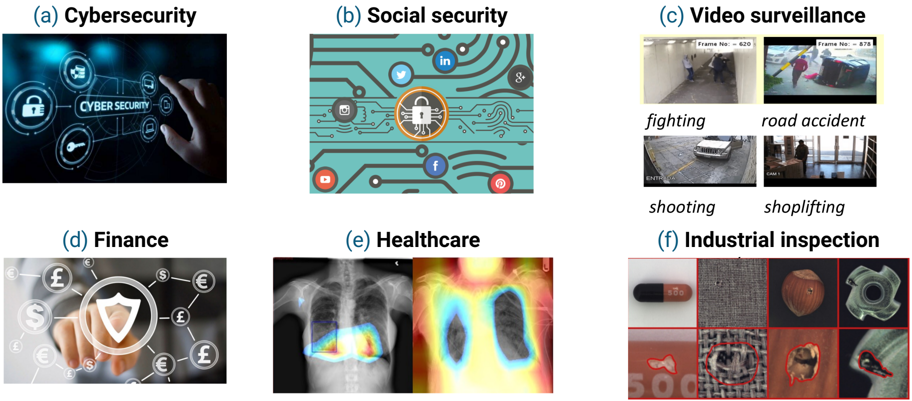


Figure: Real-world application domains[2]

Motivation

- Lack of a unified benchmark for industrial anomaly detection
- Existing frameworks are difficult to configure
- Cross-platform support is usually lacking
- Visualization is required for benchmarking and inference

- **Visualize Data**: Shows data distribution in feature space.
- **Statistical Basis**: Uses PDFs or similar methods.
- **Likelihood**: Indicates likelihood of data values.
- **Identify Anomalies**: Compare new points to known distribution.
- **Set Thresholds**: Determine anomalies based on thresholds.

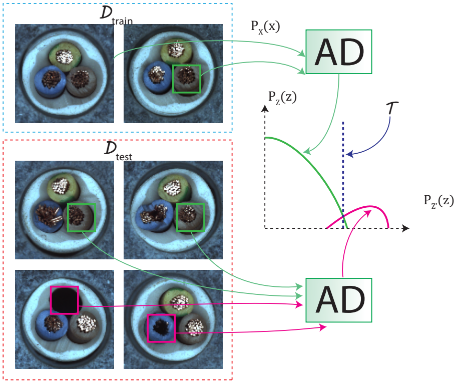


Figure: Distribution Map [3]

- **Assumption**: Normal data can be accurately reconstructed; anomalies cannot.
- **Training**: Use normal data to train a model (e.g., autoencoder).
- **Encoding**: Model learns to map input data to a low-dimensional representation.
- **Testing**: Model attempts to reconstruct new data points.
- **Anomaly Detection**: Large reconstruction error indicates a probable anomaly.

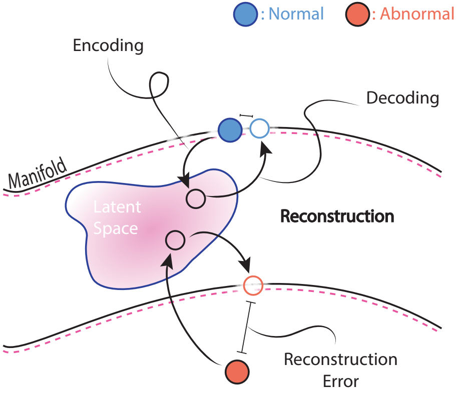


Figure: Reconstruction based [4]

SoTA Topics
Memory Bank

- **Definition**: A memory bank stores feature representations of normal instances, helping to detect anomalies by comparing new data with these stored features.
- **Process**: During training, it learns and stores normal data features. At inference, it compares new data with these stored features to identify anomalies.

SoTA Topics
Teacher Student

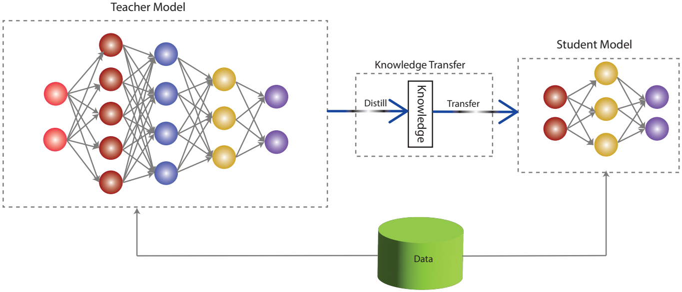


Figure: Teacher Student [5]

- **Definition**: Transferring knowledge from a large model to a smaller one.
- **Importance**: Addresses computational inefficiency of large models.
- **Process**: Small student model learns from larger teacher model.
- **Objective**: Student model achieves comparable or enhanced accuracy using teacher's knowledge.

SoTA Topics
One Class Classification

- One-class classification (OCC) models a single class, often the normal or inlier class.
- Anomaly detection (AD) aims to identify normal instances accurately.
- Anomalies are outliers that significantly deviate from normal characteristics.
- OCC methods detect anomalies by comparing distribution or behavior to the normal class.

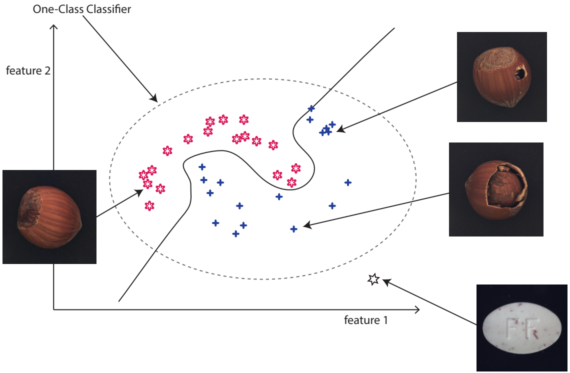


Figure: One Class Classification [6]

Dataset

Comparison of Datasets

TECHNISCHE UNIVERSITÄT DARMSTADT

Datasets	Sample Number	Categories
MVTec [7]	5354	15
MVTec3D [8]	5312	10
Btech [9]	2830	3
Kolektor [10]	399	1
VisA [11]	10821	12

Table: Comparison of Datasets

Models

Evaluated Models according to Topics

Topics	Models
Distribution Map	CFLOW [3], CSFlow [12], DFKDE [13], DFM [14], FastFlow [15], RKDE [16]
Reconstruction-based	DRAEM [17], DSR [18], GANomaly [19], FRE [20]
Memory Bank	PaDiM [21], PatchCore [22], CFA [23]
Teacher Student	Reverse Distillation [24], STFPM [25]
One Class Classification	UFLOW [26]

Table: Evaluated Models

Models

Evaluated Models according to Tasktype

Tasktype	Models
Segmentation	CFLOW [3], CSFlow [12], DFM [14], FastFlow [15], DRAEM [17], DSR [18], PaDiM [21], PatchCore [22], CFA [23], Reverse Distillation [24], STFPM [25], UFLOW [26], FRE [20]
Detection	RKDE [16]
Classification	DFKDE [13], GANomaly [19]

Table: Task Types

Metrics

TECHNISCHE UNIVERSITÄT DARMSTADT

Metrics	Formular	Usage
Precision (P)	$P = TP/(TP + FP)$	True Positive (TP), False Positive (FP)
Recall (R)	$R = TP/(TP + FN)$	False Negative (FN)
True Positive Rate (TPR)	$TPR = TP/(TP + FN)$	Classification
False Positive Rate (FPR)	$FPR = FP/(TN + FP)$	True Negative (TN)
F1 Score	$\frac{2 \times P \times R}{P + R}$	Indicates a balance between precision and recall
AUROC	$\int_0^1 (TPR) d(FPR)$	Classification
PRO	$\frac{1}{N} \sum_i \sum_k \frac{P_i \cap C_{i,k}}{C_{i,k}}$	Total ground-truth number (N), Predicted abnormal pixels (P), Defect ground-truth regions (C), Segmentation
AUPR	$\int_0^1 (P) d(R)$	Localization, Segmentation

Table: Used Metrics

Architecture
Anomalib Workflow

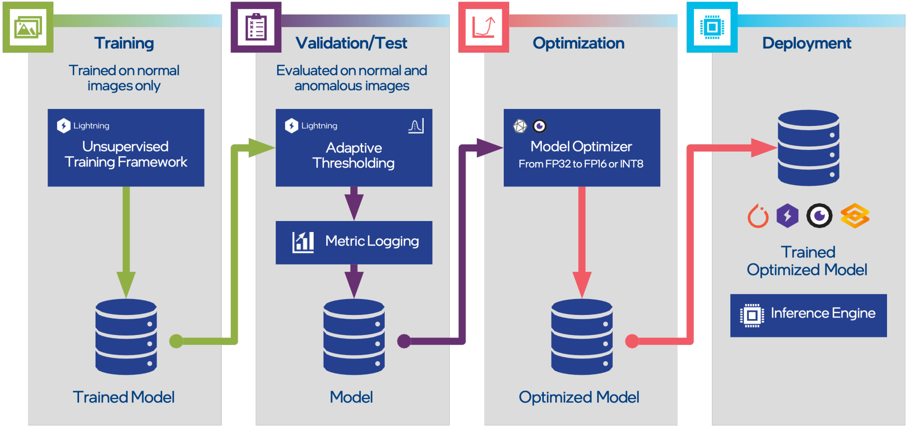


Figure: Workflow [27]

Architecture

Overview of Anomalib Architecture

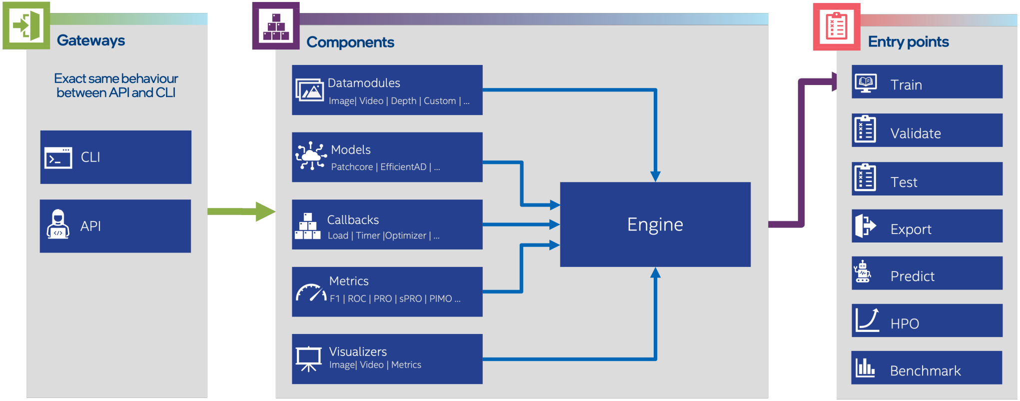


Figure: Overview of Architecture [27]

Architecture

IADBE Architecture

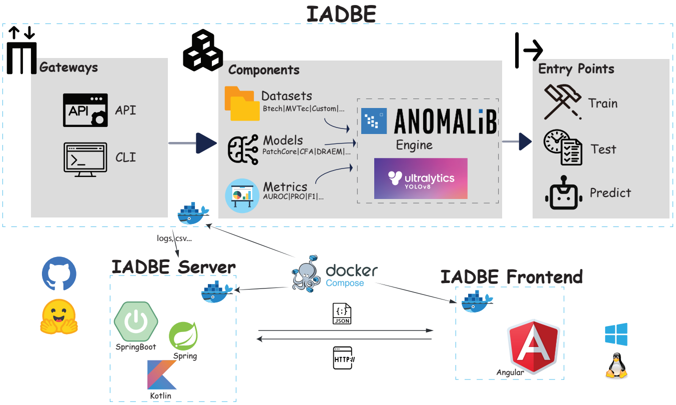


Figure: IADBE Architecture

Sample Code

Training and Testing using API

```python
datasets = ['screw', 'pill', 'capsule', 'carpet', 'grid', 'tile', 'wood', 'zipper',
            'cable', 'toothbrush', 'transistor', 'metal_nut', 'bottle', 'hazelnut',
            'leather']

for dataset in datasets:
    model = Padim()
    datamodule = MVTec(category=dataset, num_workers=0, train_batch_size=256,
                       eval_batch_size=256)
    engine = Engine(pixel_metrics=["AUROC", "PRO"], image_metrics=["AUROC", "PRO"],
    task=TaskType.SEGMENTATION)

    engine.fit(model=model, datamodule=datamodule)

    test_results = engine.test(
        model=model,
        datamodule=datamodule,
        ckpt_path=engine.trainer.checkpoint_callback.best_model_path,
    )
```

Sample Code

Training and Testing using bash

```bash
datasets=('screw' 'pill' 'capsule' 'carpet' 'grid' 'tile' 'wood' 'zipper' 'cable'
'toothbrush' 'transistor' 'metal_nut' 'bottle' 'hazelnut' 'leather')
config_file="./configs/models/padim.yaml"

for dataset in "${datasets[@]}"
do
    command = anomalib train --data anomalib.data.MVTec --data.category $dataset --
        config $config_file
    # Execute command
    $command
done
```

Environment Settings
Hardware Settings

TECHNISCHE UNIVERSITÄT DARMSTADT

- **CPU**: Ryzen Threadripper 3970X
- **GPU**: GeForce RTX 3090
- **OS**: Ubuntu 22.04

- **CPU**: Ryzen 5 3600
- **GPU**: GeForce RTX 4060
- **OS**: Windows 11

Training Monitoring

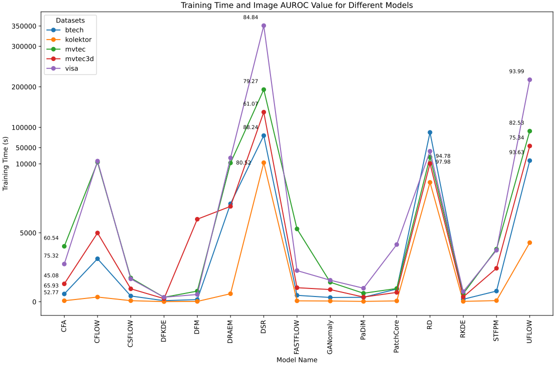


Figure: The figure illustrates the results of the training time of different models. The x-axis depicts the models, while the y-axis illustrates the training time in units of seconds. Partial AUROC values are displayed in proximity to the data points.

Training Monitoring

Model	BTech Time	BTech AUROC	Kolelktor Time	Kolelktor AUROC	MVTec Time	MVTec AUROC	MVTec3D Time	MVTec3D AUROC	Visa Time	Visa AUROC
CFA	568s	65.93	75s	52.77	4023s	60.54	1299s	45.08	2737s	75.32
CFLOW	3120s	94.06	342s	86.65	14165s	95.71	4991s	76.38	16927s	87.12
CSFLOW	419s	42.53	79s	58.70	1728s	72.74	940s	54.14	1660s	52.46
DFKDE	78s	90.47	10s	67.84	323s	75.72	229s	62.14	327s	72.09
DFM	158s	95.29	24s	92.05	773s	92.80	5987s	58.42	528s	87.21
DRAEM	7106s	85.84	582s	51.92	12291s	78.44	6919s	58.93	24690s	80.52
DSR	79740s	88.24	12954s	57.72	193188s	79.27	137580s	61.07	350887s	84.84
FASTFLOW	467s	92.03	67s	77.36	5282s	83.77	1020s	68.48	2261s	89.11
GANomaly	310s	50.05	55s	56.68	1417s	43.62	884s	46.76	1562s	55.25
PaDiM	322s	95.52	26s	80.25	617s	89.97	340s	73.41	982s	84.19
PatchCore	925s	93.59	62s	86.76	961s	98.91	679s	83.61	4149s	91.66
RD	87622s	95.28	8659s	88.24	26800s	97.98	11021s	95.06	41113s	94.78
RKDE	181s	83.81	27s	64.08	587s	75.43	347s	52.92	734s	67.95
STFPM	775s	92.23	85s	66.24	3815s	72.66	2427s	71.21	3736s	84.95
UFLOW	18066s	93.63	4288s	90.01	90776s	82.53	54160s	75.34	217558s	93.99

Table: This table presents a comparison of various machine learning models based on their performance metrics across different datasets. It includes columns for execution time and AUROC scores for five datasets.

Training Monitoring

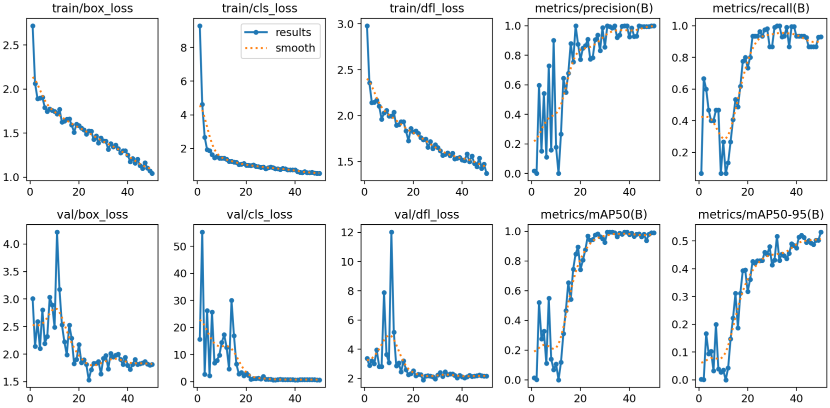


Figure: The figure illustrates the results of the training process in the custom dataset cardboard.

Visualization
MVTec Bottle

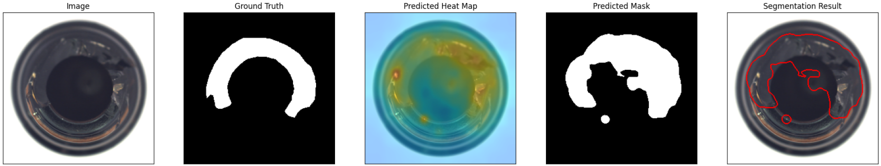


Figure: The figure illustrates the results of the inference process for the ”bottle” category in the MVTec dataset [7].

Visualization
VisA Chewing Gum

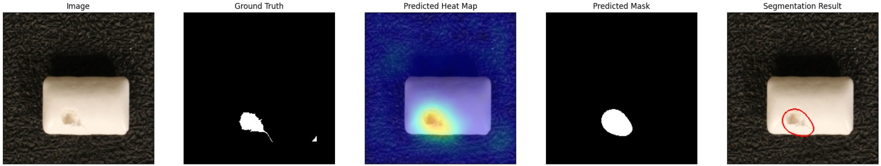


Figure: The figure illustrates the results of the inference process for the ”chewing gum” category in the VisA dataset [11].

Visualization

MVTec3D Cookies

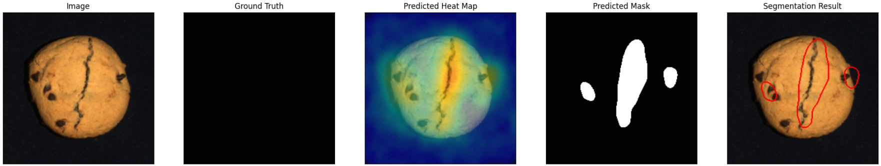


Figure: The figure illustrates the results of the inference process for the ”cookie” category in the MVTec3D dataset [8].

Visualization

Btech 03

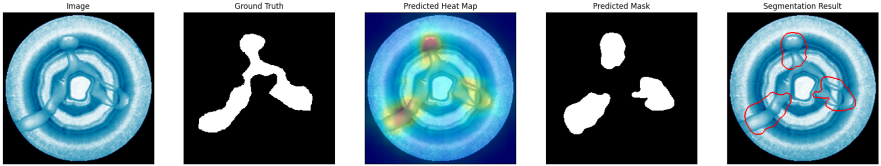


Figure: The figure illustrates the results of the inference process for the ”03” (Cover) category in the Btech dataset [9].

Visualization

Kolektor

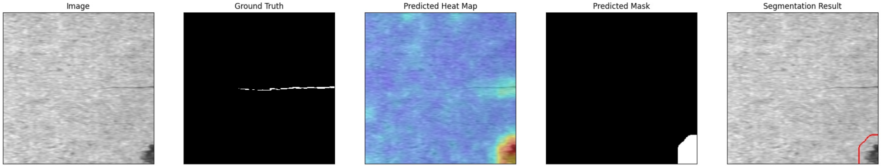


Figure: The figure illustrates the results of the inference process for the Kolektor dataset [10].

Visualization

Custom Dataset Cardboard (Anomalib Model)

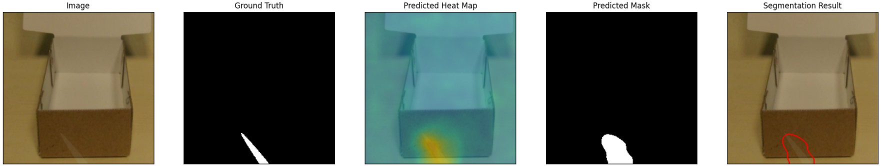


Figure: This figure shows the results of the inference process for the custom dataset Cardboard [28].

Visualization

Custom Dataset Cardboard (YOLO Model)

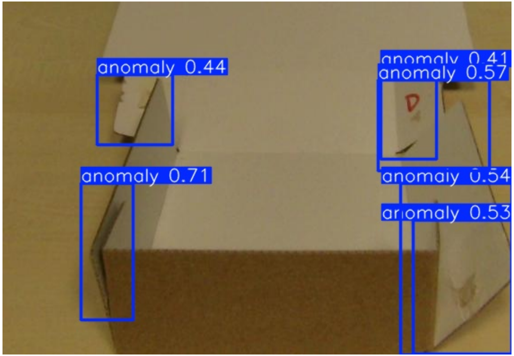


Figure: This figure shows the results of the inference process for the Cardboard custom dataset [28] with the model (YOLOv8m).

Benchmark

MVTec Image Level AUROC

	Screw	Pill	Capsule	Carpet	Grid	Tile	Wood	Zipper	Cable	Toothbrush	Transistor	Metal Nut	Bottle	Hazelnut	Leather	Average	Source
CFA	87.68	93.10	48.17	50.00	84.04	37.50	62.50	40.34	61.96	**98.06**	45.00	39.78	50.00	60.00	50.00	60.54	**99.30**
CFLOW	81.57	92.91	**95.45**	**97.27**	88.47	**100.00**	**99.39**	**97.64**	94.15	94.17	94.88	**100.00**	**100.00**	**99.68**	**100.00**	**95.71**	**96.31**
CSFLOW	41.58	38.09	60.77	**99.16**	78.36	90.44	**95.75**	88.93	67.31	44.72	50.54	74.46	89.92	71.02	**100.00**	72.74	93.50
DFKDE	69.17	64.18	72.04	68.42	46.69	92.50	81.32	88.26	68.67	78.89	81.12	76.20	92.94	77.57	77.79	75.72	77.40
DFM	77.79	**96.89**	92.98	90.93	65.41	**98.67**	**97.81**	**97.51**	94.27	**96.39**	94.79	91.72	**100.00**	**96.82**	**100.00**	92.80	94.30
DRAEM	30.01	74.55	77.02	72.47	80.45	86.80	**95.96**	78.90	63.98	70.42	90.21	93.55	**98.10**	77.50	86.65	78.44	**98.00**
DSR	56.67	70.16	72.72	43.82	**97.08**	80.70	90.75	78.97	76.96	**96.94**	91.04	81.04	86.59	81.64	83.97	79.27	**98.20**
FASTFLOW	65.85	76.60	69.64	93.38	**96.32**	93.36	**98.16**	72.69	67.62	72.50	89.62	81.33	**99.68**	79.86	**99.90**	83.77	**99.40**
FRE	58.11	83.82	76.59	94.18	64.75	**98.99**	**98.16**	93.91	85.53	**95.00**	90.92	84.12	**99.37**	92.29	**99.93**	87.71	**98.40**
GANomaly	32.79	59.96	26.57	21.71	57.23	54.42	60.88	41.05	52.47	49.17	33.46	26.30	47.78	53.86	36.68	43.62	42.10
PaDiM	78.95	79.95	86.48	**97.99**	87.47	94.55	**97.46**	77.46	85.96	82.50	94.54	**98.34**	**99.52**	88.32	**100.00**	89.97	**96.70**
PatchCore	**98.11**	94.76	**97.85**	**99.12**	**98.08**	**98.81**	**98.77**	**99.21**	**99.10**	**100.00**	**100.00**	**99.80**	**100.00**	**100.00**	**100.00**	**98.91**	**98.10**
RD	**98.03**	**97.63**	**97.93**	**99.36**	**95.49**	**100.00**	**99.39**	**97.16**	**95.45**	91.39	**97.87**	**100.00**	**100.00**	**100.00**	**100.00**	**97.98**	**98.50**
RKDE	50.58	68.77	51.93	-	75.36	67.72	62.54	75.37	85.83	77.17	65.00	90.63	85.07	**100.00**	**100.00**	75.43	-
STFPM	77.62	40.78	60.59	**98.88**	59.23	**97.08**	**98.95**	75.39	91.34	47.50	61.88	40.22	43.97	**96.50**	**100.00**	72.66	**97.00**
UFLOW	49.19	94.84	56.72	**100.00**	**99.33**	**99.39**	**95.09**	89.73	62.67	64.17	81.46	55.77	**99.21**	90.39	**100.00**	82.53	**98.74**

Table: Image Level AUROC Scores for MVTec [7] dataset categories - The table presents the area under the receiver operating characteristic curve (AUROC) scores for 16 different methods across 15 different categories from the MVTec dataset.

Benchmark

MVTec with Multi-class Unified

	Screw	Pill	Capsule	Carpet	Grid	Tile	Wood	Zipper	Cable	Toothbrush	Transistor	Metal Nut	Bottle	Hazelnut	Leather	Average	Chosen
CFA	56.26	93.24	29.84	56.14	50.00	67.21	73.86	50.00	48.37	48.33	50.00	44.70	50.00	92.96	62.23	58.21	Pill
CFLOW	92.45	19.59	12.90	**98.31**	11.99	**99.03**	35.57	30.25	37.44	91.67	87.00	94.62	1.23	6.86	55.84	51.65	Tile
CSFLOW	72.18	57.72	55.05	92.78	58.81	85.19	77.89	56.45	56.20	61.25	35.50	34.41	70.60	77.86	94.46	65.76	Carpet
DFKDE	50	50	50	50	50	92.50	49.17	50	50	50	50	50	50	50	50	52.78	Tile
DFM	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	**100.00**	50.00	50.00	53.33	Bottle
DRAEM	51.01	39.77	53.77	60.23	51.55	71.25	89.12	61.11	53.17	44.44	43.00	28.25	**97.54**	71.39	77.24	59.52	Bottle
DSR	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	91.67	50.00	50.00	50.00	50.00	50.00	52.78	Toothbrush
FASTFLOW	39.74	40.83	57.68	53.97	59.57	63.56	53.51	58.56	42.71	71.94	51.83	30.45	49.76	51.39	**100.00**	55.03	Leather
FRE	50.00	50.00	50.00	50.00	50.00	50.00	**98.77**	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	53.25	Wood
GANomaly	46.14	59.71	50	41.05	88.72	45.88	65.79	34.80	49.23	47.50	49.08	48.34	50	55.39	32.78	50.96	Wood
PaDiM	43.80	52.21	50.46	54.82	60.19	74.21	**96.84**	48.23	49.46	50.00	47.50	50.00	50.00	42.46	**99.93**	58.01	Leather
PatchCore	50.00	43.77	50.00	**96.91**	50.00	88.42	**97.28**	50.00	50.00	50.00	50.00	50.00	50.00	79.86	**100.00**	63.75	Leather
RKDE	43.19	41.13	49.54	56.34	51.13	51.88	51.23	48.44	48.91	51.11	50	50.24	50	85.00	27.82	50.40	Hazelnut
STFPM	50.00	50.00	50.00	**99.16**	76.44	85.28	93.42	50.00	50.00	50.00	50.00	50.00	50.00	50.00	**95.31**	63.31	Carpet
UFLOW	50.00	50.00	50.00	**100.00**	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	53.33	Carpet

Table: Image Level AUROC Scores for MVTec [7] dataset categories based on a single unified model - The table presents the area under the receiver operating characteristic curve (AUROC) scores for 15 different methods across 15 different categories from the MVTec dataset.

Benchmark

Custom dataset cardboard (Anomalib vs YOLO)

	CFA	CFLOW	CSFLOW	DFM	DRAEM	DSR	FASTFLOW	FRE	PADIM	PATCHCORE	RD	RKDE	STFPM	UFLOW	YOLOV3	YOLOV5	YOLOV6	YOLOV8	YOLOV9	YOLOV10
Cardboard	77.78	70.59	0	69.57	78.26	47.06	59.26	69.23	73.68	85.71	77.78	66.67	76.92	62.50	75.43	90.00	81.57	**96.28**	**96.36**	92.43

Table: Image Level F1Score for custom dataset cardboard - The table presents F1 scores for 14 Anomalib models and 6 YOLO models across one category from the cardboard dataset [28].

Summary

- Metrics: Image Level/Pixel Level PRO, AUROC, AUPR, F1Score ✅
- Evaluation of Performance of 22 popular models on 5 datasets and custom datasets ✅
- Benchmark of MVTec dataset with Multi-class Unified model ✅
- Training and Testing using API/Bash ✅
- Inference using API/Bash/Gradio ✅
- Integration to Huggingface ✅
- Conda, Docker, Docker Compose Support ✅
- Cross Plattform: Linux (Ubuntu22/20), Windows (Win11/10) ✅
- Publish projects with detailed documentationon on open source platform ✅

Discussion

- **Benchmark**
- Almost no discrepancy: CFLOW, DFM, PatchCore, GANomaly, DFKDE and RD
- Minor discrepancy: CSFLOW, DREAM, DSR, FASTFLOW, FRE, PaDiM, STFPM and UFLOW
- Notable discrepancy: CFA

- **Reason**
- Hardware settings
- Hyper-parameter configurations
- Randomness of training
- Flawed Model re-implementation from Anomalib

Future Work

TECHNISCHE UNIVERSITÄT DARMSTADT

- Extending the functionality of user interaction with Gradio
- Extending integration to Huggingface
- Using Streamlit to simplify Frontend and Backend
- Beautify the visualization of Benchmarking
- Implementing pipeline of IADBE system
- Benchmarking Multi-class Unified model
- Re-implementation of high performance models based on Anomalib
- Enhance the support of video dataset based on Anomalib

[1] IBM. *Anomaly Detection*. https://www.ibm.com/topics/anomaly-detection. Accessed: 2024-06-14.
[2] CVPR 2023 Tutorial Organizers. *CVPR 2023 Tutorial on Anomaly Detection*. https://sites.google.com/view/cvpr2023-tutorial-on-ad/. Accessed: 2024-06-14. 2023.
[3] Denis Gudovskiy, Shun Ishizaka, and Kazuki Kozuka. "CFLOW-AD: Real-Time Unsupervised Anomaly Detection With Localization via Conditional Normalizing Flows". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Jan. 2022, pp. 98–107. URL: https://openaccess.thecvf.com/content/WACV2022/html/Gudovskiy_CFLOW-AD_Real-Time_Unsupervised_Anomaly_Detection_With_Localization_via_Conditional_Normalizing_WACV_2022_paper.html.
[4] Woosang Shin et al. "Anomaly Detection using Score-based Perturbation Resilience". en. In: *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris, France: IEEE, Oct. 2023, pp. 23315–23325. ISBN: 9798350307184. DOI: 10.1109/ICCV51070.2023.02136. URL: https://ieeexplore.ieee.org/document/10376577/ (visited on 03/01/2024).
[5] S. Amit. *Everything You Need to Know About Knowledge Distillation (aka Teacher-Student Model)*. https://amit-s.medium.com/everything-you-need-to-know-about-knowledge-distillation-aka-teacher-student-model-d6ee10fe7276. Accessed: 2024-06-18. 2023.
[6] David Tax. *One-Class Classification*. https://www.tudelft.nl/ewi/over-de-faculteit/afdelingen/intelligent-systems/pattern-recognition-bioinformatics/pattern-recognition-bioinformatics/people/david-tax/one-class-classification. Accessed: 2024-06-18.
[7] Paul Bergmann et al. "MVTec AD – A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2019. URL: https://openaccess.thecvf.com/content_CVPR_2019/html/Bergmann_MVTec_AD_--_A_Comprehensive_Real-World_Dataset_for_Unsupervised_Anomaly_CVPR_2019_paper.html.
[8] Paul Bergmann et al. "The MVTec 3D-AD Dataset for Unsupervised 3D Anomaly Detection and Localization". In: *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. SCITEPRESS - Science and Technology Publications, 2022. DOI: 10.5220/0010865000003124. URL: http://dx.doi.org/10.5220/0010865000003124.

[9] Pankaj Mishra et al. "VT-ADL: A Vision Transformer Network for Image Anomaly Detection and Localization". In: *2021 IEEE 30th International Symposium on Industrial Electronics (ISIE)*. 2021, pp. 01–06. DOI: 10.1109/ISIE45552.2021.9576231.

[10] Domen Tabernik et al. "Segmentation-Based Deep-Learning Approach for Surface-Defect Detection". In: *Journal of Intelligent Manufacturing* (May 2019). ISSN: 1572-8145. DOI: 10.1007/s10845-019-01476-x. URL: https://link.springer.com/article/10.1007/s10845-019-01476-x.

[11] Yang Zou et al. "SPot-the-Difference Self-supervised Pre-training for Anomaly Detection and Segmentation". In: *Computer Vision – ECCV 2022*. Ed. by Shai Avidan et al. Cham: Springer Nature Switzerland, 2022, pp. 392–408. ISBN: 978-3-031-20056-4. URL: https://link.springer.com/chapter/10.1007/978-3-031-20056-4_23.

[12] Marco Rudolph et al. "Fully Convolutional Cross-Scale-Flows for Image-Based Defect Detection". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Jan. 2022, pp. 1088–1097. URL: https://scholar.google.com/scholar?hl=zh-CN&as_sdt=0%2C5&q=Fully+Convolutional+Cross-Scale-Flows+for+Image-Based+Defect+Detection&btnG=.

[13] Samet Akcay et al. "Anomalib: A deep learning library for anomaly detection". In: *2022 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2022, pp. 1706–1710. URL: https://ieeexplore.ieee.org/abstract/document/9897283.

[14] Nilesh A. Ahuja et al. *Probabilistic Modeling of Deep Features for Out-of-Distribution and Adversarial Detection*. 2019. arXiv: 1909.11786 [stat.ML].

[15] Jiawei Yu et al. *FastFlow: Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows*. 2021. arXiv: 2111.07677 [cs.CV].

[16] Philip Adey et al. "Region Based Anomaly Detection with Real-Time Training and Analysis". In: *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*. 2019, pp. 495–499. DOI: 10.1109/ICMLA.2019.00092.

[17] Vitjan Zavrtanik, Matej Kristan, and Danijel Skocaj. "DRAEM - A Discriminatively Trained Reconstruction Embedding for Surface Anomaly Detection". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Oct. 2021, pp. 8330–8339. URL: https://openaccess.thecvf.com/content/ICCV2021/html/Zavrtanik_DRAEM_-_A_Discriminatively_Trained_Reconstruction_Embedding_for_Surface_Anomaly_ICCV_2021_paper.html.

[18] Vitjan Zavrtanik, Matej Kristan, and Danijel Skočaj. “DSR – A Dual Subspace Re-Projection Network for Surface Anomaly Detection”. In: *Computer Vision – ECCV 2022*. Ed. by Shai Avidan et al. Cham: Springer Nature Switzerland, 2022, pp. 539–554. ISBN: 978-3-031-19821-2. URL: https://link.springer.com/chapter/10.1007/978-3-031-19821-2_31.
[19] Samet Akcay, Amir Atapour-Abarghouei, and Toby P. Breckon. “GANomaly: Semi-supervised Anomaly Detection via Adversarial Training”. In: *Computer Vision – ACCV 2018*. Ed. by C. V. Jawahar et al. Cham: Springer International Publishing, 2019, pp. 622–637. ISBN: 978-3-030-20893-6. URL: https://link.springer.com/chapter/10.1007/978-3-030-20893-6_39.
[20] Ibrahima Ndiour et al. “FRE: A Fast Method For Anomaly Detection And Segmentation”. In: *arXiv preprint arXiv:2211.12650* (2022). URL: https://arxiv.org/abs/2211.12650.
[21] Thomas Defard et al. “PaDiM: A Patch Distribution Modeling Framework for Anomaly Detection and Localization”. In: *Pattern Recognition. ICPR International Workshops and Challenges*. Ed. by Alberto Del Bimbo et al. Cham: Springer International Publishing, 2021, pp. 475–489. ISBN: 978-3-030-68799-1. URL: https://link.springer.com/chapter/10.1007/978-3-030-68799-1_35.
[22] Karsten Roth et al. *Towards Total Recall in Industrial Anomaly Detection*. 2021. arXiv: 2106.08265 [cs.CV].
[23] Sungwook Lee, Seunghyun Lee, and Byung Cheol Song. “CFA: Coupled-Hypersphere-Based Feature Adaptation for Target-Oriented Anomaly Localization”. In: *IEEE Access* 10 (2022), pp. 78446–78454. DOI: 10.1109/ACCESS.2022.3193699.
[24] Hanqiu Deng and Xingyu Li. “Anomaly Detection via Reverse Distillation From One-Class Embedding”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2022, pp. 9737–9746. URL: https://openaccess.thecvf.com/content/CVPR2022/html/Deng_Anomaly_Detection_via_Reverse_Distillation_From_One-Class_Embedding_CVPR_2022_paper.html.
[25] Guodong Wang et al. *Student-Teacher Feature Pyramid Matching for Anomaly Detection*. 2021. arXiv: 2103.04257 [cs.CV].
[26] Matías Tailanian, Álvaro Pardo, and Pablo Musé. *U-Flow: A U-shaped Normalizing Flow for Anomaly Detection with Unsupervised Threshold*. 2023. arXiv: 2211.12353 [cs.CV].
[27] Anomalib Developers. *SDD: Developer Guide*. https://anomalib.readthedocs.io/en/latest/markdown/guides/developer/sdd.html. Accessed: 2024-06-19. 2024.

[28] T3Lab. *Anomaly Detection Dataset*. https://universe.roboflow.com/t3lab/anomaly-detection-zbp45/dataset/2. Accessed: 2024-10-03. 2024.

Appendix
Explanation of Abbreviation

shortcut	explanation
AUROC	Area under ROC Curve
PRO	Per Region Overlap
AUPR	Aarea under the PR curve
CFA	Coupled Hypersphere-based Feature Adaptation
CFLOW	Conditional Normalizing Flows
CSFLOW	Cross-Scale-Flows
DFKDE	Deep Feature Kernel Density Estimation
DFM	Probabilistic Modeling of Deep Features for Out-of-Distribution and Adversarial Detection
DRAEM	A Discriminatively Trained Reconstruction Embedding for Surface Anomaly Detection
DSR	Dual Subspace Re-Projection
FastFlow	Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows
GANomaly	Semi-Supervised Anomaly Detection via Adversarial Training
PaDiM	A Patch Distribution Modeling
RD	Reverse Distillation
RKDE	Region-Based Kernel Density Estimation
STFPM	Student-Teacher Feature Pyramid Matching
UFLOW	U-shaped Normalizing Flow
FRE	A Fast Method For Anomaly Detection And Segmentation
YOLO	You Only Look Once
IADBE	Industrial Anomaly Detection Benchmark Engine

Links

- IADBE Core: `https://github.com/cjy513203427/IADBE`
- IADBE Server: `https://github.com/cjy513203427/IADBE_Server`
- IADBE Frontend: `https://github.com/cjy513203427/IADBE_Frontend`
- IADBE Pre-trained Models: `https://huggingface.co/gt111lk/IADBE_Models`
- IADBE Custom Dataset: `https://huggingface.co/datasets/gt111lk/IADBE_Custom_Dataset`
- Inference with Gradio deployed on HF spaces: `https://huggingface.co/spaces/gt111lk/IADBE_Inference`
- Final Version Paper: `https://github.com/cjy513203427/IADBE/blob/master/papers/Evaluating_State_of_the_art_Methods_for_Industrial_Anomaly_Detection_final_version.pdf`

Configuration Parameter	Value	Description
backbone	wide_resnet50_2	Backbone CNN network
gamma_c	1	gamma_c value from the paper
gamma_d	1	gamma_d value from the paper
num_nearest_neighbors	3	Number of nearest neighbors
num_hard_negative_features	3	Number of hard negative features
max_epochs	30	Maximum number of training epochs
radius	1.0e-05	Radius of the hypersphere to search the soft boundary
patience	5	Number of checks with no improvement
mode	max	One of 'min', 'max'. In 'min' mode, training will stop when the quantity monitored has stopped decreasing and in 'max' mode it will stop when the quantity monitored has stopped increasing.

Table: CFA Parameters Settings

Appendix
PaDiM Configuration Parameters

Configuration Parameter	Value	Description
layers	layer1, layer2, layer3	Layers used for feature extraction
backbone	resnet18	Pre-trained model backbone
pre_trained	true	Boolean to check whether to use a pre_trained backbone
n_features	null	Number of features to retain in the dimension reduction step. Default values from the paper are available for: resnet18 (100), wide_resnet50_2 (550). Defaults to “None“.

Table: PaDiM Configuration Parameters

Appendix

CSFLOW Parameters Settings

Configuration Parameter	Value	Description
cross_conv_hidden_channels	1024	Number of hidden channels in the cross convolution
n_coupling_blocks	4	Number of coupling blocks in the model
clamp	3	Clamp value for glow layer
num_channels	3	Number of channels in the model
max_epochs	240	Maximum number of training epochs
patience	3	Number of checks with no improvement

Table: CSFLOW Parameters Settings

Appendix
PatchCore Parameters Settings

Configuration Parameter	Value	Description
backbone	wide_resnet50_2	Backbone CNN network
layers	layer2, layer3	Layers to extract features from the backbone CNN
pre_trained	true	Boolean to check whether to use a pre_trained backbone
coreset_sampling_ratio	0.1	Coreset sampling ratio to subsample embedding
num_neighbors	9	Number of nearest neighbors

Table: PatchCore Parameters Settings

Appendix

Reverse Distillation Parameters Settings

Configuration Parameter	Value	Description
backbone	wide_resnet50_2	Backbone of CNN network
layers	layer1, layer2, layer3	Layers to extract features from the backbone CNN
anomaly_map_mode	ADD	Mode to generate anomaly map
pre_trained	true	Boolean to check whether to use a pre_trained backbone
patience	3	Number of checks with no improvement
check_val_every_n_epoch	200	Frequency of validation checking (in epochs)
max_epochs	200	Maximum number of training epochs

Table: Reverse Distillation Parameters Settings

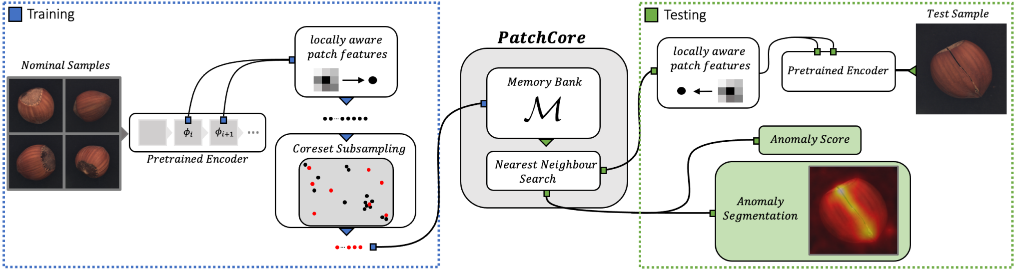


Figure: This is an overview of PatchCore. Nominal samples are broken down into a memory bank of neighborhood-aware patch-level features. In order to reduce redundancy and inference time, this memory bank is downsampled via greedy coreset subsampling. At test time, images are classified as anomalies if at least one patch is anomalous, and pixel-level anomaly segmentation is generated by scoring each patch feature [22].

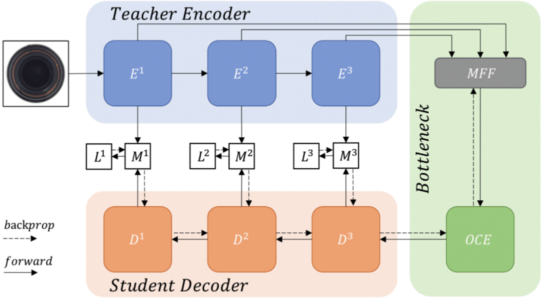


Figure: This figure provides an overview of a reverse distillation framework for anomaly detection and localization. The model comprises a pre-trained teacher encoder *E*, a trainable one-class bottleneck embedding module (OCBE), and a student decoder *D* [24].

Appendix
MVTec Pixel Level AUROC

TECHNISCHE UNIVERSITÄT DARMSTADT

	Screw	Pill	Capsule	Carpet	Grid	Tile	Wood	Zipper	Cable	Toothbrush	Transistor	Metal Nut	Bottle	Hazelnut	Leather	Average
CFA	**98.54**	**98.21**	14.96	22.27	**97.30**	31.83	35.07	23.30	24.63	**98.58**	26.06	35.98	24.52	17.44	29.79	45.23
CFLOW	**97.40**	**97.48**	**98.84**	**99.04**	**97.38**	**96.09**	**95.02**	**97.45**	**96.06**	**98.30**	87.90	**96.65**	**98.58**	**98.72**	**99.54**	**96.96**
CSFLOW	67.61	63.83	79.20	70.34	59.42	67.07	64.05	59.02	64.62	70.36	52.93	67.12	69.03	73.10	66.89	66.31
DFKDE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DFM	90.41	**96.91**	94.24	**95.51**	83.43	91.75	84.55	89.11	**96.06**	**95.95**	**97.79**	94.73	93.88	**97.14**	**95.43**	93.13
DRAEM	65.04	62.63	72.93	55.56	42.77	77.31	70.12	65.49	44.90	69.65	57.17	73.83	74.22	67.57	52.11	63.42
DSR	62.75	43.83	56.51	54.86	68.48	71.69	72.80	60.69	58.96	79.58	57.58	55.36	65.17	73.15	65.55	63.13
FASTFLOW	89.65	92.15	93.43	**97.65**	**97.27**	88.51	94.53	88.90	70.57	90.58	93.54	84.14	**97.05**	93.78	**99.56**	91.42
FRE	93.13	**95.36**	**96.80**	**98.58**	89.74	94.09	89.53	94.18	**95.99**	**98.25**	**98.56**	**97.06**	**97.49**	**98.14**	**98.40**	**95.69**
GANomaly	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
PaDiM	**97.85**	**95.14**	**98.46**	**98.99**	94.18	92.74	92.95	**98.20**	**96.98**	**98.67**	**97.56**	**96.57**	**98.42**	**98.02**	**99.02**	**96.92**
PatchCore	**98.91**	**97.67**	**98.76**	**98.87**	**97.96**	94.75	92.94	**97.94**	**98.01**	**98.64**	**97.16**	**98.22**	**98.16**	**98.45**	**98.99**	**97.70**
RD	**99.52**	**97.21**	**98.87**	**99.05**	**99.06**	**95.08**	94.82	**98.44**	**96.72**	**98.85**	90.44	**96.85**	**98.59**	**98.78**	**99.28**	**97.44**
RKDE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
STFPM	22.10	55.29	61.46	**98.61**	78.55	86.76	91.93	88.39	94.38	64.33	62.27	68.92	66.13	**97.09**	**97.90**	75.61
UFLOW	85.49	**97.79**	82.39	**99.40**	**97.86**	**96.04**	92.04	**96.69**	78.14	70.14	86.52	77.19	**97.98**	**98.71**	**99.37**	90.38

Table: Pixel Level AUROC Scores for MVTec [7] dataset categories - The table presents the area under the receiver operating characteristic curve (AUROC) scores for 16 different methods across 15 different categories from the MVTec dataset.

Appendix
MVTec Pixel Level PRO

	Screw	Pill	Capsule	Carpet	Grid	Tile	Wood	Zipper	Cable	Toothbrush	Transistor	Metal Nut	Bottle	Hazelnut	Leather	Average
CFA	40.24	75.15	**100.00**	**100.00**	49.28	**99.99**	**100.00**	**100.00**	**99.01**	33.04	**100.00**	**100.00**	**100.00**	**100.00**	**99.99**	86.45
CFLOW	26.35	71.89	36.84	82.32	47.11	86.31	81.82	67.94	53.86	32.35	57.84	90.76	84.26	82.56	77.02	65.28
CSFLOW	3.40	73.62	19.70	32.76	43.65	60.71	47.55	69.04	34.64	19.96	23.17	61.62	28.38	56.45	77.36	43.47
DFKDE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DFM	10.06	7.48	16.21	24.63	12.65	72.37	44.84	50.03	30.50	13.54	36.63	38.99	67.90	50.28	25.14	33.42
DRAEM	11.76	43.98	36.83	13.35	4.37	37.52	23.73	18.33	15.26	19.60	32.27	62.74	42.19	56.31	10.19	28.56
DSR	15.21	31.27	18.63	8.89	29.96	22.24	37.75	18.75	20.65	21.06	25.73	54.51	32.74	43.73	20.00	26.74
FASTFLOW	22.78	65.64	43.42	68.76	52.59	60.29	83.90	48.77	45.20	26.01	55.02	74.10	79.29	72.87	83.67	58.82
FRE	15.73	46.38	14.87	57.38	28.18	82.31	58.72	44.08	34.92	17.93	64.24	72.58	69.08	59.99	46.74	47.54
GANomaly	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
PaDiM	30.00	80.57	33.36	75.64	41.35	80.18	83.89	70.17	74.36	55.86	72.39	91.40	83.55	69.49	90.14	68.82
PatchCore	39.05	73.99	42.25	66.39	52.92	75.70	61.65	70.79	66.17	21.35	90.73	79.43	81.62	80.82	59.28	64.14
RD	48.35	82.31	49.69	70.49	56.88	87.72	70.25	73.66	60.30	27.36	66.67	91.35	82.73	82.93	68.25	67.93
RKDE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
STFPM	1.57	57.32	23.66	73.19	22.65	64.49	74.15	36.58	60.27	24.96	12.37	92.29	84.00	66.46	69.99	50.93
UFLOW	2.03	69.63	11.97	79.76	41.02	86.69	73.60	60.55	16.37	22.36	50.78	66.00	71.60	67.67	70.85	52.73

Table: Pixel Level PRO Scores for MVTec [7] dataset categories - The table presents the Per-Region Overlap (PRO) scores for 16 different methods across 15 different categories from the MVTec dataset. Higher PRO scores indicate superior performance in distinguishing between normal and anomalous images.

Appendix

VisA Image Level AUROC

	Candle	Capsules	Cashew	Chewinggum	Fryum	Macaroni1	Macaroni2	PCB1	PCB2	PCB3	PCB4	Pipe Fryum	Average
CFA	85.49	12.67	19.12	**96.12**	19.12	88.20	59.00	94.91	93.31	90.27	**99.04**	19.12	75.32
CFLOW	89.24	94.22	**95.60**	**99.66**	80.48	74.59	66.63	87.86	88.12	74.47	**96.50**	**98.12**	87.12
CSFLOW	36.69	55.96	41.39	67.73	39.54	52.00	38.75	63.75	49.53	55.11	75.28	31.83	52.46
DFKDE	85.91	59.15	86.32	82.86	75.94	65.47	45.47	67.19	66.20	69.61	83.19	77.79	72.09
DFM	**95.23**	57.63	**95.86**	**98.46**	**95.22**	76.92	70.17	93.07	86.92	85.45	**97.47**	90.04	87.21
DRAEM	77.33	73.80	89.16	72.54	79.02	77.44	66.49	72.96	90.77	88.99	**95.67**	82.06	80.52
DSR	65.80	77.63	88.80	91.94	83.54	81.14	65.81	**95.65**	**96.31**	94.81	**98.29**	78.35	84.84
FASTFLOW	**95.24**	78.87	89.00	**98.28**	92.82	88.41	75.74	88.93	88.10	85.57	**96.34**	92.00	89.11
FRE	90.80	65.43	89.22	94.22	86.08	70.77	70.22	81.07	79.35	80.97	**97.58**	77.68	81.95
GANomaly	69.67	67.23	80.96	56.78	89.52	64.24	49.48	27.38	31.19	22.16	54.83	49.58	55.25
PaDiM	81.06	64.85	91.30	**97.66**	87.90	83.42	69.71	86.34	84.92	75.91	**95.91**	91.28	84.19
PatchCore	**98.34**	72.48	**96.32**	**99.22**	**95.60**	86.97	71.52	94.20	93.86	93.11	**99.07**	**99.24**	91.66
RD	94.22	87.52	**96.90**	**99.42**	90.62	**96.21**	84.77	**95.79**	**97.02**	**96.51**	**99.89**	**98.46**	94.78
RKDE	73.97	60.90	85.16	64.46	77.98	62.51	46.59	68.51	78.08	74.20	49.55	73.43	67.95
STFPM	76.44	88.17	59.22	91.32	93.80	74.04	89.62	**95.22**	73.44	90.73	92.13	**95.30**	84.95
UFLOW	92.87	87.45	94.94	**99.34**	93.62	**97.80**	78.58	**95.85**	**96.27**	**96.83**	**96.90**	**97.46**	93.99

Table: Image Level AUROC Scores for VisA [11] dataset categories - The table presents the area under the receiver operating characteristic curve (AUROC) scores for 16 different methods across 12 different categories from the VisA dataset.

Appendix
VisA Pixel Level AUROC

	Candle	Capsules	Cashew	Chewinggum	Fryum	Macaroni1	Macaroni2	PCB1	PCB2	PCB3	PCB4	Pipe Fryum	Average
CFA	**98.92**	27.59	90.96	**99.07**	9.93	**99.54**	**97.93**	**99.33**	**97.59**	**98.31**	**98.50**	84.44	83.51
CFLOW	**98.88**	**97.61**	**98.55**	**99.01**	**95.54**	**97.86**	**97.10**	**99.37**	**96.85**	**96.85**	**97.82**	**98.27**	**98.05**
CSFLOW	50.12	62.73	69.37	61.21	64.98	53.68	47.45	66.91	52.07	68.52	61.46	71.78	60.86
DFKDE	-	-	-	-	-	-	-	-	-	-	-	-	-
DFM	94.70	86.39	**97.79**	**95.01**	**95.64**	82.55	76.89	**97.26**	68.03	90.83	94.42	**98.73**	89.85
DRAEM	64.50	47.57	34.86	30.21	50.06	48.08	57.57	44.42	57.54	63.24	79.05	81.09	54.85
DSR	65.63	74.78	82.29	70.97	55.04	70.43	55.11	67.83	67.78	77.37	57.86	68.34	67.15
FASTFLOW	**97.45**	**97.21**	**98.50**	**98.24**	88.20	**97.41**	94.26	**99.46**	**96.82**	**96.67**	**96.79**	**95.77**	**97.07**
FRE	**96.28**	87.90	**98.85**	**97.41**	**96.31**	90.32	87.91	**98.30**	94.00	92.90	**97.06**	**99.30**	94.71
GANomaly	-	-	-	-	-	-	-	-	-	-	-	-	-
PaDiM	**97.77**	92.45	**98.51**	**98.56**	**96.41**	**97.67**	**96.12**	**98.87**	**98.41**	**98.32**	**97.23**	**99.05**	**97.45**
PatchCore	**99.06**	**97.90**	**98.94**	**98.86**	94.49	**97.66**	**96.12**	**99.54**	**97.68**	**98.02**	**98.02**	**98.97**	**97.94**
RD	**99.01**	**99.55**	**95.37**	**98.82**	**96.19**	**99.36**	**99.23**	**99.69**	**98.78**	**99.18**	**98.28**	**98.96**	**98.54**
RKDE	-	-	-	-	-	-	-	-	-	-	-	-	-
STFPM	93.47	**96.88**	92.95	**96.55**	94.36	89.63	**97.88**	**99.36**	94.61	**98.29**	86.40	**97.79**	**95.72**
UFLOW	**99.16**	**99.14**	**99.62**	**99.44**	**96.64**	**99.72**	90.93	**99.74**	**98.83**	**99.00**	**98.33**	**99.30**	**98.32**

Table: Pixel Level AUROC Scores for VisA [11] dataset categories - The table presents the area under the receiver operating characteristic curve (AUROC) scores for 16 different methods across 12 different categories from the VisA dataset.

Appendix
VisA Pixel Level PRO

TECHNISCHE UNIVERSITÄT DARMSTADT

	Candle	Capsules	Cashew	Chewinggum	Fryum	Macaroni1	Macaroni2	PCB1	PCB2	PCB3	PCB4	Pipe Fryum	Average
CFA	43.81	0.00	0.00	28.43	0.00	28.60	5.29	10.78	15.50	10.82	36.28	0.00	14.96
CFLOW	33.84	24.14	62.06	34.91	28.97	14.30	7.52	12.50	19.57	6.22	29.28	60.67	27.83
CSFLOW	0.23	1.31	80.21	45.18	71.91	0.44	81.06	43.46	80.71	56.41	78.15	9.49	45.71
DFKDE	-	-	-	-	-	-	-	-	-	-	-	-	-
DFM	1.09	3.36	6.11	18.81	10.59	1.22	0.51	4.31	2.90	5.03	3.71	1.65	4.94
DRAEM	8.97	12.51	17.93	1.92	17.80	7.72	4.84	3.79	4.64	19.33	27.70	47.94	14.59
DSR	8.54	18.66	28.47	9.31	20.27	18.45	3.61	21.67	6.27	17.73	9.94	9.47	14.37
FASTFLOW	19.34	25.89	63.64	48.93	39.36	14.42	6.10	16.65	20.37	13.55	17.17	52.29	28.14
FRE	2.30	4.81	6.99	19.24	13.02	2.91	30.42	8.21	18.42	5.75	17.98	24.78	12.90
GANomaly	-	-	-	-	-	-	-	-	-	-	-	-	-
PaDiM	30.11	15.54	49.10	27.16	28.22	7.52	4.41	29.74	38.05	6.99	24.39	40.63	25.16
PatchCore	41.62	23.44	35.42	23.29	31.71	9.54	10.11	13.26	30.65	6.74	33.23	45.93	25.41
RD	42.79	34.65	51.22	31.30	47.60	18.38	21.95	33.47	23.62	7.45	20.64	57.32	32.53
RKDE	-	-	-	-	-	-	-	-	-	-	-	-	-
STFPM	40.91	43.52	48.15	19.84	49.25	10.05	17.16	19.84	14.73	9.57	14.67	53.44	28.43
UFLOW	44.11	30.63	41.94	24.96	49.15	18.79	17.07	11.09	17.69	9.61	31.25	62.83	29.93

Table: Pixel Level PRO Scores for VisA [11] dataset categories - The table presents the Per-Region Overlap (PRO) scores for 16 different methods across 12 different categories from the VisA [11] dataset.

Appendix

MVTec3D Image Level AUROC

	Bagel	Cable Gland	Carrot	Cookie	Dowel	Foam	Peach	Potato	Rope	Tire	Average
CFA	8.52	53.97	19.61	7.91	46.15	86.94	78.81	56.77	92.07	0.00	45.08
CFLOW	91.43	69.35	73.48	67.06	**98.82**	68.56	70.39	55.73	94.79	74.16	76.38
CSFLOW	53.33	61.03	43.41	48.16	38.31	70.69	56.62	33.45	72.55	63.84	54.14
DFKDE	64.36	71.67	73.25	49.10	72.97	53.87	58.24	52.57	83.11	42.30	62.14
DFM	92.51	66.78	45.38	31.52	45.23	33.19	59.51	65.81	84.83	59.45	58.42
DRAEM	88.02	71.10	56.78	50.87	44.79	14.03	71.12	70.65	59.78	62.16	58.93
DSR	81.87	76.25	71.21	36.51	70.58	53.37	60.14	55.26	48.19	57.29	61.07
FASTFLOW	83.99	66.28	78.09	71.67	67.68	72.62	48.40	40.51	88.27	67.31	68.48
FRE	85.49	59.44	63.22	59.15	77.09	73.87	59.11	46.59	91.53	62.02	72.75
GANomaly	36.78	65.96	50.22	2.84	62.72	47.87	59.00	56.87	40.49	44.83	46.76
PaDiM	**95.71**	73.56	64.83	62.48	89.68	64.50	72.50	53.51	76.36	80.97	73.41
PatchCore	92.72	91.95	86.84	74.48	**97.86**	72.56	84.94	60.13	94.66	79.95	83.61
RD	**96.18**	90.75	91.44	66.40	**99.33**	83.56	84.43	68.38	-	-	**95.06**
RKDE	51.81	71.43	47.00	60.85	56.32	53.87	53.92	41.90	35.85	56.28	52.92
STFPM	91.74	84.35	64.20	49.38	82.66	73.44	68.65	52.52	93.98	51.13	71.21
UFLOW	**95.40**	88.40	84.34	51.28	**96.89**	62.44	60.49	57.02	**96.24**	60.92	75.34

Appendix

MVTec3D Pixel Level AUROC

	Bagel	Cable Gland	Carrot	Cookie	Dowel	Foam	Peach	Potato	Rope	Tire	Average
CFA	23.18	36.11	30.20	24.75	8.73	**99.82**	**99.09**	**99.29**	**98.98**	34.95	55.51
CFLOW	**95.53**	**98.22**	**98.42**	94.44	**99.70**	**99.90**	**97.79**	**98.37**	**98.80**	**97.52**	**97.87**
CSFLOW	64.73	70.88	75.32	60.59	81.40	76.70	69.86	80.40	62.41	82.90	72.52
DFKDE	-	-	-	-	-	-	-	-	-	-	-
DFM	72.92	89.84	84.44	57.06	59.32	38.87	72.14	51.32	61.05	20.98	60.79
DRAEM	69.74	72.89	35.05	49.03	85.19	21.98	67.83	67.16	60.50	43.56	57.29
DSR	60.65	56.68	72.45	45.68	64.70	64.01	47.89	51.01	55.15	52.04	57.03
FASTFLOW	**98.01**	92.24	87.66	88.10	**97.15**	**99.91**	**97.97**	**97.40**	**97.76**	90.57	94.68
FRE	91.34	93.59	94.35	90.16	**96.15**	**95.41**	**97.26**	**97.32**	**95.61**	89.40	94.06
GANomaly	-	-	-	-	-	-	-	-	-	-	-
PaDiM	**98.35**	**96.26**	**97.23**	**95.25**	**98.94**	**99.77**	**98.96**	**98.04**	94.75	**95.22**	**97.28**
PatchCore	94.89	**98.34**	**98.30**	**96.16**	**99.68**	**99.57**	**98.57**	**98.45**	**97.49**	**99.20**	**98.07**
RD	**98.17**	**99.06**	**99.26**	**96.83**	**99.80**	**99.86**	**99.29**	**99.28**	-	-	**98.94**
RKDE	-	-	-	-	-	-	-	-	-	-	-
STFPM	**97.63**	**97.51**	77.11	**95.23**	**98.88**	**99.76**	**99.04**	**99.43**	**99.12**	**95.71**	**95.94**
UFLOW	**95.17**	**98.70**	**98.13**	87.15	**99.74**	**99.95**	**98.11**	**96.96**	**99.43**	88.23	**96.16**

Appendix
MVTec3D Pixel Level PRO

	Bagel	Cable Gland	Carrot	Cookie	Dowel	Foam	Peach	Potato	Rope	Tire	Average
CFA	0.00	0.00	0.00	0.00	**99.95**	45.36	42.68	20.00	18.66	0.00	22.67
CFLOW	34.00	22.17	19.21	21.65	47.07	65.67	18.21	14.27	25.80	20.63	28.87
CSFLOW	3.66	8.54	76.48	2.24	70.07	37.25	20.37	80.35	39.44	30.23	37.86
DFKDE	-	-	-	-	-	-	-	-	-	-	-
DFM	14.74	33.62	10.44	3.29	11.23	34.69	4.04	6.99	6.90	2.42	13.44
DRAEM	8.07	4.17	5.95	2.70	7.25	1.34	6.08	4.24	8.81	3.34	5.30
DSR	13.04	8.86	6.29	0.70	17.62	17.57	6.50	3.03	3.04	1.98	8.56
FASTFLOW	30.62	30.50	17.77	22.75	25.62	64.30	27.49	6.40	22.38	7.28	25.51
FRE	41.07	4.11	4.93	18.50	26.36	8.33	17.02	2.69	4.04	1.74	12.88
GANomaly	-	-	-	-	-	-	-	-	-	-	-
PaDiM	17.41	42.26	11.42	28.61	32.16	46.11	33.23	54.18	12.65	12.31	29.83
PatchCore	46.07	41.54	19.42	36.18	60.30	59.81	30.05	14.57	2.72	22.48	33.31
RD	38.78	39.68	18.11	34.88	50.54	49.48	30.58	21.72	-	-	35.47
RKDE	-	-	-	-	-	-	-	-	-	-	-
STFPM	33.75	42.29	20.81	34.09	33.28	55.04	16.59	34.61	11.70	14.13	31.03
UFLOW	30.62	30.50	17.77	22.75	25.62	64.30	27.49	6.40	22.38	7.28	25.51

	CFA	CFLOW	CSFLOW	DFKDE	DFM	DRAEM	DSR	FASTFLOW	FRE	GANomaly	PaDiM	PatchCore	RD	RKDE	STFPM	UFLOW
Kolektor	52.77	86.65	58.70	67.84	92.05	51.92	57.72	77.36	83.13	56.68	80.25	86.76	88.24	64.08	66.24	90.91

Table: Image Level AUROC Scores for Kolektor [10] dataset categories - The table presents the area under the receiver operating characteristic curve (AUROC) scores for 16 different methods across one category from the Kolektor dataset. Higher AUROC scores indicate superior performance in distinguishing between normal and anomalous images.

Appendix
Kolektor Pixel Level AUROC

	CFA	CFLOW	CSFLOW	DFKDE	DFM	DRAEM	DSR	FASTFLOW	FRE	GANomaly	PaDiM	PatchCore	RD	RKDE	STFPM	UFLOW
Kolektor	56.17	85.41	46.09	-	86.04	53.13	64.76	86.47	84.55	-	84.87	88.30	89.96	-	84.12	**97.04**

Table: Pixel Level AUROC Scores for Kolektor [10] dataset categories - The table presents the Per-Region Overlap (PRO) scores for 16 different methods across one category from the Kolektor dataset. Higher PRO scores indicate superior performance in distinguishing between normal and anomalous images.

Appendix
Kolektor Pixel Level PRO

TECHNISCHE UNIVERSITÄT DARMSTADT

	CFA	CFLOW	CSFLOW	DFKDE	DFM	DRAEM	DSR	FASTFLOW	FRE	GANomaly	PaDiM	PatchCore	RD	RKDE	STFPM	UFLOW
Kolektor	0	9.59	81.47	-	8.25	9.24	2.49	9.71	6.83	-	9.28	7.33	7.87	-	10.45	23.22

Table: Pixel Level PRO Scores for Kolektor [10] dataset categories - The table presents the Per-Region Overlap (PRO) scores for 16 different methods across one category from the Kolektor dataset. Higher PRO scores indicate superior performance in distinguishing between normal and anomalous images.