Real3D-AD: A Dataset of Point Cloud Anomaly Detection

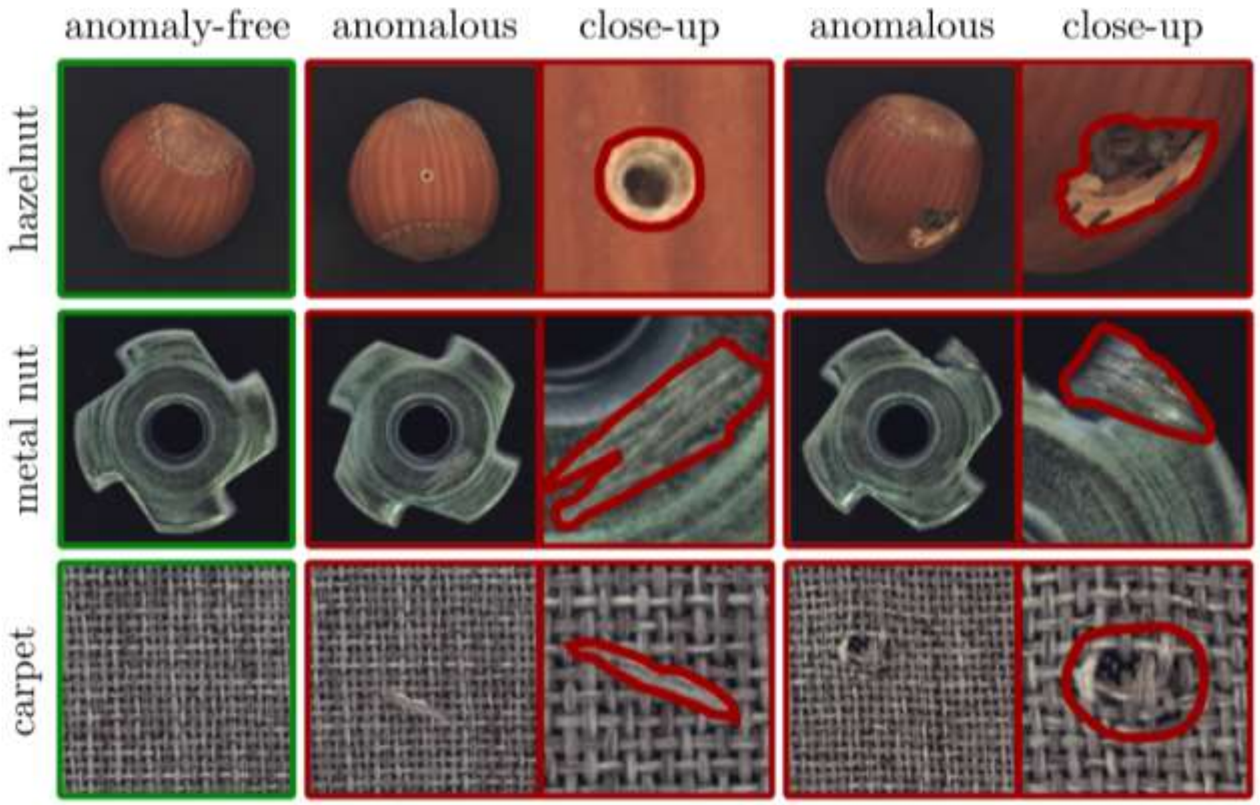
Presenter: Jiaqi Liu

Jiaqi Liu [1,*], Guoyang Xie [1,2,*], Ruitao Chen[1,*], Xinpeng Li[1], Jinbao Wang[1,†], Yong Liu[3], Chengjie Wang[3,4], Feng Zheng[1,†]
1: Department of Computer Science and Engineering,
Southern University of Science and Technology, Shenzhen 518055, China
2: NICE Group, University of Surrey, Guildford, GU2 7YX, United Kingdom
3: Tencent Youtu Lab, Shenzhen 518057, China
4: Shanghai Jiao Tong University, Shanghai 200240, China

⋆ Contributed Equally, † Corresponding Authors

Background

- **Anomaly Detection**
  - Identify abnormal samples.
  - Localize abnormal regions.
  - The training set usually includes only normal samples.
  - Anomalies are varied and unpredictable.



[1] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Mvtec ad — a comprehensive real-world dataset for unsupervised anomaly detection. 2019 IEEE/CVF Conference on ComputerVision and Pattern Recognition (CVPR), pp. 9584–9592, 2019.
[2] Paul Bergmann, Kilian Batzner, Michael Fauser, David Sattlegger, and Carsten Steger. Beyond dentsand scratches: Logical constraints in unsupervised anomaly detection and localization. International Journal of Computer Vision, 130(4):947–969, 2022.

Background

- **Current 3D Anomaly Detection**
  - Extend RGB anomaly detection to RGBD.
  - Only single side.

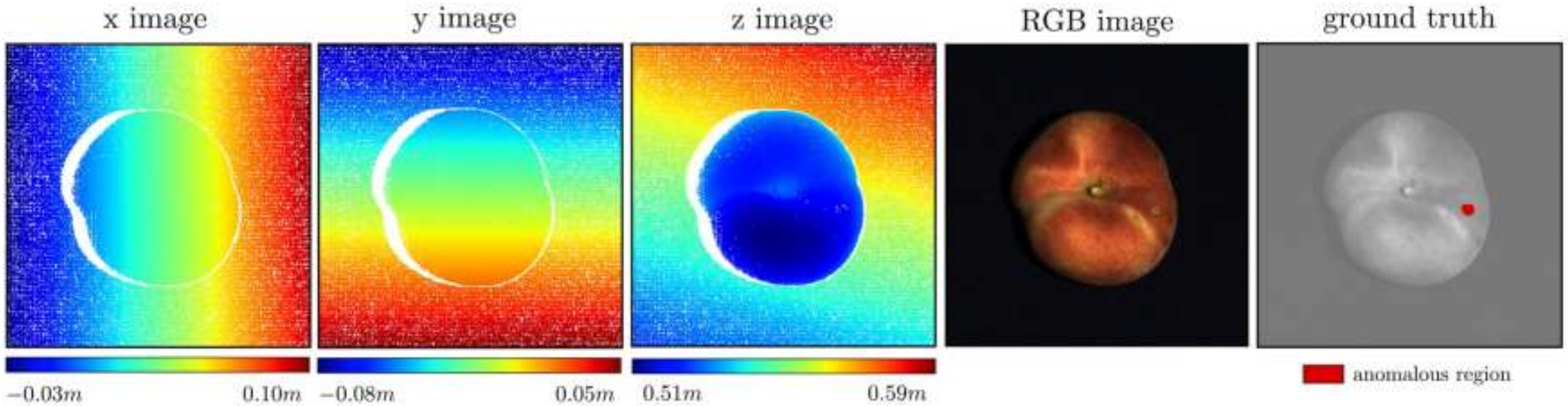


Figure 3: Visualization of the provided data for one anomalous test sample of the dataset category *peach*. In addition to three images that encode the 3D coordinates of the object, RGB information as well as a pixel-precise ground-truth image are provided.

[3] Paul Bergmann, Xin Jin, David Sattlegger, Carsten Steger: 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 - Volume 5: VISAPP, 202-213, 2022, DOI: 10.5220/0010865000003124.

Motivation

- **Enabling the model to recognize anomalies like humans do**

- Humans rely on a complete product prototype to infer defects in other products.
- Providing complete product prototypes for the training set, eliminating factors such as object poses and shooting angles.
- During testing, only one side is observed, just like manual inspection on the actual production line.

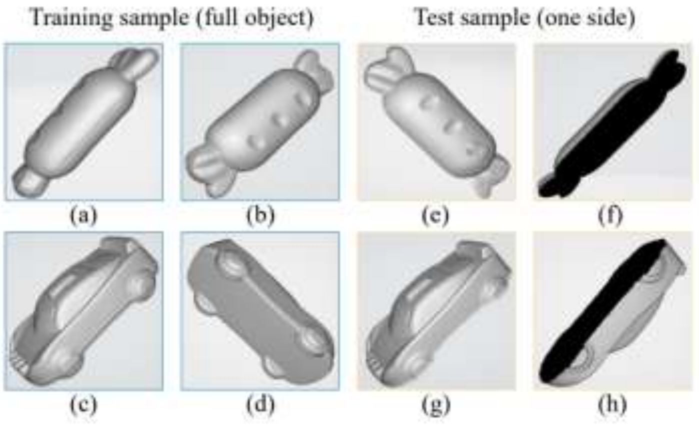


Figure 7: Examples of training and test samples in Real3D-AD.

Dataset

• Dataset: data collection

- The training prototype samples used for training are obtained through multiple scans and manual stitching.

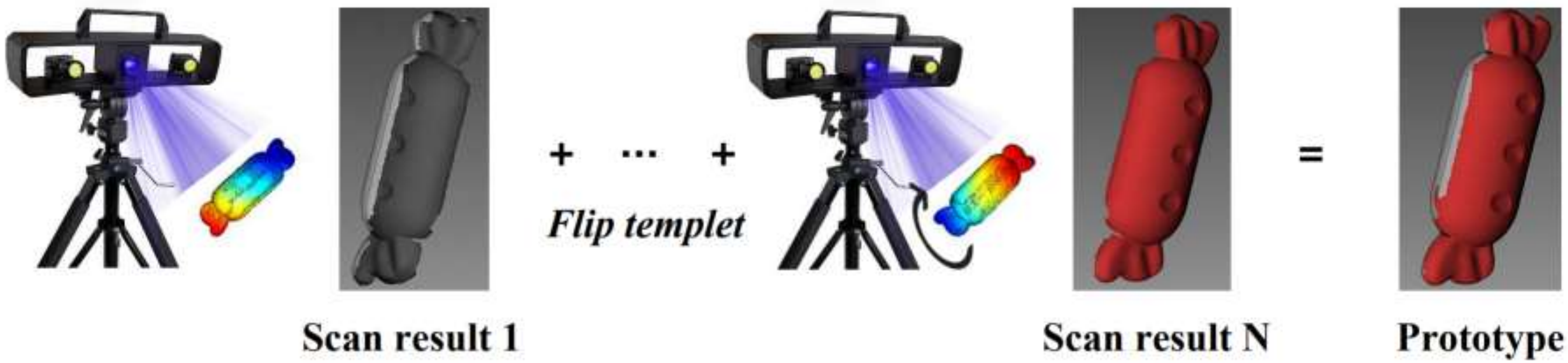


Figure 3: A prototype in the training set is made from two or more scan results.

Dataset

• Dataset: data collection

- We used CloudCompare to label anomalies and output the raw data in pcd format.

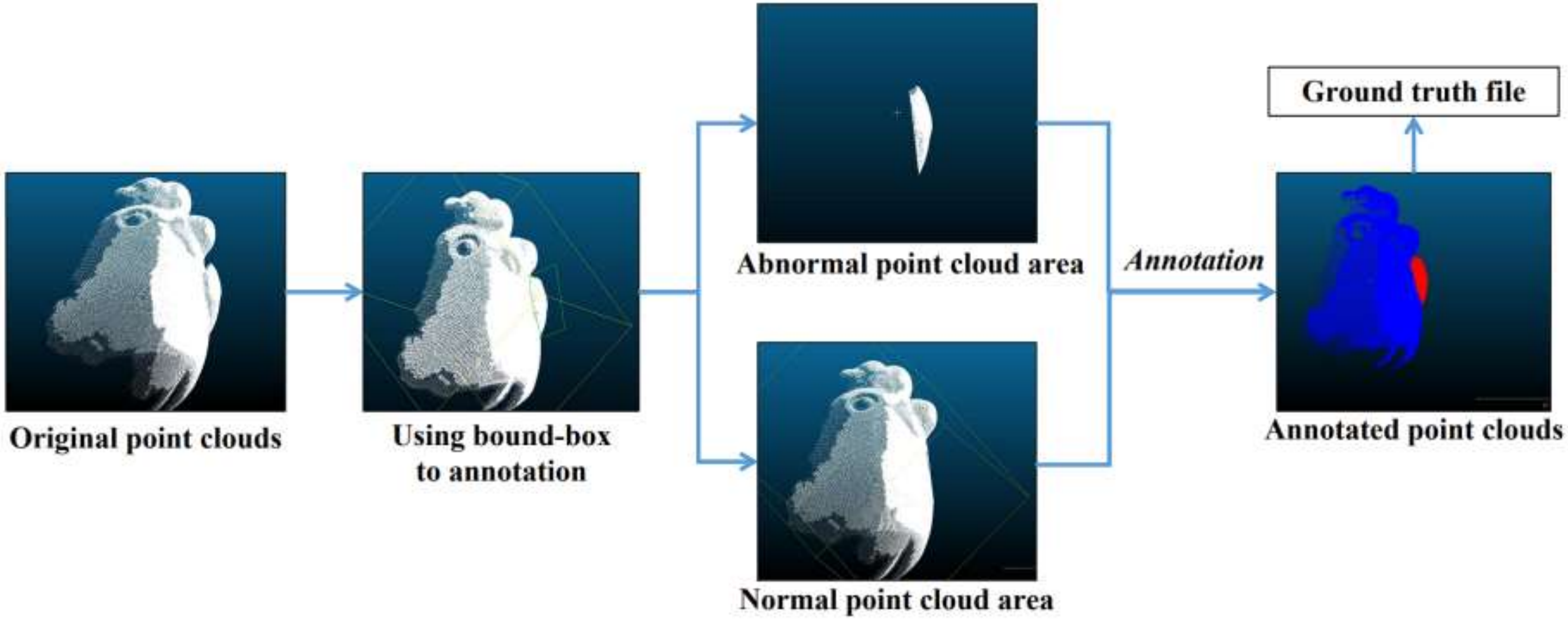


Figure 4: Anomalies annotation in Real3D-AD.

[4] CloudCompare Community. Cloudcompare - a 3d pointcloud and mesh software. 2016.

Dataset

• Dataset: data

- Our dataset contains 12 different objects.

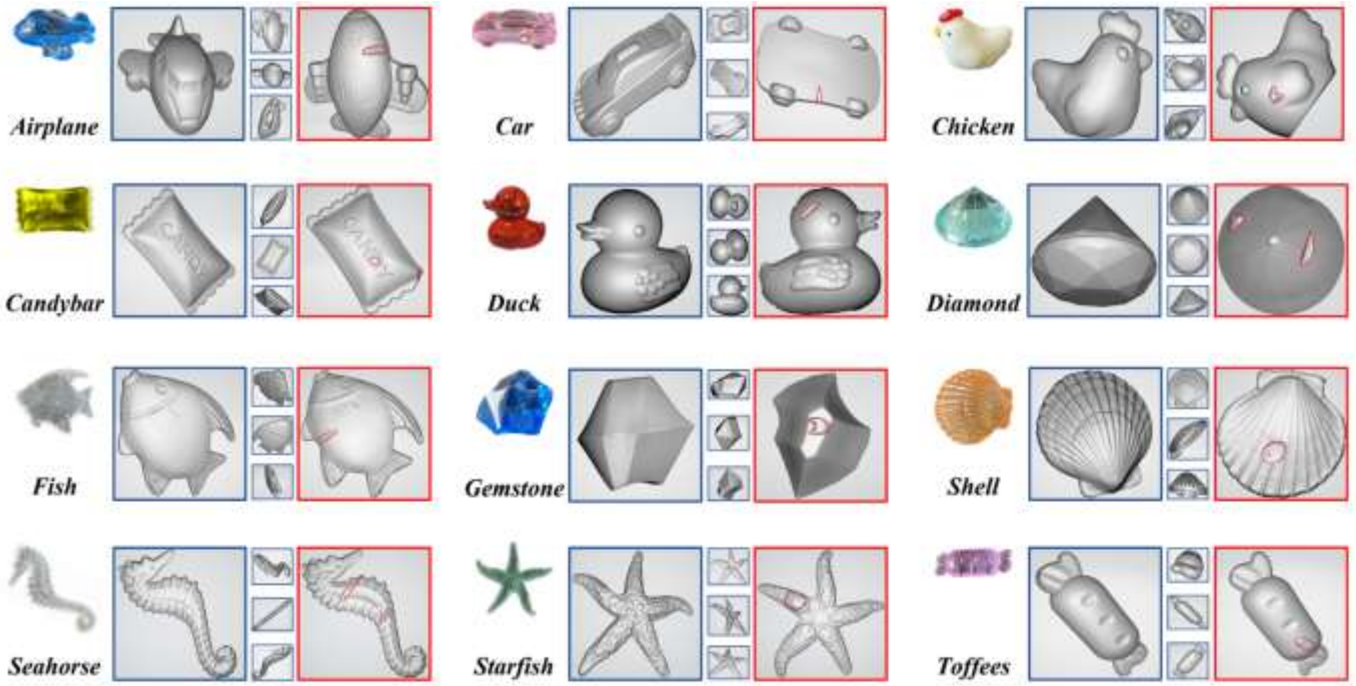


Figure 1: Real3D-AD dataset examples for each category. The blue box indicates the normal images in the training dataset. The red box denotes the abnormal images in the test dataset. There are no blind spots in Real3D-AD since our dataset are achieved by scanning all the views of the object instead of the single view photoed by RGBD camera.

Dataset

• Dataset: data statistics

- The objects we captured are relatively small, and the proportion of defect points is low, which brings some challenges to the detection.

Category	Real Size [mm]			Attribute	Training	Test		Total	Anomaly Point Ratio
	Length	*Width*	*Height*		*Normal*	*Normal*	*Abnormal*		Δ
Airplane	34.0	14.2	31.7	Transparency	4	50	50	104	1.18%
Car	35.0	29.0	12.5	Transparency	4	50	50	104	1.99%
Candybar	33.0	20.0	8.0	Transparency	4	50	50	104	2.37%
Chicken	25.0	14.0	20.0	White	4	52	54	110	4.39%
Diamond	29.0	29.0	18.7	Transparency	4	50	50	104	5.41%
Duck	30.0	22.2	29.4	Transparency	4	50	50	104	2.00%
Fish	37.7	24.0	4.0	Transparency	4	50	50	104	2.86%
Gemstone	22.5	18.8	17.0	Transparency	4	50	50	104	2.06%
Seahorse	38.0	11.2	3.5	Transparency	4	50	50	104	4.57%
Shell	21.7	22.0	7.7	Transparency	4	52	48	104	2.25%
Starfish	27.4	27.4	4.8	Transparency	4	50	50	104	4.47%
Toffees	38.0	12.0	10.0	Transparency	4	50	50	104	2.46%
Mean	30.9	20.3	13.9	—	4	50	50	104	3.00%
Total	—	—	—	—	48	604	602	1254	—

Dataset

• Dataset: data statistics

- We used CloudCompare to label anomalies and output the raw data in pcd format.

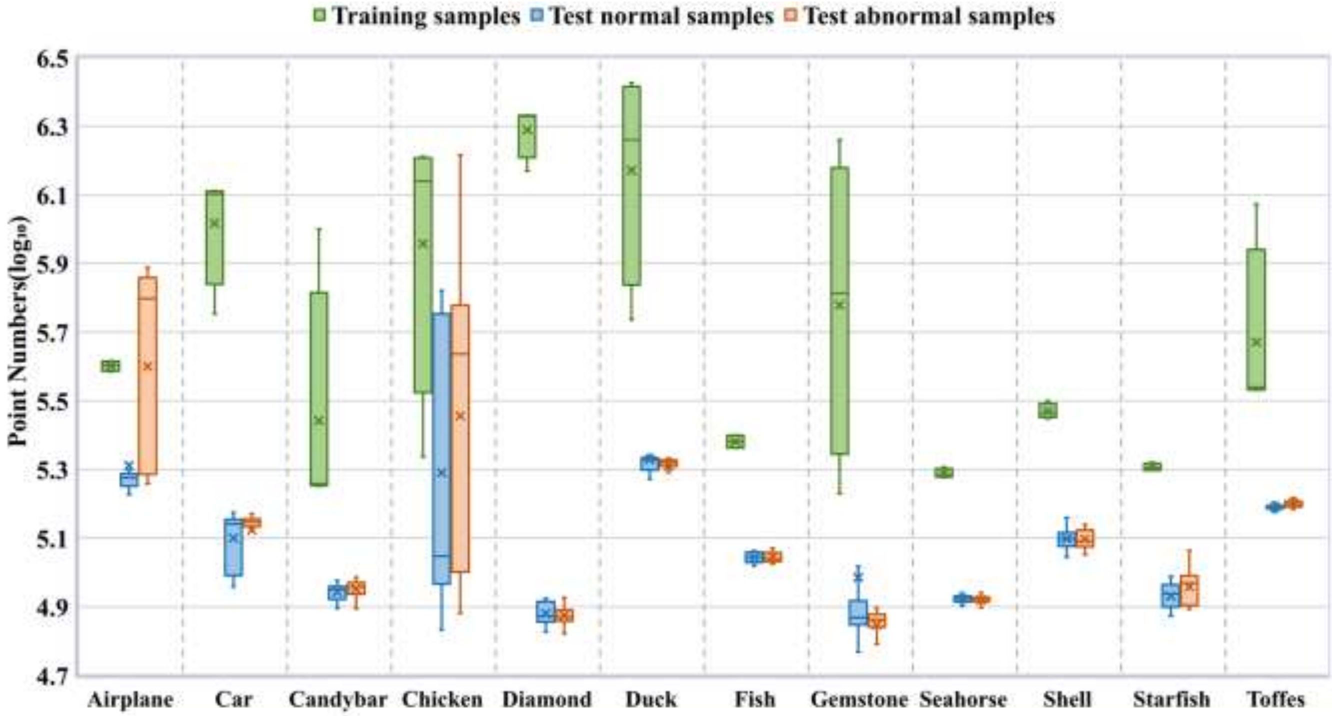


Figure 6: Point numbers for all samples on a logarithmic scale, visualized by a box-and-whisker plot.

ADBench-3D

• Benchmark: M3DM[5] , BTF[6] and PatchCore[7]

- We adapted some feature-based retrieval methods to our dataset to establish a benchmark.

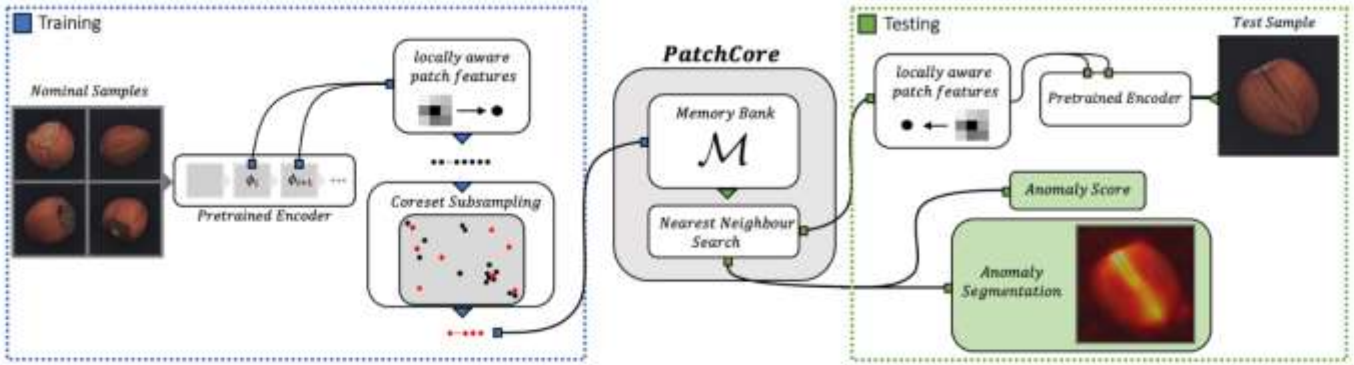


Figure 2. Overview of *PatchCore*. Nominal samples are broken down into a memory bank of neighbourhood-aware patch-level features. For reduced redundancy and inference time, this memory bank is downsampled via greedy coreset subsampling. At test time, images are classified as anomalies if at least one patch is anomalous, and pixel-level anomaly segmentation is generated by scoring each patch-feature.

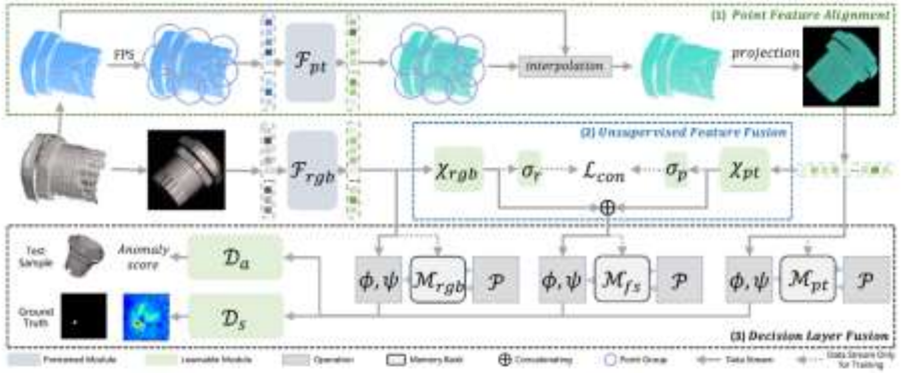


Figure 2. **The pipeline of Multi-3D-Memory (M3DM).** Our M3DM contains three important parts: (1) ***Point Feature Alignment*** (PFA) converts Point Group features to plane features with interpolation and project operation, FPS is the farthest point sampling and $\mathcal{F}_{pt}$ is a pretrained Point Transformer; (2) ***Unsupervised Feature Fusion*** (UFF) fuses point feature and image feature together with a patch-wise contrastive loss $\mathcal{L}_{con}$, where $\mathcal{F}_{rgb}$ is a Vision Transformer, χ_{rgb}, χ_{pt} are MLP layers and σ_r, σ_p are single fully connected layers; (3) ***Decision Layer Fusion*** (DLF) combines multimodal information with multiple memory banks and makes the final decision with 2 learnable modules $\mathcal{D}_a, \mathcal{D}_s$ for anomaly detection and segmentation, where $\mathcal{M}_{rgb}, \mathcal{M}_{fs}, \mathcal{M}_{pt}$ are memory banks, ϕ, ψ are score function for single memory bank detection and segmentation, and $\mathcal{P}$ is the memory bank building algorithm.

[5] Wang, Y., Peng, J., Zhang, J., Yi, R., Wang, Y., & Wang, C. (2023). Multimodal Industrial Anomaly Detection via Hybrid Fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8032-8041).
[6] Horwitz, E., & Hoshen, Y. (2023). Back to the feature: classical 3d features are (almost) all you need for 3d anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2967-2976).
[7] Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., & Gehler, P. (2022). Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 14318-14328).

ADBench-3D

• Benchmark: M3DM[4] , BTF[5] and PatchCore[6]

- We adapted some feature-based retrieval methods to our dataset to establish a benchmark.

Table 4: ADBENCH-3D for Real3D-AD. The score indicates object-level AUROC ↑. The best results are highlighted in bold.

Category	BTF		M3DM		PatchCore			Reg3D-AD
	Raw	*FPFH*	*PointMAE*	*PointBERT*	*FPFH*	*FPFH+Raw*	*PointMAE*	
Airplane	0.730	0.520	0.434	0.407	**0.882**	0.848	0.726	0.716
Car	0.647	0.560	0.541	0.506	0.590	**0.777**	0.498	0.697
Candybar	0.539	0.630	0.552	0.562	0.541	0.570	0.663	**0.685**
Chicken	0.789	0.432	0.683	0.673	0.837	**0.853**	0.827	0.852
Diamond	0.707	0.545	0.602	0.627	0.574	0.784	0.783	**0.900**
Duck	0.691	**0.784**	0.433	0.466	0.546	0.628	0.489	0.584
Fish	0.602	0.549	0.540	0.556	0.675	0.837	0.630	**0.915**
Gemstone	**0.686**	0.648	0.644	0.617	0.370	0.359	0.374	0.417
Seahorse	0.596	**0.779**	0.495	0.494	0.505	0.767	0.539	0.762
Shell	0.396	**0.754**	0.694	0.577	0.589	0.663	0.501	0.583
Starfish	0.530	**0.575**	0.551	0.528	0.441	0.471	0.519	0.506
Toffees	0.703	0.462	0.450	0.442	0.565	0.626	0.585	**0.827**
Average	0.635	0.603	0.552	0.538	0.593	0.682	0.594	**0.704**

[4] Wang, Y., Peng, J., Zhang, J., Yi, R., Wang, Y., & Wang, C. (2023). Multimodal Industrial Anomaly Detection via Hybrid Fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8032-8041).
[5] Horwitz, E., & Hoshen, Y. (2023). Back to the feature: classical 3d features are (almost) all you need for 3d anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2967-2976).
[6] Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., & Gehler, P. (2022). Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 14318-14328).

ADBench-3D

• Baseline: Reg3D-AD

- Based on the characteristics of the dataset, we combined point cloud registration with feature-based retrieval for anomaly detection.

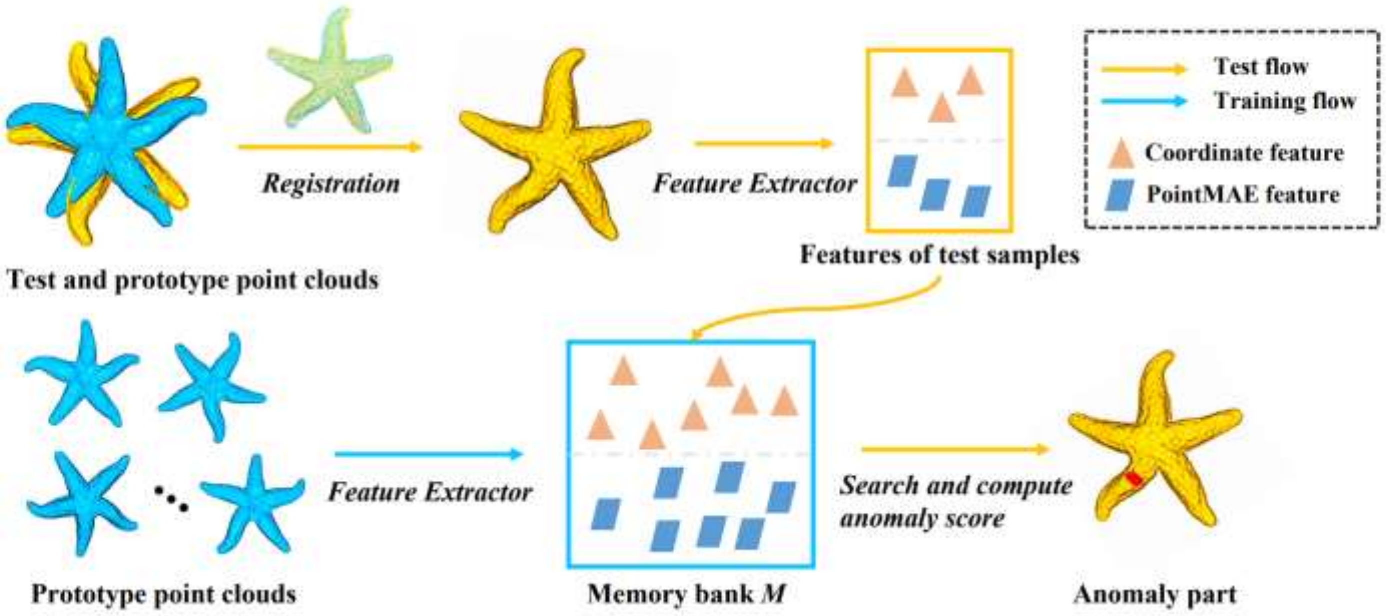


Figure 8: Pipeline of our baseline method. We extract features from the training set and sample the most representative features to the memory bank during training. During inference, we use the prototype as the target to calibrate the test sample and then extract the characteristics of the test sample to compare with the memory bank. We compute the anomaly score for each point according to the distance between test features and the memory bank.

[4] Wang, Y., Peng, J., Zhang, J., Yi, R., Wang, Y., & Wang, C. (2023). Multimodal Industrial Anomaly Detection via Hybrid Fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8032-8041).
[5] Horwitz, E., & Hoshen, Y. (2023). Back to the feature: classical 3d features are (almost) all you need for 3d anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2967-2976).
[6] Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., & Gehler, P. (2022). Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 14318-14328).

NEURAL INFORMATION PROCESSING SYSTEMS

SUSTech Southern University of Science and Technology

UNIVERSITY OF SURREY

Tencent YouTu Lab

SHANGHAI JIAO TONG UNIVERSITY

THANK YOU

[1] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Mvtec ad — a comprehensive real-world dataset for unsupervised anomaly detection. 2019 IEEE/CVF Conference on ComputerVision and Pattern Recognition (CVPR), pp. 9584–9592, 2019.

[2] Paul Bergmann, Kilian Batzner, Michael Fauser, David Sattlegger, and Carsten Steger. Beyond dentsand scratches: Logical constraints in unsupervised anomaly detection and localization. International Journal of Computer Vision, 130(4):947–969, 2022.

[3] Paul Bergmann, Xin Jin, David Sattlegger, Carsten Steger: 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 - Volume 5: VISAPP, 202-213, 2022, DOI: 10.5220/0010865000003124.

[4] CloudCompare Community. Cloudcompare - a 3d pointcloud and mesh software. 2016.

[5] Wang, Y., Peng, J., Zhang, J., Yi, R., Wang, Y., & Wang, C. (2023). Multimodal Industrial Anomaly Detection via Hybrid Fusion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8032-8041).

[6] Horwitz, E., & Hoshen, Y. (2023). Back to the feature: classical 3d features are (almost) all you need for 3d anomaly detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 2967-2976).

[7] Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., & Gehler, P. (2022). Towards total recall in industrial anomaly detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 14318-14328).