MC3D-AD: A Unified Geometry-aware Reconstruction Model for Multi-category 3D Anomaly Detection——Appendix

1 Additional Experiments

1.1 More ablation experiments

(a)

$LGE_{w/o\ AGMA}$	LGE_{AGMA}	$GQD_{w/o\ AGMA}$	GQD_{AGMA}	O-AUROC	P-AUROC
✓				0.658	0.650
✓	✓			0.752	0.709
✓		✓		0.741	0.733
✓	✓	✓		0.756	0.755
✓	✓	✓	✓	**0.782**	**0.768**

(b)

Sampling	O-AUROC	P-AUROC	Perturbation	O-AUROC	P-AUROC
FPS(1024)	0.729	0.734	0(no perturbaiton)	0.753	0.738
FPS(2048)	0.747	0.757	5	0.758	0.735
FPS(4096)	**0.782**	**0.768**	10	0.771	0.756
Random(1024)	0.719	0.724	15	0.774	0.761
Random(2048)	0.738	0.746	20	**0.782**	**0.768**
Random(4096)	0.775	0.756	25	0.766	0.752

Table 1: More ablation results on Real3D-AD. (a) Ablation on key modules, where **w/o AGMA** means model without the AGMA module. (b) Ablation on Farthest Point Sampling (FPS) and perturbation.

Table 1 provides additional ablation results to further validate the effectiveness of the proposed modules. Due to space constraints, the ablation study of key components is clearly and concisely illustrated in (a). In (b), we investigate the impact of point cloud sampling strategies and feature perturbation levels. The results indicate that FPS better preserves structural information than random sampling, and increasing the number of groups enhances feature completeness. Moreover, moderate noise perturbation promotes the learning of normal feature representations, improving model robustness.

1.2 Experiment Results

Results Anomaly-ShapeNet. The pixel-level AUROC results of MC3D-AD on the Anomaly-ShapeNet dataset [Li *et al.*, 2024] are presented in Table 2. It can be observed that MC3D-AD achieves an 8.0% improvement in pixel-level AUROC compared to the second single-category method. This further confirms the effectiveness of MC3D-AD in multi-category anomaly localization.

1.3 Visualization

Figure 1 shows the visualized results of our method on Anomaly-ShapeNet. It is clear that MC3D-AD can accurately detect and locate small anomalies within the point cloud from different categories.

1.4 Generalization of Our Method for Classification

To examine the generalization of our method, extra classification experiments are conducted on Real3D-AD [Liu *et al.*, 2023], and the results are shown in Table 3. It can be observed in Table 3 that, without losing anomaly detection performance, the classification capability remains satisfactory even when each class contains only four training samples.

The Anomaly-ShapeNet dataset contains 40 categories, with only 4 training samples per class, making the classification task significantly more challenging. Therefore, experiments are conducted on a subset of the dataset with 10 categories and the experimental results are shown in Table 4. It is evident that our MC3D-AD achieved an accuracy of 0.901, which further demonstrates the effectiveness of our MC3D-AD in dealing with multi-task.

1.5 Extensibility of the Proposed AGMA

To evaluate the extensibility of AGMA, it was integrated into PointMAE [Pang *et al.*, 2022] to perform point cloud classification and segmentation tasks on ModelNet40 [Wu *et al.*, 2015] and ShapeNet-part [Yi *et al.*, 2016]. The methods selected for comparison include PointNet [Charles *et al.*, 2017], PointNet++ [Qi *et al.*, 2017], DGCNN [Wang *et al.*, 2019],

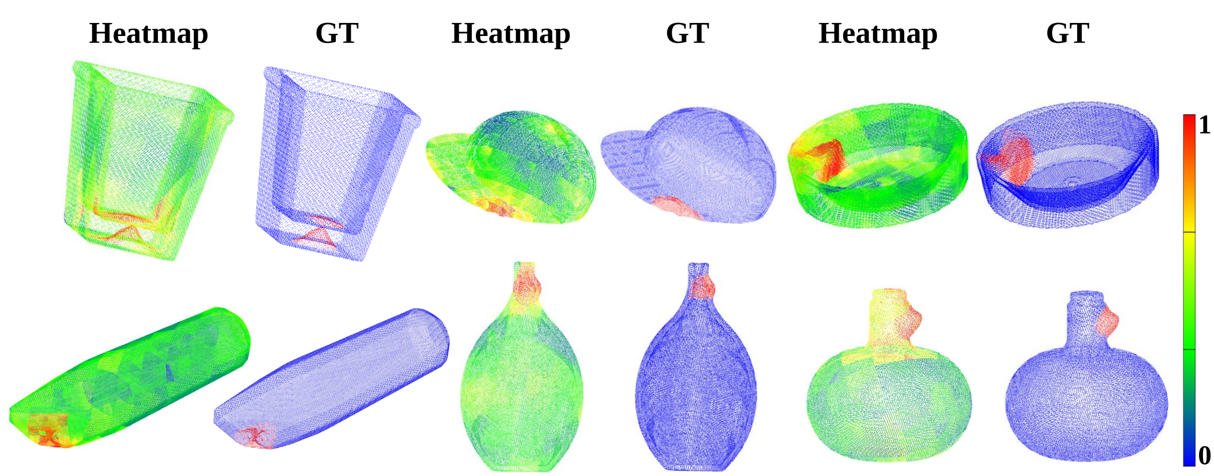

Figure 1: Point heatmap comparison of our MC3D-AD with the Ground Truth (GT) on Anomaly-ShapeNet.

P-AUROC(↑)														
Method	**cap0**	**cap3**	**helmet3**	**cup0**	**bowl4**	**vase3**	**headset1**	**eraser0**	**vase8**	**cap4**	**vase2**	**vase4**	**helmet0**	**bucket1**
BTF(Raw)	0.524	0.687	0.700	0.632	0.563	0.602	0.475	0.637	0.550	0.469	0.403	0.613	0.504	0.686
BTF(FPFH)	**0.730**	0.658	**0.724**	**0.790**	0.679	**0.699**	0.591	0.719	0.662	0.524	0.646	0.710	0.575	0.633
M3DM	0.531	0.605	0.655	0.715	0.624	0.658	0.585	0.710	0.551	0.718	0.737	0.655	0.599	0.699
Patchcore(FPFH)	0.472	0.653	**0.737**	0.655	**0.720**	0.430	0.464	**0.810**	0.575	0.595	0.721	0.505	0.548	0.571
Patchcore(PointMAE)	0.544	0.488	0.615	0.510	0.501	0.465	0.423	0.378	0.364	0.725	**0.742**	0.523	0.580	0.574
CPMF	0.601	0.551	0.520	0.497	0.683	0.582	0.458	0.689	0.529	0.553	0.582	0.514	0.555	0.601
Reg3D-AD	0.632	**0.718**	0.620	0.685	**0.800**	0.511	**0.626**	0.755	**0.811**	**0.815**	0.405	**0.755**	**0.600**	0.752
IMRNet	0.715	0.706	0.663	0.643	0.576	0.401	0.476	0.548	0.635	0.753	0.614	0.524	0.598	**0.774**
Ours	**0.854**	**0.903**	0.585	**0.763**	0.670	**0.800**	**0.592**	**0.820**	**0.874**	**0.858**	**0.781**	**0.772**	**0.749**	**0.868**

Method	**bottle3**	**vase0**	**bottle0**	**tap1**	**bowl0**	**bucket0**	**vase5**	**vase1**	**vase9**	**ashtray0**	**bottle1**	**tap0**	**phone**	**cup1**
BTF(Raw)	**0.720**	0.618	0.551	0.564	0.524	0.617	0.585	0.549	0.564	0.512	0.491	0.527	0.583	0.561
BTF(FPFH)	0.622	0.642	0.641	0.596	0.710	0.401	0.429	**0.619**	0.568	0.624	0.549	0.568	0.675	0.619
M3DM	0.532	0.608	0.663	0.712	0.658	**0.698**	0.642	0.602	0.663	0.577	0.637	0.654	0.358	0.556
Patchcore(FPFH)	0.512	0.655	0.654	**0.768**	0.524	0.459	0.447	0.453	0.663	0.597	0.687	**0.733**	0.488	0.596
Patchcore(PointMAE)	0.653	**0.677**	0.553	0.541	0.527	0.586	0.572	0.551	0.423	0.495	0.606	**0.858**	**0.886**	**0.856**
CPMF	0.435	0.458	0.521	0.657	0.745	0.486	**0.651**	0.486	0.545	0.615	0.571	0.458	0.545	0.509
Reg3D-AD	0.525	0.548	**0.886**	**0.741**	**0.775**	0.619	0.624	0.602	**0.694**	**0.698**	0.696	0.589	0.599	**0.698**
IMRNet	0.641	0.535	0.556	0.699	**0.781**	0.585	**0.682**	**0.685**	0.691	0.671	**0.702**	0.681	0.742	0.688
Ours	**0.902**	**0.897**	**0.902**	0.584	**0.775**	**0.902**	0.588	0.608	**0.762**	**0.807**	**0.867**	0.502	**0.891**	0.694

Method	**vase7**	**helmet2**	**cap5**	**shelf0**	**bowl5**	**bowl3**	**helmet1**	**bowl1**	**headset0**	**bag0**	**bowl2**	**jar**	**Mean**
BTF(FPFH)	0.540	0.643	0.586	0.619	**0.699**	0.590	**0.749**	**0.768**	0.620	**0.746**	0.518	0.427	0.628
M3DM	0.517	0.623	0.655	0.554	0.489	0.657	0.427	0.663	0.581	0.637	**0.694**	0.541	0.616
Patchcore(FPFH)	**0.693**	0.455	**0.795**	0.613	0.358	0.327	0.489	0.531	0.583	0.574	0.625	0.478	0.580
Patchcore(PointMAE)	0.651	0.651	0.545	0.543	0.562	0.581	0.562	0.524	0.575	0.674	0.515	0.487	0.577
CPMF	0.504	0.515	0.551	**0.783**	0.684	0.641	0.542	0.488	**0.699**	0.655	0.635	0.611	0.573
Reg3D-AD	**0.881**	**0.825**	0.467	**0.688**	0.691	0.654	**0.624**	0.615	0.580	0.715	0.593	0.599	**0.668**
IMRNet	0.593	0.644	0.742	0.605	**0.715**	0.599	0.604	**0.705**	**0.705**	0.668	**0.684**	**0.765**	0.650
Ours	0.576	**0.818**	**0.882**	0.625	0.562	**0.779**	0.591	0.562	0.666	**0.857**	0.597	**0.847**	**0.748**

Table 2: The pixel-level AUROC experimental results for anomaly location across 40 categories of Anomaly-ShapeNet. The best and the second-best results are highlighted in **red** and **blue**, respectively. The results of the baselines are excerpted from their papers.

Category	**Accuracy**	**O-AUROC**	**P-AUROC**
car	1.000	0.700	0.816
shell	0.881	0.820	0.792
fish	0.843	0.863	0.943
chicken	0.623	0.722	0.607
diamond	0.910	0.756	0.831
candybar	0.765	0.839	0.969
starfish	0.897	0.758	0.661
toffees	0.442	0.794	0.896
duck	0.918	0.724	0.847
seahorse	0.983	0.687	0.667
airplane	0.932	0.864	0.621
gemstone	1.000	0.475	0.385
mean	0.849	0.750	0.753

Table 3: Object classification and anomaly detection performance of our MC3D-AD on Real3D-AD.

and PointMAE [Pang *et al.*, 2022]. The experimental results are shown in Table 5 and Table 6, respectively.

The experimental results show that the performance improvement is highly related to the number of point cloud groups. For ShapeNet_part, where the group parameter is set to 128, the improvement of instance average Intersection over Union (IoU) is modest, while for ModelNet40, with a group size of 512, the accuracy is clearly improved. This is because sparse group centers contain insufficient geometric information, making it challenging for AGAM to capture the spatial structure of the point cloud.

Category	**Accuracy**	**O-AUROC**	**PAUROC**
bowl4	0.879	0.641	0.555
cup0	0.931	0.924	0.694
bucket0	0.528	0.911	0.639
bottle0	1.000	0.800	0.775
tap1	1.000	0.944	0.522
headset1	1.000	0.838	0.571
vase3	0.973	0.824	0.699
helmet3	0.838	0.976	0.624
shelf0	0.895	0.783	0.592
cap0	0.970	0.737	0.763
mean	0.901	0.838	0.643

Table 4: Object classification and anomaly detection performance of our MC3D-AD on Anomaly-ShapeNet

Method	**Accuracy**
PointNet	0.892
PointNet++	0.907
DGCNN	0.929
PointMAE	0.931
$\text{PointMAE}_{\text{AGMA}}$	0.934

Table 5: Point cloud classification performance on ModelNet40.

1.6 Inference Speed

In real-world scenarios, inference speed is very important for model deployment, so relevant experiments are conducted

Method	IoU
PointNet	0.837
PointNet++	0.851
DGCNN	0.852
PointMAE	0.860
$\text{PointMAE}_{\text{AGMA}}$	0.861

Table 6: Point cloud segmentation performance on ShapeNet-part.

and the experimental results are shown in Fig 2.

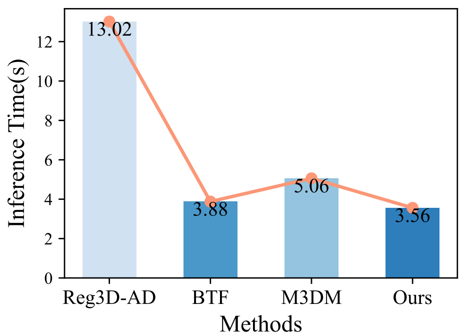

Figure 2: Average inference time (↓) per object on Real3D-AD

It is clear that our method outperforms BTF [Horwitz and Hoshen, 2023], M3DM [Wang *et al.*, 2023], and Reg3D-AD [Liu *et al.*, 2023] in terms of inference speed.Our method with one RTX 3090 achieved an average inference time of 3.560 seconds for each object in high-precision Real3D-AD, which is better than open-source methods such as BTF (3.882), M3DM (5.061), and Reg3D-AD (13.022). Although M3DM and Reg3D-AD achieve good anomaly detection performance, their inference speeds are considerably slow. BTF demonstrates good efficiency, but its anomaly detection performance needs more improvement. The proposed MC3D-AD effectively balances anomaly detection performance with inference speed, highlighting its promising potential for real-world industrial applications.

References

[Charles *et al.*, 2017] R. Qi Charles, Hao Su, Mo Kaichun, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 77–85, 2017.

[Horwitz and Hoshen, 2023] Eliahu Horwitz and Yedid Hoshen. 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 (CVPR) Workshops*, pages 2968–2977, June 2023.

[Li *et al.*, 2024] Wenqiao Li, Xiaohao Xu, Yao Gu, Bozhong Zheng, Shenghua Gaol, and Yingna Wu. Towards scalable 3d anomaly detection and localization: A benchmark via 3d anomaly synthesis and a self-supervised learning network. In *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 22207–22216, 2024.

[Liu *et al.*, 2023] Jiaqi Liu, Guoyang Xie, Ruitao Chen, Xinpeng Li, Jinbao Wang, Yong Liu, Chengjie Wang, and Feng Zheng. Real3d-ad: A dataset of point cloud anomaly detection. In *Proceedings of the Advances in Neural Information Processing Systems*, volume 36, pages 30402–30415, 2023.

[Pang *et al.*, 2022] Yatian Pang, Wenxiao Wang, Francis E. H. Tay, Wei Liu, Yonghong Tian, and Li Yuan. Masked autoencoders for point cloud self-supervised learning. In *Proceedings of the 16th European Conference Computer Vision (ECCV)*, pages 604–621, 2022.

[Qi *et al.*, 2017] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J. Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *Proceedings of the Advances in Neural Information Processing Systems*, pages 5099–5108, 2017.

[Wang *et al.*, 2019] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M. Solomon. Dynamic graph CNN for learning on point clouds. *ACM Trans. Graph.*, 38(5):146:1–146:12, 2019.

[Wang *et al.*, 2023] Yue Wang, Jinlong Peng, Jiangning Zhang, Ran Yi, Yabiao Wang, and Chengjie Wang. Multimodal industrial anomaly detection via hybrid fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8032–8041, June 2023.

[Wu *et al.*, 2015] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1912–1920, 2015.

[Yi *et al.*, 2016] Li Yi, Vladimir G. Kim, Duygu Ceylan, I-Chao Shen, Mengyan Yan, Hao Su, Cewu Lu, Qixing Huang, Alla Sheffer, and Leonidas J. Guibas. A scalable active framework for region annotation in 3d shape collections. *ACM Trans. Graph.*, 35(6):210:1–210:12, 2016.