A. Expanded Details of the Real-IAD D³ Dataset

Table A1 presents the dimensions of the materials included in the Real-IAD D³ dataset. In comparison with existing datasets, such as MVTec 3D-AD and Real3D-AD, the components in Real-IAD D³ are characterized by significantly smaller dimensions, which introduces unique challenges for anomaly detection tasks. Specifically, the materials in this dataset have lengths ranging from 7 mm to 27 mm, widths from 5 mm to 25 mm, and heights predominantly below 15 mm. These compact dimensions pose additional challenges for detecting subtle defects, as the anomalies often occupy only a small fraction of the material's surface, typically less than 3% and in some cases as small as 0.46%.

Furthermore, the materials in the dataset are sourced from real-world industrial components, including electronic devices, mechanical parts, and connectors. Examples include humidity sensors, audio jack sockets, fork crimp terminals, and ethernet connectors. This diversity in material types and geometries ensures the practical relevance of the dataset for industrial applications, reflecting real-world conditions where anomalies can vary significantly in appearance and location.

The combination of small material sizes and fine-grained defects, such as scratches, dents, and pits, considerably amplifies the difficulty of the anomaly detection task. These defects, which are often barely perceptible, demand high-resolution imaging and precise algorithms to capture the subtle variations in surface texture and geometry. The Real-IAD D³ dataset thus provides a rigorous benchmark for advancing multimodal anomaly detection in complex industrial settings.

B. Analysis of Additional Defects and Modalities in Real-IAD D³ Dataset

Figure A1 provides examples of defects and their corresponding masks for the first ten product categories. These examples demonstrate the diversity of materials and the high accuracy of defect annotations in the dataset. The displayed components, ranging from electronic connectors to mechanical parts, contain various types of surface anomalies such as scratches, dents, and cracks. The provided masks precisely delineate the defective regions, which are essential for both supervised training and objective evaluation of anomaly detection models.

Figure A2 complements the previous set by presenting additional examples of defects and masks from another ten product categories. These categories feature a broader variety of geometries and textures, making the detection task more complex. The annotations continue to exhibit a high level of precision, supporting robust training and reliable benchmarking of detection algorithms. The combination of detailed annotations and diverse materials makes this dataset an excellent benchmark for evaluating anomaly detection methods in realistic industrial scenarios.

Table A1. Visualization of additional defects and corresponding products across 2D, pseudo-3D, and 3D modalities, showcasing the complementary strengths of each modality in capturing diverse defect characteristics.

Material Name	Length (mm)	Width (mm)	Height (mm)
humidity_sensor	23	8	3
fuse_holder	27	10	7
ferrite_bead	23	10	10
lego_pin_connector_plate	15	8	3
fork_crimp_terminal	22	5	5
purple_clay_pot	20	20	8
ethernet_connector	17	13	10
miniature_lifting_motor	23	20	4
dc_power_connector	25	22	7
lego_propeller	25	25	10
limit_switch	17	8	6
headphone_jack_socket	18	9	5
audio_jack_socket	15	12	15
connector_housing-female	15	12	5
common-mode-filter	10	10	12
lattice_block_plug	16	12	15
knob_cap	7	7	5
telephone_spring_switch	23	14	10
power_jack	15	12	17
crimp_st_cable_mount_box	15	10	15

Figure A3 highlights the multimodal approach of the dataset, showing the integration of 2D images, pseudo-3D data, and 3D point clouds. The 2D images provide essential visual details such as surface texture and color variation, which are effective for identifying shallow defects like surface scratches. Pseudo-3D data captures depth variations, making it suitable for detecting surface irregularities such as dents that are difficult to perceive in standard 2D images. Finally, the 3D point clouds offer precise geometric information, which is invaluable for localizing structural defects such as cracks or deformations. Together, these modalities complement each other, providing a comprehensive framework for detecting a wide range of anomalies in industrial applications.

C. Imaging Report Analysis

Figures A4 and A5 present the imaging report generated from the experiments conducted using the proposed four-eye structured light system and its comparison with alternative imaging modalities. These reports comprehensively evaluate the system's capability in capturing surface details, resolving occlusions, and reconstructing accurate 3D models of industrial components.

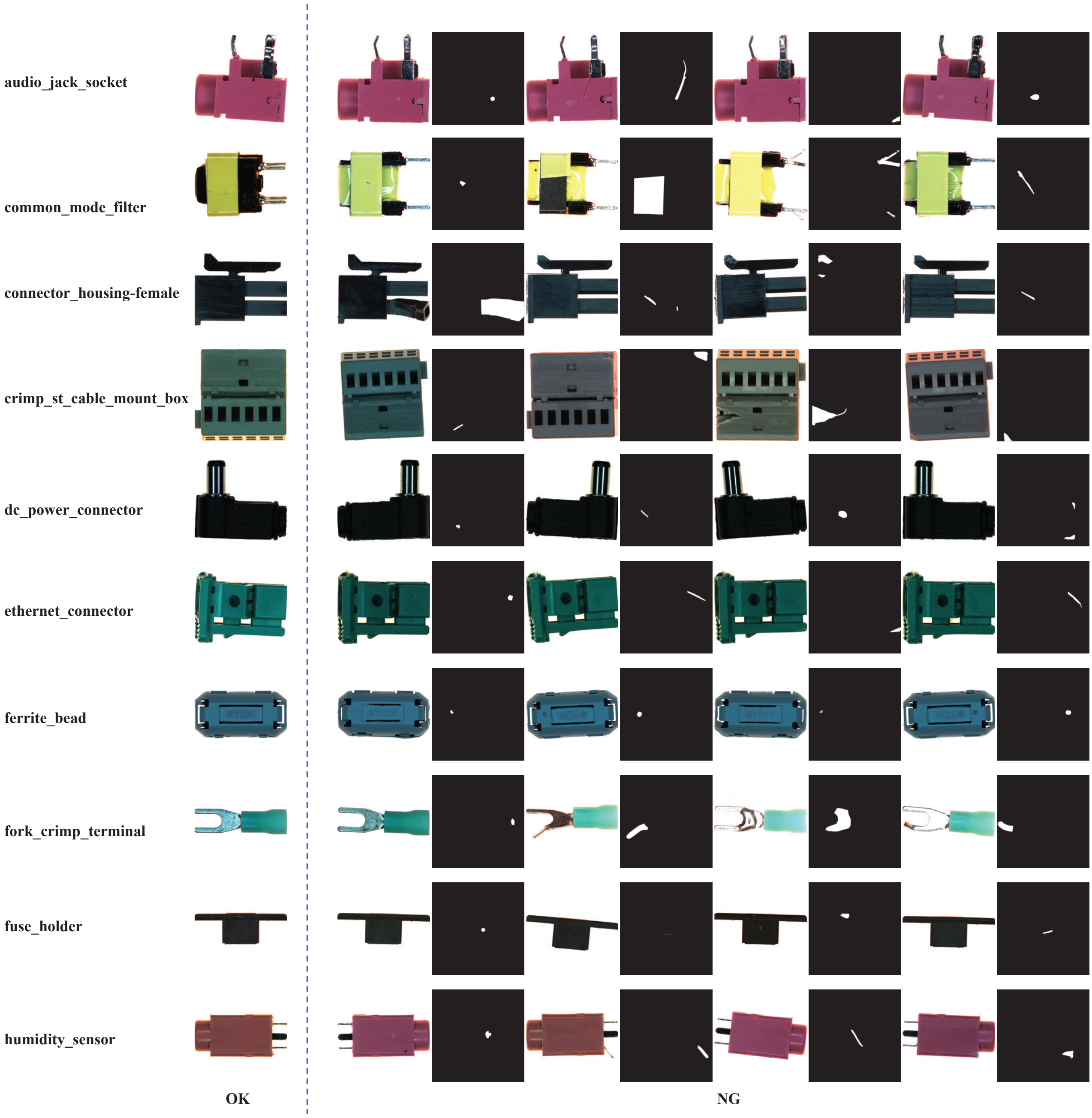


Figure A1. Visualization of additional defects and their corresponding masks for the first ten product categories in the Real-IAD D³ dataset, showcasing the dataset's diversity and precision.

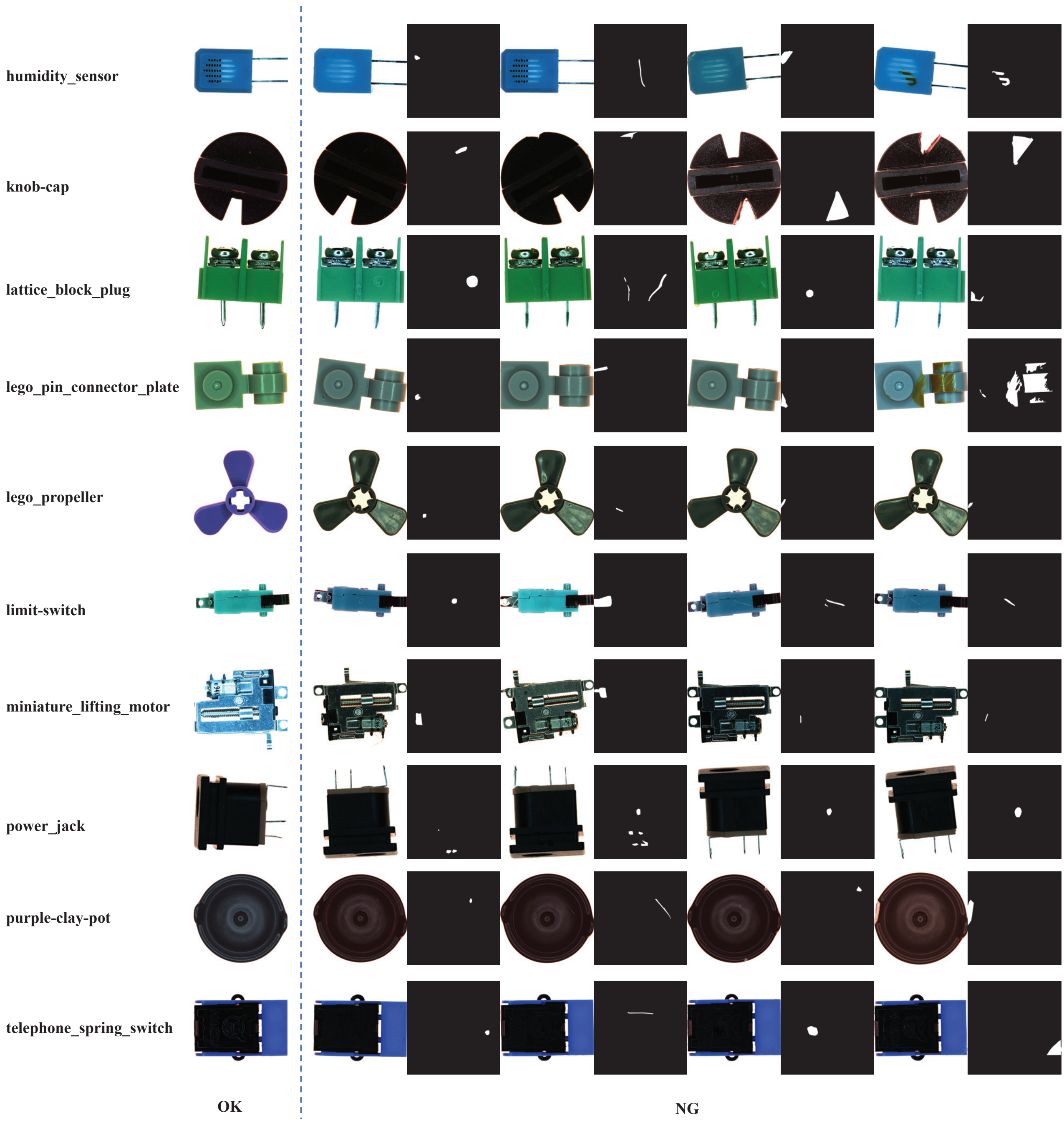


Figure A2. Visualization of additional defects and corresponding masks for the second group of ten product categories in the Real-IAD D³ dataset, further illustrating its diversity and precision.

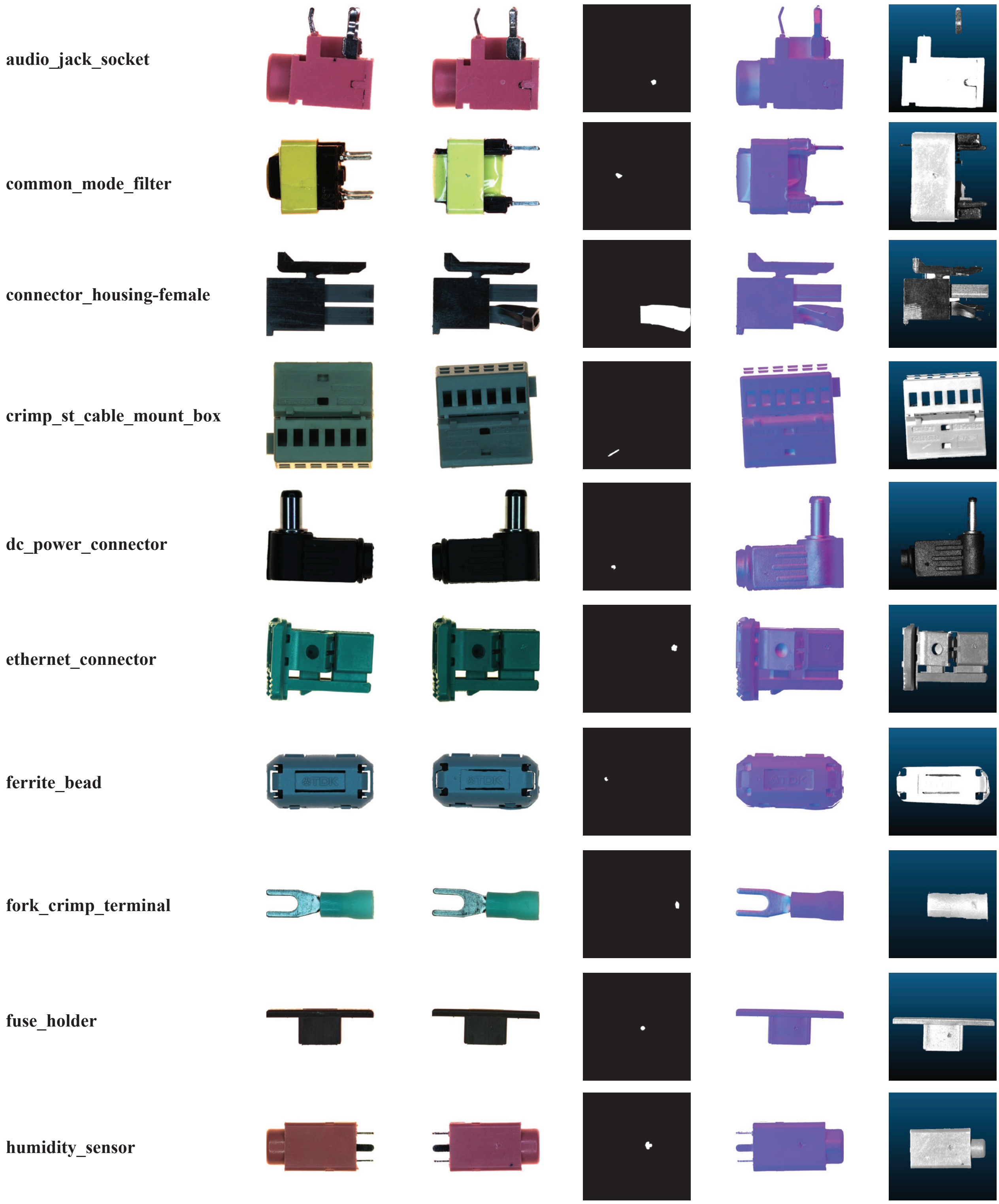
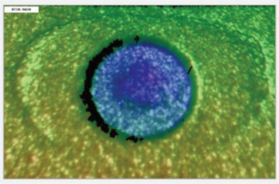
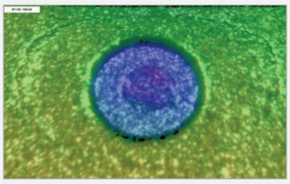
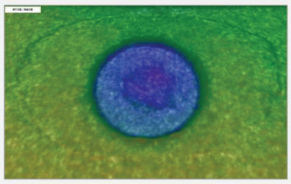
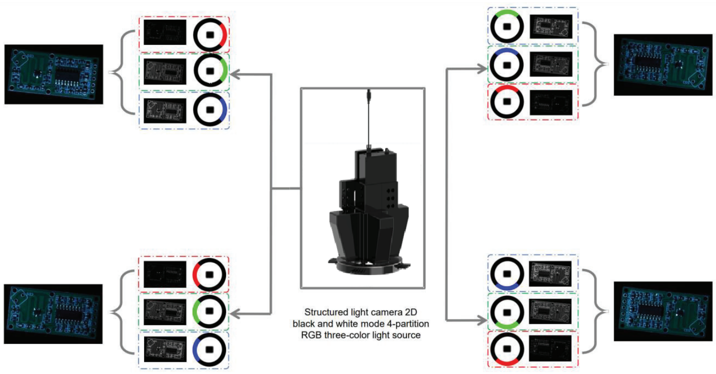


Figure A3. Visualization of additional defects and corresponding products across 2D, pseudo-3D, and 3D modalities, showcasing the complementary strengths of each modality in capturing diverse defect characteristics.

Figure A4. Imaging report generated using the four-eye structured light system, demonstrating the captured raw data and its corresponding structured light patterns. The report showcases the effectiveness of the four-view system in capturing surface details and resolving occlusions.

名称	Monocular structured light	Binocular structured light	Four-Eye Structured Light
Model number	S162060	ST162053	SQ162053
FOV（mm)	60*34.2	53*30.2	53*30.2
Point Precision	0.55	<0.3	<0.3
Spatial Distance	0.011	0.010	0.010
photometric stereo	×	×	√
Image			
Superiority statement	Single projection mode: the bottom imaging in the groove is complete, and the side wall imaging is occluded, resulting in more missing imaging. The depth information needs to be calculated based on the reference plane, and the measurement accuracy is limited. In some scenes without features or	Double projection mode: the bottom image in the groove is complete, and the side wall image is slightly occluded, resulting in missing. Binocular matching may be ambiguous due to unobvious surface texture, repeated texture or uneven illumination, which leads to inaccurate depth calculation.	Four projection mode: the bottom of the groove imaging complete, side wall imaging without dead Angle. Through the acquisition of image information from multiple views, the features and spatial information of the object can be more comprehensively captured, which makes the depth calculation more

- **Visual Imaging Effects** | Photometric stereoscopic parameter



Photometric Stereo Parameters			
	SQ162053	Resolution	5328x3040
		Quantity	1
		Angle	90°
		Working Distance	200mm
		Z-axis Accuracy	<1um
		2D Image Accuracy	0.009mm/pixel
Lens1	XF-10MDT0275-200	Angle	90°
		Quantity	1
		View	53x30mm
Butterfly Light Source	MPD-RL274	Quantity	1
		Angle	90°

Figure A5. Comparison of 3D imaging and pseudo-3D imaging results. The report highlights the differences in depth reconstruction and surface detail representation between the two modalities, illustrating the complementary strengths of pseudo-3D imaging for fine surface textures and 3D imaging for volumetric features.