Towards Dual Transparent Liquid Level Estimation in Biomedical Lab: Dataset, Methods and Practices

Xiayu Wang[1*], Ke Ma[1*], Ruiyun Zhong[1*], Xinggang Wang[2], Yi Fang[3,4], Yang Xiao[5], and Tian Xia[6†]

[1] School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, 430074, China
[2] School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, 430074, China
[3] Center for Artificial Intelligence and Robotics, New York University Abu Dhabi, Saadiyat Island, Abu Dhabi, 129188, United Arab Emirates
[4] Embodied AI and Robotics (AIR) Lab, New York University, 6 MetroTech Center, Brooklyn, New York 11201, NY, USA
[5] National Key Laboratory of Multispectral Information Intelligent Processing Technology, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, 430074, China
[6] The School of Software Engineering, Huazhong University of Science and Technology, Wuhan, 430074, China
{wangxyu,make,zhongry,xgwang,Yang_Xiao,tianxia}@hust.edu.cn yfang@nyu.edu

The supplementary material is organized as follows: Appendix 1 provides a comprehensive specification of the DTLD dataset; Appendix 2 presents a comparative analysis of the impact resulting from two different training objectives; Appendix 3 illustrate the reasons for choosing Bézier Curves; Appendix 5 evaluates the 6D pose estimation performance of the FFB6D model on the DTLD dataset; Appendix 6 showcases additional visualized contact line detection results using BCRM and TCLD modules.

1 Dataset Analysis

We present a supplementary specification of the DTLD to facilitate a comprehensive understanding of the dataset's distribution characteristics, encompassing object instances, viewpoints, lighting conditions, and liquid levels. Analyses are elucidated using a series of visualization charts. As illustrated in Fig. 8, the dataset includes a rendered image of four transparent objects: T175 Cell Flask, T75 Cell Flask, T25 Cell Flask, and G-rex Bioreactor. The dataset composition is further detailed, showing the number of images across three laboratory settings, with a pattern of incremental composition. Although all instances are represented across the dataset, certain instances may not be captured in every image due to limitations in camera coverage.

*Contributed equally as co-first authors.
†Corresponding author.

A set of parameters configures the camera viewpoint. In Fig. 9 (a), x_m, y_m, z_m represent the target object coordinate system, while x_c, y_c, z_c denote the camera coordinate system. The parameter d corresponds to the distance from the coordinate system origin. In Fig. 9 (b), the origin of the object coordinate system is aligned with that of the camera coordinate system, where θ_x, θ_y, θ_z represent the angles between the x, y, z axes, respectively. Fig. 10 (c) provides an example image showing exact parameters for each object instance.

Fig. 10 presents a statistical analysis of the dataset's relative orientations and distances between the camera and the objects. Specifically, Fig. 10 (a), (b), and (c) detail the distribution of the camera's orientation relative to the objects, considering the x, y, and z axes of the object coordinate system (as defined in Fig. 9 (b)). These subfigures also include the corresponding counts of images for each orientation category. Furthermore, Fig. 10 (d) illustrates the distribution of distances from the camera to the objects concerning the baseline measurements shown in Fig. 9 (a). This distribution spans from a minimum distance of 300mm to a maximum of 1050mm, covering a total range of 750mm.

Fig. 11 displays the distribution of image samples under seven lighting conditions, categorized and analyzed across three laboratory environments. Additionally, Fig. 12 includes two charts: the first chart enumerates the image count corresponding to each liquid level. In contrast, the second chart delineates the range of liquid levels observed in the DTLD. In total, fifty-six measurements of liquid levels are documented within the dataset. These measurements are further categorized into five ranges, each spanning 20mm, to facilitate a detailed analysis of liquid level variations.

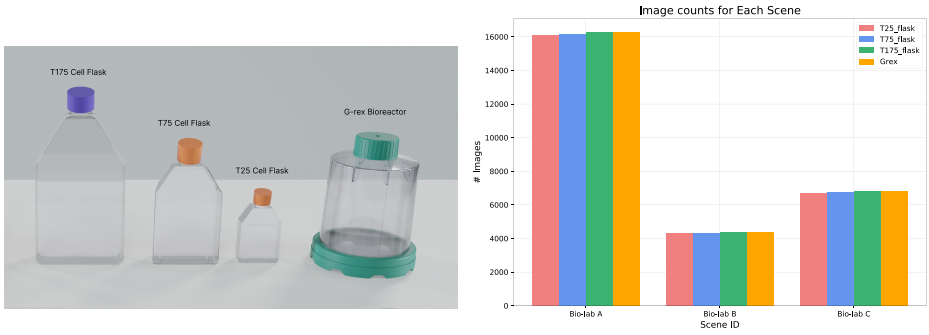

Fig. 8: Dataset Analysis. Left: a rendered image of four objects in the DTLD dataset. Right: the number of images for each object across three ABL scenes.

2 Training Objective Comparison

Section 4.1.1 introduces $\mathcal{L}_{Bezier}$ used to train the BCRM, obtaining the loss function by calculating the L1 Norm between the predicted Bézier curve control

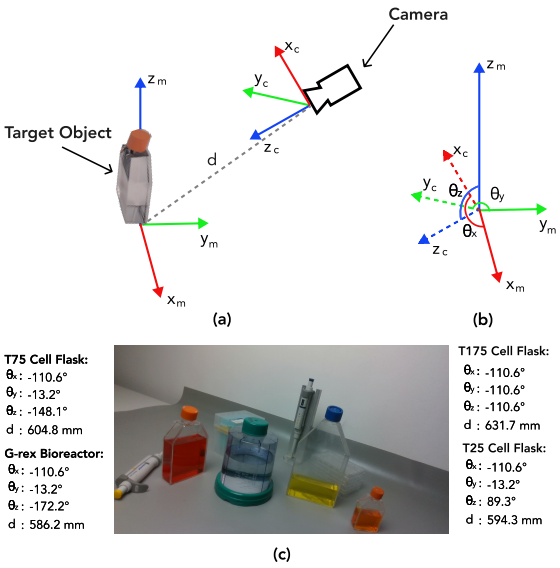

Fig. 9: The object-camera transformation diagram.

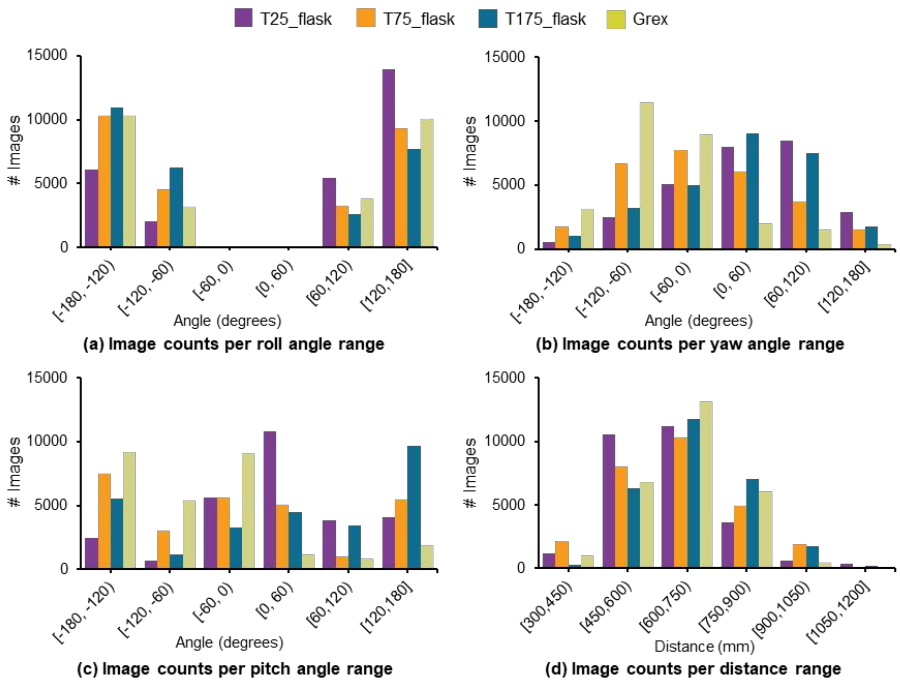

Fig. 10: Visualizing the Distribution of Camera and Object Positions in the Dataset.

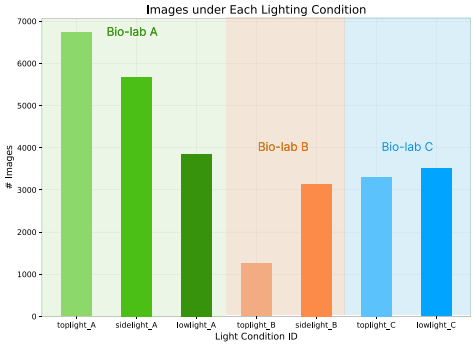

Fig. 11: Sample distribution of seven lighting conditions in three lab scenes.

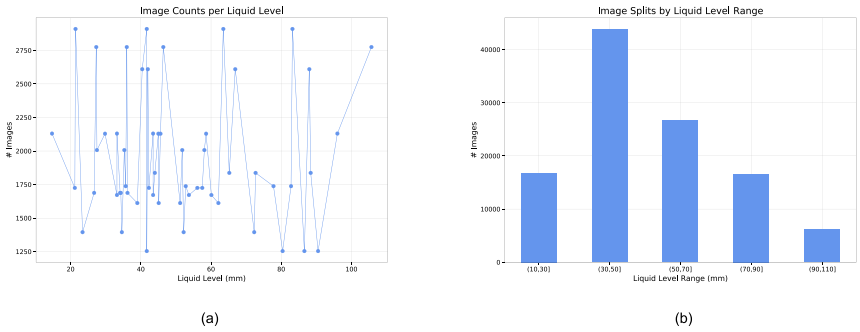

Fig. 12: Sample distribution of liquid levels, presented regarding accurate values and interval categorization.

points and the corresponding ground truth control points. To investigate the impact of two training objectives on contact line detection performance, an alternative loss function is employed to supervise our BCRM and TCLD modules.

We obtain N points by uniformly sampling the ground truth and predicted Bézier curve and calculate the L1 norm between these sampled points to compute the loss,

$$\mathcal{L}_{reg\_sp} = \frac{1}{N}\sum_{i}^{N} \|p_i - \hat{p}_i\|_1. \tag{13}$$

where $\hat{p}_i$ and p_i represent the sampled points on the ground truth and the predicted Bézier curve, respectively.

For the probability of each proposal curve, we use binary cross-entropy loss and retain other training settings. The retrained and original results are compared in Tab. 6.

We observe that differences between models trained with different objectives are marginal, and the alternative loss function only achieves a slight advantage in MAPE (less than 0.002 points). Therefore, all reported experiment results are based on the control point loss function because of its reduction capacity of training time expenditure.

Table 6: Comparative results of contact line detection using two training objectives.

Method	Average CL-IoU	Precision@0.50	Recall@0.50	F1@0.50	F1@0.75	mF1	MAPE
BCRM-GT	0.526	0.569	0.555	0.562	0.269	0.292	0.203
BCRM-GT(sampled)	0.529	0.576	0.558	0.567	0.283	0.301	0.205
TCLD-GT	0.590	0.675	0.651	0.663	0.356	0.361	0.186
TCLD-GT(sampled)	0.592	0.675	0.655	0.665	0.352	0.362	0.187

3 Reasons for choosing Bézier Curves

Contact line detection requires predicting a smooth curve while using Polynomials may result in excessive fitting, particularly with high-order Polynomials. Furthermore, Bézier curves can more accurately fit the contact line than Polynomials, as shown in Tab. 7. The smoothness and accuracy of Bézier curves became the basis for our ultimate selection.

4 Implementation Details

ResNet-18 and ResNet-34 [2] serve as the backbones for the CRM and BCRM modules. TCLD is trained using the Adam optimizer [3] with a batch size of 48 and an initial learning rate of 0.0006, coupled with a Cosine Annealing Learning Rate schedule [4]. Object instances within the ROI obtained from the 6D estimation result are resized to a resolution of 512×512 pixels before the network input.

Table 7: Here, RMSE denotes Root Mean Square Error, and AME represents Average Max Error.

model	Polynomial			Bézier		
	2-ord	3-ord	4-ord	2-ord	3-ord	4-ord
RMSE	14.833	4.724	4.229	5.280	3.587	2.762
AME	26.768	11.080	9.981	15.181	9.274	7.061

The PyTorch framework implements the training procedure, spanning 36 epochs on a single GPU. Data augmentation techniques such as random affine transformations and horizontal flips are employed further to augment the robustness and diversity of the dataset.

5 6D Pose Estimation Evaluation

We train the FFB6D model on the DTLD dataset for thirty epochs using an NVIDIA RTX 3090 GPU. Evaluating the 6D estimation performance of the FFB6D model, we compute the Area Under the Curve (AUC) percentage based on the ADD-S metrics:

$$ADD\text{-}S = \frac{1}{\mathcal{M}} \sum_{c_1 \in \mathcal{M}} \min_{c_2 \in \mathcal{M}} \left\| (Rc_1 + T) - \left(\hat{R}c_2 + \hat{T} \right) \right\|, \tag{14}$$

where, c_1, c_2 are points in the point cloud model $\mathcal{M}$ of the object, R, T are the predicted 6D pose matrix and $\hat{R}$, $\hat{T}$ are the ground truth. The ADD-S metric evaluates the 6D pose by averaging the closest point distance. ADD-S metrics for each object instance are reported in Tab. 8. Our analysis reveals that T25 flasks, characterized by their smaller volume, are more susceptible to occlusion by other objects, resulting in inferior ADD-S performance.

Table 8: The ADD-S AUC for each object reported for evaluating 6D pose estimation accuracy on the DTLD Dataset.

Object	T25 Cell Flask	T75 Cell Flask	T175 Cell Flask	G-rex bioreactor
ADD-S	56.88	79.45	80.95	90.77

6 Evaluation of detection and estimation performance

We present additional evidence to support the accuracy evaluation of contact line detection and liquid level estimation. Specifically, Fig. 13 offers visualizations

of the contact lines predicted by various modules, along with comparisons to the ground truth data. Additionally, Tab. 10 presents the estimated liquid level for each object instance within an image depicted in Fig. 13. The estimations conducted by PBLTs are also listed for comparison.

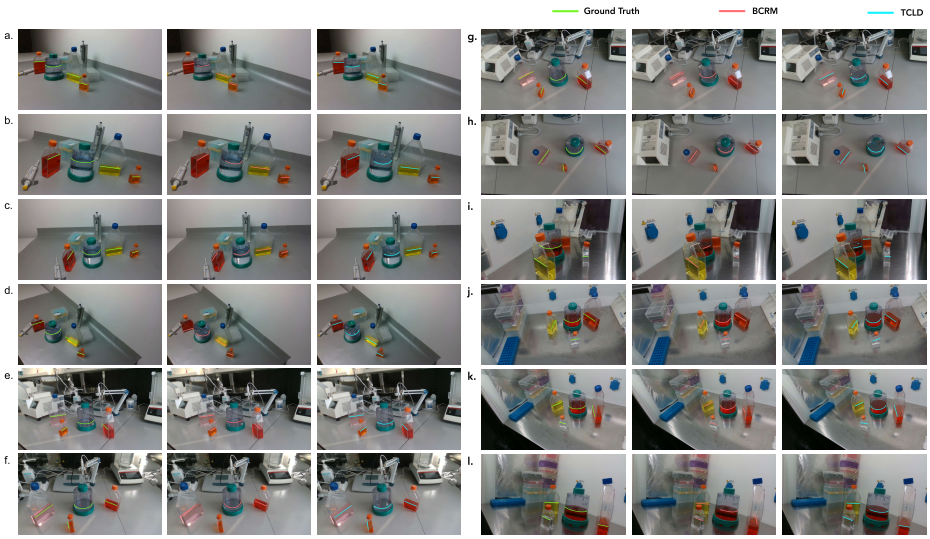

Fig. 13: Visualizations of contact line predictions.

7 Comparative Analysis with Thermal Imaging Methods

Several existing papers [7] have employed Thermal Imaging (TI) to detect liquid level positions. To compare the advantages of our proposed method over TI methods, we conducted a new experiment on TI data. We observed that, in many scenarios, thermal signatures of air and liquid in thermal images are **nearly indistinguishable**. We collected a test set of 64 samples using a TI camera (Dahua BF3241) and an RGB-D camera at **37°C** and **25°C**, shown in Fig. 14. We selected these two temperatures because of the stringent temperature protocol of biomedical experiments. **Maintaining 37°C** in incubators and gradually decreasing to a **controlled room temperature (18∼26 °C)** during short operations is one of the crucial conditions for ABL experiments [1, 6, 8].

We conducted another experiment by implementing TI method that references and adapts the mask segmentation approach based on threshold segmentation proposed by [7] as cited in the related work section to **segment the container and liquid masks**. Similar to the CMU-Liquid [5], we calculate the **height ratio** of two masks to estimate the liquid level. Evaluating TI method and our TCLD on the test set, we observed that MAPE results in Tab. 9 indicate TCLD, with an average of **0.219**, outperforms the TI method (**0.502**).

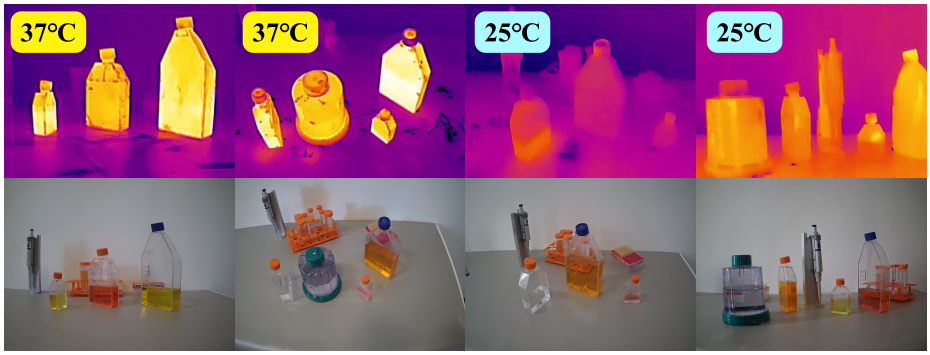

Fig. 14: Random selected TI and RGB image samples.

Table 9: MAPE result of thermal imaging approach.

Method	T25 Cell Flask	T75 Cell Flask	T175 Cell Flask	G-rex Bioreactor	Average
TI method	0.435	0.554	0.423	0.596	0.502
TCLD-FFB6D	0.132	0.119	0.173	0.453	0.219

8 DTLD's limitations

During the data collection, we aimed to reduce potential biases, ensure broad coverage of scenes and challenging factors, and balance these feature distributions. We acknowledge current limitations and propose extending DTLD based on your advice to increase coverage and diversity by including common container types that differ in structure, material, size, thickness, and four challenging scenarios depicted in Fig. 15, including 1) severe occlusion, 2) coexistence of multiple identical objects, 3) label occlusion, and 4) extreme low-light. We propose curating a large-scale realistic dataset of new objects and scenarios to reduce biases and the gap between DTLD and real ABLs.

Fig. 15: Four challenging scenarios to extend DTLD.

Table 10: Liquid levels for each object type estimated by different approaches and annotated ground-truth.

Index	T25 Cell Flask				T75 Cell Flask				T175 Cell Flask				G-rex Bioreactor			
	PBLT	BCRM	TCLD	GT	PBLT	BCRM	TCLD	GT	PBLT	BCRM	TCLD	GT	PBLT	BCRM	TCLD	GT
a	40.00	40.89	40.23	40.40	95.00	74.00	88.14	88.00	50.00	44.62	44.99	42.00	87.00	70.82	64.83	66.90
b	45.00	37.59	37.77	40.40	93.50	90.55	89.46	88.00	45.00	47.01	46.19	42.00	85.00	70.72	68.61	66.90
c	45.00	39.72	38.71	40.40	95.00	79.38	84.07	88.00	40.00	44.19	41.65	42.00	83.00	71.44	66.16	66.90
d	40.00	38.53	39.21	40.40	96.00	14.38	87.65	88.00	38.00	27.80	42.49	42.00	80.00	78.36	73.56	66.90
e	37.50	32.83	34.17	34.60	60.00	55.38	53.69	52.20	70.00	71.39	71.72	72.30	32.50	28.62	22.12	23.40
f	35.00	38.72	35.40	34.60	60.00	54.05	53.26	52.20	63.00	78.96	70.94	72.30	30.00	33.08	20.55	23.40
g	40.00	38.34	35.55	34.60	65.00	49.49	50.89	52.20	80.00	65.30	70.25	72.30	32.00	30.65	25.62	23.40
h	30.00	19.34	19.09	34.60	45.00	55.74	55.02	52.20	70.00	79.97	75.18	72.30	22.00	91.01	99.53	23.40
i	60.00	43.56	44.33	45.10	68.00	65.77	63.04	62.10	60.00	54.70	52.41	51.20	45.00	43.10	40.18	39.00
j	0.00	53.46	44.40	45.10	68.00	82.12	64.94	62.10	55.00	53.39	54.98	51.20	45.00	39.68	39.31	39.00
k	55.00	46.11	47.91	45.10	70.00	63.47	63.84	62.10	58.00	52.56	52.39	51.20	45.00	57.82	56.04	39.00
l	46.00	49.51	45.94	45.10	60.00	60.52	61.91	62.10	40.00	60.89	50.06	51.20	40.00	39.78	39.17	39.00

References

1. Han, Y., Li, X., Zhang, Y., Han, Y., Chang, F., Ding, J.: Mesenchymal stem cells for regenerative medicine. Cells **8**(8), 886 (2019)
2. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
3. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
4. Loshchilov, I., Hutter, F.: Sgdr: Stochastic gradient descent with warm restarts. arXiv preprint arXiv:1608.03983 (2016)
5. Narasimhan, G., Zhang, K., Eisner, B., Lin, X., Held, D.: Self-supervised transparent liquid segmentation for robotic pouring. In: 2022 International Conference on Robotics and Automation (ICRA). pp. 4555–4561. IEEE (2022)
6. Pitoiset, F., Barbié, M., Monneret, G., Braudeau, C., Pochard, P., Pellegrin, I., Trauet, J., Labalette, M., Klatzmann, D., Rosenzwajg, M.: A standardized flow cytometry procedure for the monitoring of regulatory t cells in clinical trials. Cytometry Part B: Clinical Cytometry **94**(5), 777–782 (2018)
7. Schenck, C., Fox, D.: Towards learning to perceive and reason about liquids. In: 2016 International Symposium on Experimental Robotics. pp. 488–501. Springer (2017)
8. Xu, S., Liu, B., Fan, J., Xue, C., Lu, Y., Li, C., Cui, D.: Engineered mesenchymal stem cell-derived exosomes with high cxcr4 levels for targeted sirna gene therapy against cancer. Nanoscale **14**(11), 4098–4113 (2022)