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

Abstract. "Dual Transparent Liquid" refers to a liquid and its container, both being transparent. Accurately estimating the levels of such a liquid from arbitrary viewpoints is fundamental and crucial, especially in AI-guided autonomous biomedical laboratories for tasks like liquid dispensing, aspiration, and mixing. However, current methods for estimating liquid level focus on scenarios of a single instance captured from a fixed view. We propose a new dual transparent liquid level estimation paradigm, including a dataset, methods, and practices. The dual transparent liquid dataset, named DTLD, comprises 27,458 images with four object instances captured from multiple views across three biomedical lab scenes. Based on DTLD, we propose an end-to-end learning method for detecting the liquid contact line, followed by an approach to estimate the liquid level. To enhance contact line detection, a color rectification module is proposed to stabilize the color distribution at the local region of the air-liquid interface. Additionally, our method surpasses the current best approach, reducing the mean absolute percentage error by a percentage of 43.4. The dataset and code are available at https://github.com/dualtransparency/TCLD.

Keywords: Liquid level estimation · Dual transparent liquid · Contact line detection · Autonomous biomedical laboratory

---

*Contributed equally as co-first authors.
†Corresponding author.
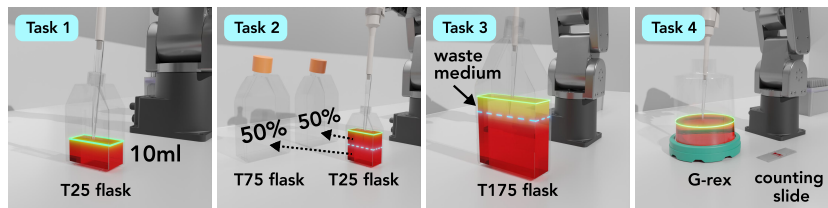
Fig. 1: Four universal liquid handling tasks in ABLs that require liquid level estimation.

1   Introduction

Recent work by Szymanski *et al.* [31], Boiko *et al.* [2], Triantafyllidis *et al.* [33], Xie *et al.* [35], and Burger *et al.* [3] showcases the evolution of biomedical laboratories towards incorporating autonomous or humanoid-like robots for laboratory AI applications. This integration of advanced computer vision and robotics heralds a new era of lab systems characterized by flexibility and efficiency. Unlike traditional automation setups with fixed coordinates for container placement, these innovative lab systems allow for the dynamic positioning of containers. In AI-guided autonomous biomedical laboratories (ABLs), liquid handling such as liquid dispensing, aspiration, and mixing emerges as a fundamental yet critical task [20, 37]. The first phase of liquid handling—accurate liquid level estimation—is vital for various laboratory procedures. We give examples of four universal biomedical liquid handling tasks conducted in ABLs in Fig. 1, including 1) adding a precise amount of cell culture medium, 2) equally dividing cell suspension, 3) aspirating waste liquid, 4) mixing and sampling for cell counting [1, 7, 14, 28]. We routinely use robotic pipettes to perform these essential tasks in our ABLs, which rely on simultaneous liquid level estimates for sequential executions. These tasks are unfeasible without a precise estimation method due to the lack of liquid level. This necessity underscores the importance of our research in tackling the challenges of liquid level estimation in ABL scenarios.

"Dual Transparent Liquid" refers to a liquid and its container, both being transparent. Accurately estimating levels of dual transparent liquid is a further challenging task in ABLs. Current research concentrates on cases of a single instance captured from a fixed camera view, which do not adequately reflect the complexity of actual laboratory conditions. For instance, a study by Gautham *et al.* [26] (named CMU-liquid) introduces a method for estimating liquid levels within a single transparent cup under one fixed camera perspective. However, real ABL environments often involve multiple objects organized in complex and flexible arrangements, rendering it difficult for a camera to capture the proper image of each object from a direct frontal viewpoint. Environmental settings differ across ABLs, including background objects, textures, and object arrangments. Current single-instance methods are ineffective in estimating liquid level for multiple cluttered and occlusive object instances in real laboratory scenes. Multiple object instances captured from arbitrary viewpoints, representing a more realistic setting, have yet to be thoroughly investigated. Our ability to
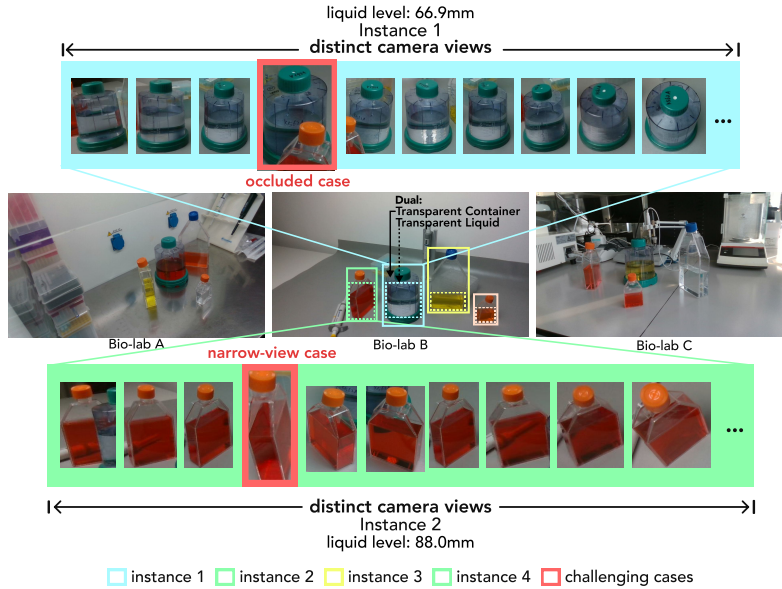
Fig. 2: Critical challenges in the "Dual Transparent Liquid" scenario across three biomedical laboratory scenes, representing realistic complexities of multi-instance arrangements captured from multiple views. "Dual" refers to the transparent liquid within its transparent container, visualized by a dashed box inside the object box. We present Region of Interest (ROI) images captured from distinct camera views for two selected object instances: a cylindrical bioreactor and a cubic cell flask. We highlight two examples of challenging cases with red boxes: occluded and narrow-view cases.

estimate liquid levels in flexibly placed containers from arbitrary viewpoints is a fundamental prerequisite for implementing laboratory AI. Therefore, we conduct research towards dual transparent liquid level estimation to bridge the gap between the demanding requirements of biomedical experiments and existing methods. Fig. 2 demonstrates the challenges of multiple instances under multiple camera views within the "Dual Transparent Liquid" context.

In response to this gap, we propose a new paradigm for dual transparent liquid level estimation including the dataset, methods, and practices. We first introduce the Dual Transparent Liquid Dataset (DTLD), which captures multiple instances in one image from multiple views across three ABL settings, comprising 27,458 images. DTLD enhances the scope beyond the current CMU-liquid dataset [26] regarding viewpoints, object instances, laboratory scenes, and the number of image samples. It also enriches complexity by incorporating seven lighting conditions, five liquid colors, and three background textures. DTLD thus establishes the benchmark dataset, characterized by its multi-instance, multiple viewpoints, and multi-factorial nature. It also poses unprecedented challenges and demands for liquid level estimation methods.

Object boundaries in images are stable features for detection tasks [16]. We use the Bézier curve to represent the boundary of the air-liquid interface (ALI), termed as the contact line. We further propose the Transparent Contact Line Detection (TCLD) method to detect contact lines in linear or curved forms within the DTLD. The parameterized Bézier curves, defined by multiple control points, provide an insightful advancement in liquid level estimation.

We observe that imaging characteristics at ALIs are influenced by varying optical refractions and reflections under different conditions. Some affected ALI cases are shown in Fig. 2. To enhance the TCLD's resilience to these variations, we present a Color Rectification Module (CRM) to stabilize the color distribution at ALI region for local image enhancement [30] and restoriont [15, 22].

Experiments with the DTLD dataset confirm TCLD's superiority in tackling the multifaceted challenges of dual transparent liquid. Combining CRM increases contact line detection precision by 10.33 points. Our liquid level estimation approach surpasses the SOTA approach, reducing the mean absolute percentage error by a percentage of 43.4. Our primary contributions are summarized as:

- We formally define and investigate liquid level estimation in the context of the dual transparent liquid, making a foundational contribution to the field.
- We create the DTLD, a comprehensive dataset encompassing 27,458 images from three ABLs, with multiple instances and captured from multiple viewpoints, offering more realistic and challenging scenarios.
- We propose the TCLD method, leveraging a CRM and Bézier curve detection module, enhancing liquid level estimation performance.

2   Related Work

2.1   Methods for Estimating Liquid Levels

Conventional methods for estimating liquid level that rely on thermal [29], optical [36], acoustic [34], and depth sensors [8] encounter challenges in accurately estimation liquid levels within dual transparent systems. Although image processing algorithms [9, 11], including edge detection, threshold segmentation, and region-based segmentation, have been employed, they exhibit reduced adaptability and robustness in dynamic or complex conditions. Deep learning studies [23, 26, 39] in this domain utilize Convolutional Neural Networks (CNNs) to segment the liquid area from an RGB image. Nonetheless, these models' ability to generalize is compromised by variables such as multiple camera perspectives and fluctuating illumination. Moreover, the output is a pixel-based liquid level, which does not translate to practical liquid measurements needed for liquid handling tasks. The conversion from a 2D-pixel representation to a 3D measurement remains elusive in multi-instance scenarios under arbitrary viewpoints.

2.2   Bézier Curve Modeling

The Bézier curve, a primary model for curve delineation, excels at portraying boundary characteristics within an image, encompassing salient features like

edges, corners, and textures. Motivated by this advantage, we integrate the Bézier curve to define liquid contact lines as linear or curved contours.

$$c(t) = \sum_{i=0}^{n} B_i^k b_i^k(t), 0 \leq t \leq 1. \tag{1}$$

k denotes the degree, B_i^k symbolizes the i-th control point, and $b_i^k(t)$ signifies the Bernstein basis polynomial where C_k^i is a binomial coefficient, noted as:

$$b_i^k(t) = C_k^i t^i \left(1 - t\right)^{k-i}, i = 0, \ldots, k. \tag{2}$$

Recent research [12, 13, 21, 25] has incorporated the Bézier curve within deep learning frameworks. ABCNet [25] first introduces the idea of amalgamating the Bézier curve, establishing the Adaptive Bézier-Curve Network equipped with a BézierAlign layer for feature extraction, aimed at detecting curved scene text. Drawing on ABCNet's innovation, Feng *et al.* [12] adapts the Bézier curve for parametrically modeling lanes in landscape imagery. Our methodology extends the application of Bézier curve modeling to the context of dual transparent liquid, addressing unique challenges of contact line detection.

Two key factors motivate our choice of Bézier curves. Firstly, 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 verified by an experiment in the Supplementary Material.

3   The DTLD Benchmark

Estimating liquid level with dual transparency presents significant complexities in practical laboratory settings due to a range of factors such as variations in camera angles, number of objects, lighting, the color of the liquids, the shape of the containers, the background textures, and the presence of distracted objects. The distinct characteristics of the contact line are often weakened and obscured by unpredictable phenomena in optical refraction, such as irregular highlights and shadows, distorted textures, and projected object frames. Notably, existing large-scale datasets for transparent objects [6, 10, 24] have not incorporated liquids within transparent containers, thereby limiting the development of liquid level estimation methods. Although two datasets of transparent liquids (CMU Liquids) have been proposed by Gautham Narasimhan *et al.* [26] for robotic pouring tasks, their scope is confined to a single container type and captured from a fixed frontal view. This dataset underestimates the challenges encountered in real ABLs, creating a disparity between settled conditions and realistic scenarios. We propose the labeled dataset DTLD, designed to reflect the multifaceted complexity experienced in real-world ABL settings. DTLD introduces several novel features: 1) multiple viewpoints, 2) multiple objects, 3) diverse lighting conditions, and 4) multi-colored liquids. Importantly, we measure liquid level as a dynamic continuous variable rather than as discrete levels. The analytical data and comparisons are listed in Tab. 1.

3.1   Data Collection

We employ high-resolution cameras to capture the necessary visual data from multiple perspectives. Four objects including three cell flasks (T25, T75, T125), and one cylinder bioreactor (G-rex) are arranged on a tabletop, each containing liquids of different colors and levels. Our data collection utilizes a RealSense D435i stereo depth camera to simultaneously capture RGB and depth data at a resolution of 1280×720 at 30 FPS. The camera is manually maneuvered around the objects, ensuring a comprehensive range of spatial trajectories encompassing various distances and angles. We collect data at three ABLs with distinct background textures and distractors shown in Fig. 2, capturing seven, four, and three sequences of images.

Each sequence combines unique object arrangement, lighting conditions, and liquid color. The object arrangement represents the three-dimensional transitions and orientations of the object, producing diverse poses of the object and the contained liquid. The lighting configuration—both in intensity and direction—affects the liquid's internal refractions and reflections, altering properties like location, shape, brightness, and specular effects of objects in images. Liquids of different colors may have different refractive indices, leading to specific imaging variations as light passes through. Such challenges are inherently associated with dual transparency and are neither discerned nor addressed in scenarios restricted to transparent objects without liquids. While the DTLD dataset currently contains four objects, it represents a pioneering effort in ABL research. We have carefully curated the dataset by selecting four representative objects to mimic complex feature variations due to optical reflections and refractions typical in ABL settings. DTLD showcases challenging cases to validate TCLD's accuracy for detecting contact lines and estimating liquid level.

3.2   Data Annotation

Every sample in DTLD is annotated with two critical pieces of information: each container's 6D pose—which encompasses the three-dimensional position and orientation—and the precise contact line within. The rationale for labeling the 6D pose is twofold: firstly, it allows for adjusting the 6D pose estimation method to our specific experimental setting. Secondly, understanding the container's pose is crucial for accurately determining the liquid level, as the container's orientation can significantly influence the perceived level of the liquid within.

Contact Line Annotation We employ the LabelMe [32] to pinpoint critical points on the contact line of cubic flasks and the cylindrical bioreactor. We map out a rectangle of ALI for the cubic flask by annotating the four vertices. The cylindrical bioreactor's elliptical ALI is marked with sixteen evenly spaced points along its contour. These annotations capture precise coordinate data that researchers can use to interpolate further points. Following established protocols, we produce a comprehensive set of ground truth annotations, including contact lines, ALI masks, and contact line segmentation masks.

Table 1: Statistics of DTLD dataset and the comparison with CMU Liquids.

Dataset	# objects	# scenes	# images	# liquid color	# lighting condition	multi -instance	multiple viewpoints	continuous level	distractors
CMU Liquids [26]	1	1	4,601	2	1	×	×	✓	×
DTLD (Ours)	4	3	27,458	5	7	✓	✓	✓	✓
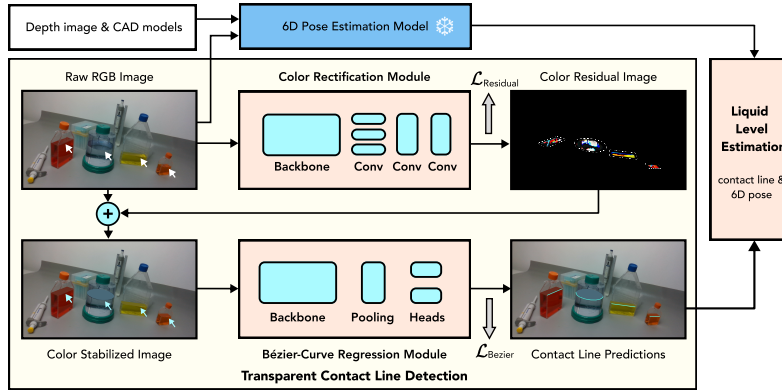
Fig. 3: Overview of the liquid level estimation framework: The process involves 6D pose estimation using RGB-D input and CAD models, along with color rectification and Bezier curve regression modules for contact line prediction, ultimately leading to liquid level estimation. $\mathcal{L}_{residual}$ and $\mathcal{L}_{Bezier}$ represent color residual loss and Bézier curve detection loss, respectively. The snowflake symbol denotes a pre-trained model. The arrows pinpoint the ALI regions, and the dashed boxes depict residual regions.

6D Pose Annotation Annotating the 6D pose of each object instance requires an efficient pipeline. We leverage ProgressLabeller [5], a tool specifically designed to annotate the 6D poses of image sequences involving transparent objects. The program integrates image sequence and CAD models of objects into the workspace. Subsequently, the Kinectfusion [27] and ORB-SLAM3 [4] are utilized for accomplishing 3D reconstruction. We calibrate re-projections to ensure the seamless alignment between each object and the imported CAD model. Finally, 6D pose annotations including 6D poses, segmentation masks, and depth samples are generated and exported in BOP [19] format.

3.3 Data Analysis

We comprehensively analyze the DTLD across various data distributions, examining object instances, viewpoints, lighting conditions, and scene types. The Supplementary Material demonstrates complete data visualization results.
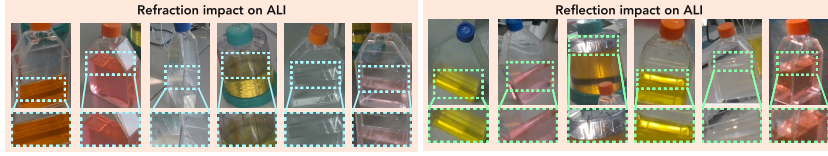
Fig. 4: Complex ALI features caused by optical refraction and reflection.
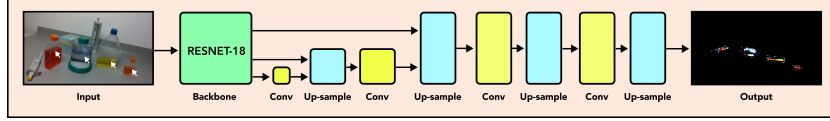

Fig. 5: The network architecture of CRM.

4 Method

4.1 Transparent Contact Line Detection (TCLD)

The TCLD method integrates CRM and BCRM to perform contact line detection efficiently. The comprehensive framework is delineated in Fig. 3.

The training phase begins with a one-to-one assignment between the predictions y and the ground truths, which are detailed in the subsequent annotations

$$\hat{y} = \left\{ \hat{B}_i^k, \hat{m}, \hat{I}_{res}, \hat{l} \right\}, \tag{3}$$

where $\hat{B}_i^k$ is the i-th control points of k-order Bézier curve, $\hat{m}$ is the liquid curve mask, $\hat{I}_{res}$ is liquid surface color residual and $\hat{l}$ is liquid height.

Bézier-Curve Regression Module (BCRM) The BCRM is introduced to process an RGB image input $I \in R^{3 \times H \times W}$, where H and W define the spatial dimensions, and directly yield the Bézier curve control points $\left\{ B_i^k \right\}_{i=1}^{k+1}$. Drawing upon the principles of ABCNet [25], the BCRM refines the architecture into a single-stage, end-to-end model specifically optimized for contact line detection tasks. We initially employ ResNet-34 [17] as backbone network to extract the feature representation $F_I \in R^{C \times H^{'} \times W^{'}}$, where C denotes the channel count, and $H^{'}$ and $W^{'}$ ($H^{'} = H/16$, $W^{'} = W/16$) represent the reduced spatial dimensions.

Given feature map F_I, the output can be donated as:

$$\left\{ \left\{ B_i^k \right\}, \ p \ \right\}_j^{W^{'}} = \left\{ MLP_B\left(F_o\right), \ MLP_p\left(F_o\right) \right\} , \tag{4}$$

where, MLP stands for Multilayer Perceptron, and $F_o \in R^{C \times H^{'} \times W^{'}}$ represent the feature obtained by downsampling F_I through an average pooling layer as Fig. 3, BCRM predict W' Bézier curve proposals $\left\{ \left\{ B_i^k \right\}_{i=1}^{k+1}, \ p \ \right\}_j^{W^{'}}$, MLP_B

and MLP_p denote the regression head and logit head for control points and the probability of each curve. The regression head generates an output of $W' \times 8$ to estimate four control points of the Bézier curve, while the logit head outputs a W'-dimensional vector denoting the probability.

We define Bézier control points regression loss $\mathcal{L}_{reg}$ as:

$$\mathcal{L}_{reg} = \frac{1}{k+1} \sum_i \left\| B_i^k - \hat{B}_i^k \right\|_2.$$ (5)

We use binary cross-entropy loss $\mathcal{L}_{logit}$ for binary classification tasks to determine the presence or absence of proposals. The Bézier curve detection loss is:

$$\mathcal{L}_{Bezier} = \lambda_1 \mathcal{L}_{reg} + \lambda_2 \mathcal{L}_{logit}.$$ (6)

Color Rectification Module (CRM) Examining the visualized outcomes revealed that a significant proportion of incorrect predictions coincides with the ALI region. It is suggested that complex ALI features, influenced by optical refractions and reflections, could adversely affect contact line detection. Several patterns of affected ALI cases are illustrated in Fig. 4. A manual review analysis confirmed that around 48.3% of test images exhibited such complications. This observation led to the formulation of a research question to mitigate the effects of variable optical phenomena on the ALI.

Residual learning is a fundamental deep learning approach widely used for image enhancement [30] and image restoration [15, 22]. Inspired by these principles, we propose a local image rectification module, named CRM, for stabilizing the color distribution within ALI to achieve image enhancement and restoration.

We conceptualize the color rectification process as combining the raw RGB image I and the color residual image I_{res}. Thus, the resultant color-stabilized image $\bar{I}$ can be donated as:

$$\bar{I}(x,y) = I(x,y) + I_{res}(x,y), x \in [0,W], y \in [0,H].$$ (7)

Given an image sample $I \in R^{3 \times H \times W}$ with H and W as spatial dimensions, the CRM leverages a ResNet-18 based backbone network [17] to extract features from three distinct layers, represented as $F_i \in R^{C_i \times H_i' \times W_i'}$ for $i \in \{1, 2, 3\}$, where C_i, H_i' and W_i' denote the channel count and spatial dimensions of the i-th layer's features, respectively. These features are then fed into a specialized decoder that upsamples F_i to $F_o \in R^{C_o \times H \times W}$. A convolutional head processes F_o and outputs the final color residual predict result $\hat{I} \in R^{3 \times H \times W}$. See Fig. 5.

We employ Mean Square Error(MSE) as residual loss $\mathcal{L}_{residual}$ to supervise CRM learning progress. We obtain $\bar{I}$ by averaging RGB values masked by ALI in the raw image I. The ground truth I_{res} is calculated as $I_{res} = \bar{I} - I$.

$$\mathcal{L}_{residual} = \frac{1}{N} \sum \left\| I_{res}(x,y) - \hat{I}_{res}(x,y) \right\|^2$$ (8)

Overall loss The total loss $\mathcal{L}$ is a weighted sum of the above losses. Specifically, following [38], we introduce line segmentation loss $\mathcal{L}_{seg}$ aims at accuracy at the pixel level within line segmentation tasks using weighted binary cross-entropy loss. The overall loss function is defined as:

$$\mathcal{L} = \lambda_1\mathcal{L}_{reg} + \lambda_2\mathcal{L}_{logit} + \lambda_3\mathcal{L}_{seg} + \lambda_4\mathcal{L}_{residual}. \tag{9}$$

Weights are preset to: $\lambda_1 = 1.0$, $\lambda_2 = 0.1$, $\lambda_3 = 0.75$ and $\lambda_4 = 1.0$. Default values are empirically chosen to ensure that each component contributes proportionally, fostering an equilibrium in model performance.

4.2   Liquid Level Estimation

Given that TCLD detects the contact line for each object instance, we need additional information to estimate the liquid level. 6D pose information can help understand the rotation and transition of the object instance. We use a state-of-the-art 6D pose estimation method, FFB6D [18] based on RGB-D images to estimate the 6D poses for all object instances. Based on the estimated translation and rotation matrix, the 2D projections of the object's point cloud model are computed within the current image. Subsequently, we resample N points from the detected contact line represented by the Bézier curve. The nearest M points included in the 2D projection points are selected and mapped to the 3D model to determine the liquid level for each sample point. We take the average of levels from N sampled points as the liquid level of the container. The FFB6D model is pre-trained on DTLD to estimate the 6D poses and outputs results to liquid level estimation module. The workflow is illustrated in Fig. 3.

5   Experiment

We conduct experiments to evaluate our model's capabilities in contact line detection and liquid level estimation, as detailed in 5.3 and 5.4, respectively, using benchmark dataset DTLD. In 5.5, we benchmark liquid level estimation results against both the SOTA method and the performance of professional technicians. Furthermore, we perform ablation studies in 5.6 to justify the selected modules' effectiveness and examine the impact of loss functions.

5.1   Dataset

We divide the dataset into training and test sets. We employ a ten-fold cross-validation approach to segregate a subset for validation from the training dataset. The test set provides an unbiased platform to assess the model's efficacy on new, unseen data. The training set consists of eleven sequences with 18,935 images, while the test set includes three sequences from different laboratory settings, amassing 5,614 images. Sequences selected for training are not included in the test set to eliminate data leakage risk.
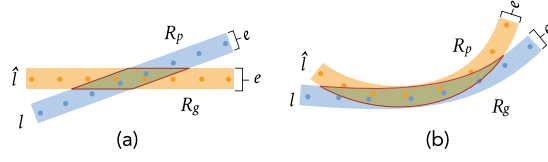
Fig. 6: Illustration of CL-IoU. (a) and (b) represent two different patterns of the Bézier curve to predict results in DTLD. While $\hat{l}$ represents ground truth points and l represents points on the predicted curve. CL-IoU can be calculated using the ratio of the intersection (area within the red boundary lines) of the region R_g and the region R_p, which are generated by extending $\hat{l}$ and l with a width of e, to their union.

5.2 Baseline and Evaluation Metrics

We compare our method with the SOTA method [26]. Contact Line Intersection over Union (CL-IoU) metric is introduced as a standard to determine the threshold range for identifying true positive samples. Utilizing the CL-IoU, we can compute standard evaluation metrics, including recall, precision, and F1 score. Points on both the predicted and actual Bézier curves are sampled to construct two line sets: $\hat{l} = \{\hat{p}_1, \ldots, \hat{p}_N\}$ and $l = \{p_1, \ldots, p_N\}$ (see Fig. 6).

By assigning a minimal, fixed width e, based on the spatial distribution of sampled points, two lines are extrapolated into regions R_g and R_p. The CL-IoU is computed as the ratio of the intersected to the united area of these regions.

$$CL\text{-}IoU = \frac{R_g \cap R_p}{R_g \cup R_p} \tag{10}$$

For validation purposes, a CL-IoU threshold of 0.5 is adopted to discern true positives, and performance is quantified using recall, precision, and the F1 score.

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{11}$$

Additionally, $mF1$ is calculated by averaging F1 from 0.5 to 0.95, separated by 0.05, $mF1 = (F1@0.50 + F1@0.55 + \cdots + F1@0.95)/10$.

The mean absolute percentage error (MAPE) is employed for liquid level estimation to quantify the relative deviation between the predicted and actual liquid levels. The MAPE is expressed as:

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{l_i - \hat{l}_i}{\hat{l}_i} \right| \tag{12}$$

5.3 Contact Line Detection Accuracy

To ascertain the contribution of the CRM to prediction performance, a comparative analysis is conducted between BCRM alone and the integrated TCLD

Table 2: Contact line detection performance on DTLD evaluated by different metrics. Comparison results between TCLD and BCRM alone are listed to present the improvement resulting from CRM.

Model	Average CL-IoU	Precision@0.50	Recall@0.50	F1@0.50	F1@0.75	mF1
BCRM-GT	0.526	0.569	0.555	0.562	0.269	0.292
TCLD-GT	**0.590**	0.675	**0.651**	**0.663**	**0.356**	**0.361**
BCRM-FFB6D	0.442	0.677	0.407	0.508	0.219	0.235
TCLD-FFB6D	0.469	**0.709**	0.445	0.547	0.279	0.312

approach. Metrics such as CL-IoU, precision, recall, F1 score, and mF1 are tabulated in Tab. 2. TCLD method exhibits a precision of 10.33 points superior to that of BCRM alone, signifying a substantial enhancement due to the CRM's ability to stabilize features in the ALI region, thereby facilitating improved contact line detection. This advancement also addresses the complexity of optical feature ambiguity caused by the dual transparency in the DTLD.

We repeat the manual review analysis to evaluate the contact lines impacted by the obscured ALI features as predicted by TCLD. The findings reveal a notable reduction in the previously reported rate of 48.34% to 21.72%, which coherently reflects the CRM's augmented performance in contact line detection. Image samples with distinct viewpoints across three scenes are selected to visualize the detected contact lines in Fig. 7. These samples demonstrate the prediction outcomes of BCRM alone on raw images and the results of TCLD on images where the ALI region has been color-stabilized by CRM.
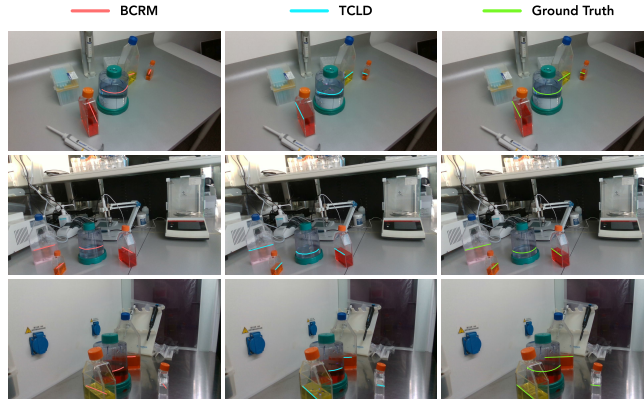
5.4   Liquid Level Estimation Accuracy

The objective of the experiments is to evaluate liquid level estimation performance. MAPE results are reported in Tab. 3. The experiments involve four module combinations: two for contact line detection (BRCM and TCLD) and two for 6D pose estimation (FFB6D and GT6D). The combination of TCLD with GT6D yields the best results, achieving the lowest average MAPE of 0.186. When GT6D is replaced by FFB6D, the average MAPE increases by 0.1, indicating that FFB6D's pose estimation accuracy affects liquid level estimation accuracy. TCLD surpasses BCRM, suggesting that CRM's image enhancement contributes to increased estimation accuracy.

5.5   Comparative studies

Two comparative studies are performed to evaluate our liquid level estimation method against the SOTA method and professional biomedical laboratory technicians (PBLTs) assessments.

First, we benchmark our method against the SOTA method [26], with both methods undergoing training and testing on identical dataset splits. The results, as depicted in Tab. 3, show that our method based on TCLD and FFB6D [18]

Fig. 7: Visualizations of contact lines predicted by BCRM (left in red curves) and TCLD (medium in blue curves) methods. Green lines represent ground truth.

Table 3: Liquid level estimation MAPE results of our methods and the SOTA CMU-Liquid [26] method on DTLD. Our methods respectively use 6D pose and object bounding box estimated by FFB6D [18] or provided by ground truth.

Method	T25 Cell Flask	T75 Cell Flask	T175 Cell Flask	G-rex Bioreactor	Average
CMU-Liquid [26]	0.611	0.212	1.232	0.808	0.720
BCRM-FFB6D	0.152	0.151	0.198	0.487	0.305
BCRM-GT6D	0.104	0.091	0.146	0.457	0.203
TCLD-FFB6D	0.145	0.137	0.193	0.473	**0.286**
TCLD-GT6D	**0.094**	**0.074**	**0.117**	**0.451**	**0.186**

outstrips the SOTA method, evidenced by a substantial 0.434 reduction in MAPE. The lowest MAPE recorded with our approach is in the context of the T75 cell flask, corresponding to an actual liquid level deviation of 5.5 mm.

Second, five PBLTs, averaging eight years of experiment experience, are provided with twenty images with distinct viewpoints from the DTLD dataset to estimate liquid levels in the transparent containers. The PBLTs encounter challenges due to the varying object angles. Tab. 4 presents the MAPE for each PBLT, with the collective average MAPE at 0.539. Human assessment outperforms CMU-liquid [26] and performs comparable capability with our method.

Table 4: Comparararive MAPE results with PBLTs.

PBLT ID	T25 Cell Flask	T75 Cell Flask	T175 Cell Flask	G-rex Bioreactor	Average
PBLT-A	0.911	0.571	0.328	0.938	0.687
PBLT-B	0.735	0.733	0.239	0.997	0.676
PBLT-C	0.334	0.420	0.473	0.552	0.445
PBLT-D	0.514	0.411	0.411	0.170	0.377
PBLT-E	0.506	0.760	0.422	0.359	0.512
PBLT-Avg	0.600	0.579	0.375	0.632	0.539
TCLD-FFB6D	0.145	0.137	0.193	0.473	0.286

Table 5: Three comparisons to testify the effectiveness of BCRM and CRM based on two TCLD models: FFB6D [26] and GT6D.

prior source	$\mathcal{L}_{Bezier}$	$\mathcal{L}_{residual}$	Precision@0.50	Recall@0.50	F1@0.50	F1@0.75	mF1	MAPE
			0.633	0.340	0.442	0.165	0.187	0.346
FFB6D	✓		0.677	0.407	0.508	0.219	0.235	0.305
	✓	✓	0.709	0.445	0.547	0.279	0.312	0.286
			0.830	0.318	0.460	0.270	0.269	0.238
GT6D	✓		0.569	0.555	0.562	0.269	0.292	0.203
	✓	✓	0.675	**0.651**	**0.663**	**0.356**	**0.361**	**0.186**

5.6 Ablation Study

We construct three comparisons based on two TCLD models (FFB6D and GT6D): 1) without using BCRM and CRM during training, 2) using only BCRM, and 3) using both BCRM and CRM to analyze the effectiveness of two modules in improving TCLD metrics. We report the performance of these counterparts in terms of F1/mF1/MAPE on the DTLD dataset. The results of the ablation study in Tab. 5 reveal the optimal model performance when leveraging both BCRM and CRM implemented by the composite loss function that includes all individual loss components. The performance diminishes when the model is trained without the supervision of $\mathcal{L}_{Bezier}$ or $\mathcal{L}_{residual}$.

6 Conclusions and Limitations

This study pioneers liquid level estimation for dual transparent scenes by introducing a dataset, theoretical framework, and practices. We outline the multifaceted challenges associated with multi-instance scenarios typical in ABLs. We establish the DTLD dataset for dual transparent objects and introduce the TCLD method, which enhances liquid level estimation by predicting Bézier curve representations of contact lines. TCLD's novelty stems from its approach to addressing these new challenges, utilizing modules that, while possibly proposed in other contexts, are uniquely applied here. Specifically, we introduce the Bézier Curve model for enhanced contact line detection accuracy and the 6D pose estimation model, transforming 2D contact lines into 3D for the precise location of the ALI. Our experimental results affirm the superior performance of TCLD in both contact line detection and liquid level estimation, outperforming SOTA and comparable to PBLTs. Nevertheless, considerable room remains to improve the model's performance in addressing optical complexity under varying illumination. We plan to extend the DTLD with new scenes, lighting conditions, objects, and materials to ensure broad coverage, balance feature distributions, and reduce potential biases.

Acknowledgments

References

1. Baudequin, T., Nyland, R., Ye, H.: Objectives, benefits and challenges of bioreactor systems for the clinical-scale expansion of t lymphocyte cells. Biotechnology Advances **49**, 107735 (2021)
2. Boiko, D.A., MacKnight, R., Kline, B., Gomes, G.: Autonomous chemical research with large language models. Nature **624**(7992), 570–578 (2023)
3. Burger, B., Maffettone, P.M., Gusev, V.V., Aitchison, C.M., Bai, Y., Wang, X., Li, X., Alston, B.M., Li, B., Clowes, R., et al.: A mobile robotic chemist. Nature **583**(7815), 237–241 (2020)
4. Campos, C., Elvira, R., Rodríguez, J.J.G., Montiel, J.M., Tardós, J.D.: Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam. IEEE Transactions on Robotics **37**(6), 1874–1890 (2021)
5. Chen, X., Zhang, H., Yu, Z., Lewis, S., Jenkins, O.C.: Progresslabeller: visual data stream annotation for training object-centric 3d perception. In: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 13066–13073. IEEE (2022)
6. Chen, X., Zhang, H., Yu, Z., Opipari, A., Chadwicke Jenkins, O.: Clearpose: Large-scale transparent object dataset and benchmark. In: European Conference on Computer Vision. pp. 381–396. Springer (2022)
7. Costariol, E., Rotondi, M.C., Amini, A., Hewitt, C.J., Nienow, A.W., Heathman, T.R., Rafiq, Q.A.: Demonstrating the manufacture of human car-t cells in an automated stirred-tank bioreactor. Biotechnology Journal **15**(9), 2000177 (2020)
8. Do, C., Burgard, W.: Accurate pouring with an autonomous robot using an rgb-d camera. In: Intelligent Autonomous Systems 15: Proceedings of the 15th International Conference IAS-15. pp. 210–221. Springer (2019)
9. Eppel, S., Kachman, T.: Computer vision-based recognition of liquid surfaces and phase boundaries in transparent vessels, with emphasis on chemistry applications. arXiv preprint arXiv:1404.7174 (2014)
10. Fang, H., Fang, H.S., Xu, S., Lu, C.: Transcg: A large-scale real-world dataset for transparent object depth completion and a grasping baseline. IEEE Robotics and Automation Letters **7**(3), 7383–7390 (2022)
11. Feng, F., Wang, L., Tan, M., Yu, Z.: Liquid surface location of transparent container based on visual analysis. In: 2017 First International Conference on Electronics Instrumentation & Information Systems (EIIS). pp. 1–4. IEEE (2017)
12. Feng, Z., Guo, S., Tan, X., Xu, K., Wang, M., Ma, L.: Rethinking efficient lane detection via curve modeling. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17062–17070 (2022)
13. Fujio, M., Nakazaki, K., Miura, N., Kaga, Y., Takahashi, K.: Finger region estimation by boundary curve modeling and bezier curve learning. In: ICPRAM. pp. 371–378 (2023)
14. Ganeeva, I., Zmievskaya, E., Valiullina, A., Kudriaeva, A., Miftakhova, R., Rybalov, A., Bulatov, E.: Recent advances in the development of bioreactors for manufacturing of adoptive cell immunotherapies. Bioengineering **9**(12), 808 (2022)
15. Guo, S., Yan, Z., Zhang, K., Zuo, W., Zhang, L.: Toward convolutional blind denoising of real photographs. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1712–1722 (2019)
16. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge university press (2003)

17. 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)
18. He, Y., Huang, H., Fan, H., Chen, Q., Sun, J.: Ffb6d: A full flow bidirectional fusion network for 6d pose estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3003–3013 (2021)
19. Hodan, T., Michel, F., Brachmann, E., Kehl, W., GlentBuch, A., Kraft, D., Drost, B., Vidal, J., Ihrke, S., Zabulis, X., et al.: Bop: Benchmark for 6d object pose estimation. In: Proceedings of the European conference on computer vision (ECCV). pp. 19–34 (2018)
20. Holland, I., Davies, J.A.: Automation in the life science research laboratory. Frontiers in Bioengineering and Biotechnology **8**, 571777 (2020)
21. Hu, L., Zhang, G.: A geometry-based deep learning feature extraction method for airfoils. In: International Conference on Computer, Artificial Intelligence, and Control Engineering (CAICE 2023). vol. 12645, pp. 691–697. SPIE (2023)
22. Kiku, D., Monno, Y., Tanaka, M., Okutomi, M.: Beyond color difference: Residual interpolation for color image demosaicking. IEEE Transactions on Image Processing **25**(3), 1288–1300 (2016)
23. Lin, H., Fu, Y., Xue, X.: Pourit!: Weakly-supervised liquid perception from a single image for visual closed-loop robotic pouring. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 241–251 (2023)
24. Liu, X., Jonschkowski, R., Angelova, A., Konolige, K.: Keypose: Multi-view 3d labeling and keypoint estimation for transparent objects. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11602–11610 (2020)
25. Liu, Y., Chen, H., Shen, C., He, T., Jin, L., Wang, L.: Abcnet: Real-time scene text spotting with adaptive bezier-curve network. In: proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 9809–9818 (2020)
26. 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)
27. Newcombe, R.A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A.J., Kohi, P., Shotton, J., Hodges, S., Fitzgibbon, A.: Kinectfusion: Real-time dense surface mapping and tracking. In: 2011 10th IEEE international symposium on mixed and augmented reality. pp. 127–136. Ieee (2011)
28. Palmerini, P., Dalla Pietà, A., Sommaggio, R., Ventura, A., Astori, G., Chieregato, K., Tisi, M.C., Visco, C., Perbellini, O., Ruggeri, M., et al.: A serum-free protocol for the ex vivo expansion of cytokine-induced killer cells using gas-permeable static culture flasks. Cytotherapy **22**(9), 511–518 (2020)
29. Schenck, C., Fox, D.: Towards learning to perceive and reason about liquids. In: 2016 International Symposium on Experimental Robotics. pp. 488–501. Springer (2017)
30. Sun, K., Meng, F., Tian, Y.: Underwater image enhancement based on noise residual and color correction aggregation network. Digital Signal Processing **129**, 103684 (2022)
31. Szymanski, N.J., Rendy, B., Fei, Y., Kumar, R.E., He, T., Milsted, D., McDermott, M.J., Gallant, M., Cubuk, E.D., Merchant, A., et al.: An autonomous laboratory for the accelerated synthesis of novel materials. Nature pp. 1–6 (2023)
32. Torralba, A., Russell, B.C., Yuen, J.: Labelme: Online image annotation and applications. Proceedings of the IEEE **98**(8), 1467–1484 (2010)

33. Triantafyllidis, E., Acero, F., Liu, Z., Li, Z.: Hybrid hierarchical learning for solving complex sequential tasks using the robotic manipulation network roman. Nature Machine Intelligence **5**(9), 991–1005 (2023)

34. Wilson, J., Sterling, A., Lin, M.C.: Analyzing liquid pouring sequences via audio-visual neural networks. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 7702–7709. IEEE (2019)

35. Xie, Y., Feng, S., Deng, L., Cai, A., Gan, L., Jiang, Z., Yang, P., Ye, G., Liu, Z., Wen, L., et al.: Inverse design of chiral functional films by a robotic ai-guided system. Nature Communications **14**(1), 6177 (2023)

36. Yamaguchi, A., Atkeson, C.G.: Stereo vision of liquid and particle flow for robot pouring. In: 2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids). pp. 1173–1180. IEEE (2016)

37. Zhang, J., Wan, W., Tanaka, N., Fujita, M., Harada, K.: Integrating a manual pipette into a collaborative robot manipulator for flexible liquid dispensing. arXiv preprint arXiv:2207.01214 (2022)

38. Zheng, T., Huang, Y., Liu, Y., Tang, W., Yang, Z., Cai, D., He, X.: Clrnet: Cross layer refinement network for lane detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 898–907 (2022)

39. Zhu, F., Hu, S., Leng, L., Bartsch, A., George, A., Farimani, A.B.: Pour me a drink: Robotic precision pouring carbonated beverages into transparent containers. arXiv preprint arXiv:2309.08892 (2023)