Data in Brief 52 (2024) 110054

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Comprehensive image dataset for enhancing object detection in chemical experiments

Ryosuke Sasaki [a], Mikito Fujinami [b], Hiromi Nakai [a,b,*]

[a] *Department of Chemistry and Biochemistry, School of Advanced Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan*
[b] *Waseda Research Institute for Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan*

A R T I C L E I N F O

Article history:
Received 15 December 2023
Revised 4 January 2024
Accepted 5 January 2024
Available online 9 January 2024

Dataset link: Annotated Chemical Apparatus Image Dataset (Original data)

Keywords:
Chemical apparatus
Image recognition
Image classification
Deep learning

A B S T R A C T

The application of image recognition in chemical experiments has the potential to enhance experiment recording and risk management. However, the current scarcity of suitable benchmarking datasets restricts the applications of machine vision techniques in chemical experiments. This data article presents an image dataset featuring common chemical apparatuses and experimenter's hands. The images have been meticulously annotated, providing detailed information for precise object detection through deep learning methods. The images were captured from videos filmed in organic chemistry laboratories. This dataset comprises a total of 5078 images including diverse backgrounds and situations surrounding the objects. Detailed annotations are provided in accompanying text files. The dataset is organized into training, validation, and test subsets. Each subset is stored within independent folders for easy access and utilization.

© 2024 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

* Corresponding author.
E-mail address: nakai@waseda.jp (H. Nakai).

https://doi.org/10.1016/j.dib.2024.110054
2352-3409/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Specifications table

Subject	Computer Vision and Pattern Recognition, Artificial Intelligence
Specific subject area	Object detection of chemical apparatuses
Data format	Images: JPG, Annotations: TXT
Type of data	Images and annotations
Data collection	Images of chemical experiments and object annotations are included. The images were captured from videos recorded in organic chemistry laboratories at Waseda University, Tokyo, Japan, using a smartphone (iPhone 13). Filming took place between November 2021 and December 2022. The extraction and annotation of images were performed using Visual Object Tagging Tool (VoTT, version 2.2.0, developed by Microsoft). The image format is JPG with a size of 1920 × 1080, 1280 × 720, or 960 × 540 pixels. The image dataset was divided into training, validation, and test subsets. The total size of the dataset folder is 4.50 GB, and for convenient downloading, a compressed file in zip file is provided.
Data source location	Institution: Waseda University City: Shinjuku, Tokyo Country: Japan Latitude and longitude: 35.706555°N,139.705096°E
Data accessibility	Repository name: Mendeley Data Data identification number: 10.17632/8p2hvgdvpn.1 Direct URL to data: https://doi.org/10.17632/8p2hvgdvpn.1

1. Value of the Data

- The annotated image dataset serves as a valuable resource for object detection algorithms employing deep learning in real organic chemistry laboratories, covering diverse scenes and offering annotations for seven distinct types of objects.
- Annotations cover six chemical apparatuses and the experimenter's hands. Special attention is given to annotations when apparatuses are held or when objects overlap, enhancing the dataset's applicability for recognizing real-world chemical experiments.
- While previous publications have focused on image datasets of experimental apparatuses [1], semantic segmentation of chemical laboratories [2], and the artificial synthesis of images with multiple experimental figures [3], this data article presents an image dataset containing the largest number of original images and object samples for chemical apparatuses.
- The dataset is organized into subsets for training, evaluation, and testing, establishing benchmark image datasets. The different characteristics in each subset, sampled from independent videos, are expected to contribute to diverse and robust evaluations.

2. Background

Automatic recognition of chemical experiments through image recognition based on deep learning holds significant potential for digitizing manual experiments, including the automation of experiment recording and the detection of hazardous activities. Essential to this automation is the identification of both experimental apparatuses and experimenters present in images. Object detection, recognizing the location and type of objects in images, and segmentation, detecting objects at the pixel level, are effective methods for automatic recognition. Previous researches of creating chemical experiment datasets are limited for object detection [1] and segmentation [2], evaluating their recognition accuracy. Another study focuses on augmenting image data of chemical apparatuses for object detection through the artificial combination of diverse images [3]. While these studies have demonstrated the potential of applying image recognition to chemical experiments, the practical implementation of automatic recognition requires additional types and quantities of image datasets for benchmarking. This study introduces a dataset comprising images from actual laboratory scenes. This dataset provides the largest number of images

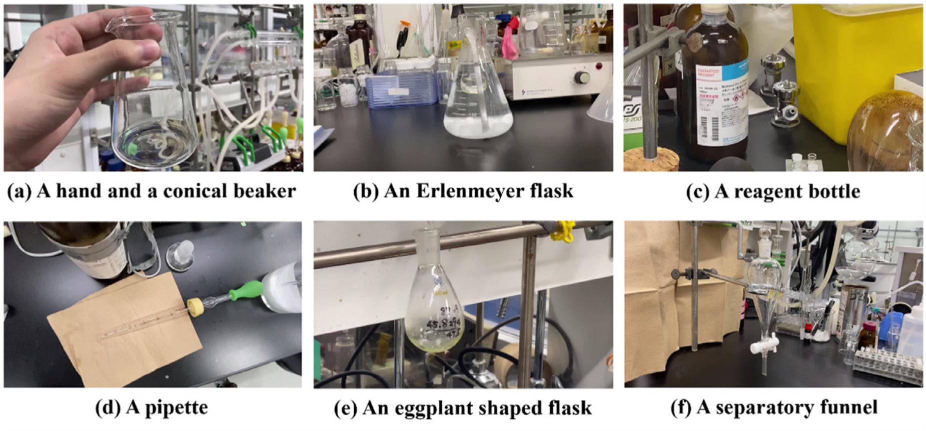

Fig. 1. Representative examples of annotated objects. Except for the hand and conical beaker in (a), each image captures a single object.

and annotated objects, featuring six distinct experimental apparatuses and the experimenter's hands.

3. Data Description

The image dataset consists of 5078 images captured within organic chemistry laboratories [4]. Each image has been annotated with coordinates and labels to facilitate object detection. Seven objects were selected for annotation, including an experimenter's hand and six common chemical apparatuses: a conical beaker, an eggplant-shaped flask, an Erlenmeyer flask, a pipette, a reagent bottle, and a separatory funnel. Fig. 1 shows representative examples of these objects. In Fig. 1(a), a hand and a conical beaker are annotated. Fig. 1(b)-(F) display an Erlenmeyer flask, a reagent bottle, a pipette, an eggplant-shaped flask, and a separatory funnel, respectively. The images were extracted from videos of chemical experiments with diverse backgrounds and situations surrounding objects. The videos were filmed using a smartphone (iPhone 13). All images have a resolution of 96 dpi, with three varying dimensions: 1920 × 1080, 1280 × 720, or 960 × 540 pixels in width and height, respectively. Manual annotation using the Visual Object Tagging Tool (VoTT, version 2.2.0, developed by Microsoft) [5] was performed. The annotations were formatted according to the You Only Look Once (YOLO) [6,7] dataset format.

Fig. 2 displays samples of the images, including multiple annotations. In Fig. 2(a), an eggplant shaped flask, a conical beaker, and an Erlenmeyer flask are featured. Fig. 2(b) shows a conical beaker, an Erlenmeyer flask, a separatory funnel, and an eggplant shaped flask. Fig. 2(c) includes two hands and an eggplant shaped flask, with partial overlap between the hand and flask. Fig.2(d) showcases an Erlenmeyer flask and a reagent bottle. Fig. 2(e) displays an Erlenmeyer flask and an eggplant shaped flask with solution. Fig. 2(f) exhibits two different types of pipettes. Objects were labeled if more than 40 % of the object was visible. Glass containers were labeled whether they contained solution or not.

The image dataset is divided into training, validation, and test subsets, each containing images extracted from independent, non-overlapping videos. Table 1 details the quantity of images and objects in the dataset. In Table 1, "Image" represents the count of images containing objects, and "Object" denotes the number of annotations. The training set comprises 4303 images with 12,041 objects (2.80 objects per image), the validation set includes 570 images with 1775 objects (3.11 objects per image), and the test set consists of 205 images with 405 objects (1.98 objects per image).

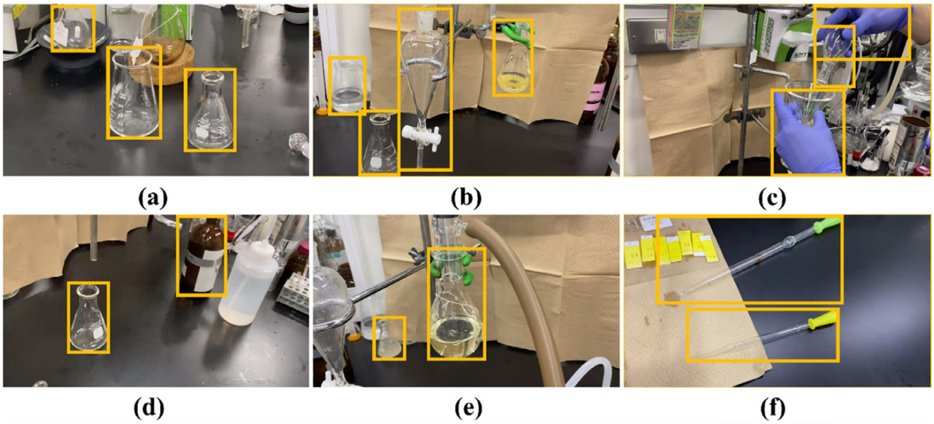

Fig. 2. Sample images from the dataset. Annotated objects are surrounded by yellow bounding boxes.

Table 1
Numbers of images and objects for seven classes in the training, validation, and test subsets.

Class	Training		Validation		Test	
	Image	Object	Image	Object	Image	Object
Hand	1321	1594	105	114	51	59
Conical beaker	993	1263	137	162	61	64
Erlenmeyer flask	1469	1880	219	351	84	101
Reagent bottle	1915	3633	300	612	32	60
Pipette	944	1206	192	270	40	44
Eggplant shaped flask	1675	2052	162	194	49	52
Separatory funnel	413	413	71	72	25	25
Total	4303	12,041	570	1775	205	405

The dataset structure, illustrated in Fig. 3, is organized into three folders: training (Train), validation (Valid), and test (Test) sets. Additionally, a "Classes.names" text file defines the list of seven objects to be detected. Each folder includes an "Images" subfolder containing the corresponding subsets of images and a "Labels" subfolder with YOLO format annotations for the images. The "Images" folders include the original images before being labeled. The "Train", "Valid", and "Test" folders contain 4303, 570, and 205 images, respectively.

The "Labels" folders contain annotation information stored in text files for each image. Objects in images are enclosed in rectangular areas. The annotations describe the object class ID and coordinates of the rectangular area in YOLO format. The class IDs are represented as integers ranging from 0 to 6, determined by the order specified in "Classes.names" (0: hand, 1: conical beaker, 2: Erlenmeyer flask, 3: reagent bottle, 4: pipette, 5: eggplant-shaped flask, 6: separatory funnel).

Eqs. (2.1) to (2.4) calculate the coordinates (x, y) and dimensions (w, h) of the rectangular area, expressed as ratios to the image size in the range of 0 to 1, as follows:

$$x = \text{absolute\_x}/\text{image\_width} \tag{2.1}$$

$$y = \text{absolute\_y}/\text{image\_height} \tag{2.2}$$

$$w = \text{absolute\_width}/\text{image\_width} \tag{2.3}$$

$$h = \text{absolute\_height}/\text{image\_height} \tag{2.4}$$

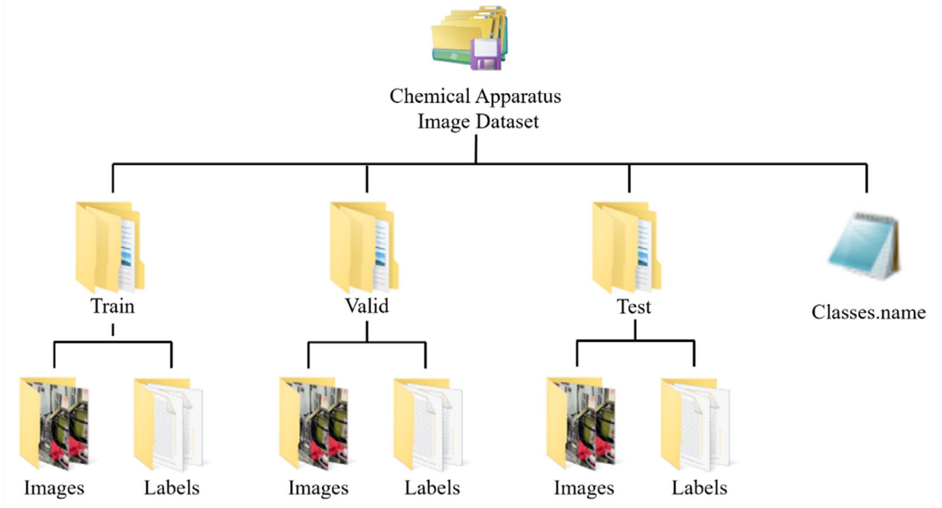

Fig. 3. Illustrative representation of the dataset structure. Images and annotations are divided into three subsets: Train, Valid, and Test sets. Class definitions are provided in the text file “Classes.names”.

Here, “absolute_x” and “absolute_y” are the coordinates of the center of the rectangular area, while “absolute_width” and “absolute_height” represent the width and height of the rectangular area, respectively. Furthermore, “image_width” and “image_height” are the width and height of the whole image, respectively.

Fig. 4(a) illustrates the terms used in the formulas. Figs. 4(b) and 4(c) provide an example of an annotated image and its corresponding annotation, respectively. In Fig. 4(b), three conical

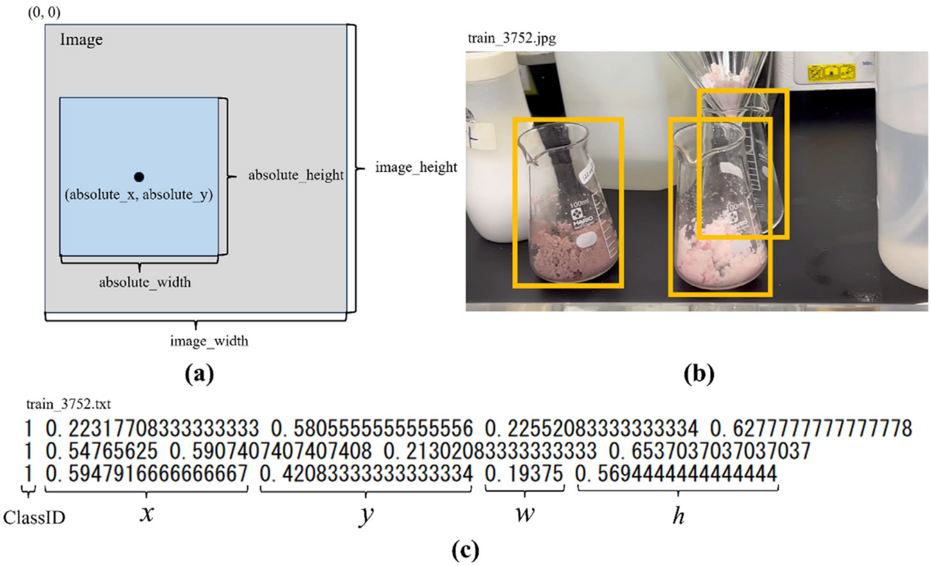

Fig. 4. (a) Schematic definition of the terms necessary for calculating coordinates and length information in YOLO format. (b) Sample image with three conical beakers annotated. (c) Example of YOLO format annotation corresponding to the sample image.

beakers are annotated. As shown in Fig. 4(c), annotations for a single object are described on a single text line, where each line includes class ID, *x*, *y*, *w*, and *h* separated by single-byte spaces.

Limitations

Not applicable.

Ethics Statement

The authors have confirmed that this study does not involve human subjects, animal experiments, or any data collected from social media platforms, and follows the ethical requirement for publication in Data in Brief.

Data Availability

Annotated Chemical Apparatus Image Dataset (Original data) (Mendeley Data).

CRediT Author Statement

Ryosuke Sasaki: Methodology, Software, Validation, Formal analysis, Investigation, Visualization, Data curation; **Mikito Fujinami:** Validation, Data curation, Writing – original draft, Funding acquisition; **Hiromi Nakai:** Conceptualization, Validation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

We acknowledge Prof. Kanomata, Prof. Shibata, and their group members for their cooperation in recording videos of chemical experiments. One of the authors (M.F.) is grateful for the financial support from The Grant-in-Aid for Transformative Research Areas (A) JP21A204 Digitalization-driven Transformative Organic Synthesis (Digi-TOS) (KAKENHI Grant Number JP22H05380) from the Ministry of Education, Culture, Sports, Science & Technology (MEXT), Japan.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

References

[1] Z. Ding, S. Ran, Z. Wu, Z. He, Q. Chen, Y. Wei, X. Wang, L. Zou, A new benchmark data set for chemical laboratory apparatus detection, in: N.H.T. Dang, YD. Zhang, J.M.R.S. Tavares, BH. Chen (Eds.), Artificial Intelligence in Data and Big Data Processing. ICABDE 2021, Lecture Notes On Data Engineering and Communications Technologies, 124, 2022, pp. 201–210, doi:10.1007/978-3-030-97610-1_17.

[2] S. Eppel, H. Xu, M. Bismuth, A Aspuru-Guzik, Computer vision for recognition of materials and vessels in chemistry lab settings and the vector-labpics data set, ACS Cent. Sci. 6 (10) (2020) 1743–1752, doi:10.1021/acscentsci.0c00460.

[3] S. Rostianingsih, A. Setiawan, C.I. Halim, COCO (creating common object in context) dataset for chemistry apparatus, Procedia Comput. Sci. 171 (2020) 2445–2452, doi:10.1016/j.procs.2020.04.264.

[4] R. Sasaki, M. Fujinami, H. Nakai, Annotated chemical apparatus image dataset, Mendeley Data (2023), doi:10.17632/8p2hvgdvpn.1.

[5] Github, https://github.com/microsoft/VoTT, 2021 (Accessed 14 December 2023).
[6] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: unified, real-time object detection, in: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 2016, pp. 779–788, doi:10.1109/CVPR.2016.91.
[7] Github, https://github.com/AlexeyAB/Yolo_mark/issues/60, 2018 (accessed 14 December 2023).