

Detection of odor-related objects in images based on everyday odors in Japan

Yuki Eda¹, Haruka Matsukura^{1,*}, Yuji Nozaki¹ and Maki Sakamoto¹

¹The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan

Abstract

This paper reports on detection of odor-related objects in images. An image data set of odor-related objects classified into 12 categories is built to train an object detection model. The results are presented to show around 60% accuracy is obtained with the trained network. The detection system of odor-related objects has potential to be applied to not only to entertainment purposes but also olfactory sensing system which may be related to human well-being.

Keywords

odor, object detection, object recognition, well-being

1. Introduction

This paper reports on detection of odor-related objects in images. Object detection has received great attention and been widely used for various purposes[1]. Generally speaking, object detection is technique for recognizing and locating objects in images based on visual features. Meanwhile, the authors have attempted to detect objects based on whether they emit smell or not. Therefore, we build an image data set of odor-related objects to train an object detection model which is built with an algorithm called YOLOv7[2] as shown in fig. 1.

The data set of odor-related images is built, referring to classification of odors[3]. The perception of odors varies by various factors including culture and environment. In this research, we focus on the odors perceived in everyday life in Japan. There is already a report in which detection of odor-related objects was attempted[4]. This research, however, addressed only four specific types of categories: aqua, coffee, orange, and rose. Our research tries to handle 12 types of categories, which are broader and vaguer.

Many of animals living on land including humans have a keen sense of odors which is sophisticated and are able to distinguish numerous odors by detecting faint chemical substances in the air. There are various researches reporting sensing system of chemical substances mimicking olfactory mechanism in animals so as to differentiate odors[5]. This kind of sensing system is called electric nose, e-nose, and is applied to wide variety of situations. Some researchers are thinking of using e-nose as a tool

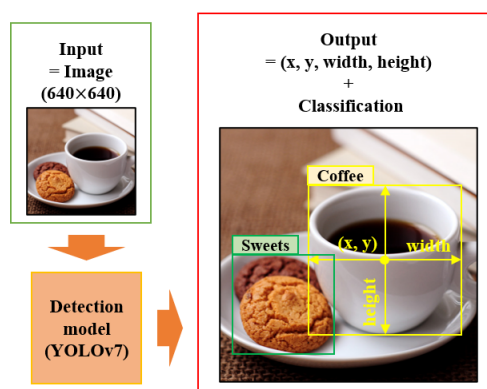


Figure 1: Schematic diagram of detection of odor-related objects.

to overcome anosmia which is known as smell blindness. However, e-nose systems developed so far have not reached a level to substitute for animals' olfaction. Our research has potential to assist the sensing system in enhancing accuracy of odor discrimination.

2. Methods

2.1. Image data set of odor-related objects

Our data set consists of 12 classes: coffee, curry, flowers, fruits, sweets, vinegar, garlic, soy sauce, butter, menthol, incense, and woods. These classes are decided based on research in which 98 everyday odors perceived in Japan are classified into multiple groups according to relative distance between each group.[3]. The word groups with a relative distance of 0.85 or more were adopted in our research. However, there are also seven groups which means unpleasant odors and chemical odors: fishy odor,

AAAI 2023 Spring Symposia, Socially Responsible AI for Well-being, March 27–29, 2023, USA

*Corresponding author.

✉ matsukura@uec.ac.jp (H. Matsukura); maki.sakamoto@uec.ac.jp (M. Sakamoto)

ORCID 0000-0002-6390-0861 (H. Matsukura)

© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License

Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

rottenness/feces, sulfur, dust, burning odor, gasoline/rubber, and thinner. These seven groups were excluded from our data set in the current stage because they may cause discomfort feeling when the odors are actually presented to the users in an application of our detection system.

The images of coffee, flowers, fruits, menthol, incense, and woods were collected from Google Images using a python library called google-images-download¹. The images of sweets were collected from ImageNet². The images of curry, vinegar, garlic, soy sause, and butter were collected from UECFOOD-256, a meal image data set that includes many representative images of Japanese dishes[6]. The collected images are annotated with bounding boxes using annotation tools. The number of collected images varies for each odor group. The minimum number is 292 for menthol and the maximum number is 1559 for soy source.

As for the images of vinegar, garlic, soy sause, and butter, we collected images of dishes including them as seasoning and spice rather than images of themselves. For example, for garlic, images of various things smelling of garlic such as Chinese dumplings and fried rice were selected. Therefore, website of Japanese cooking recipes called Delish Kitchen³ was used to find dishes including the four classes. Four to ten different dishes were extracted for each class and find images corresponding to the extracted dishes from UECFOOD-256.

2.2. Training and Evaluation

Among various object detection models, the YOLOv7 model was employed in this paper because it has a high analysis speed and enables real-time detection. Evaluation of the precidion of object detection is conducted with mean average precision (mAP) value, which is an index generally used in the research field.

3. Experiments and Discussion

The YOLOv7 model uses ELAN and E-ELAN in its basic architecture, which allows for faster processing. In addition, the model uses a label assignment strategy called auxiliary loss, which aims to improve Recall.

This detection model was trained over 300 epochs, dividing 6,356 data into 8 batch sizes and the pre-trained model used "yolov7-e6.pt"⁴. The percentage of correct answers for each class is summarised in the confusion matrix shown in fig. 2. Finally, the average of AP values (mAP) ended up at 0.66 under the condition that intersection over union (IoU) value was 0.5.

¹<https://pypi.org/project/google-images-download/>

²<https://www.image-net.org>

³<https://delishkitchen.tv>

⁴<https://github.com/WongKinYiu/yolov7/releases/>.

		True												
Predicted		Coffee	Curry	Flowers	Fruits	Sweets	Vinegar	Garlic	Soy sause	Butter	Menthol	Incense	Woods	Backgr-ound FN
	Coffee		0.82											
Curry			0.87				0.02							
Flowers				0.70										0.18
Fruits					0.02	0.80								0.01
Sweets		0.01	0.04	0.01		0.50	0.03				0.03			0.02
Vinegar			0.02			0.01	0.62	0.05	0.01					0.02
Garlic						0.01	0.09	0.40	0.24	0.05				0.12
Soy sause			0.02					0.35	0.60	0.08				0.14
Butter		0.01					0.01	0.09	0.05	0.61				0.03
Menthol		0.01									0.82	0.02		0.02
Incense												0.67	0.01	0.01
Woods		0.01											0.72	0.06
Backgr-ound FN		0.16	0.04	0.26	0.20	0.48	0.23	0.11	0.10	0.24	0.18	0.31	0.23	

Figure 2: Percentage of correct answers for each class with confusion matrix.

4. Conclusions

This paper reported on object detection of odor-related objects in image. The detection model trained with data sets of 12 classes of image including odor-related objects. The accuracy of 66% was obtained as a result. In future work, improvement of detection accuracy will be addressed by using techniques such as label Smoothing. This work was supported by JSPS KAKENHI Grant Numbers 22K12124.

References

- [1] L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, R. Qu, A survey of deep learning-based object detection, *IEEE Access* 7 (2019) 128837–128868. doi:10.1109/ACCESS.2019.2939201.
- [2] C.-Y. Wang, A. Bochkovskiy, H.-Y. M. Liao, Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, *arXiv preprint arXiv:2207.02696* (2022).
- [3] S. Saito, Expressions of offensive odors and everyday odors using words, (in japanese), *Journal of Japan Association on Odor Environment* 44 (2013) 363–379. doi:10.2171/jao.44.363.
- [4] S. Kim, J. Park, J. Bang, H. Lee, Seeing is smelling: Localizing odor-related objects in images, in: *Proceedings of the 9th Augmented Human International Conference*, 2018.
- [5] S. Lekha, S. M, Recent advancements and future prospects on e-nose sensors technology and machine learning approaches for non-invasive diabetes diagnosis: A review, *IEEE Reviews in Biomedical Engineering* 14 (2021) 127–138. doi:10.1109/RBME.2020.2993591.
- [6] Y. Kawano, K. Yanai, Automatic expansion of a food image dataset leveraging existing categories with domain adaptation, in: *Proc. of ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision (TASK-CV)*, 2014.