

# An Efficient Learning Method for Image Recognition of Dental Instruments

Shintaro OKA (✉ [oka.shintaro.dent@osaka-u.ac.jp](mailto:oka.shintaro.dent@osaka-u.ac.jp))

Osaka University Dental Hospital

Kazunori NOZAKI

Osaka University Dental Hospital

Mikako HAYASHI

Osaka University Graduate School of Dentistry

---

## Article

**Keywords:** Dentistry, Image Recognition, Deep Learning, Medical Safety

**Posted Date:** June 10th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1668911/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

**Version of Record:** A version of this preprint was published at Scientific Reports on January 4th, 2023.

See the published version at <https://doi.org/10.1038/s41598-022-26372-y>.

# Abstract

## Purpose

To prevent needlestick injury and leftover instruments, and to perform efficient dental treatment, it is important to know which and how many instruments are present during dental treatment. Therefore, we will obtain a dataset for image recognition of dental treatment instruments, develop a system for detecting dental treatment instruments during treatment by image recognition, and evaluate the performance of the system to establish a method for detecting instruments during treatment.

## Methods

We created an image recognition dataset using 23 types of instruments commonly used in the Department of Restorative Dentistry and Endodontology at Osaka University Dental Hospital and a surgeon's hands as detection targets. Two types of datasets were created: one annotated with only the characteristic parts of the instruments, and the other annotated with the entire parts of instruments. YOLOv4 was used as the image recognition system. The performance of the system was evaluated in terms of two metrics: detection accuracy (DA), which indicates the probability of correctly detecting the number of target instruments in an image, and the average precision (AP).

## Results

The mean DA and AP for the instruments used in this study were 89.3% and 70.9%, respectively, when the characteristic parts of the instruments were annotated and 85.3% and 59.9%, respectively, when the entire parts of the instruments were annotated.

## Conclusion

The detection of dental instruments can be performed efficiently by targeting the parts characterizing them.

## Introduction

As several instruments and devices are used in medical care, they should be operated correctly and safely. However, due to staff shortages and fluctuations in the number of patients, it is extremely difficult to ensure thorough management of instruments and equipment. Therefore, incidents affecting the lives of patients and the quality of medical care are likely to occur in the general operations of healthcare workers. A specific example of a life-threatening incident is leaving instruments or gauze inside a patient's body after a surgical operation [1]. A needlestick injury (NSI) is one of the most frequent medical incidents in dental treatment, and an NSI can occur not only during treatment but also during cleaning

procedures. NSIs tend to occur intraoperatively in oral and periodontal surgery, whereas, in general dentistry, NSIs tend to occur during cleanup due to various treatment methods and frequent changes during instrument preparation [2]. Thus, to prevent NSIs, not only the surgeon or assistant, but also the cleaner, must be aware of all available instruments. In dentistry, a variety of instruments are used according to the nature of the treatment. For example, in caries treatment, an excavator is used to remove caries, and a composite instrument is used to fill composite resin; in root canal treatment, an excavator is used to remove temporary seals, and a reamer is used to form root canals; in periodontal treatment, a probe is used to measure periodontal pockets, and a scaler is used to remove calculus. It is critical that these instruments are well prepared and that their presence is known before treatment for smooth and safe treatment. In other words, to prevent medical accidents and improve the efficiency of dental treatment, it is necessary to monitor changes in the number of instruments present during treatment in detail and in real time and to implement operations that use this information.

In recent years, various technologies based on deep learning (DL) [3] have been used in various fields. One of such technologies is object detection via image recognition [4–7] and semantic segmentation [8, 9]. Object detection is the estimation of an area in an image, where the target object exists, and semantic segmentation is the estimation of the object type for each pixel in an image. Both technologies use a computational method called convolutional neural networks (CNNs) [10]. DL-based image recognition has been used for human safety. For example, object detection is used in automobiles for collision prevention and automatic driving [11, 12] based on detection of vehicles, people, and road signs and used in trains for detection of passengers falling from platforms [13]. Object detection is also used in the medical field, for instance, to detect lesions using endoscope images [14] and to detect abnormal areas in X-ray images [15]. Semantic segmentation is also used to extract instruments in endoscopic images [16].

According to a study on medical instrument detection during a procedure, the accuracy of detecting instruments using shape recognition based on edge detection using the Canny filter for instruments used in otolaryngology surgery was 84.9% [17]. Simple shape recognition via contour extraction has a limitation of being unable to detect instruments because their contours change when they overlap. Therefore, although instrument recognition via contour extraction is effective in operating rooms and other environments, where operating room nurses can arrange instruments, it is unsuitable in a general dental practice, where a surgeon inserts in and removes instruments as instruments tend to overlap. In addition, training data for image recognition systems are annotated with rectangles to detect targets; however, since instruments overlap in dental treatment, there is a high possibility that another instrument may exist within a rectangle when annotating instruments with rectangles, making efficient training difficult.

Therefore, in this study, to realize real-time monitoring of instrument selection during dental treatment, we constructed a system that detects instruments placed on a tray during dental treatment using a CNN and compared the detecting accuracy of the instruments for different annotation methods.

## Methods

In this study, to clarify the difference in the detection accuracy (DA) of instruments due to different annotation methods, we annotated images taken in a laboratory and clinic through two different approaches to create two training datasets for image recognition. The two datasets were created by annotating only parts characterizing the target instruments and by annotating the entire target instruments, respectively. The image recognition system was trained using each dataset, the weights obtained were used to detect instruments in a clinic, and the results were compared and evaluated. The DL-based object detection software YOLOv4 [7] was used as the object detection method.

## 1. Two Types of Annotation Methods to Create Datasets for Training and Evaluation

In DL-based object detection, the labels and coordinates of the bounding boxes (BBs) of a target object in an image are estimated and used to train the detector. Therefore, to obtain images used for training and estimating the detector, a device was developed to capture images of a paper tray (size: 16 cm × 25 cm) on which instruments are placed during an actual dental procedure.

The device was equipped with a Raspberry Pi 3 Model B (Raspberry Pi Foundation, Cambridge, USA) and a Raspberry Pi 3 Model B (Raspberry Pi Foundation, Cambridge, USA) to capture images of the tray and its surroundings in H.264 format, 1920 pixels wide by 1080 pixels high, 25-fps frame rate, and 16.67 million color resolutions. This device can be fixed to a dental treatment table with a digital camera stand (Hakuba Photo Industry, Tokyo, Japan). In this study, images of 23 types of instruments/objects commonly used in the Department of Restorative Dentistry and Endodontology, Osaka University Dental Hospital, and a surgeon's hands were used for image recognition (Fig. 1). From August 13, 2018, to September 25, 2018, the treatment table was photographed 64 times using this device during the treatment of consenting patients at the Hospital, and 508 images without duplication were selected by eye examination. Since the number of images that can be taken in the clinic is limited to the number of images that can be taken in a clinic room during an actual examination, we used an iPhone7 (Apple, California, USA) to capture 1–3 of the 23 different instruments on the tray, obtaining 1425 images, which were augmented to create 1943 images used in this study (Table 1).

For the 1943 images used for training, the types of instruments present and BB coordinate information about the instruments were labeled. Here labeling was performed using two different annotation methods. The first annotation method (Annotation A: AA) annotates an instrument-specific part (Fig. 2). The "instrument-specific part" refers to a part characterizing the instrument, excluding parts common to other instruments, e.g., the gripping part, the mirror surface at the tip of a dental mirror, or the scale at the tip of a probe, (Fig. 1). The second annotation method (Annotation B: AB) annotates the entire instrument (Fig. 3). In this method, "condenser," "condenser\_disk," and "condenser\_round" are treated as the same label. Therefore, the number of labeling types was 22. In addition, it was difficult defining some specific parts, such as "clamp\*," "dish\*," "finger\_ruler\*," "reamer\*," "reamer\_guard\*," "hand\*," and "cotton\*," as a characteristic part of an instrument, so, in such cases, the entire apparatus was annotated using either annotation method. LabelImg [18] was used for labeling, realizing the training dataset.

Similarly, to create the evaluation dataset, 200 images without duplication were selected by eye examination from images taken during 98 examinations of consenting patients between September 26, 2018, and January 22, 2020, and these images were annotated (Table 2).

## 2. Training and Evaluation of Image Recognition System

YOLOv4 [7] was used as the image recognition system. YOLOv4 is a one-stage detector that estimates the position and label of an existing object using a single CNN network. For the parameters of the YOLOv4 neural network, the input size was changed to (832 × 832). For AA, the number of outputs was changed to 24, and for AB, the number of outputs was changed to 22.

YOLOv4 performs object detection using a predetermined size anchor box. The appropriate size of the anchor box for training and inference differs between annotating a specific part of an instrument and annotating the entire instrument because the size of the target object differs between the two cases. Therefore, we used the k-means method to calculate the appropriate anchor box size based on the size of the BB in each image [7]. This resulted in anchor boxes of {(17, 23), (21, 39), (78, 38), (76, 71), (74, 117), (118, 187), (210, 118), (228, 260), (360, 568)} for AA, {(16, 23), (26, 26), (17, 40), (104, 53), (69, 106), (138, 196), (364, 92), (381, 220), (342, 426)} for AB.

To evaluate the accuracy of the detection of the number of instruments present in a clinic using the trained image recognition system, the number of each instrument detected via image recognition was set as true if it was correct and false if otherwise, and the percentage of true recognition for each instrument was calculated as the DA. In addition, as a performance evaluation metric of the image recognition system, average precision (AP) at the intersection over union (IoU) = 50% was obtained for each instrument using the same method as the PASCAL VOC Challenge [19].

For “condenser,” “condenser\_disk,” and “condenser\_round” in AA, the results were averaged and summarized as “condenser.”

A desktop PC with Intel Xeon Gold 6226R CPU, 96 GB RAM, NVIDIA Quadro RTX6000 GPU, and Ubuntu 18.04 OS was used for training and evaluating YOLOv4.

This study was conducted following the Ethics Review Committee approval (H29-E23) of the Osaka University Graduate School of Dentistry and Dental Hospital, and was conducted in accordance with the “Ethical Guidelines for Medical and Biological Research Involving Human Subjects”. Although the data obtained in this study do not contain identifying information, the conditions of the instruments during the examination were photographed only after explaining the study to the patients and obtaining their informed consent in advance.

## Results

Table 3 shows the DA and DA change for each instrument for the two annotation methods (AA and AB), listed in descending order. The mean DA of instruments for AA was 89.3%, whereas that for AB was

85.3%, decreasing the mean DA by 4.0%. The six instruments for which changing the annotation method from AA to AB improved DA were “dish\*,” “excavator\_spoon,” “cotton\*,” “tweezers,” “reamer\_guard\*,” and “spreader,” and otherwise, DA remained the same or declined.

Table 4 shows the AP and AP change for each instrument for the two annotation methods (AA and AB), listed in descending order. The mean AP of instruments for AA was 70.9%, whereas that for AB was 59.9%, decreasing the mean AP by 11.0%. The six instruments with improved AP were “clamp\_forceps,” “tweezers,” “dish\*,” “reamer\_guard\*,” “clamp\*,” and “syringe,” and otherwise, the DA remained the same or declined.

## Discussion

### 1. Effectiveness of the Proposed Method

The accuracy of detecting instruments during otolaryngology surgery via shape recognition based on contour extraction is 84.9% [17]. The CNN used in this study extracts feature maps using a kernel for local regions to be explored from images and evaluates the similarity of feature vectors [20]. Therefore, it is more suitable for detecting deformable objects such as “clamp\_forceps” and “syringes” than shape recognition because CNN-based image recognition is robust to deformation and size changes. Although the evaluation method used in this study is different from that used in previous studies and thus cannot be uniformly evaluated, the number of instruments present was detected with a certain degree of accuracy, even when the instrumentation nurse did not align those instruments and when they overlapped. This indicates that the proposed CNN-based instrument detection method proposed is effective in detecting dental treatment instruments placed on top of each other on a narrow tray during dental treatment.

In addition, to improve the DA of dental instruments by image recognition using a dataset containing a few images, it was shown that DA and AP were improved by explicitly annotating images of the target part of an instrument, instead of annotating images of the entire instrument. This annotation method is considered reasonable and efficient because dentists also discriminate instruments by identifying the tips where the instruments perform their function. However, since it is possible to narrow down the candidates of instrument types from the grasping part that is common to multiple instruments, it cannot be assured that excluding the common parts of each instrument from the annotation will necessarily improve DA when learning instrument types from a significant number of images.

### 2. Relationship between DA and AP

Although AP is used as a general image recognition performance indicator, it has the problem that it cannot evaluate omissions of detection (false negatives) because its evaluation is based on the reliability and correctness of the results obtained via image recognition. Since the focus of this study was to correctly detect the number of instruments present during dental treatment, we defined a DA index to evaluate the correctness of the number of instruments detected.

This evaluation index tends to be suitable for instruments with a reasonably high frequency of occurrence: (e.g., “reamer\_guard,” “tweezers,” “probe,” “explorer,” and “dental\_mirror”). However, there is a problem that the number of true negatives for infrequently occurring instruments is high, resulting in high apparent values, even if there are many false negatives. For example, instruments that appeared less than 20 times in the evaluation dataset (“clamp\*,” “clamp\_forceps,” “excavator\_spoon,” “spreader,” and “plugger”) all had above-average DA of more than 90%. However, the AP of all other instruments except “clamp” was below average for each annotation method. Conversely, for instruments placed in large quantities at a time, even one false positive or negative results in false detection, reducing the DA. For example, instruments with a frequency of 600 or more occurrences (“reamer” and “cotton”) in the evaluation dataset both had AP above 90%, but their DA was below average.

### 3. Factors Affecting DA and AP

Instruments included in the basic set of dental practice in the Department of Restorative Dentistry and Endodontology are “dental\_mirror,” “explorer,” “tweezer,” and “probe.” The DA for most of these instruments was below the average of the results for each annotation method, but only “dental\_mirror,” which was annotated with instrument-specific parts, was above the average. For AP, “dental\_mirror” also showed a high value of 95.4% when annotating instrument-specific parts. This result may be due to the larger mirror surface area of “dental\_mirror” and its relative area in the BB. The one-stage detector has a tradeoff with its high detection speed, which reduces the recognition accuracy of small objects [21]. Among the instruments in the basic dental set, the recognition accuracy of “explorers,” “tweezers,” and “probes” was lower. This may be because the parts characterizing the instruments were sharper and smaller, and the parts discriminating the instruments were relatively smaller when the entire body of the instrument was annotated.

On the other hand, a common trend among all types of instruments with high AP was that they had a monotonous shape and a large size in the BB. “Hand,” “dish,” “reamer,” “reamer\_guard,” and “cotton” are typical examples, as there are no sharp edges as the shape of the instruments and the area of the instruments occupying the BB is large. However, when the entire part of an instrument is annotated, the shape of the BB becomes more complex, including the gripping part, and AP is considered to have decreased.

A common tendency of instruments with low AP was the small number of annotations required and shape complexity. For example, “excavator\_spoon,” “composite\_instrument,” “clamp\_forceps” and “excavator” all had less than 300 annotations in the training data, and their shapes were also complex. Since augmenting the training dataset generally improves the training effect on image recognition systems, it is important to increase the number of training images for instruments with complex shapes.

### 4. Effect of Annotation Method on DA and AP

For instruments where their entire parts were targeted for detection using both annotation methods (instruments with \* appended to the instrument name), there were differences in results for both DA and

AP. This may be due not only to differences in the initial anchor box size but also to differences in the number of output layers and the annotations of other instruments used as negative examples during training.

Instruments for which annotating their entire parts improved DA were “excavator\_spoon” and “tweezers,” but the change was smaller than that of “dish\*,” which was annotated in the same way. Therefore, changing the annotation method improved DA. Conversely, annotating an entire instrument (“condenser,” “explorer,” “dental\_mirror,” “probe,” etc.) resulted in a larger than average decrease in DA, probably because annotating the entire body of an instrument made it difficult to detect parts that identify the instrument, as described above.

The AP of “clamp\_forceps” and “tweezers” was more improved than that of “dish\*.” For these instruments, annotating their entire parts, rather than characteristic parts, may improve DA and allow for more efficient learning. On the other hand, for “probe,” “dental\_mirror,” and “explorer,” DA decreased by more than 10%. This is a larger change than 4.5% for “finger\_ruler\*,” which did not change its annotation method. AP of all instruments was improved, except for instruments for which the annotation method was not changed, namely, “clamp\_forceps,” “tweezers,” and “syringe.” Annotating an entire instrument decreases DA, which may reduce the learning efficiency.

## 5. Medical Applications and Future Prospects for System Development

In the medical field, instruments and gauze are frequently left behind in patients’ bodies during laparotomies. When an instrument or gauze is left behind, not only reoperation is required, but in the worst-case scenario, the patient may die. To avoid such risks, it is considered reasonable that all patients should be radiographed after surgery [1]. However, a problem exists in that patients are exposed to radiation, which is essentially unnecessary. By introducing our image recognition-based method for identifying the number of instruments present in the surgical field, the risk of leaving instruments behind in a patient’s body can be mitigated and minimally invasive medical care can be provided to patients during surgery.

In addition, by analyzing the history of instrument use and procedures automatically generated from object detection results, operational efficiency can be improved, for instance, by suggesting combinations of instruments when preparing instruments necessary for a procedure, or by suggesting when to replace instruments based on their frequency of use, thereby preventing accidents involving instrument damage. This will reduce the workload of medical personnel, which in turn will improve medical safety.

## Conclusions

In this study, we developed a real-time dental instrument detection system using CNNs. When creating the dataset used to train the CNN, it was found that by annotating only those parts of the dataset that characterize medical instruments rather than annotating the entire body of the instruments, the mean DA,

the percentage of the number of instruments correctly detected, was improved by 4%, and the mean AP was improved by 11%.

## Declarations

### Acknowledgment

This research was funded by NEC, J. MORITA CORP., J. MORITA MFG. CORP., and J. MORITA TOKYO MFG. CORP.

The authors would like to thank Enago ([www.enago.jp](http://www.enago.jp)) for the English language review.

### Data availability

The datasets used during the current study available from the corresponding author on reasonable request.

### Competing Interests

OKA Shintaro declares that he has no conflict of interest for this study.

NOZAKI Kazunori declares that he has no conflict of interest for this study

HAYASHI Mikako declares that she has no conflict of interest for this study.

### Funding

This research was funded by NEC, J.MORITA CORP., J.MORITA MFG. CORP., and J.MORITA TOKYO MFG. CORP.

## References

1. Gawande, A. A., Studdert, D. M., Orav, E. J., Brennan, T. A., Zinner, M. J. Risk factors for retained instruments and sponges after surgery. *Obstet. Gynecol. Surv.* **58**, 250–251 (2003). <https://doi.org/10.1097/01.ogx.0000058682.18516.48>
2. Lee, J. J., Kok, S. H., Cheng, S. J., Lin, L. D., Lin, C. P. Needlestick and sharps injuries among dental healthcare workers at a university hospital. *J. Formos. Med. Assoc.* **113**, 227–233 (2014). <https://doi.org/10.1016/j.jfma.2012.05.009>
3. Hinton, G. E., Osindero, S., Teh, Y. W. A fast learning algorithm for deep belief nets. *Neural Comput.* **18**, 1527–1554 (2006). <https://doi.org/10.1162/neco.2006.18.7.1527>
4. Girshick, R. Fast R-CNN. *Proc. IEEE Int. Conf. Comput. Vis. 2015*, 1440–1448 (2015). <https://doi.org/10.1109/ICCV.2015.169>

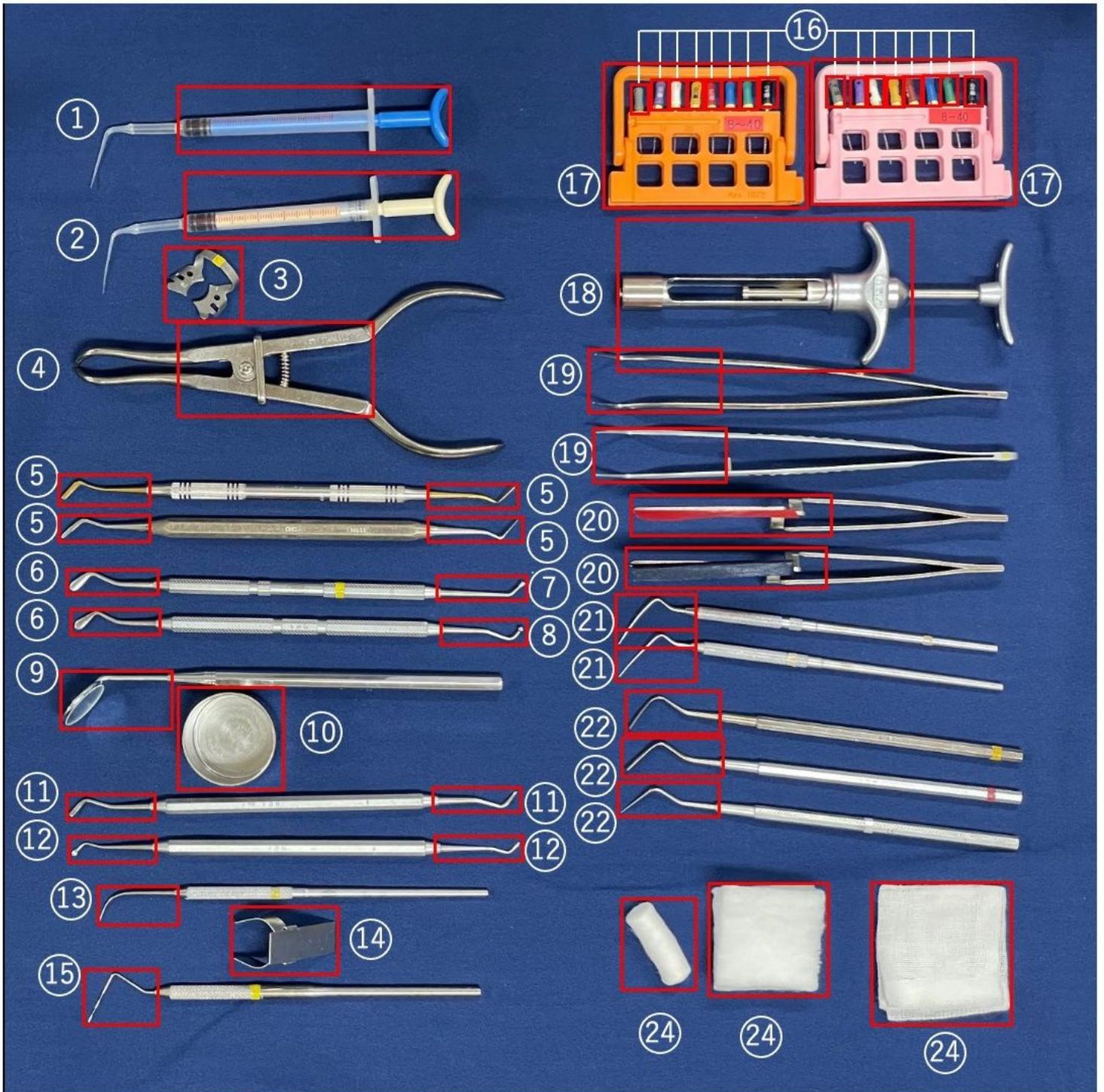
5. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., Berg, A. C. SSD: Single shot multibox detector. *Lect. Notes. Comput. Sci.* **9905**, 21–37 (2016). [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2)
6. Redmon, J., Farhadi, A. YOLOv3: An incremental improvement. Preprint at <https://arxiv.org/abs/1804.02767> (2018). Accessed 23 March 2022  
<https://doi.org/10.48550/arXiv.1804.02767>
7. Bochkovskiy, A., Wang, C. Y., Liao, H. Y. M. YOLOv4: Optimal speed and accuracy of object detection. Preprint at <http://arxiv.org/abs/2004.10934> (2020). Accessed 23 March 2022  
<https://doi.org/10.48550/arXiv.2004.10934>
8. Ronneberger, O., Fischer, P., Brox, T. U-Net: Convolutional networks for biomedical image segmentation. *MICCAI 2015*, 234–241 (2015). <https://doi.org/10.1109/ACCESS.2021.3053408>
9. He, K., Gkioxari, G., Dollár, P., Girshick, R. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 386–397 (2020). <https://doi.org/10.1109/TPAMI.2018.2844175>
10. Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* **36**, 193–202 (1980).  
<https://doi.org/10.1007/BF00344251>
11. Geiger, A., Lenz, P., Urtasun, R. Are we ready for autonomous driving? the KITTI vision benchmark suite. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 3354–3361 (2012).  
<https://doi.org/10.1109/CVPR.2012.6248074>
12. Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C., Winner, H. Three decades of driver assistance systems: Review and future perspectives. *IEEE Intell. Transp. Syst. Mag.* **6**, 6–22 (2014).  
<https://doi.org/10.1109/MITS.2014.2336271>
13. Delgado, B., Tahboub, K., Delp, E. J. Automatic detection of abnormal human events on train platforms. *NAECON 2014*, 169–173 (2014). <https://doi.org/10.1109/NAECON.2014.7045797>
14. Min, J.K., Kwak, M. S., Cha, J. M. Overview of deep learning in gastrointestinal endoscopy. *Gut. Liver* **13**, 388–393 (2019). <https://doi.org/10.5009/gnl18384>
15. McBee, M. P. *et al.* Deep learning in radiology. *Acad. Radiol.* **25**, 1472–1480 (2018).  
<https://doi.org/10.1016/j.acra.2018.02.018>
16. Ross, T. *et al.* Robust medical instrument segmentation challenge 2019. Preprint at <http://arxiv.org/abs/2003.10299> (2020) Accessed 23 March 2022.  
<https://doi.org/10.48550/arXiv.2003.10299>
17. Glaser, B., Dänzer, S., Neumuth, T. Intra-operative surgical instrument usage detection on a multi-sensor table. *Int. J. Comput. Assist. Radiol. Surg.* **10**, 351–362 (2015).  
<https://doi.org/10.1007/s11548-014-1066-0>
18. GitHub - tzutalin/labellmg: Labellmg is a graphical image annotation tool and label object bounding boxes in images. <https://github.com/tzutalin/labellmg>. Accessed 21 December 2020
19. Everingham, M., Gool, L. V., Williams, C. K. I., Winn, J., Zisserman, A. The Pascal visual object classes (VOC) challenge. *Int. J. Comput. Vis.* **88**, 303–338 (2010). <https://doi.org/10.1007/s11263-009-0275->

20. Bishop, C. M. Pattern recognition and machine learning. (Springer, 2006)
21. Pham, M., Courtrai, L., Friguet, C., Lefèvre, S., Baussard, A. YOLO-Fine: One-stage detector of small objects under various backgrounds in remote sensing images. *Remote. Sens.***12**, 2501–2516 (2020). <https://doi.org/10.3390/rs12152501>

## Tables

Tables 1 to 4 are available in the Supplementary Files section.

## Figures



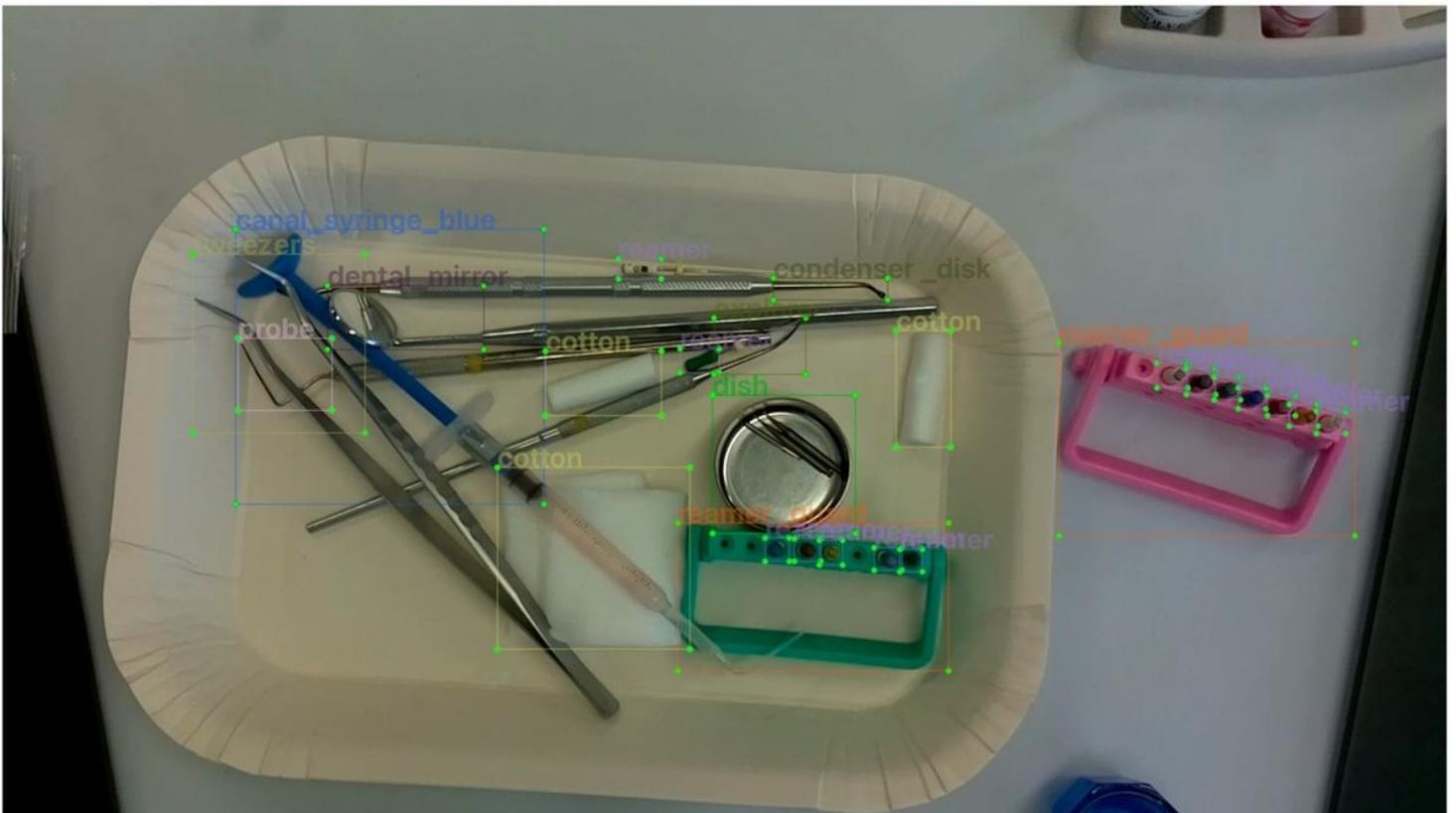
**Figure 1**

23 Types of Instruments/Objects Targeted for Image Recognition

For the “instrument-specific part,” in the case of a twin-headed instrument, each double-headed part was treated as a separate type of instrument. The red rectangles indicate the respective “instrument-specific part.” For instruments (6), (7), and (8), since different combinations constitute a single instrument, they were treated as the same instrument when labeling the entire part of the instrument.

canal\_syringe\_blue canal\_syringe\_white clamp\* clamp\_forceps  
composite\_instrument condenser condenser\_disk condenser\_round  
dental\_mirror dish\* excavator excavator\_spoon explorer finger\_ruler\*  
probe reamer\* reamer\_guard\* syringe tweezers articulatin\_paper\_holder  
spreader plugger hand\*(omitted) cotton\*

An instrument with an asterisk (\*) in its name denotes an instrument whose entire part has been annotated using the two annotation methods.



**Figure 2**

Example of Labeling of Instrument-Specific Part (AA)

Only the characteristic parts of the instrument to be detected were surrounded by BBs and labeled to record the type of instrument.

