



OPEN

## Automated recognition of objects and types of forceps in surgical images using deep learning

Yoshiko Bamba<sup>1</sup>✉, Shimpei Ogawa<sup>1</sup>, Michio Itabashi<sup>1</sup>, Shingo Kameoka<sup>2</sup>, Takahiro Okamoto<sup>3</sup> & Masakazu Yamamoto<sup>1</sup>

Analysis of operative data with convolutional neural networks (CNNs) is expected to improve the knowledge and professional skills of surgeons. Identification of objects in videos recorded during surgery can be used for surgical skill assessment and surgical navigation. The objectives of this study were to recognize objects and types of forceps in surgical videos acquired during colorectal surgeries and evaluate detection accuracy. Images (n = 1818) were extracted from 11 surgical videos for model training, and another 500 images were extracted from 6 additional videos for validation. The following 5 types of forceps were selected for annotation: ultrasonic scalpel, grasping, clip, angled (Maryland and right-angled), and spatula. IBM Visual Insights software was used, which incorporates the most popular open-source deep-learning CNN frameworks. In total, 1039/1062 (97.8%) forceps were correctly identified among 500 test images. Calculated recall and precision values were as follows: grasping forceps, 98.1% and 98.0%; ultrasonic scalpel, 99.4% and 93.9%; clip forceps, 96.2% and 92.7%; angled forceps, 94.9% and 100%; and spatula forceps, 98.1% and 94.5%, respectively. Forceps recognition can be achieved with high accuracy using deep-learning models, providing the opportunity to evaluate how forceps are used in various operations.

Recently, artificial Intelligence (AI) has been extensively utilized in many fields<sup>1</sup> and has contributed tremendously to improvements and advancements of technology. In this context, development using deep-learning technology<sup>2,3</sup> has shared in the contribution. Deep learning is based on computer programs that automatically conduct repetitive learning from provided data and identify appropriate rules based on this process<sup>4,5</sup>. In the medical field, convolutional neural networks (CNNs)<sup>6,7</sup> have also been extensively used in recent years not only for saving and archiving endoscopic surgical videos but also for analyzing the data from operations. The object recognition model used in this study has been commonly used to diagnose retinal diseases<sup>8,9</sup>, skin cancer<sup>10-13</sup>, colorectal neoplasms in endoscopy<sup>14,15</sup>, and arrhythmia in electrocardiography<sup>16-18</sup>. This research is expected to improve surgeons' knowledge and professional skills<sup>19</sup>.

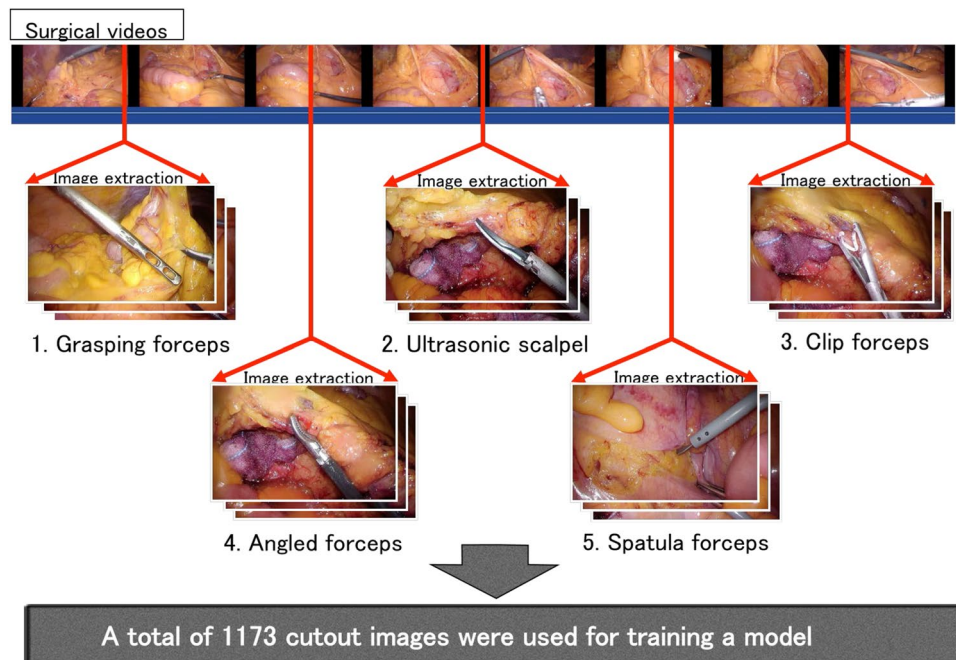
By analyzing preoperative images and intraoperative procedures and returning useful information to the surgeon during an operation, optimal surgery for patients that avoids risk through surgical navigation is the ultimate ideal. As a first step in the analysis of surgical procedures, an object recognition model is required to identify objects in surgical videos that require surgical skill assessment and surgical navigation. Attempts to develop such an object recognition model have been made, but sufficient results have not yet been obtained<sup>8</sup>. Herein, we constructed a model to recognize the object and types of forceps in surgical videos acquired during colorectal surgeries and evaluated its accuracy.

### Materials and methods

**Institutional approval.** The protocol for this study was reviewed and approved by the Tokyo Women's Medical University Review Board (Protocol No: 5380) and conducted according to the principles of the Declaration of Helsinki. All datasets were encrypted, and the identities of the patients were protected.

**Consent to participate.** Oral consent was obtained from all study subjects. Informed consent forms that include information on the purpose of the study and study methods, the subject, the name of the implementing organization, the name of the person in charge, and how to handle personal information were obtained and

<sup>1</sup>Department of Surgery, Institute of Gastroenterology, Tokyo Women's Medical University, 8-1, Kawadacho Shinjuku-ku, Tokyo 162-8666, Japan. <sup>2</sup>Ushiku Aiwa Hospital, Ibaraki, Japan. <sup>3</sup>Department of Surgery 2, Tokyo Women's Medical University, Tokyo, Japan. ✉email: bamba.yoshiko@twmu.ac.jp



**Figure 1.** Extraction of still images from surgical videos for data labeling. Five types of forceps, namely, grasping forceps, ultrasonic scalpel clip forceps, angled forceps (Maryland and right angle), and spatula forceps, were annotated in these images.

Forceps	Number of annotated forceps
Grasping forceps	903
Ultrasonic scalpel	311
Clip forceps	138
Angled forceps	185
Spatula forceps	281
Total	1818

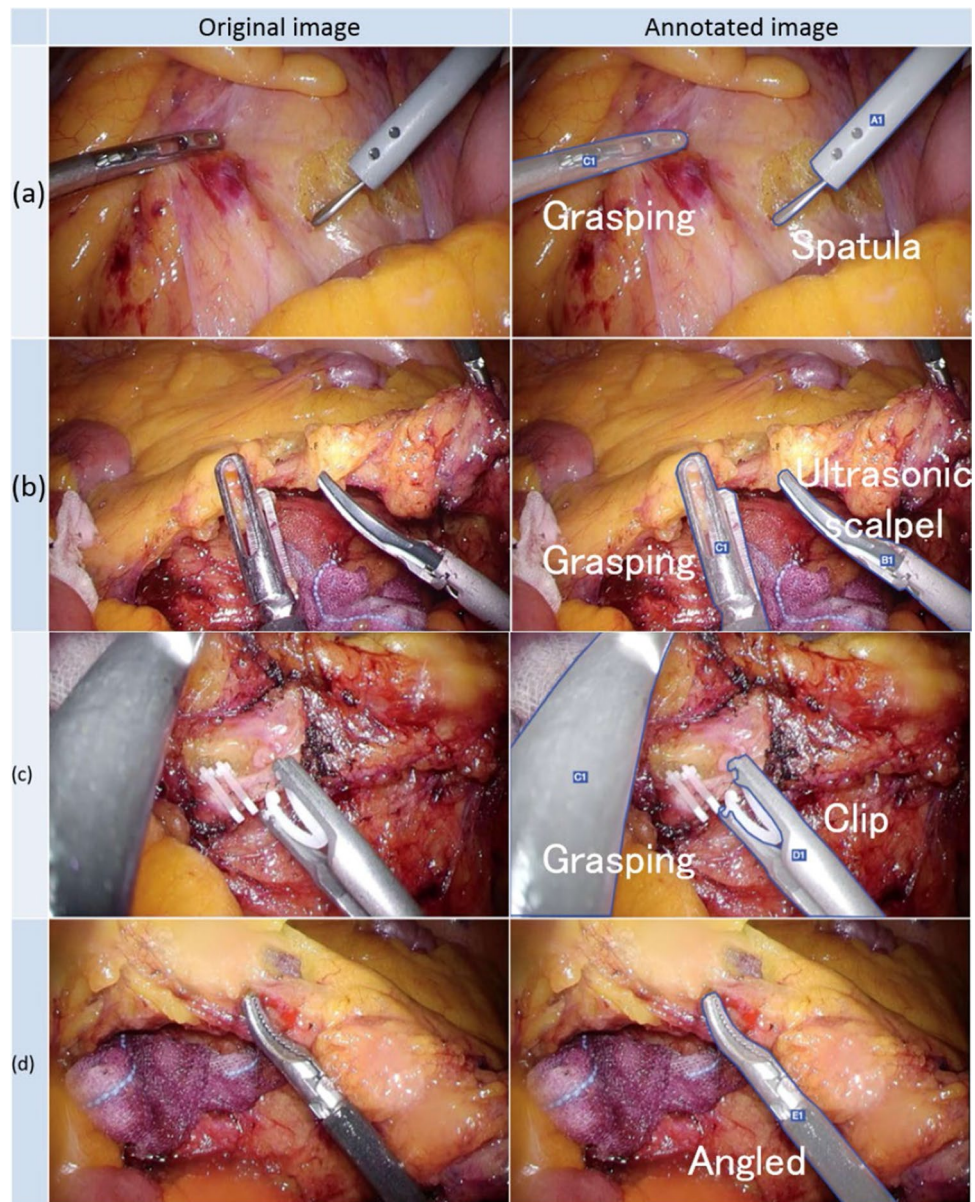
**Table 1.** Number of annotated forceps.

captured in the electronic medical records. For all other research subjects, information will also be disclosed by posting a document approved by the Ethics Committee on the Tokyo Women's Medical University website; this posting will also mention the possibility to refuse to participate as a research subject.

**Datasets.** The colorectal surgical videos used for annotation were recorded during surgeries conducted at the Tokyo Women's Medical University. A total of 1173 images were extracted from 11 surgical videos for model training, and another 500 images were extracted from 6 additional videos for validation. The following 5 types of forceps in the videos were selected for annotation: grasping, ultrasonic, clip, angled (Maryland and right-angled), and spatula forceps. A surgical video with a 60 s run time was extracted from the other videos and used to verify the model.

**Analysis.** The software IBM Visual Insights<sup>20</sup> (Power SystemAC922; NVIDIA Tesla V100 GPU, 32 GB) was used for the CNN for deep learning. It includes the most popular open-source deep-learning framework and tools, and is built for easy and rapid deployment. The modeling types included in the software are GoogLeNet, Faster R-CNN, tiny YOLO V2, YOLO V3, Detectron, Single Shot Detector (SSD) and Structured segment network (SSN). Detectron was selected for use in this study. IBM Visual Insights automatically splits the dataset for internal validation of the model's performance during training. The default value of 80/20 will result in the use of 80% of the test data (at random) for training and the use of the remaining 20% for measurements/validation.

**Imaging data and model deployment.** Abdominal endoscopic images were extracted from surgical videos (Fig. 1). In total, 1173 images were extracted to train a forceps-type recognition model. Five types of forceps were selected for manual annotation by only 1 researcher. The selected types of forceps were grasping forceps, ultrasonic scalpel, clip forceps, angled forceps, and spatula forceps (Table 1 and Fig. 2). The model was



**Figure 2.** Representative images of labeled forceps. Five types of forceps, namely, grasping forceps, ultrasonic scalpel, clip forceps, angled forceps, and spatula forceps, were selected and labeled in the extracted images to create a forceps-type recognition model. The images on the left side are original, and the images on the right side show the annotated forceps.

deployed, and the other 500 test images of various different angles of forceps with different patterns were input into the deployed model to verify its diagnostic accuracy (Fig. 3).

**Performance metrics.** Accuracy: percentage of correct image labels.

Mean average precision (mAP): calculated mean of precision for each object.

Precision: percentage of images with a correctly labeled object out of all labeled images that contain an object.

Recall: percentage of images that are labeled to contain an object out of all tested images that contain an object.

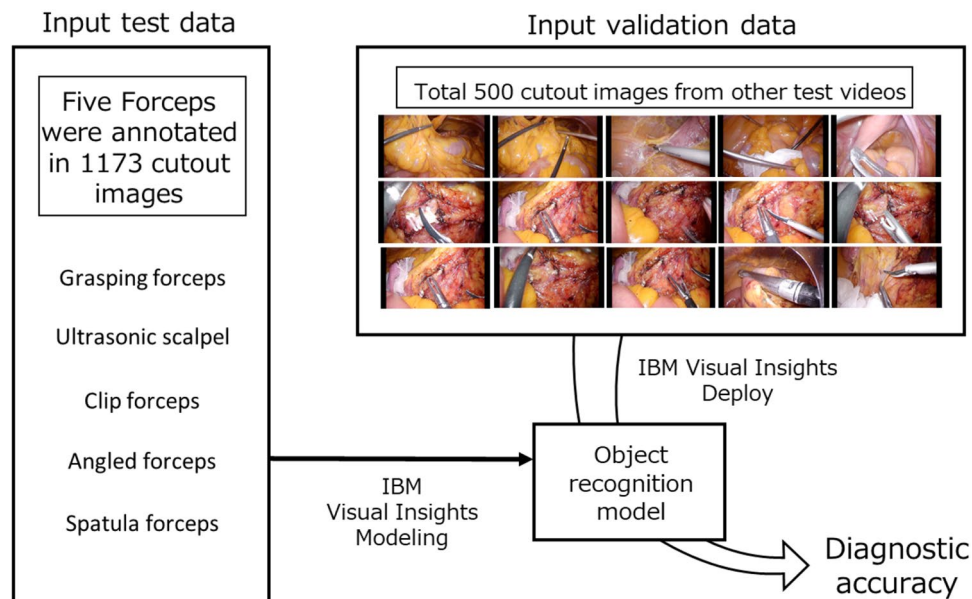
Intersection over Union (IoU): location accuracy of the image label boxes.

Confidence score: event probability.

## Results

The accuracy, mAP, precision, recall, and IoU of the model were 90%, 100%, 92%, 100%, and 77%, respectively (Fig. 4).

The total number of forceps identified in 500 test images was 1062. Of these, the number of correctly detected forceps was 1039 (97.8%). The number of false positives was 31. The recall and precision of each type of forceps calculated from the outcome values were as follows: grasping forceps, 98.1% and 98.0%; ultrasonic scalpel,



**Figure 3.** Flow of analysis using IBM Visual Insights. The five selected types of forceps were labeled in 1173 extracted images to create a forceps-type recognition model. Another 500 cutout images used for validation were input in the model to verify whether each type of forceps was recognized accurately.

99.4% and 93.9%; clip forceps, 96.2% and 92.7%; angled forceps, 94.9% and 100%; and spatula forceps, 98.1% and 94.5%, respectively (Table 2).

A surgical video with a 60 s run time was used to test the model, with the results indicating that the object was detected accurately (Supplementary Information).

## Discussion

In the field of surgery, AI-based decision support systems have provided a broad range of technological approaches to augment the information available to surgeons that have accelerated intraoperative pathology and surgical step recommendations<sup>19</sup>. Accurate and efficient object representation and segmentation are necessary for multilabel object classification in surgery based on the annotation of objects and frameworks<sup>21</sup>. Further, skill and motion assessments in surgical videos using CNN have been reported in recent years<sup>22–24</sup>.

In this study, we demonstrated the recognition of forceps (including type of forceps) from surgical images using CNN. In most test results, all 5 types of forceps were detected correctly with high confidence scores. Correspondingly, we obtained positive results in terms of the corresponding recall and precision values. The trained model was able to accurately detect the forceps at various angles (Fig. 4a–i). These results indicate that the model recognized the shapes and colors of each type of forceps with high precision.

Although small in number, some forceps were not detected, or the outcomes yielded false positives. Based on the incorrect outcome images, we found that errors arose when only part of the forceps was observed in the images (Fig. 5a,b) or when the shapes of the forceps were similar to those of other types of forceps (Fig. 5c,d). Additionally, the results suggest that image resolution affects the validation outcome considerably. Because the forceps are in motion during surgeries, they are sometimes blurred in surgical videos or are closed in the cutout images. As a result, the model could not identify them or would recognize them as another type of forceps.

The potential of automatic video indexing and surgical skill assessment has been reported with the use of 300 laparoscopic sigmoidectomy videos from multiple institutions in Japan<sup>25</sup>. In the present study, the recall and precision values were good despite the limited learning because of the mixed frameworks of deep learning based on the use of the commercial software IBM Visual Insights.

The results of our study will aid the development of a system that will manage, deliver, and retrieve surgical instruments for surgeons upon request. The object recognition model in surgery has reached feasible performance levels for widespread clinical use. The object recognition of forceps could be used to provide real-time object information during surgeries upon further development based on the results of this study. By integrating and developing these technologies, the digitalization of surgical scenes and techniques becomes possible. The ability to evaluate how and what procedure was performed is significant. Moreover, these innovations will enable surgical technique evaluation and surgical navigation. Utilization of AI is largely expected not only in medical treatments, such as the prevention and diagnosis of diseases, but also in cases associated with insufficient resources and in risk management to prevent medical accidents.

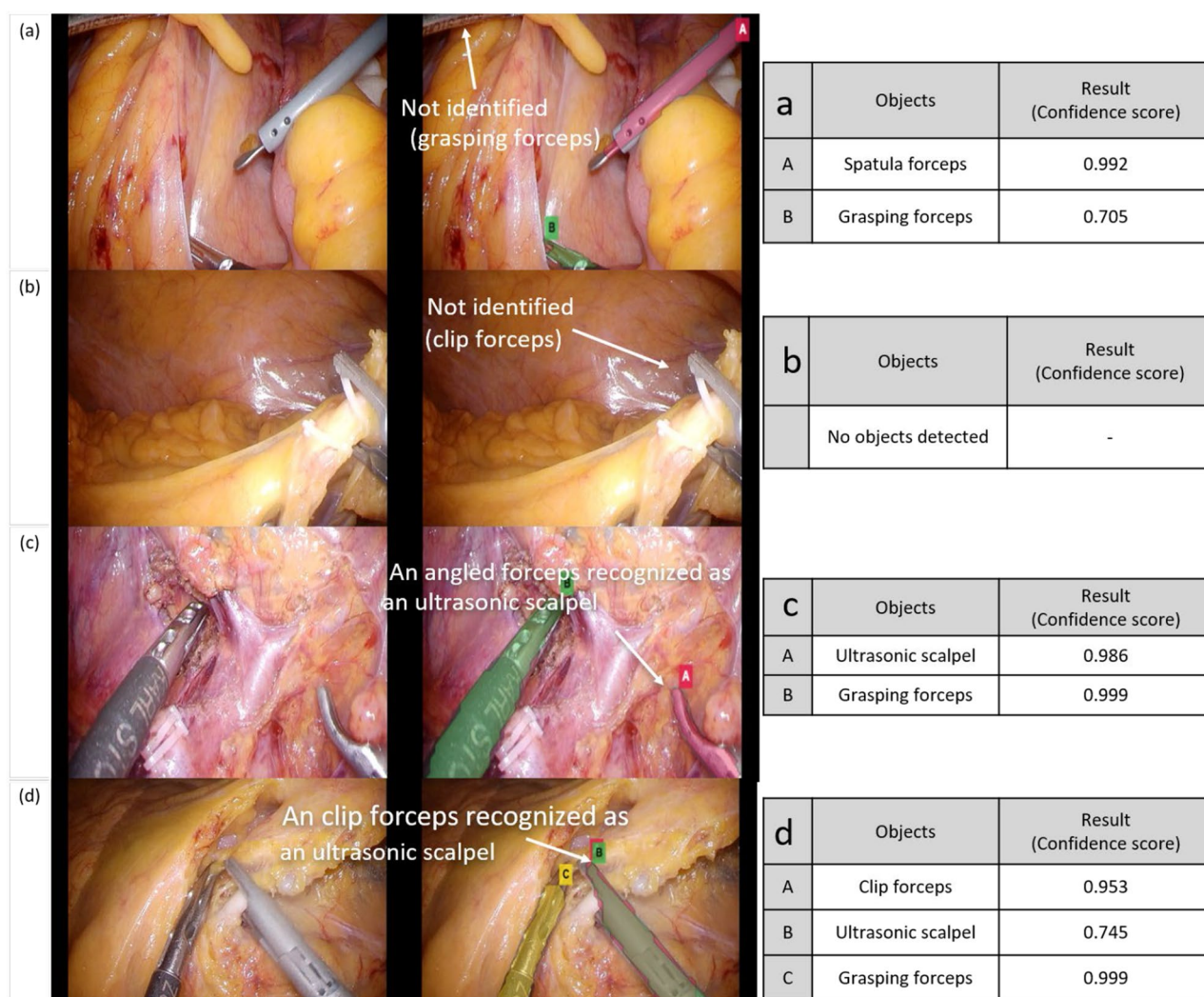
This study had some limitations. First, it is difficult to modify the model itself via tuning other than by changing the training data, because the model was made using IBM Visual Insights. Further, there were only limited types of forceps created from colorectal cancer videos of a single facility.



**Figure 4.** Representative images demonstrating accurate results. The images on the right side are original. The images in the middle are test results. The images on the right side show the confidence scores of each result.

Forceps	Number of forceps in images	Number of forceps identified correctly	Number of forceps not identified	False positive	Recall (95% CI)	Precision (95% CI)
Grasping	648	636	12	13	98.1% (97.1–99.2)	98.0% (96.9–99.1)
Ultrasonic scalpel	170	169	1	11	99.4% (98.3–100.6)	93.9% (90.4–97.4)
Clip	53	51	2	4	96.2% (91.1–101.4)	92.7% (85.9–99.6)
Angled	138	131	7	0	94.9% (91.3–98.6)	100% (100–100)
Spatula	53	52	1	3	98.1% (94.5–101.8)	94.5% (88.5–100.5)
Total	1062	1039	23	31	97.8% (97.0–98.7)	97.1% (96.1–98.1)

**Table 2.** Test results for each type of forceps, and corresponding recall and precision.



**Figure 5.** Representative images demonstrating inaccurate results. The images on the right side are original. The images in the middle are test results. The images on the right show the confidence scores of each result. (a) A grasping forceps and 1 spatula forceps were detected accurately, but 1 of the 2 grasping forceps in the image was not detected correctly; (b) the clip forceps was not identified correctly; (c) the angled forceps was recognized as an ultrasonic scalpel incorrectly; and (d) the clip forceps was identified correctly but was also recognized as an ultrasonic scalpel.

### Conclusion

In this study, we evaluated the recognition of different types of forceps using CNN and obtained positive results with high accuracy. Results of this study demonstrate the opportunity to evaluate use and navigation of forceps in surgeries.

Received: 15 June 2021; Accepted: 26 October 2021

Published online: 19 November 2021

## References

1. Topol, E. J. High-performance medicine: The convergence of human and artificial intelligence. *Nat. Med.* **25**, 44–56 (2019).
2. Litjens, G. *et al.* A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017).
3. Greenspan, H., van Ginneken, B. & Summers, R. M. Deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Trans. Med. Imaging.* **35**, 1153–1159 (2016).
4. Bejnordi, B. E. *et al.* Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *J. Am. Med. Assoc.* **318**, 2199–2210 (2017).
5. Lakhani, P. & Sundaram, B. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* **284**, 574–582 (2017).
6. Tajbakhsh, N. *et al.* Convolutional neural networks for medical image analysis: Full training or fine tuning?. *IEEE Trans. Med. Imaging.* **35**, 1299–1312 (2016).
7. Anwar, S. M. *et al.* Medical image analysis using convolutional neural networks: A review. *J. Med. Syst.* **42**, 226 (2018).
8. Gulshan, V. *et al.* Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *J. Am. Med. Assoc.* **316**, 2402–2410 (2016).
9. Milea, D. *et al.* Artificial intelligence to detect papilledema from ocular fundus photographs. *N. Engl. J. Med.* **382**, 1687–1695 (2020).
10. Esteve, A. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **546**, 686–686 (2017).
11. Haenssle, H. A. *et al.* Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann. Oncol.* <https://doi.org/10.1093/annonc/mdy166>, August 1, 2018 (2018).
12. Yu, L. Q., Chen, H., Dou, Q., Qin, J. & Heng, P. A. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Trans. Med. Imaging.* **36**, 994–1004 (2017).
13. Cui, X. *et al.* Assessing the effectiveness of artificial intelligence methods for melanoma: A retrospective review. *J. Am. Acad. Dermatol.* **81**, 1176–1180 (2019).
14. Misawa, M. *et al.* Characterization of colorectal lesions using a computer-aided diagnostic system for narrow-band imaging endocytoscopy. *Gastroenterology* **150**, 1531–1532 (2016).
15. Kudo, S. *et al.* Artificial intelligence-assisted system improves endoscopic identification of colorectal neoplasms. *Clin. Gastroenterol. Hepatol.* **18**, 1874–1881 (2020).
16. Halcox, J. P. J. *et al.* Assessment of remote heart rhythm sampling using the AliveCor heart monitor to screen for atrial fibrillation: The REHEARSE-AF study. *Circulation* **136**, 1784–1794 (2017).
17. Ramkumar, S. *et al.* Atrial fibrillation detection using single lead portable electrocardiographic monitoring: A systematic review and meta-analysis. *BMJ Open* **8**, 16 (2018).
18. Yildirim, O., Plawiak, P., Tan, R. & Acharya, U. R. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Comput. Biol. Med.* **102**, 411–420 (2018).
19. Navarrete-Welton, A. J. & Hashimoto, D. A. Current applications of artificial intelligence for intraoperative decision support in surgery. *Front. Med.* **14**, 369–381 (2020).
20. Xiao, C., Choi, E. & Sun, J. Opportunities and challenges in developing deep learning models using electronic health records data: A systematic review. *J. Am. Med. Inform. Assoc.* **25**, 1419–1428 (2018).
21. Loukas, C. & Sgouros, N. P. Multi-instance multi-label learning for surgical image annotation. *Int. J. Med. Robot. Comput. Assist. Surg.* **16**, 12 (2020).
22. Wang, Z. H. & Fey, A. M. Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery. *Int. J. Comput. Assist. Radiol. Surg.* **13**, 1959–1970 (2018).
23. Kowalewski, K. F. *et al.* Sensor-based machine learning for workflow detection and as key to detect expert level in laparoscopic suturing and knot-tying. *Surg. Endosc. Other Interv. Tech.* **33**, 3732–3740 (2019).
24. Zia, A., Sharma, Y., Bettadapura, V., Sarin, E. L. & Essa, I. Video and accelerometer-based motion analysis for automated surgical skills assessment. *Int. J. Comput. Assist. Radiol. Surg.* **13**, 443–455 (2018).
25. Kitaguchi, D. *et al.* Automated laparoscopic colorectal surgery workflow recognition using artificial intelligence: Experimental research. *Int. J. Surg.* **79**, 88–94 (2020).

## Acknowledgements

We thank Ms. Junko Machida for performing analyses.

## Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Y.B. The first draft of the manuscript was written by Y.B, and all authors commented on subsequent versions of the manuscript. All authors read and approved the final manuscript.

## Funding

This work was supported in part by a research grant from the TWMU Career Development Center for Medical Professionals and by a NAKAYAMA KOMEI Research Fellowship Grant.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-01911-1>.

**Correspondence** and requests for materials should be addressed to Y.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021