Assessment of changes in vessel area during needle manipulation in microvascular anastomosis using a deep learning-based semantic segmentation algorithm: A pilot study

Minghui Tang, PhD[1,2,3], Taku Sugiyama, MD, PhD[3,4*†], Ren Takahari, RT, Msc[5†], Hiroyuki Sugimori, RT, PhD[2,3,6], Takaaki Yoshimura, RT, PhD[7,8], Katsuhiko Ogasawara, MBA, PhD[2,7], Kohsuke Kudo, MD, PhD[1,2,3], Miki Fujimura, MD, PhD[4]

[†]These two authors contributed equally to this work.

[1] Department of Diagnostic Imaging, Hokkaido University Faculty of Medicine and Graduate School of Medicine, Sapporo, Japan

[2] Clinical AI Human Resources Development Program, Hokkaido University Graduate School of Biomedical Science and Engineering, Sapporo, Japan

[3] Medical AI Research and Development Center, Hokkaido University Hospital, Sapporo, Japan

[4] Department of Neurosurgery, Hokkaido University Graduate School of Medicine, Sapporo, Japan

[5] Graduate School of Health Sciences, Hokkaido University, Sapporo, Japan

[6] Department of Biomedical Science and Engineering, Faculty of Health Sciences, Hokkaido University, Sapporo, Japan

[7] Department of Health Sciences and Technology, Faculty of Health Sciences, Hokkaido University, Sapporo, Japan

[8] Department of Medical Physics, Hokkaido University Hospital, Sapporo, Japan

***Corresponding author:** Taku Sugiyama, M.D., Ph.D.

Department of Neurosurgery, Hokkaido University Graduate School of Medicine, North 15 West 7, Kita-ku, Sapporo 060-8638, Japan

Tel: +81-11-706-5987, Fax: +81-11-708-7737

E-mail: takus1113@med.hokudai.ac.jp

Abstract

Appropriate needle manipulation to avoid abrupt deformation of fragile vessels is a critical determinant of the success of microvascular anastomosis. However, no study has yet evaluated the area changes in surgical objects using surgical videos. The present study therefore aimed to develop a deep learning-based semantic segmentation algorithm to assess the area change of vessels during microvascular anastomosis for objective surgical skill assessment with regard to the "respect for tissue." The semantic segmentation algorithm was trained based on a ResNet-50 network using microvascular end-to-side anastomosis training videos with artificial blood vessels. Using the created model, video parameters during a single stitch completion task, including the coefficient of variation of vessel area (CV-VA), relative change in vessel area per unit time (ΔVA), and the number of tissue deformation errors (TDE), as defined by a ΔVA threshold, were compared between expert and novice surgeons. A high validation accuracy (99.1%) and Intersection over Union (0.93) were obtained for the auto-segmentation model. During the single-stitch task, the expert surgeons displayed lower values of CV-VA ($p<0.05$) and ΔVA ($p<0.05$). Additionally, experts committed significantly fewer TDEs than novices ($p<0.05$), and completed the task in a shorter time ($p<0.01$). Receiver operating curve analyses indicated relatively strong discriminative capabilities for each video parameter and task completion time, while the combined use of the task completion time and video parameters demonstrated complete discriminative power between experts and novices. In conclusion, the assessment of changes in the vessel area during microvascular anastomosis using a deep learning-based semantic segmentation algorithm is presented as a novel concept for evaluating microsurgical performance. This will be useful in future computer-aided devices to enhance surgical education and patient safety.

Keywords: artificial intelligence, EC-IC bypass, machine learning, semantic segmentation, surgical

skill, tissue deformation

Introduction

The skill level of surgeons is a critical determinant of postoperative outcomes in terms of complications, morbidity, and mortality [3, 8]. The use of a comprehensive surgical skill evaluation can provide an accurate index for predicting postoperative outcomes, leading to better postoperative care among patients [3]. Notably, the education of surgical residents depends on case observations and practice, highlighting the importance of detailed and objective evaluations of surgical skills to identify and address shortcomings as feedback. Furthermore, this can contribute to accrediting the competence of surgeons.

The traditional approach to evaluate the surgical skills of surgeons relies on the visual assessment of technical skills by experienced surgeons [13, 21, 23]. However, as these evaluation methods employ subjective judgment scales, they may not accurately assess actual surgical skills [6, 31]. Subsequently, several criteria-based scores, including checklists, have been developed and utilized for surgical skill assessments [1, 17], thus reducing the subjective aspect to a certain extent. Nevertheless, the challenge is that various biases cannot be entirely avoided [6]. Recently, efforts have been made to develop objective and quantitative evaluation methods to overcome the challenge of subjectivity in traditional surgical skills. For example, evaluation methods have been developed using measurement devices such as sensors attached to surgical instruments to detect force [27, 30, 36], or electromagnetic sensors attached to the body or instruments of the surgeon for motion tracking [12, 14, 15]. Although this type of quantitative evaluation allows an objective comparison of the surgical skill levels of trainee surgeons, it requires many specialized devices and sensors for data measurement, making it difficult to introduce into many facilities, as well as posing challenges to its generalizability and reproducibility [31, 34].

In recent years, the focus has shifted towards analytical methods utilizing video recordings of surgical procedures [7, 24, 31]. This approach is gaining popularity because of its applicability in diverse operating room environments. Previous studies have demonstrated correlations between comprehensive motion analysis scores of surgeons' hands or surgical instruments and skill levels across various surgical procedures [4, 20]. Moreover, the use of artificial intelligence (AI) for surgical skill evaluation through video analysis is becoming increasingly common [16, 25, 37]. However, most of these methods have primarily focused on aspects directly related to the surgeon.

Recent studies have developed a novel indicator for the objective and quantitative evaluation of surgical skills by focusing on tissue movement during surgery. In previous studies, the motion parameters of target tissues during surgery (e.g., acceleration and velocity) were analyzed in videos of carotid endarterectomy [28, 31]. The results of these studies showed that novice surgeons tended to have higher values for the motion parameters of target tissues than expert surgeons, and the motion parameters of surgical videos were proposed as an alternative index for evaluating surgical skills. It is believed that the less movement of tissues during surgery, the more quantitative evaluation marker of "gentleness" in handling tissues during surgery [31, 37]. By focusing on the motion parameters of tissues during surgery, it has become evident that an objective evaluation of surgical skills is possible and is now a feasible evaluation marker.

Cerebral revascularization surgery, specifically the extracranial-intracranial (EC-IC) bypass procedure, involves surgical anastomosis of two blood vessels, and is thus considered a challenging neurosurgical procedure [9, 11, 29]. Inadvertent mishandling of the fragile cerebral blood vessels can easily damage, leading to bypass occlusion, stroke, or intracranial hemorrhage [9, 10, 33]. Thus, it is critical to ensure gentle handling of these vessels during anastomosis, and to ensure a quantitative and objective assessment of this gentleness to enhance both surgical education and patient safety [1, 18, 20].

In this pilot study, we focused on vessel deformation to objectively assess the surgical performance under the hypothesis that avoiding excessive or abrupt tissue (vessel) deformation could serve as a novel metric for gentle maneuvers in microvascular anastomosis. Although several methods and tools for microvascular training have been developed, the quantification of these aspects remains limited. To achieve this, a segmentation AI model was constructed for simulated blood vessels, and expert and novice surgical skill levels were compared by analyzing the change in the area of the vessel during needle manipulation.

Methods

Collection of the microsurgical training video

This study was approved by Hokkaido University Hospital (No. 018-0291). As our facility holds periodic off-the-job microvascular training sessions for educational purposes, we were able to collect video records of microsurgical training from these sessions.

In the training session, the routine surgical task consisted of an end-to-side anastomosis of 2.0-mm artificial blood vessels (Blood vessel model, WetLab Incorporated, Shiga, Japan) using a 10-0 surgical suture. A microscope (EMZ-10B; Meiji Techno, Saitama, Japan) equipped with a camera with an electronic shutter (MKC-230HD; Ikegami, Tokyo, Japan) recorded the procedures using a recorder (HVO-1000MD; Sony, Tokyo, Japan). Two artificial vessels (diameter, 2 mm; material, polyvinyl alcohol; length, 60 mm) were cut and prepared for the donor and recipient, and the recipient vessels were stabilized using microvascular clips or pins. To create a wide fish-mouth opening, the end of the donor was sharply cut across the artery at a 60° angle and then cut along the donor artery at the same length. The recipient was then cut linearly. Two stay sutures are placed on the heel and toe. Interrupted sutures were placed on both sides to complete the side-to-end anastomosis. Each stitch was secured using three knots.

In this study, video recordings including the process of interrupted suturing after two stay sutures were used as training data. The interlaced videos were output as Moving MPEG-2 (Picture Experts Group 2) video stream (M2V) files (resolution: 1920×1080) at a rate of 30 frames per second (fps).

Data pre-processing

Each video is separated into frames. Each frame was then resized to a resolution of 512×512 pixels by padding the shorter side of the images with zeros (i.e., black) and saved as Joint Photographic Experts Group (JPEG) files (resolution: 1280×720 pixels) using an in-house program. A pre-processing flowchart is shown in **Figure 1**. Given the variability in background colors and zoom rates across videos, we visually categorized the JPEGs into 20 distinct scenes. For each scene, 70 images were randomly selected to label artificial blood vessels. The area visually identified as a vessel was labeled. To minimize interference from the surgeon's hands or tools (e.g., surgical tweezers) obstructing the microscopic view, the areas covered by these were also labeled. All labeling was performed by a non-surgical healthy researcher with no visual impairment. The labeled images were randomly separated into training, validation, and test datasets at a ratio of 4:1:2 for validation. The training dataset was augmented by image rotation from -175° to 180° in 5 °increments. The total number of images for the training, validation, and test dataset was 57600, 200, and 400, respectively. All processing was performed using MATLAB (version R2021a, MathWorks, Natick, MA, USA) and labeling was performed using ImageLabeler, a MATLAB tool.

Deep learning

The training and test datasets were inputted into a Residual Network 50 (ResNet-50) model using MATLAB. The training parameters were as follows: optimizer stochastic gradient descent with momentum (Momentum SGD) [22], maximum training epochs of 30, mini-batch size of 192, initial learning rate of 0.001, momentum of 0.9, L2 regulation of 0.05, and validation patient of 4. Validation patience specifies the number of times the loss in the validation set can be larger than or equal to the previously smallest loss before the network training stops.

After training, the accuracy of the AI semantic segmentation model was assessed by calculating the Intersection over Union (IoU) of the test dataset between manual labeling (Ml) and AI-segmented (As) using the following equation [35]:

$$\mathrm{IoU} = \frac{Ml \cap As}{Ml \cup As}.$$

The model was subsequently applied to all the data from all videos, and the area of the vessel in each JPEG image was depicted using AI. Training and validation were performed on an Ubuntu 18.04 long-term support (LTS)-based workstation (Deeplearning box II, GDEP Advance, Tokyo, Japan) with a Core i9 10980XE 18core/36thread 3.0-GHz central processing unit (CPU), four NVIDIA Quadro RTX8000 graphics processing unit (GPU) cards (Nvidia Corporation, Santa Clara, CA, USA), and 128-GB (16GB×8) DDR4-2933 quad-channel memory.

Experimental study on surgical performance evaluation

To validate the semantic segmentation model to assess the surgical performance level, surgeon participants were recruited from a microvascular training session and categorized into experts (instructors) and novices (trainees). The expert group included board-certified neurosurgeons who had independently performed more than 20 EC-IC bypass surgeries (postgraduate years [PGY] 12–28, n=5). The novice group was primarily composed of junior neurosurgical residents (PGY 1-3, n=7) who had not previously performed EC-IC bypass surgery. The details of the participating surgeons are summarized in **Table 1.**

Prior to the experimental study, novice surgeons (trainees) received a pre-education lecture from expert surgeons (instructors) using a standardized textbook in addition to a surgical video focusing on microvascular anastomosis techniques. The lecture extensively covered surgical techniques, including needle manipulation to prevent injury to the target vessel, and strategies to minimize task completion time. In this study, the designated task for each surgeon was to perform interrupted suturing following the two-stay sutures. A "surgical trial" was defined as the completion of a single suturing process, which involved grasping the needle with forceps, inserting and extracting the needle, pulling the threads, tying knots, and cutting the threads. To eliminate design effects in the anastomosis, tasks such as stabilizing the two vessels, cutting the donor artery, arteriostomy of the recipient vessel, and preparation of two stay sutures were confirmed by a single instructor in a unified fashion. The surgical field was maintained under semi-wet conditions using a spray.

To mitigate the learning effect on tissue deformation during task repetition among the participating surgeons, the first two trials from each surgeon were selected for the performance evaluation. Subsequently, 10 trials conducted by five experts and 14 trials conducted by seven novices were further analyzed to assess the validity of the semantic segmentation model in evaluating surgical performance levels.

Video parameters

To evaluate the surgical performance across different phases of a “surgical trial, ” each trial was manually divided into four distinct phases by non-surgeon researchers as follows [19, 32]: *phase A*: from grasping the needle to its insertion; *phase B*: pushing and extraction of the needle; *phase C*: from pulling the threads to the first knot; *phase D*: completion of three knots and cutting the threads.

In each trial and phase, the vessel area (VA), defined as the total pixel count, was calculated for each labeled JPEG image, and predicted using the developed autosegmentation AI model (**Figure 2**). Consequently, time-series data was obtained for VA.

Because the variability of VA within each trial and phase (*phase A-D*) was considered an important parameter to represent a consistent surgical maneuver by surgeons, the coefficient of variation (CV) of all VA readings was calculated as follows [30]:

$$\mathrm{CV} = \sigma/\mu$$

σ: standard deviation (SD) of VA, μ: mean value of VA

To identify abrupt deformations in the blood vessels, sudden changes in the VA, defined as the relative rate of area change per unit time (ΔVA), was calculated based on the initially predicted VA (**Figure 3**). The maximum absolute values of ΔVA (Max-ΔVA) were adopted to represent data in each trial and phase.

To define technical errors made by surgeons during procedures, the 95% confidence intervals (CI, the mean ± 1.96×SD) of ΔVA for all trials by all surgeons were calculated and used as thresholds. Any values of ΔVA violating the thresholds were regarded as a *tissue deformation error (TDE),* which were counted for each trial across all procedures and phases (**Figure 3**) [30].

Statistical analysis

We assumed that parameters including task completion time, CV of VA (CV-VA), Max-ΔVA, and number of TDEs, would be lower in experts than novices; as such, the null hypothesis posited that there would be no difference in these parameters between experts and novices. The normality of each video parameter was tested using the Shapiro-Wilk test. Normally distributed parameters were compared between experts and novices using the Welch *t*-test, while non-normally distributed data were compared using the Mann-Whitney *U* test.

Receiver operating characteristic (ROC) curves based on the video parameters (task completion time, CV-VA, Max-ΔVA, and number of TDEs) were calculated. The area under the curve (AUC) with a 95% confidence interval (CI) was calculated and used to evaluate the discriminative capacity of these parameters to distinguish experts and novices.

Statistical analyses were performed using JMP Pro (version 16.2.0; SAS Institute Inc., Cary, North Carolina, USA). Statistical significance was set at $p<0.05$.

Results

Evaluation of the created semantic segmentation model

The learning time was 162 min and 5 s, and the learning process ended at the 7th epoch, with a validation accuracy of 99.07%. The IoU was 0.93. Although a high validation accuracy and IoU were obtained, misdepiction occurred in 4% of the average total number of images. All videos were carefully reviewed by non-surgical researchers, and misdepicted data were removed from subsequent analyses after a consensus was reached by all researchers.

Experimental surgical performance analysis between experts and novices

All of the video parameters for both the experts and novices within each trial and phase (*phase A-D*) are shown in **Figure 4.** The task completion time was significantly shorter for experts than for novices across the entire trial ($p<0.01$), while significant differences between the two groups were observed in phases C ($p<0.01$) and D ($p<0.01$).

Regarding CV-VA, experts showed a lower value than novices in one trial ($p<0.05$), while significant differences between the two groups were found in phases B ($p<0.05$) and D ($p<0.05$). Regarding Max-ΔVA, experts showed lower values than novices across the trial ($p<0.05$), while significant differences between two groups were also found in phases B ($p<0.05$), and D ($p<0.05$).

The mean ± 1.96×SD of ΔVA for all trials by all surgeons was calculated as ± 1.13 (**Figure 3**), and used as a threshold for subsequent analysis. Overall, experts had significantly fewer TDEs than the novices in a trial ($p<0.01$), and significant differences between the two groups were found in phases B ($p<0.05$), C ($p<0.05$), and D ($p<0.05$).

The ROC curves for each parameter in a trial in the discrimination between experts and novices were shown in **Figure 5A**, with the AUC values showing relatively strong associations for each parameter, with 0.94 (95% CI 0.81–1.00) for the procedural time, 0.71 (95% CI 0.41–1.00) for the CV-VA, 0.80 (95% CI 0.53–1.00) for the Max-ΔVA, and 0.83 (95% CI 0.58–1.00) for the number of TDEs. The combined use of procedural time and CV-VA, Max-ΔVA, or number of TDE could completely distinguish between experts and novices (**Figure 5B**).

Discussion

In this pilot study, we constructed an auto-segmentation AI model capable of assessing vessel deformation (i.e., time-course change in area of vessel on two-dimensional video) during microvascular end-to-side anastomosis. We further applied this model to microvascular training videos and evaluated the surgical performances of both expert and novice surgeons. Through the application of this model, we were able to show differences between experts and novices in terms of video parameters including CV-VA, ΔVA, and the number of TDEs. Interestingly, the values of these video parameters were lower and the task completion time was shorter for experts than for novices, indicating that both the speed and gentleness of surgical maneuvers are associated with microsurgical proficiency.

Although microvascular anastomosis involves a variety of elements, including instrument and suture handling, technique, quality of knot, and operation flow [20], respect for the tissue is an important component of microsurgical proficiency as the target tissue (i.e., cerebral arteries) is anatomically fragile and easily torn by rough manipulation of needles, sutures, or the vascular wall [9, 10, 33]. Tissue deformation and motion are related to the force applied by surgeons, and the assessment of these aspects can be considered as a representation of the gentleness in tissue handling. Although such aspects have not yet been assessed in the literature, many experts and facilities have emphasized the importance of gentle and careful handling of fragile cerebral arteries in EC-IC bypass procedures [1, 18, 20]. To the best of our knowledge, this is the first study to utilize an auto-segmentation deep-learning-based model to assess this aspect.

The concept of the present study can be achieved using only surgical videos, without the need for physical sensors or devices. Therefore, this method is highly applicable for future clinical studies. One of the expected targets is AI-aided devices that can promptly provide feedback to trainees to enhance their surgical training [2]. In addition, the accumulation of clinical data involving adverse events such as vascular injury, bypass occlusion, and stroke, could allow the implementation of a real-time warning system such that operation staff and surgeons are warned when the safety threshold of tissue deformation is violated [5].

Regarding the technical aspect, novice surgeons experienced a significant increase in the average total number of TDEs (3.86 fold higher), while taking a significantly longer time (1.93 fold longer) to complete the single-stitch task compared to expert surgeons. Interestingly, in phase B (needle extraction), all video parameter values were lower for experts than for novices; however, the task completion time did not differ between the two groups. This may be attributed to the fact that needle extraction often requires special technical finesse [15]. During this process, the needle should be extracted along its curve with the appropriate application of counterpressure using forceps not used for grasping the needle [26]. This result highlights the importance of paying special attention during needle extraction during microvascular training. In contrast, no parameters differed between the two groups during phase A (from grasping to inserting the needle), implying that mastering this process was the easiest part of microvascular anastomosis.

This study had some limitations. Although the segmented AI model exhibited a high level of accuracy, misdepictions occurred for images with shadows from surgical tweezers or hands in the temporal area series. In this study, misdepictions were excluded from the analysis of the number of errors because of their small proportion (approximately 4%) in the temporal area series. To solve this problem, it could be beneficial to purposely increase the training datasets for deep learning to contain the shadows of surgical tweezers or hands. Additionally, because only ResNet-50 was used in this study, the optimization of the network type should be investigated further. As this pilot study evaluated only limited stages of microvascular anastomosis, other surgical tasks, such as the design and alignment of the donor and recipient arteries, arteriotomy, and placement of two-stay sutures on both the toe and heel, could also contribute to the success of microvascular anastomosis. Although these aspects were beyond the scope of this pilot study, future research comprising a more comprehensive analysis, including a patency or leakage test of microvascular anastomoses, should be conducted to validate our results. Additionally, we did not evaluate data from repeated trials to determine whether novice surgeons improved after receiving feedback on their techniques. Further studies are necessary to evaluate whether this approach can add value to training efficiency.

Conclusion

Herein, we present assessment of vessel deformation during microvascular anastomosis using a deep learning-based semantic segmentation algorithm as a novel concept for evaluating surgical performance. This research approach could be introduced in future computer-aided surgeries for both education and surgical success.

Acknowledgements: We would like to thank Yasuhiro Ito, MD, Haruto Uchino, MD, Masayuki Gekka, MD, and Masaki Ito, MD, together with the Hokkaido University Neurosurgery Residency Program for their participation in this study.

Declarations

Ethical approval: This study was approved by the institutional review board of Hokkaido University Hospital, Sapporo, Japan.

Competing interests: The authors have no relevant financial or non-financial interests to disclose.

Authors' contributions: Minghui Tang, Taku Sugiyama, Hiroyuki Sugimori, and Katsuhiko Ogasawara contributed to the study conception and design. Data collection was performed by Minghui Tang, Ren Takahari, and Taku Sugiyama. Data analysis was performed by Minghui Tang, Taku Sugiyama, Ren Takahari, Takaaki Yoshimura, and Hiroyuki Sugimori. The first draft of the manuscript was written by Minghui Tang and Ren Takahari. Taku Sugiyama critically revised the manuscript and all authors commented on the previous version of the manuscript. Taku Sugiyama contributed to grant acquisition. All authors read and approved the final manuscript. Kohsuke Kudo and Miki Fujimura contributed to supervision of the study.

Funding: This work was supported by JSPS KAKENHI Grant Number JP21K09091.

Availability of data and materials: The data and materials that support the findings of this study are available from the corresponding author upon reasonable request.

Consent to participate: Informed consent was obtained from all individual participants included in this study.

Consent to publish: The authors affirm that human research participants provided informed consent for publication.

Reference

1. Aoun SG, El Ahmadieh TY, El Tecle NE, Daou MR, Adel JG, Park CS, Batjer HH, Bendok BR (2015) A pilot study to assess the construct and face validity of the Northwestern Objective Microanastomosis Assessment Tool. J Neurosurg 123:103–109. doi: 10.3171/2014.12.JNS131814
2. Baghdadi A, Lama S, Singh R, Hoshyarmanesh H, Razmi M, Sutherland GR (2021) A data-driven performance dashboard for surgical dissection. Sci Reports 2021 111 11:1–13. doi: 10.1038/s41598-021-94487-9
3. Birkmeyer JD, Finks JF, O'Reilly A, Oerline M, Carlin AM, Nunn AR, Dimick J, Banerjee M, Birkmeyer NJO (2013) Surgical skill and complication rates after bariatric surgery. N Engl J Med 369:1434–1442. doi: 10.1056/NEJMsa1300625
4. Castillo-Segura P, Fernández-Panadero C, Alario-Hoyos C, Muñoz-Merino PJ, Delgado Kloos C (2021) Objective and automated assessment of surgical technical skills with IoT systems: A systematic literature review. Artif. Intell. Med. 112:102007
5. Curtis NJ, Dennison G, Brown CSB, Hewett PJ, Hanna GB, Stevenson ARL, Francis NK (2019) Clinical Evaluation of Intraoperative Near Misses in Laparoscopic Rectal Cancer Surgery. Ann Surg XX:1. doi: 10.1097/sla.0000000000003452

6. Darzi A, Smith S, Taffinder N (1999) Assessing operative skill. Br. Med. J. 318:887–888

7. Davids J, Makariou SG, Ashrafian H, Darzi A, Marcus HJ, Giannarou S (2021) Automated Vision-Based Microsurgical Skill Analysis in Neurosurgery Using Deep Learning: Development and Preclinical Validation. World Neurosurg 149:e669–e686. doi: 10.1016/j.wneu.2021.01.117

8. Fecso AB, Szasz P, Kerezov G, Grantcharov TP (2017) The effect of technical performance on patient outcomes in surgery. Ann Surg 265:492–501. doi: 10.1097/SLA.0000000000001959

9. Fujimura M, Kaneta T, Tominaga T (2008) Efficacy of superficial temporal artery-middle cerebral artery anastomosis with routine postoperative cerebral blood flow measurement during the acute stage in childhood moyamoya disease. Child's Nerv Syst 24:827–832. doi: 10.1007/s00381-007-0551-y

10. Fujimura M, Tominaga T (2021) Flow-augmentation bypass for moyamoya disease. J. Neurosurg. Sci. 65:277–286

11. Fujimura M, Tominaga T, Kuroda S, Takahashi JC, Endo H, Ogasawara K, Miyamoto S (2022) 2021 Japanese Guidelines for the Management of Moyamoya Disease: Guidelines from the Research Committee on Moyamoya Disease and Japan Stroke Society. Neurol Med Chir (Tokyo) 62:165–170. doi: 10.2176/jns-nmc.2021-0382

12. Ghasemloonia A, Maddahi Y, Zareinia K, Lama S, Dort JC, Sutherland GR (2017) Surgical Skill Assessment Using Motion Quality and Smoothness. J Surg Educ 74:295–305. doi: 10.1016/j.jsurg.2016.10.006

13. Goff BA, Lentz GM, Lee D, Houmard B, Mandel LS (2000) Development of an objective structured assessment of technical skills for obstetric and gynecology residents. Obstet Gynecol 96:146–150. doi: 10.1016/S0029-7844(00)00829-2

14. Grober ED, Roberts M, Shin EJ, Mahdi M, Bacal V (2010) Intraoperative assessment of technical skills on live patients using economy of hand motion: establishing learning curves of

surgical competence. Am J Surg 199:81–85. doi: 10.1016/j.amjsurg.2009.07.033

15. Harada K, Morita A, Minakawa Y, Baek YM, Sora S, Sugita N, Kimura T, Tanikawa R, Ishikawa T, Mitsuishi M (2015) Assessing Microneurosurgical Skill with Medico-Engineering Technology. World Neurosurg 84:964–971. doi: 10.1016/j.wneu.2015.05.033

16. Jo K, Choi Y, Choi J, Chung JW (2019) Robust Real-Time Detection of Laparoscopic Instruments in Robot Surgery Using Convolutional Neural Networks with Motion Vector Prediction. Appl Sci 9:2865. doi: 10.3390/app9142865

17. Martin JA, Regehr G, Reznick R, Macrae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg 84:273–278. doi: 10.1002/bjs.1800840237

18. Mascitelli JR, Yoon S, Cole TS, Kim H, Lawton MT (2019) Does eloquence subtype influence outcome following arteriovenous malformation surgery? J Neurosurg 131:876–883. doi: 10.3171/2018.4.JNS18403

19. McBeth PB, Louw DF, Yang F, Sutherland GR (2005) Quantitative measures of performance in microvascular anastomoses. Comput Aided Surg 10:173–180. doi: 10.3109/10929080500229694

20. McGoldrick RB, Davis CR, Paro J, Hui K, Nguyen D, Lee GK (2015) Motion analysis for microsurgical training: Objective measures of dexterity, economy of movement, and ability. Plast Reconstr Surg 136:231e-240e. doi: 10.1097/PRS.0000000000001469

21. Neilsen PE, Foglia LM, Mandel LS, Chow GE (2003) Objective structured assessment of technical skills for episiotomy repair. Am J Obstet Gynecol 189:1257–1260. doi: 10.1067/S0002-9378(03)00812-3

22. Qian N (1999) On the momentum term in gradient descent learning algorithms. Neural Networks 12:145–151. doi: 10.1016/S0893-6080(98)00116-6

23. Reznick RK (1993) Teaching and testing technical skills. Am J Surg 165:358–361. doi:

10.1016/S0002-9610(05)80843-8

24. Sarkiss CA, Philemond S, Lee J, Sobotka S, Holloway TD, Moore MM, Costa AB, Gordon EL, Bederson JB (2016) Neurosurgical Skills Assessment: Measuring Technical Proficiency in Neurosurgery Residents Through Intraoperative Video Evaluations. World Neurosurg 89:1–8. doi: 10.1016/j.wneu.2015.12.052

25. Sugimori H, Sugiyama T, Nakayama N, Yamashita A, Ogasawara K (2020) Development of a Deep Learning-Based Algorithm to Detect the Distal End of a Surgical Instrument. Appl Sci 10:4245. doi: 10.3390/app10124245

26. Sugiyama T (2017) Mastering Intracranial Microvascular Anastomoses -Basic Techniques and Surgical Pearls. MEDICUS SHUPPAN, Publishers Co., Ltd

27. Sugiyama T, Gan LS, Zareinia K, Lama S, Sutherland GR (2017) Tool-Tissue Interaction Forces in Brain Arteriovenous Malformation Surgery. World Neurosurg 102:221–228. doi: 10.1016/j.wneu.2017.03.006

28. Sugiyama T, Ito M, Sugimori H, Tang M, Nakamura T, Ogasawara K, Matsuzawa H, Nakayama N, Lama S, Sutherland GR, Fujimura M (2023) Tissue Acceleration as a Novel Metric for Surgical Performance During Carotid Endarterectomy. Oper Neurosurg 25:343–352. doi: 10.1227/ons.0000000000000815

29. Sugiyama T, Kazumata K, Asaoka K, Osanai T, Shimbo D, Uchida K, Yokoyama Y, Nakayama N, Itamoto K, Houkin K (2015) Reappraisal of microsurgical revascularization for anterior circulation ischemia in patients with progressive stroke. World Neurosurg 84:1579–1588. doi: 10.1016/j.wneu.2015.07.053

30. Sugiyama T, Lama S, Gan LS, Maddahi Y, Zareinia K, Sutherland GR (2018) Forces of tool-tissue interaction to assess surgical skill level. JAMA Surg 153:234–242. doi: 10.1001/jamasurg.2017.4516

31. Sugiyama T, Nakamura T, Ito Y, Tokairin K, Kazumata K, Nakayama N, Houkin K (2019) A

Pilot Study on Measuring Tissue Motion During Carotid Surgery Using Video-Based Analyses for the Objective Assessment of Surgical Performance. World J Surg 43:2309–2319. doi: 10.1007/s00268-019-05018-7

32. Sugiyama T, Sugimori H, Tang M, Ito Y, Gekka M, Uchino H, Ito M, Ogasawara K, Fujimura M (2024) Deep learning-based video-analysis of instrument motion in microvascular anastomosis training. Acta Neurochir (Wien) 166:6. doi: 10.1007/s00701-024-05896-4

33. Takagi Y, Kikuta KI, Nozaki K, Hashimoto N (2007) Histological features of middle cerebral arteries from patients treated for moyamoya disease. Neurol Med Chir (Tokyo) 47:1–4. doi: 10.2176/nmc.47.1

34. Tang M, Kawahira S, Nomura N, Yamamoto T (2020) Torque property of titanium alloy cerebral aneurysm clips in a magnetic resonance scanner. J Mater Sci Mater Med 31:6. doi: 10.1007/s10856-019-6329-4

35. Tanimoto TT (1958) An Elementary Mathematical Theory of Classification and Prediction. Proc IBM Intern Rep 1–11

36. Zareinia K, Maddahi Y, Gan LS, Ghasemloonia A, Lama S, Sugiyama T, Yang FW, Sutherland GR (2016) A Force-Sensing Bipolar Forceps to Quantify Tool-Tissue Interaction Forces in Microsurgery. IEEE/ASME Trans Mechatronics 21:2365–2377. doi: 10.1109/TMECH.2016.2563384

37. Zhao Z, Cai T, Chang F, Cheng X (2019) Real-time surgical instrument detection in robot-assisted surgery using a convolutional neural network cascade. In: Healthcare Technology Letters. pp 275–279

Figure Legends

Fig. 1 Flow chart of the data pre-processing.

Fig. 2 Overview of the user interface of the home-made software showing that the created auto-segmentation model correctly depicted the vessel area. The plot of the depicted vessel area (y-axis: number of pixels) and procedural time (x-axis: image number) are shown.

Fig. 3 A plot of image number versus the relative rate of vessel area change per unit time (ΔVA) during an exemplar surgical trial. The dot lines (± 1.13) represent the mean ± 1.96 × standard deviation of ΔVA for all trials, which define the tissue deformation error (TDE) by surgeon. Arrows indicate TDEs.

Fig. 4 Comparison of the parameters assessed within the total trial and each phase, including the task completion time, the coefficient of variation of vessel area (CV-VA), the maximum absolute value of ΔVA (Max-ΔVA), and the number of TDEs, between expert and novice surgeons.
*:p<0.05. **:p<0.01.

Fig. 5

A: Receiver operating curves for each parameter within a trial.

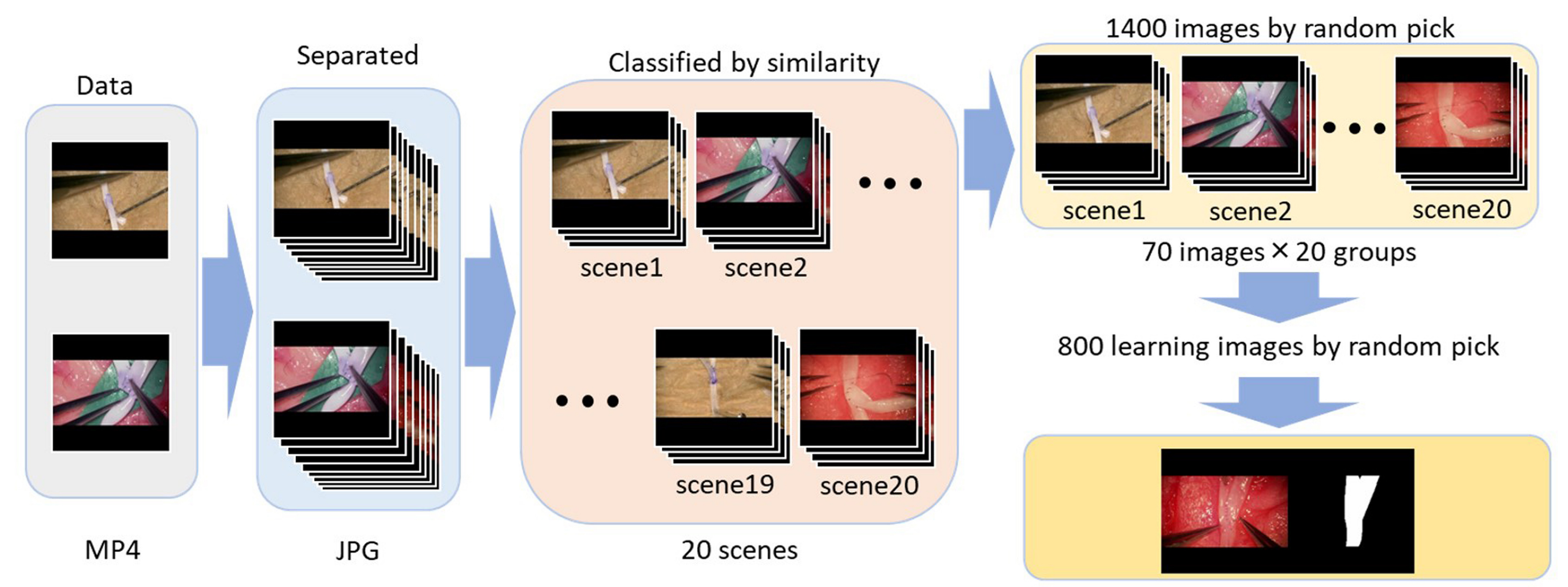
B: Average task completion time versus video parameters (CV-VA, Max-ΔVA, No. for TDEs) of each surgeon.

No.	Group	Board certification	PGY	Experience (N)		
				Microsurgery	EC-IC bypass	Laboratory training
1		yes	28	>2000	>300	>50
2		yes	20	>1500	>150	>50
3	Expert	yes	14	>250	>20	>30
4		yes	13	>900	>90	>30
5		yes	12	>600	>80	>30
6		no	3	<30	0	<20
7		no	3	<40	0	<15
8		no	2	<30	0	<5
9	Novice	no	2	<20	0	<5
10		no	2	<10	0	<5
11		no	1	<5	0	<2
12		no	1	<5	0	<10

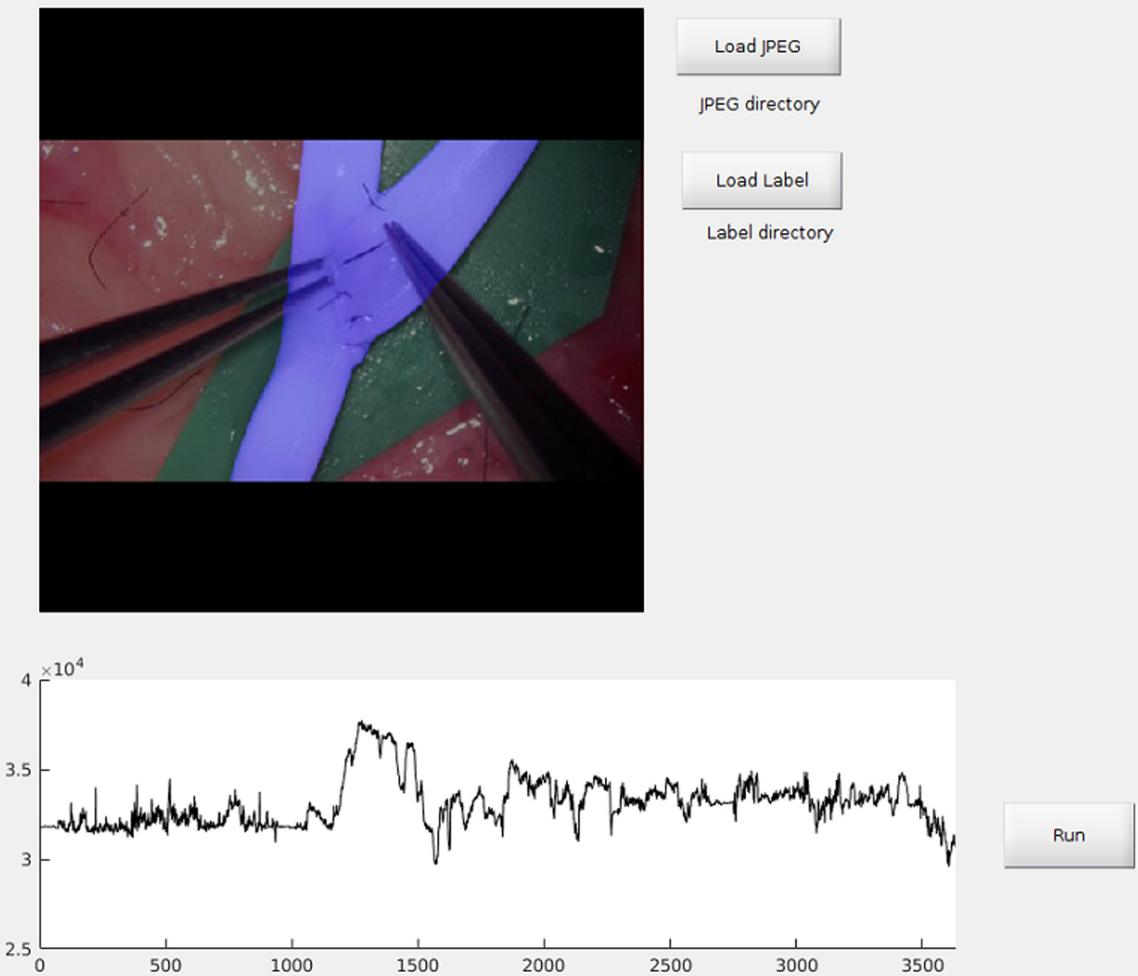
Table 1 Participating surgeons

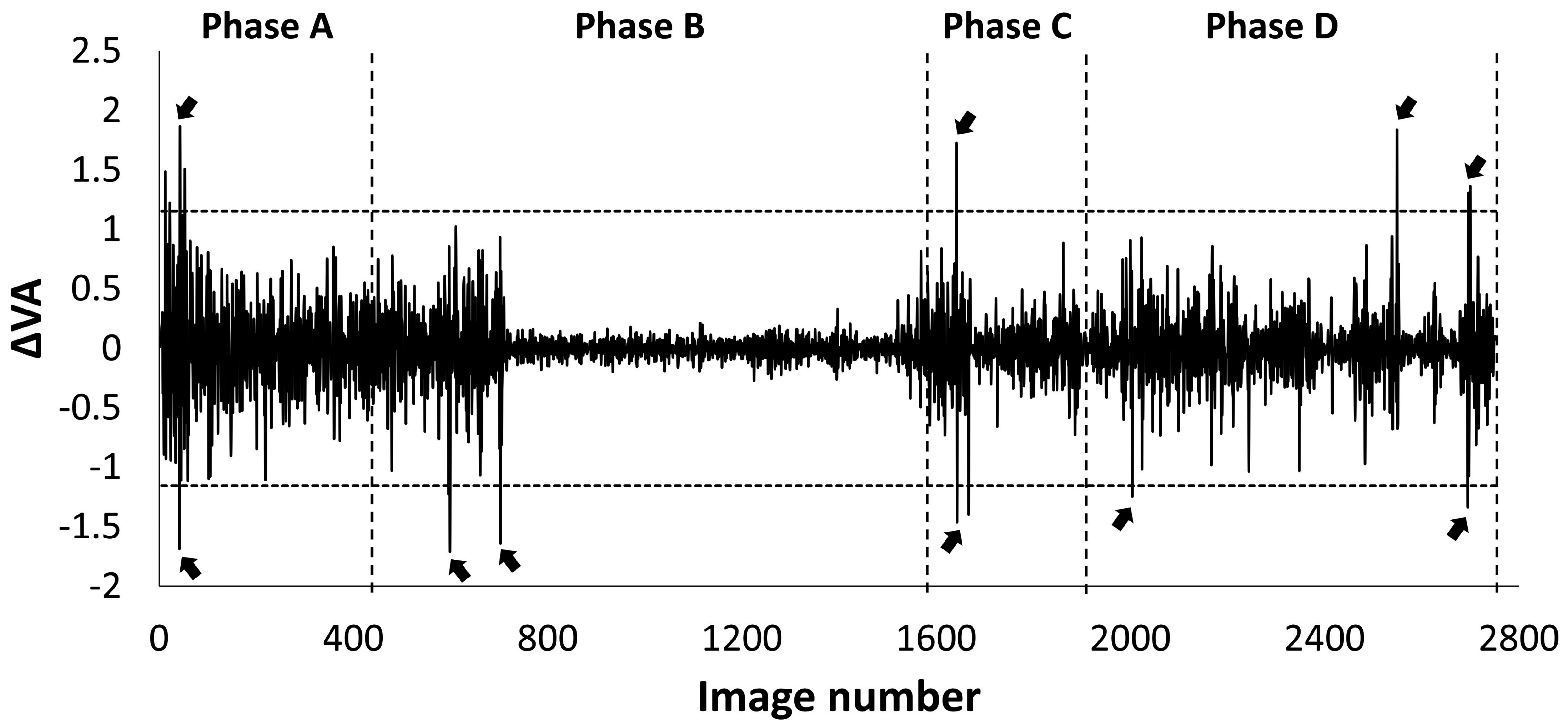
PGY, postgraduate year; N, number; EC-IC, extracranial-intracranial

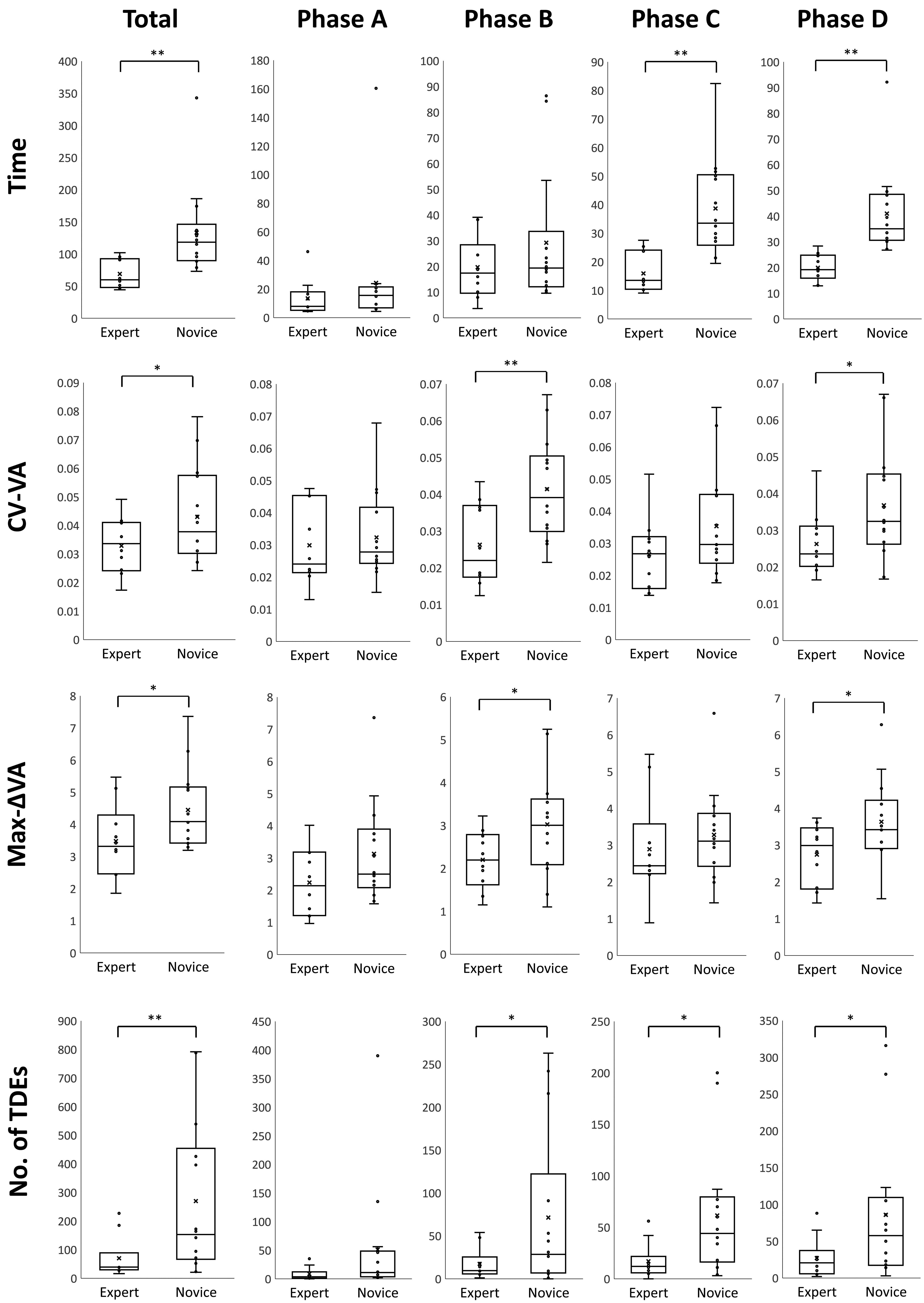
*Board certification: The Japan Neurosurgical Society and Japanese Society on Surgery for Cerebral Stroke

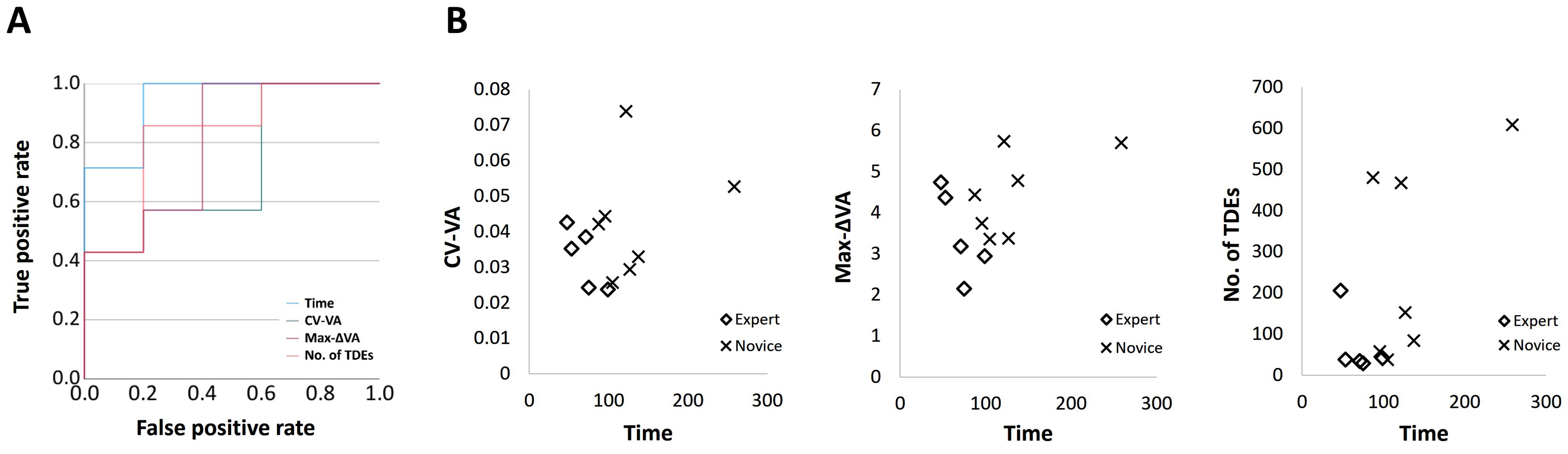


Data
MP4
Separated
JPG
Classified by similarity
scene1
scene2
scene19
scene20
20 scenes
1400 images by random pick
scene1
scene2
scene20
70 images × 20 groups
800 learning images by random pick


Load JPEG
JPEG directory
Load Label
Label directory
$\times 10^4$
4
3.5
3
2.5
0
500
1000
1500
2000
2500
3000
3500
Run


Phase A
Phase B
Phase C
Phase D
ΔVA
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
400
800
1200
1600
2000
2400
2800
Image number


Total
Phase A
Phase B
Phase C
Phase D
Time
CV-VA
Max-ΔVA
No. of TDEs
Expert
Novice
**
*


A
B
True positive rate
False positive rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.2 0.4 0.6 0.8 1.0
Time
CV-VA
Max-ΔVA
No. of TDEs
CV-VA
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0 100 200 300
Time
Expert
Novice
Max-ΔVA
7
6
5
4
3
2
1
0
0 100 200 300
Time
Expert
Novice
No. of TDEs
700
600
500
400
300
200
100
0
0 100 200 300
Time
Expert
Novice