Exploring Literature Surrounding Automatic Dental Segmentation

Oliver Smith

University Of Lincoln, School of Engineering and Physical Sciences

26357261@students.lincoln.ac.uk

Abstract. Recent advancements in artificial intelligence and machine learning have transformed medical imaging, including the automation of image analysis traditionally performed by trained professionals. This project focuses on creating a semantic segmentation model to classify dental structures in 2D intraoral X-rays.

This review addresses potential challenges, and explores how techniques such as transfer learning and data augmentation can be incorporated to improve model performance to meet clinical standards.

Keywords: Artificial Intelligence · Machine Learning · Semantic Segmentation · Medical Imaging · Transfer Learning · Pre-trained Models · Segment Anything Model · Domain Adaptation

Total Word Count: 1689

1 Literature Review

1.1 Background Introduction

The rapid advancement of artificial intelligence and machine learning has unlocked new possibilities across various domains, including medical imaging and diagnostics [2, 20, 24]. One application is using segmentation models to automate the analysis of radiological images typically requiring the expertise of highly trained professionals [20]. Traditional manual analysis of these images is time-consuming and prone to human error, worsened by the increasing workloads radiologists face [8].

This study explores the development of a semantic segmentation model for classifying dental structures in 2D intraoral X-rays. The goal is to create robust and reliable tool to provide dental professionals with accurate, automated image annotations. To achieve this, a pre-trained model will be fine-tuned on annotated dental images to improve performance in clinical settings [4].

1.2 Modern Medical Image Segmentation

Semantic segmentation has become a cornerstone of medical imaging [1], enabling pixel-wise extraction of anatomical structures to aid in diagnosis and treatment. Recent advancements in Convolutional Neural Networks (CNN) have revolutionised segmentation models, offering state-of-the-art performance on diverse medical datasets [29, 17, 27]. Notable architectures such as U-Net [23], ResNet [9], and vision transformers (ViT) [7] have been widely adopted for biomedical applications [2], with their unique encoder-decoder structures effectively capturing desired features.

Despite these advancements, current research has primarily focused on segmentation models for general medical imaging, with a significant gap for specialised applications in dental care [25, 19]. A potential reason for this is that the effectiveness of these models is still limited by the "quality and scale of data" [3]. While large medical imaging datasets, like those in ophthalmology, are becoming more accessible [11], publicly available dental imaging datasets remain relatively scarce [30]. Whether due to data protection concerns or the need for professional annotations, dental datasets often tend to be small, lack structure, and suffer from low variable completeness [19]. This may lead to issues like overfitting, causing poor performance when given new data [14].

1.3 Transfer Learning and Pre-Trained Models

Incorporating large pre-trained models can use knowledge gained from extensive datasets to overcome the challenges of limited data, and achieve better performance over models trained from scratch [32]. This approach, known as transfer learning, has been shown to enhance model efficiency and accuracy, especially with smaller, less diverse datasets [28, 4].

The Segment Anything Model (SAM) [12] became a significant milestone for pre-trained segmentation models by adopting a general-purpose approach rather than being tailored to any specific domain. Unlike models specifically designed for biomedical applications with carefully crafted encoder-decoder pipelines, SAM stands out for its flexibility. It can segment any image from user prompts like bounding boxes or text. Although not designed exclusively for medical imaging, SAM's extensive pretraining and ability to perform zero-shot transfer to new image distributions [12] make it a versatile tool for a wide range of segmentation challenges, requiring minimal domain-specific tuning to achieve accuracy.

SAM2 is the next iteration of SAM but with a focus on video segmentation, where it shows notable improvements in accuracy and efficiency. SAM2 achieves better results with 3x fewer interactions in video segmentation tasks and is 6x faster in image segmentation compared to SAM [22]. However, when applied to medical imaging, the performance of SAM and SAM2 is comparable, with both models demonstrating strong capabilities in handling complex medical images [26].

Attempts to apply SAM and SAM2 to medical image analysis have yielded promising results [13, 14, 10, 31, 32]. A notable example of this is the MEDSAM model, which uses domain adaptation techniques to adapt SAM to specialised medical datasets, making SAM and SAM2 "inferior to MedSAM for most 2D medical image modalities" [15]. Unfortunately, MedSAM does not include dental imaging in the training set, so it will require refinement to achieve optimal results.

1.4 Comparison to Existing Approaches

In their 2024 paper, He Zhicheng et al. [32] demonstrate that MedSAM offers significant improvements over SAM, especially in dental imaging. However, the study's relatively small training dataset may have affected the model's generalisation, as it doesn't fully represent all potential variations of impacted teeth. The authors suggest that further research is required to test MedSAM's performance on larger, more diverse datasets to ensure clinical relevance. This project addresses this by sourcing a larger, more diverse dataset and using data augmentation techniques to increase variability making the model more suitable for clinical environments.

Many studies that use SAM or its variants focus primarily on the model's generalisation and one-shot capabilities, often without tailoring the model to specific domain challenges [16, 26]. This lack of adaptation can result in suboptimal performance in specialised tasks. This project builds on these studies by directly applying SAM's adaptability to dental imaging through domain adaptation, rather than relying solely on the model's base performance.

A 2023 paper by Wu et al. [31] explores a unique method for adapting SAM to medical image segmentation. Their approach updates "only 2% of SAM's tunable parameters" while still achieving improved performance over MedSAM and other fine-tuned models. This study further proves the potential of producing highly accurate segmentations with minimal model modifications. Despite these impressive results, their study, like many others [13, 6], focuses on 3D scans rather than 2D dental X-rays, which may limit use in clinical settings. This project uses a combination of 2D panoramic, bitewing, and periapical radiographs, which are some of the most commonly available imaging formats in dental practices [5]. By focusing on the most readily available technologies, this project better suits clinical scenarios, making it more applicable to real-world dental practices.

1.5 Conclusion

Building on the strengths of pre-trained models, this project aims to utilise their feature extraction capabilities and adapt them for dental image segmentation. By fine-tuning these models on a large dataset of annotated dental X-rays, the goal is to improve segmentation performance to meet clinical standards.

This approach reflects a growing trend in AI research of using transfer learning to adapt general models for specialised applications and balance the benefits of pretraining with domain-specific needs [4].

2 Progress Update

2.1 Objective Stages

NOTE: *I have used shortened objectives here. Please find the full objectives in the project proposal.*

1. **Collect and preprocess a dataset of at least 200 images.**

   Stage: Completed.
   Approximately 900 high-quality images including seven classes have been collected to test, train, and validate the model. The images were resampled to 1024×1024 for use with pre-trained models.

2. **Apply data augmentation techniques to the dataset to improve the model's generalisation.**

   Stage: Completed.
   Augmentations were applied to all images including flips, rotations, and adjustments to saturation and exposure.

3. **Organise user testing sessions with dental professionals to receive feedback.**

   Stage: In Progress.
   This project was discussed with my dentist, who showed interest and suggested sending an email to the practice during user testing. While not a firm confirmation, it's a promising lead for obtaining professional feedback.

4. **Conduct a literature review.**

   Stage: Completed.
   The literature review addresses these points and positions this project within the context of relevant prior work.

5. **Document and present findings in a dissertation paper.**

   Stage: Incomplete.
   This project is not yet complete enough to summarise findings in a dissertation paper.

2.2 Revised Objectives:

1. **Initial Objective: Develop a U-Net-based semantic segmentation model capable of identifying and categorising dental structures with a target pixel accuracy of above 80%**

   Revised Objective: Fine-tune a pre-trained semantic segmentation model (e.g., SAM, SAM2, or MEDSAM) on a labelled dataset of dental X-ray images to identify and categorise dental structures (such as teeth, fillings, implants), with a target IoU score of above 80%.

   Justification: Pre-trained models provide a more robust foundation for accurate segmentation, offering superior segmentation and generalisation over a U-Net-based model developed from scratch; especially with limited data [28, 4]. While U-Net is reliable, using a cutting-edge pre-trained model ensures state-of-the-art performance aligning better with the project's objectives.

   IoU is used for stricter evaluation, ensuring precise feature extraction vital for medical diagnosis [18].

2. **Initial Objective: Perform hyperparameter tuning for optimised segmentation accuracy by systematically adjusting parameters such as the loss function, optimiser, and number of convolutional filters.**

   Revised Objective: Perform hyperparameter tuning for dental segmentation tasks by systematically adjusting key fine-tuning parameters, such as the learning rate, batch size, and loss function, to achieve an IoU score improvement of at least 5% compared to the baseline performance.

   Justification: Revised to focus on domain adaptation of a pre-trained model rather than a custom U-net model. Architecture-specific parameters like convolutional filters have been replaced by optimisation of relevant fine-tuning parameters.

2.3 Current Progress:

Finding a suitable dataset was challenging due to the need for high-quality images, correct and balanced class types, and properly formatted annotations. To address this, three Python scripts were created:

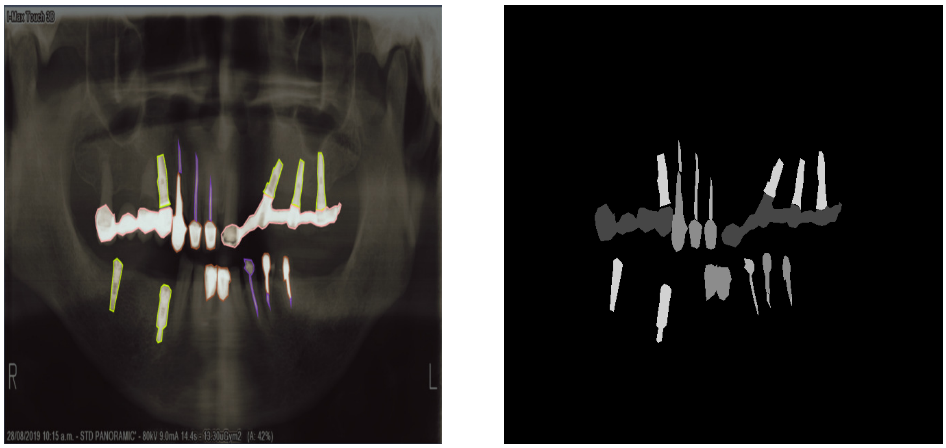
– **Image augmentation:** Iterates through each image, applying random augmentations such as horizontal/vertical flips, 90° rotations, and adjustments to saturation and exposure.

(a) Non-Augmented Image.

(b) Augmented Image.

Fig. 1: Example of an augmentation that could be applied to an image.

- **COCO JSON to Segmentation Mask Converter:** Converts polygon-based annotations to rasterized segmentation masks for seven predefined classes.
  This ensures the dataset is suitable for the project as limiting the search to pixel-wise segmentation masks isn't viable. COCO JSON was chosen for its common use.



(a) Polygon-Based Bounding boxes. (b) Generated Ground Truth Masks.

Fig. 2: Example conversion from unusable polygon-based annotations (a), to usable ground truth masks (b).

– **Dataset Analysis:** Analyses annotations to reveal class distribution, pixel coverage, and co-occurrence, visualised with Matplotlib and Seaborn to identify imbalances, guide augmentation, and understand class relationships.

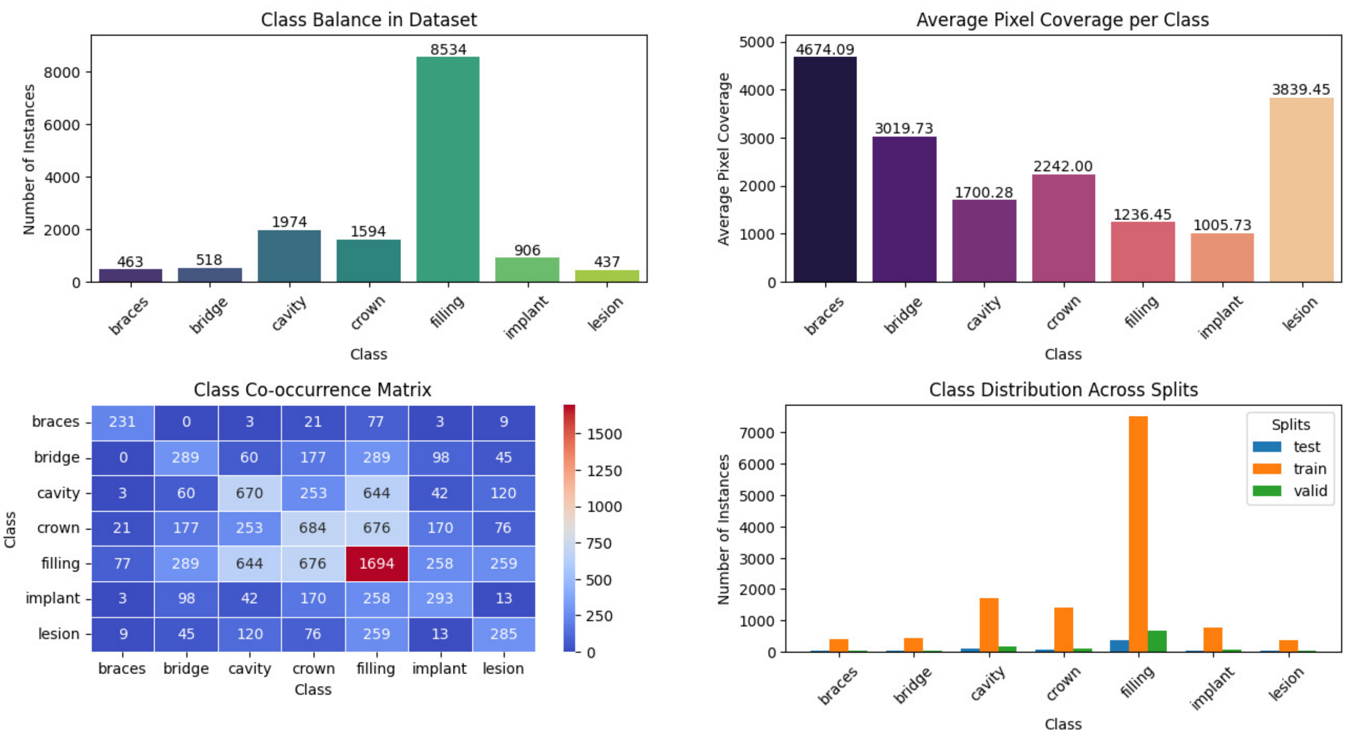


Fig. 3: A selection of graphs output by the script.

The impact of the class imbalance on model generalisation will be monitored. If issues arise, undersampling, class-weighted loss, or stratified cross-validation will ensure minority classes are adequately represented in each training and testing fold [21].

A pipeline is being developed to fine-tune a pre-trained model using the collected and augmented dataset, though further improvements are needed to achieve target performance. Currently, a PyTorch implementation of SAM is being used, but a recent study showed comparable performance between SAM and SAM2 [15], so the next iteration will directly compare both models to identify the most effective for dental X-rays.

Given its superior performance over SAM and SAM2 [15, 14], MEDSAM is planned for future iterations. However, its implementation is deferred until a later stage as it currently lacks a pre-built fine-tuning pipeline and further research is required to implement the model manually.

2.4 Reflection of supervisory engagement and project management:

Mamatha and I meet weekly for 15–30 minutes to discuss progress, upcoming tasks, and questions. I prepare notes and document key points during the meeting. She's been incredibly helpful, offering guidance and detailed answers to all my queries.

Despite balancing assignments, work, and postgraduate applications, I've managed my time effectively, dedicating free moments to research and project development. While I'm a few weeks behind on my Gantt chart, I have sufficient buffer time to catch up. To stay on track, I've created a kanban board to prioritise tasks, set daily goals, and allocate focused time for the project.

With semester one now over and more free time available, I look forward to dedicating more time to this project and getting ahead before the next exam season. I plan to improve time management by regularly reviewing my Gantt chart and adjust deadlines to reflect realistic progress.

References

1. Aggarwal, P., Vig, R., Bhadoria, S., Dethe, C.: Role of segmentation in medical imaging: A comparative study. International Journal of Computer Applications **29**(1), 54–61 (2011)
2. Azad, R., Aghdam, E.K., Rauland, A., Jia, Y., Avval, A.H., Bozorgpour, A., Karimijafarbigloo, S., Cohen, J.P., Adeli, E., Merhof, D.: Medical image segmentation review: The success of u-net. IEEE Transactions on Pattern Analysis and Machine Intelligence **46**(12), 10076–10095 (2024). https://doi.org/10.1109/TPAMI.2024.3435571
3. Chen, Q., Zhao, Y., Liu, Y., Sun, Y., Yang, C., Li, P., Zhang, L., Gao, C.: Mslp-net: multi-scale location perception network for dental panoramic x-ray image segmentation. Neural Computing and Applications **33**(16), 10277–10291 (2021). https://doi.org/10.1007/s00521-021-05790-5, https://doi.org/10.1007/s00521-021-05790-5
4. Cheplygina, V., de Bruijne, M., Pluim, J.P.: Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Medical Image Analysis **54**, 280–296 (2019). https://doi.org/https://doi.org/10.1016/j.media.2019.03.009, https://www.sciencedirect.com/science/article/pii/S1361841518307588
5. Clinic, C.: Dental x-rays (2024), https://my.clevelandclinic.org/health/diagnostics/11199-dental-x-rays
6. Cui, Z., Li, C., Chen, N., Wei, G., Chen, R., Zhou, Y., Shen, D., Wang, W.: Tsegnet: An efficient and accurate tooth segmentation network on 3d dental model. Medical Image Analysis **69**, 101949 (2021). https://doi.org/https://doi.org/10.1016/j.media.2020.101949, https://www.sciencedirect.com/science/article/pii/S1361841520303133
7. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An image is worth 16x16 words: Transformers for image recognition at scale (2021), https://arxiv.org/abs/2010.11929
8. Fitzgerald, R.: Error in radiology. Clinical radiology **56**(12), 938–946 (2001)
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition (2015), https://arxiv.org/abs/1512.03385
10. Hu, X., Xu, X., Shi, Y.: How to efficiently adapt large segmentation model(sam) to medical images (2023), https://arxiv.org/abs/2306.13731
11. Khan, S.M., Liu, X., Nath, S., Korot, E., Faes, L., Wagner, S.K., Keane, P.A., Sebire, N.J., Burton, M.J., Denniston, A.K.: A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability. The Lancet Digital Health **3**(1), e51–e66 (2021). https://doi.org/10.1016/S2589-7500(20)30240-5, https://doi.org/10.1016/S2589-7500(20)30240-5, doi: 10.1016/S2589-7500(20)30240-5
12. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al.: Segment anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4015–4026 (2023)
13. Liao, J., Wang, H., Gu, H., Cai, Y.: Ppa-sam: Plug-and-play adversarial segment anything model for 3d tooth segmentation. Applied Sciences **14**(8) (2024), https://www.mdpi.com/2076-3417/14/8/3259

14. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. Nature Communications **15**(1), 654 (2024). https://doi.org/10.1038/s41467-024-44824-z, https://doi.org/10.1038/s41467-024-44824-z
15. Ma, J., Kim, S., Li, F., Baharoon, M., Asakereh, R., Lyu, H., Wang, B.: Segment anything in medical images and videos: Benchmark and deployment. arXiv preprint arXiv:2408.03322 (2024)
16. Mazurowski, M.A., Dong, H., Gu, H., Yang, J., Konz, N., Zhang, Y.: Segment anything model for medical image analysis: An experimental study. Medical Image Analysis **89**, 102918 (Oct 2023). https://doi.org/10.1016/j.media.2023.102918, http://dx.doi.org/10.1016/j.media.2023.102918
17. Müller, D., Kramer, F.: Miscnn: a framework for medical image segmentation with convolutional neural networks and deep learning. BMC Medical Imaging **21**(1), 12 (2021). https://doi.org/10.1186/s12880-020-00543-7, https://doi.org/10.1186/s12880-020-00543-7
18. Müller, D., Soto-Rey, I., Kramer, F.: Towards a guideline for evaluation metrics in medical image segmentation. BMC Research Notes **15**(1), 210 (2022). https://doi.org/10.1186/s13104-022-06096-y, https://doi.org/10.1186/s13104-022-06096-y
19. Panetta, K., Rajendran, R., Ramesh, A., Rao, S.P., Agaian, S.: Tufts dental database: A multimodal panoramic x-ray dataset for benchmarking diagnostic systems. IEEE Journal of Biomedical and Health Informatics **26**(4), 1650–1659 (2022). https://doi.org/10.1109/JBHI.2021.3117575
20. Qureshi, I., Yan, J., Abbas, Q., Shaheed, K., Riaz, A.B., Wahid, A., Khan, M.W.J., Szczuko, P.: Medical image segmentation using deep semantic-based methods: A review of techniques, applications and emerging trends. Information Fusion **90**, 316–352 (2023). https://doi.org/https://doi.org/10.1016/j.inffus.2022.09.031, https://www.sciencedirect.com/science/article/pii/S1566253522001695
21. Raschka, S.: How can the f1-score help with dealing with class imbalance. Retrieved from Sebastian Raschka: https://sebastianraschka. com/faq/docs/computing-the-f1-score. html
22. Ravi, N., Gabeur, V., Hu, Y.T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., et al.: Sam 2: Segment anything in images and videos. arXiv preprint arXiv:2408.00714 (2024)
23. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. pp. 234–241. Springer (2015)
24. Rubin, D.L.: Artificial intelligence in imaging: the radiologist's role. Journal of the American College of Radiology **16**(9), 1309–1317 (2019)
25. Schwendicke, F., Samek, W., Krois, J.: Artificial intelligence in dentistry: Chances and challenges. Journal of Dental Research **99**(7), 769–774 (2020). https://doi.org/10.1177/0022034520915714, https://doi.org/10.1177/0022034520915714, pMID: 32315260
26. Sengupta, S., Chakrabarty, S., Soni, R.: Is sam 2 better than sam in medical image segmentation? (2024), https://arxiv.org/abs/2408.04212
27. Seo, H., Khuzani, M.B., Vasudevan, V., Huang, C., Ren, H., Xiao, R., Jia, X., Xing, L.: Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. Medical Physics **47**(5), e148–e167 (2020). https://doi.org/10.1002/mp.13649

28. Stuckner, J., Harder, B., Smith, T.M.: Microstructure segmentation with deep learning encoders pre-trained on a large microscopy dataset. npj Computational Materials **8**(1), 200 (2022). https://doi.org/10.1038/s41524-022-00878-5, https://doi.org/10.1038/s41524-022-00878-5
29. Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J.: Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging **35**(5), 1299–1312 (2016). https://doi.org/10.1109/TMI.2016.2535302
30. Uribe, S., Issa, J., Sohrabniya, F., Denny, A., Kim, N., Dayo, A., Chaurasia, A., Sofi-Mahmudi, A., Büttner, M., Schwendicke, F.: Publicly available dental image datasets for artificial intelligence. Journal of Dental Research **103**(13), 1365–1374 (2024). https://doi.org/10.1177/00220345241272052, https://doi.org/10.1177/00220345241272052, pMID: 39422586
31. Wu, J., Ji, W., Liu, Y., Fu, H., Xu, M., Xu, Y., Jin, Y.: Medical sam adapter: Adapting segment anything model for medical image segmentation (2023), https://arxiv.org/abs/2304.12620
32. Zhicheng, H., Yipeng, W., Xiao, L.: Deep learning-based detection of impacted teeth on panoramic radiographs. Biomedical Engineering and Computational Biology **15**, 11795972241288319 (2024). https://doi.org/10.1177/11795972241288319, https://doi.org/10.1177/11795972241288319