

Machine Learning-driven Histotype Diagnosis of Ovarian Carcinoma: Insights from the OCEAN AI Challenge

Maryam Asadi-Aghbolaghi^{1*}, Hossein Farahani^{1*}, Allen Zhang^{2,3*}, Ardalan Akbari^{2,3}, Sirim Kim^{2,3}, Ashley Chow⁴, Sohier Dane⁴, OCEAN Challenge Consortium, OTTA Consortium, David G Huntsman², C Blake Gilks³, Susan Ramus⁵, Martin Köbel⁶, Anthony N Karnezis^{7,8}, Ali Bashashati^{1,2,8†}

¹ School of Biomedical Engineering, University of British Columbia, Vancouver, BC, Canada

² Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, Canada

³ Vancouver General Hospital, Vancouver, BC, Canada

⁴ Kaggle Inc, Mountain View, CA, USA

⁵ School of Clinical Medicine, University of New South Wales, Australia

⁶ Department of Pathology and Laboratory Medicine, University of Calgary, Alberta, Canada

⁷ Department of Pathology and Laboratory Medicine, UC Davis, Sacramento, USA

⁸ Joint senior authors

† Corresponding author

Correspondence: Ali Bashashati. Email: ali.bashashati@ubc.ca. Mailing address: 2222 Health Sciences 25 Mall, Vancouver, BC V6T 1Z3 Canada

* These authors contributed equally to this work

Abstract

Ovarian cancer poses a significant health burden as one of the deadliest malignancies affecting women globally. Histotype assignment of epithelial ovarian cancers can be challenging due to morphologic overlap, inter-observer variability, and the lack of ancillary diagnostic techniques in some areas of the world. Moreover, rare cancers can pose particular diagnostic difficulties because of a relative lack of familiarity with them, underscoring the necessity for robust diagnostic methodologies. The

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

emergence of Artificial Intelligence (AI) has brought promising prospects to the realm of ovarian cancer diagnosis. While various studies have underscored AI's promise, its validation across multiple healthcare centers and hospitals has been limited. Inspired by innovations in medical imaging driven by public competitions, we initiated the Ovarian Cancer subtype Classification and outlier detection (OCEAN) challenge — the most extensive histopathology competition to date.

Introduction

Epithelial ovarian cancers are some of the most lethal gynecologic malignancies in North America and across the world [1]. This disease exhibits marked heterogeneity, characterized by five major histotypes: high-grade serous carcinoma (HGSC), constituting 70% of cases (and 90% of advanced-stage disease and mortality); clear cell ovarian carcinoma (CCOC), representing 12%; endometrioid (ENOC) at 11%; low-grade serous (LGSC) at 4%; and mucinous carcinoma (MUC) at 3% [2, 3], and several additional rare subtypes histotypes. Clinical management varies greatly between histotypes, with high-grade serous carcinomas treated most aggressively with combination platinum-taxane chemotherapy and a subset benefitting from PARP inhibitor therapy and MEK inhibitors [4]. Endometrioid and clear cell carcinomas are associated with mismatch repair deficiency, which can be sporadic or inherited due to Lynch syndrome. Accurate diagnosis histotyping is critical for risk assessment, hereditary cancer screening, and clinical trial enrollment. However, histological classification of ovarian carcinomas by pathologists still suffers from suboptimal interobserver and intraobserver reproducibility outside of gynecologic pathology subspecialty practice and tertiary academic centers [3, 5–8].

Initial diagnosis relies on histological examination of hematoxylin and eosin (H&E)-stained tissue sections, yet studies indicate that among pathologists lacking specialty training in gynecologic pathology, the interobserver agreement in diagnosis remains only moderate (0.54–0.67 Cohen's kappa) [9, 10]. While adjunct diagnostic techniques such as immunohistochemistry and next-generation sequencing can improve diagnostic accuracy and interobserver reproducibility [7, 10–12], these techniques are unavailable in much of the world [13, 14]. Additionally, there is a pronounced shortage of pathologists compared to the demand, leading to substantial variances in pathologist numbers among countries [15]. Even in well-equipped nations, the demand for pathologists surpasses the available supply [14, 16]. Addressing these challenges is imperative to ensure effective diagnosis and management of ovarian carcinoma patients [17].

In recent years, the integration of AI algorithms into the field of medical diagnostics has shown considerable promise [18–23], particularly in aiding pathologists with the histological classification of ovarian cancer [24, 25]. However, despite their potential, AI algorithms are not immune to biases that may arise during their development and validation processes. These biases can manifest in algorithms that perform suboptimally when applied to datasets beyond those used in their initial training, highlighting the critical need for robust methodologies to assess their generalizability. To facilitate this endeavor, we launched the Ovarian Cancer subtype Classification and outlier detection (OCEAN) challenge, a global competition that provided participants with access to the largest and most diverse public histopathology dataset of ovarian cancer to date.

The OCEAN dataset comprises approximately 2500 samples, encompassing both whole slide images and tissue microarrays sourced from over 20 centers across multiple countries. Variations in patient demographics, tissue processing, and H&E slide staining protocols across pathology labs contribute to diversity in the dataset. The extensive variation in color among H&E slide samples presents a unique opportunity for assessing the generalizability and robustness of algorithms versus those trained on slides from a single center [25]. Importantly, since the evaluation is conducted independently of algorithm development, we mitigate the risk of information leakage.

Through the OCEAN challenge, our dedication lies in expediting progress and establishing the foundation for developing AI solutions that make tangible clinical differences in diagnosing and managing ovarian cancer. By making this expansive dataset as well as top-performing AI models publicly accessible, our goal is to provide a significant resource for advancing research in ovarian cancer diagnosis and treatment, ultimately leading to improved patient outcomes, in particular for resource-limited practice settings.

Results

Dataset

The OCEAN dataset comprises 2,438 images distributed across three distinct sets: training, public test, and private test, sourced from 24 centers that were mainly a part of the ovarian tumor tissue analysis (OTTA) consortium [26]. This dataset encompasses H&E images from both Whole Slide Images (WSIs) and Tissue Microarrays (TMAs). Specifically, 538 images are allocated to the training set, while 437 and 1,488 images are designated for the public and private test sets, respectively. Within these images, the five main subtypes of ovarian carcinoma are represented: CCOC, ENOC, HGSC, LGSC, and MUC. Notably, the public and private test sets collectively contain 147 outlier images (i.e., other), comprising rare ovarian cancer subtypes along with normal tissues. For further details regarding the composition of the OCEAN dataset, please refer to Table 1.

Table 1: Overview of the OCEAN dataset.

Set	Subtype	Train	Public Test	Private Test
TMAs	CCOC	5	64	175
	ENOC	5	41	210
	HGSC	5	51	492
	LGSC	5	36	212
	MUC	5	20	56
	Other	5	31	44
WSIs	CCOC	94	23	61
	ENOC	119	35	70
	HGSC	217	68	71
	LGSC	42	10	39
	MUC	41	9	35
	Other	0	49	23
Total	All	538	437	1,488

OCEAN Competition

The OCEAN competition commenced on October 6th, 2023, and concluded on January 3rd, 2024, spanning a duration of three months. This competition was hosted on the Kaggle platform, a venue for data science challenges and collaborations. The event garnered significant interest, amassing a total of 9,247 registrations from participants worldwide. A total of 1,772 individuals comprising 1,326 teams representing 84 countries actively engaged in the competition. Throughout the duration of the competition, participants collectively submitted 35,279 entries in pursuit of refining their algorithms. Throughout the competition, participants had the chance to evaluate the effectiveness of their algorithms using the publicly provided test set. Notably, the leaderboard standings were determined based on the performance of algorithms on this public test set. Subsequently, winners were selected based on their algorithms' performance on a separate private test set after the competition closing, ensuring a fair and unbiased evaluation process.

Performance on Public and Private Datasets

Balanced accuracy served as the primary metric for assessing the efficacy of the submitted methodologies. Balanced accuracy, calculated as the average recall across all classes, addresses the issue of class imbalance by providing a comprehensive assessment of performance. Participants were required to assign each image to one of the five ovarian cancer histotypes or designate it as an "other" category within their submission files. Among the top 10 performing submissions on the public test set, balanced accuracy ranged from 61% to 68%. On the private test set, the top 10 performances exhibited balanced accuracies spanning from 58% to 66%.

Conclusion and Discussion

The OCEAN challenge represents a significant step forward in the pursuit of accurate and robust AI solutions for the classification of ovarian carcinoma histotypes and the detection of outliers. The competition, hosted on the Kaggle platform, attracted substantial global participation, highlighting the widespread interest and commitment to advancing medical diagnostics through AI technologies.

The OCEAN dataset stands as a milestone in the field, emerging as the largest competition on histopathology images to date. It offers a comprehensive collection of histopathology images of ovarian carcinoma, unparalleled in its size and diversity. By subjecting algorithms to rigorous evaluation across datasets sourced from numerous hospitals, the OCEAN challenge sought to fill a crucial gap in the field. Specifically, it aimed to develop methodologies capable of robust generalization across diverse patient demographics, and digital slide scanners, tissue processing, and staining protocols across pathology labs.

Moving forward, the insights gained from the OCEAN challenge serve as a valuable foundation for further research and development efforts aimed at refining AI algorithms for the accurate diagnosis and classification of ovarian carcinoma histotypes.

Data Availability

Data can be accessed and downloaded from the Kaggle challenge page (<https://www.kaggle.com/competitions/UBC-OCEAN>).

Acknowledgements

This work was supported by BC Cancer Foundation, CIHR (No 418734), NSERC (RGPIN-2019-04896), and Health Research BC grants to AB.

References

- [1] J. Ferlay, M. Ervik, F. Lam, M. Laversanne, M. Colombet, L. Mery, M. Piñeros, A. Znaor, I. Soerjomataram, and F. Bray, “Global cancer observatory: Cancer today,” Lyon, France: International Agency for Research on Cancer, 2024, available from: <https://gco.iarc.who.int/today>, accessed [08 04 2024].
- [2] H. Moch, “Who classification of tumours editorial board. female genital tumours: Who classification of tumours,” 2020.
- [3] M. Köbel, S. E. Kalloger, N. Boyd, S. McKinney, E. Mehl, C. Palmer, S. Leung, N. J. Bowen, D. N. Ionescu, A. Rajput *et al.*, “Ovarian carcinoma subtypes are different diseases: implications for biomarker studies,” *PLoS medicine*, vol. 5, no. 12, p. e232, 2008.
- [4] D. M. Gershenson, A. Miller, W. E. Brady, J. Paul, K. Carty, W. Rodgers, D. Millan, R. L. Coleman, K. N. Moore, S. Banerjee *et al.*, “Trametinib versus standard of care in patients with recurrent low-grade serous ovarian cancer (gog 281/logs): an international, randomised, open-label, multicentre, phase 2/3 trial,” *The Lancet*, vol. 399, no. 10324, pp. 541–553, 2022.
- [5] M. Krasovitsky, Y. C. Lee, H.-W. Sim, T. Chawla, H. Moore, D. Moses, L. Baker, C. Mandel, A. Kiellar, A. Hartery *et al.*, “Interobserver and intraobserver variability of recist assessment in ovarian cancer,” *International Journal of Gynecologic Cancer*, vol. 32, no. 5, 2022.
- [6] R. K. Matsuno, M. E. Sherman, K. Visvanathan, M. T. Goodman, B. Y. Hernandez, C. F. Lynch, O. B. Ioffe, D. Horio, C. Platz, S. F. Altekruise *et al.*, “Agreement for tumor grade of ovarian carcinoma: analysis of archival tissues from the surveillance, epidemiology, and end results residual tissue repository,” *Cancer causes & control*, vol. 24, pp. 749–757, 2013.
- [7] M. Köbel, J. Bak, B. I. Bertelsen, O. Carpen, A. Grove, E. S. Hansen, A.-M. Levin Jakobsen, M. Lidang, A. Måsbäck, A. Tolf *et al.*, “Ovarian carcinoma histotype determination is highly reproducible, and is improved through the use of immunohistochemistry,” *Histopathology*, vol. 64, no. 7, pp. 1004–1013, 2014.
- [8] M. E. Barnard, A. Pyden, M. S. Rice, M. Linares, S. S. Tworoger, B. E. Howitt, E. E. Meserve, and J. L. Hecht, “Inter-pathologist and pathology report agreement for ovarian tumor characteristics in the nurses’ health studies,” *Gynecologic oncology*, vol. 150, no. 3, pp. 521–526, 2018.

- [9] C. Patel, B. Harmon, R. Soslow, K. Garg, D. DeLair, S. Hwang, J. Liu, S. Zee, K. Shroyer, S. Burke *et al.*, “Interobserver agreement in the diagnosis of ovarian carcinoma types: Impact of sub-specialization,” in *Laboratory Investigation*, vol. 92. Nature Publishing Group 75 VARICK ST, 9TH FLR, NEW YORK, NY 10013-1917 USA, 2012, pp. 292A–292A.
- [10] M. Köbel, S. E. Kalloger, S. Lee, M. A. Duggan, L. E. Kelemen, L. Prentice, K. R. Kalli, B. L. Fridley, D. W. Visscher, G. L. Keeney *et al.*, “Biomarker-based ovarian carcinoma typing: a histologic investigation in the ovarian tumor tissue analysis consortium,” *Cancer epidemiology, biomarkers & prevention*, vol. 22, no. 10, pp. 1677–1686, 2013.
- [11] M. Köbel, K. Rahimi, P. F. Rambau, C. Naugler, C. Le Page, L. Meunier, M. De Ladurantaye, S. Lee, S. Leung, E. L. Goode *et al.*, “An immunohistochemical algorithm for ovarian carcinoma typing,” *International Journal of Gynecological Pathology*, vol. 35, no. 5, pp. 430–441, 2016.
- [12] N. Missaoui, S. Salhi, A. Bdioui, S. Mestiri, N. Abdessayed, M. Mokni, and M. T. Yacoubi, “Immunohistochemical characterization improves the reproducibility of the histological diagnosis of ovarian carcinoma,” *Asian Pacific Journal of Cancer Prevention: APJCP*, vol. 19, no. 9, p. 2545, 2018.
- [13] L. Gautron, “What is happening to immunohistochemistry?” *BioEssays*, vol. 45, no. 8, p. 2300056, 2023.
- [14] K. Patel, D. Chumba, R. M. Strother, F. Ndiangui, W. Jacobson, C. Dodson, R. W. Strate, J. W. Smith, and M. B. Resnic, “Development of immunohistochemistry services for cancer care in western kenya: Implications for low-and middle-income countries,” *African Journal of Laboratory Medicine*, vol. 5, no. 1, pp. 1–7, 2016.
- [15] M. L. Wilson, K. A. Fleming, M. A. Kutu, L. M. Looi, N. Lago, and K. Ru, “Access to pathology and laboratory medicine services: a crucial gap,” *The Lancet*, vol. 391, no. 10133, pp. 1927–1938, 2018.
- [16] Royal College of Pathologists. (2018) Meeting pathology demand: histopathology workforce census. [Online]. Available: <https://www.rcpath.org/static/952a934d-2ec3-48c9-a8e6e00fcdca700f/Meeting-Pathology-Demand-Histopathology-Workforce-Census-2018.pdf>
- [17] G. Lujan, J. C. Quigley, D. Hartman, A. Parwani, B. Roehmholdt, B. Van Meter, O. Ardon, M. G. Hanna, D. Kelly, C. Sowards *et al.*, “Dissecting the business case for adoption and implementation of digital pathology: a white paper from the digital pathology association,” *Journal of Pathology Informatics*, vol. 12, no. 1, p. 17, 2021.
- [18] A. Darbandsari, H. Farahani, M. Wiens, D. Cochrane, A. Jamieson, D. Farnell, P. Ahmadvand, M. Douglas, S. Leung, P. Abolmaesumi *et al.*, “Artificial intelligence-based histopathology image analysis identifies a novel subset of endometrial cancers with distinct genomic features and unfavourable outcome,” *medRxiv*, pp. 2023–05, 2023.

- [19] J. Boschman, H. Farahani, A. Darbandsari, P. Ahmadvand, A. Van Spankeren, D. Farnell, A. B. Levine, J. R. Naso, A. Churg, S. J. Jones *et al.*, “The utility of color normalization for ai-based diagnosis of hematoxylin and eosin-stained pathology images,” *The Journal of Pathology*, 2021.
- [20] M. Y. Lu, T. Y. Chen, D. F. Williamson, M. Zhao, M. Shady, J. Lipkova, and F. Mahmood, “Ai-based pathology predicts origins for cancers of unknown primary,” *Nature*, vol. 594, no. 7861, pp. 106–110, 2021.
- [21] H. R. Tizhoosh and L. Pantanowitz, “Artificial intelligence and digital pathology: challenges and opportunities,” *Journal of pathology informatics*, vol. 9, no. 1, p. 38, 2018.
- [22] B. E. Bejnordi, M. Veta, P. J. Van Diest, B. Van Ginneken, N. Karssemeijer, G. Litjens, J. A. Van Der Laak, M. Hermsen, Q. F. Manson, M. Balkenhol *et al.*, “Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer,” *Jama*, vol. 318, no. 22, pp. 2199–2210, 2017.
- [23] J. N. Kather, A. T. Pearson, N. Halama, D. Jäger, J. Krause, S. H. Loosen, A. Marx, P. Boor, F. Tacke, U. P. Neumann *et al.*, “Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer,” *Nature medicine*, vol. 25, no. 7, pp. 1054–1056, 2019.
- [24] Y. Wang, D. Farnell, H. Farahani, M. Nursey, B. Tessier-Cloutier, S. J. Jones, D. G. Huntsman, C. B. Gilks, and A. Bashashati, “Classification of epithelial ovarian carcinoma whole-slide pathology images using deep transfer learning,” *arXiv preprint arXiv:2005.10957*, 2020.
- [25] H. Farahani, J. Boschman, D. Farnell, A. Darbandsari, A. Zhang, P. Ahmadvand, D. Huntsman, M. Köbel, C. B. Gilks, N. Singh, and A. Bashashati, “Deep learning-based histotype diagnosis of ovarian carcinoma whole-slide pathology images,” *Modern Pathology*, 2022.
- [26] M. Köbel, E.-Y. Kang, A. Weir, P. F. Rambau, C.-H. Lee, G. S. Nelson, P. Ghatage, N. S. Meagher, M. J. Riggan, J. Alsop *et al.*, “p53 and ovarian carcinoma survival: an ovarian tumor tissue analysis consortium study,” *The Journal of Pathology: Clinical Research*, vol. 9, no. 3, pp. 208–222, 2023.