Nima Osman

No23451

Research proposal presentation

Transcript

# Transcript

**Slide 1: Title Page:** Good morning/afternoon. I'm Nima, presenting my research proposal: A Systematic and Reproducible Comparison of Vision Transformers and GANs for Breast Cancer Subtype Classification in Histopathology.

**Slide 2: Presentation Outline:** Today, I'll outline the challenges and opportunities in computational pathology. I'll highlight key literature gaps, present my research questions and objectives, and then detail the proposed methodology. This methodology emphasises reproducibility for breast cancer subtype classification using the TCGA-BRCA dataset. I will also cover the dataset selection, the project timeline, its anticipated contributions, and important ethical considerations.

**Slide 3: Introduction Title Page:** The Challenge and Opportunity in Computational Histopathology. [Pause]

**Slide 4: Introduction: Problem & Motivation:** Histopathology is foundational for cancer diagnosis but faces growing pressure from manual workload, subjectivity, and inefficiency. Deep learning, especially advanced models like Vision Transformers and Generative Adversarial Networks, offers significant potential, with many studies reporting impressive accuracy. However, translating this potential into reliable clinical tools is challenging. The research landscape is often hampered by methodological weaknesses such as poor comparability between studies, a lack of reproducibility, and evaluations that are too shallow. My proposal directly tackles these challenges, focusing on the important task of Breast Cancer Subtype Classification.

**Slide 5: The Knowledge Gap Title Page:** Critical Gaps Hindering Progress. [Pause]

**Slide 6: Literature Gaps & Motivation (Part 1): Performance & Comparison:** The literature shows these advanced models have distinct strengths. ViTs can achieve high classification accuracy, as Abimouloud reported 97 percent in 2025, and they also show robustness, even with difficult images, as Park and colleagues found in 2025. At the same time, GANs excel at data manipulation, like H&E-to-IHC translation shown by Saad in 2025, and they can improve classifier precision through data augmentation, demonstrated by Dee and colleagues in 2024. Despite these individual successes, a critical gap exists there's a lack of rigorous, systematic comparisons between these model types. It remains unclear which approach performs better for specific tasks like breast cancer subtyping, or what the crucial trade-offs are. Therefore, proposing a direct, methodologically sound comparison is essential to move the field forward.

**Slide 7: Literature Gaps & Motivation (Part 2): Methodology Deficiencies**: Beyond the comparison gap, other methodological problems also slow progress. Reproducibility is a major challenge. Essential details are often omitted from publications, source code is rarely shared – Salvi 2024 study was a rare GAN paper providing code – hyperparameters are frequently unreported, and important preprocessing steps are poorly documented. This lack of transparency makes it very hard to verify findings or build upon previous work reliably. Furthermore, evaluation methods are often too superficial. Most studies focus only on accuracy, frequently ignoring computational efficiency, even when significant costs like high VRAM usage or long training times are noted, as reported by Salvi, Dee, and Park. Metrics related to clinical relevance, like prediction calibration – crucial for trustworthy AI – are reported only in about 20% of the studies I reviewed. Assessments of model robustness and generalisability are also inconsistent. Technical hurdles like GAN training instability and the large data needs of ViTs further stress the need for more rigorous, standardised benchmarking and reporting. Taken together, these gaps create a major barrier preventing promising research from becoming reliable clinical tools.

**Slide 8: Research Questions:** These gaps lead directly to my primary research question: How do Vision Transformers compare systematically against GAN-based data augmentation approaches for breast cancer subtype classification using TCGA-BRCA data? This comparison will use a comprehensive framework covering accuracy, computational efficiency, robustness, and clinical relevance proxies, all performed under strictly reproducible conditions.

**Slide 9: Research Objectives:** To answer this question, I have five key objectives. First, I will conduct a systematic comparison of state-of-the-art ViT and GAN-based augmentation methods, ensuring reproducibility specific to Breast Cancer Subtype Classification. Second, I'll develop and use a comprehensive evaluation framework assessing not just accuracy but also efficiency, robustness, and clinical relevance proxies. Third, I plan to quantify the impact of GAN-based data augmentation on ViT performance, calibration, and robustness, while also evaluating robust stain normalisation techniques. Fourth, I will ensure full reproducibility by providing meticulously documented open-source code, detailed methods, standard reporting protocols, and containerised environments, aiming to set a higher standard for transparency. Fifth, my goal is to generate clear, evidence-based recommendations to guide model selection strategies for this task.

**Slide 10: Methodology Title Page:** Proposed Methodology. [Pause]

**Slide 11: Methodology: Dataset Selection & Preparation**: For this research, we'll use the TCGA-BRCA dataset. This choice is justified as it's large-scale, multi-institutional, serves as a standard reference (as Mansour and colleagues used for MammoViT in 2025), is clinically relevant, and provides sufficient data variation for robust testing. I recognise the challenge of stain variability within TCGA, which reflects real-world data. Therefore, implementing and evaluating robust stain normalisation techniques, perhaps exploring GAN-inspired methods like those related to Elif's 2024 StainSWIN approach, is a cornerstone of the methodology. Actively addressing data variance is necessary for generalisable insights. All preprocessing steps and standardised data splits, considering subtype balance and preventing data leakage during augmentation, will be carefully documented to ensure a fair and rigorous comparison.

**Slide 12: Methodology: Model Implementation:** My core experimental design compares representative state-of-the-art models, focusing on the value of generative augmentation. I will compare a Baseline Vision Transformer against a GAN-Augmented Vision Transformer. For the baseline ViT, I'll use a well-established architecture, perhaps adapting DINO-ViT principles discussed by Wessels et al. in 2023, or using standard ImageNet pretraining similar to Park et al.'s 2025 study. For the GAN-Augmented ViT, I will train a Conditional GAN, possibly drawing inspiration from architectures like IC-CGAN proposed by Ravi & Matt in 2025. This GAN will be trained only on separate training partitions to generate synthetic data, specifically targeting underrepresented subtypes in TCGA-BRCA. This setup allows a direct measurement of augmentation's impact on the same baseline ViT architecture. Depending on initial findings, I might optionally include a relevant Hybrid Model, perhaps inspired by MammoViT from Mansour et al. (2025), trained under identical conditions to provide broader context. All implementations will prioritise comparable architectures using consistent frameworks like PyTorch.

**Slide 13: Methodology: Training & Reproducibility Commitment**: Reproducibility is a fundamental pillar of this project, and I commit to exceeding current standards. This includes full transparency: I will document and share all hyperparameters, optimiser details, loss functions, training epochs, batch sizes, and random seeds for all model training phases. To address GAN training instability, I'll use and clearly document specific stabilisation techniques, potentially methods used by Dee et al. in 2024 like spectral normalisation or careful loss weighting, ensuring reliable generation of augmentation data. Furthermore, I commit to open source: all code for preprocessing, GAN training, ViT training, and the entire evaluation pipeline will be released under a permissive license on GitHub. Finally, regarding the environment, I'll use containerisation tools like Docker to package the exact computational environment, ensuring others can replicate my results accurately and tackling a key reproducibility challenge mentioned in the literature.

**Slide 14: Methodology: Comprehensive Evaluation Framework:** My evaluation framework is designed to be comprehensive, addressing the superficial evaluations often seen in prior work. It goes beyond simple accuracy to provide a holistic assessment for Breast Cancer Subtype Classification. I will measure task performance using standard multi-class metrics like AUC per class, F1 scores, precision, recall, and accuracy on the held-out test set. Regarding computational efficiency, I will benchmark training time, inference speed, and peak GPU memory usage, providing crucial context often missing, highlighting real-world costs discussed by Salvi, Dee, and Park. I will also assess clinical relevance proxies by evaluating prediction calibration using metrics like Brier Score and ECE, similar to the methods Dee et al. used in 2024, and performing correlation analysis with available clinical outcomes in TCGA-BRCA, akin to Wessels et al.'s 2023 analysis. If using GAN augmentation, I will assess GAN quality by evaluating the diversity and realism of generated samples. Finally, I'll evaluate robustness by checking performance consistency across cross-validation folds and potentially across different TCGA sites or simulated stain variations post-normalization.

**Slide 15: Timeline of Proposed Activities (6 Months):** This timeline shows the proposed 6-month schedule. Month one focuses on Setup, finalising the literature review and protocols, initiating ethics approval, and acquiring TCGA-BRCA data. Month two moves into intensive Data Preparation, including normalisation and patching, alongside starting model implementation. Month three is dedicated to initial model training for the baseline ViT and the augmentation GAN, also generating synthetic data. Following this, Month four involves training the comparative models, primarily the GAN-augmented ViT, and initiating cross-validation runs. Month five is focused on Evaluation and Analysis, executing the comprehensive evaluation framework and analysing results and trade-offs. Finally, Month six is allocated for the main Write-up and Dissemination, including preparing the thesis or report, releasing the open-source code, and preparing the final presentation. While ambitious, the timeline is structured to meet core objectives, utilising parallelisation and buffer periods where possible.

**Slide 16: Expected Outcomes & Significance Title Page:** Anticipated Contributions. [Pause]

**Slide 17: Expected Outcomes & Significance:** I expect this research to make several significant contributions. First, it aims to provide a much-needed systematic benchmark comparing ViT versus GAN-augmented ViT approaches for Breast Cancer Subtype Classification on TCGA-BRCA, serving as a fully reproducible reference point. Second, I will release a publicly available, meticulously documented codebase and evaluation protocol, acting as a valuable asset for the research community and promoting higher standards. Third, my study will generate clear, quantitative evidence detailing the crucial trade-offs between these modelling strategies concerning accuracy, cost, robustness, calibration, and other factors important for deployment. Fourth, based on this empirical data, I will offer practical, evidence-based guidelines to help researchers and potentially clinicians make informed model selection decisions for this task. Finally, and perhaps most importantly, this work demonstrates how to address common methodological pitfalls like reproducibility, evaluation depth, and handling real-world data variation, contributing to more reliable, translatable, and ultimately impactful AI in pathology.

**Slide 18: Ethical Considerations:** Conducting this research responsibly is a priority. I will use the publicly available, de-identified TCGA-BRCA data, which minimises patient privacy risks, although I remain aware of theoretical concerns about high-resolution images mentioned by Wessels et al. in 2023. My evaluation framework includes assessing performance disparities across different subtypes or sites, contributing to fairness analysis and mitigating risks associated with dataset bias, a concern noted even by Dee et al. in 2024. My commitment to transparency through an open-source release, following examples like Salvi et al. 2024, promotes scrutiny and responsible use. A formal risk assessment identifies minimal risks, mainly related to data security during local processing and potential code misuse, which will be mitigated through clear documentation and licensing. I commit to disseminating findings responsibly.

**Slide 19: Conclusion Title Page**: Conclusion: Towards Reliable AI in Histopathology. [Pause]

**Slide 20: Conclusion:** In conclusion, AI holds immense potential in histopathology, especially techniques like ViTs and GANs. However, realising this potential requires moving beyond fragmented studies often marked by methodological weaknesses in systematic comparison, reproducibility, and comprehensive evaluation. These gaps represent significant barriers to reliable progress and clinical translation. This proposal outlines a clear research plan to address these gaps directly by performing a systematic, deeply reproducible, and comprehensively evaluated comparison of Vision Transformers and GAN-augmented approaches, focused specifically on Breast Cancer Subtype Classification using TCGA-BRCA data. By providing a reliable benchmark, valuable open-source resources, and practical guidance based on empirical trade-offs, I hope this work will foster greater rigor and transparency in the field. Ultimately, my goal is to help accelerate the development and trustworthy deployment of AI tools that can truly benefit patient care.

**Slide 21: References**: And here is the list of the key research papers and sources cited throughout this presentation.

**Slide 22: Thank you. Title Page:**Thank you for listening.

**References:**

Abimouloud, M. L., Bensid, K., Elleuch, M., Ammar, M. B. and Kherallah, M. (2024) *TokenMixer-ViT: Advancing breast cancer diagnosis with hybrid vision transformers*. GitHub repository. Available at: <https://github.com/author/tokenmixer-vit> (Accessed: 15 June 2025).

Chaurasia, A. K., Toohey, P. W., Harris, H. C., & Hewitt, A. W. (2025). Multi-resolution vision transformer model for skin cancer subtype classification using histopathology slides. <https://doi.org/10.1101/2025.01.28.25321234>

Chen, L., Zhang, W. and Huang, F. (2024) ‘GAN-based virtual staining for frozen sections’, Nature Digital Medicine, 3(1). doi: 10.1038/s44287-023-00012-7.

Dee, W., Alaaeldin Ibrahim, R. and Marouli, E. (2024). *Histopathological domain adaptation with generative adversarial networks: Bridging the domain gap between thyroid cancer histopathology datasets*. PLOS ONE, 19(12), p.e0310417. doi: <https://doi.org/10.1371/journal.pone.0310417>.

Delaire, J., Yzet, C. et al. (2025) ‘ViT for inflammatory bowel disease’, Journal of Crohn’s and Colitis, 19(Suppl 1). doi: 10.1093/ecco-jcc/jjae190.0574.

Eghbali, N. and Ghassemi, M. M. (2025) GLoG-CSUnet: Radiomic-enhanced ViTs. arXiv. Available at: arXiv:2501.02788.

Elif Baykal Kablan and Ayas, S. (2024). *StainSWIN: Vision transformer-based stain normalization for histopathology image analysis*. Engineering Applications of Artificial Intelligence, 133. doi: <https://doi.org/10.1016/j.engappai.2024.108136>.

Guo, X., Li, Y. and Wang, H. (2024) ‘Multi-scale hierarchical vision transformers for gastric cancer classification’, *Medical Image Analysis*, 103025. doi: [10.1016/j.media.2024.103025](https://doi.org/10.1016/j.media.2024.103025).

Maheriya, K., Ganatra, A. et al. (2025) ‘XAI-ViT for prostate cancer’, *Artificial Intelligence in Oncology*. doi: [10.1201/9781032626345-12](https://doi.org/10.1201/9781032626345-12).

Mansour, A. M. A. et al. (2025) ‘MammoViT for BIRADS classification’, Diagnostics, 15(3), p. 285. doi: 10.3390/diagnostics15030285.

Mouhamed Laid Abimouloud, Khaled Bensid, Elleuch, M., Ammar, M.B. and Monji Kherallah (2025). Advancing breast cancer diagnosis: token vision transformers for faster and accurate classification of histopathology images. *Visual Computing for Industry Biomedicine and Art*, 8(1). doi:https://doi.org/10.1186/s42492-024-00181-8.

Park, S.Y., Ayana, G., Wako, B.D., Jeong, K.C., Yoon, S.D. and Choe, S.W. (2025). *Vision Transformers for Low-Quality Histopathological Images: A Case Study on Squamous Cell Carcinoma Margin Classification*. Diagnostics, 15(3), p.260. doi: 10.3390/diagnostics15030260.

Ravi, Chaitra & Matt, Siddesh. (2025). IC-CGAN: Imbalanced class-conditional generative adversarial network with weighted loss function. International Journal of Electrical and Computer Engineering (IJECE). 15. 1632. 10.11591/ijece.v15i2.pp1632-1646.

Saad, A., Ghatwary, N. and El-Mahallawy, M. S. (2025) ‘IHC-GAN: H&E-to-IHC translation’, BMC Medical Imaging, 25(1). doi: 10.1186/s12880-024-01522-y.

Salvi, M., Branciforti, F., Molinari, F. and Meiburger, K.M. (2024). *Generative models for color normalization in digital pathology and dermatology: Advancing the learning paradigm*. doi: <https://doi.org/10.1016/j.eswa.2023.123105>.

Springenberg, M., Frommholz, A., Wenzel, M., Weicken, E., Ma, J., & Strodthoff, N. (2023). From modern CNNs to vision transformers: Assessing the performance, robustness, and classification strategies of deep learning models in histopathology. Medical Image Analysis, 87, 102809. https://doi.org/10.1016/j.media.2023.102809

Sun, S., Tessier, L. and Baumgartner, C. F. (2025) *Label-free vision transformers for gigapixel histopathology*. arXiv preprint. Available at: [arXiv:2501.02922](https://arxiv.org/abs/2501.02922).

Von Wangenheim, A. and Pereira, R. (2024) *A systematic review of vision transformers in histopathology*. ResearchGate. Available at: <https://www.researchgate.net/publication/383231612> (Accessed: 15 June 2025).

Wang, T. and Li, J. (2025) ‘Grad-Attn ViT for lymph node metastasis’, *Expert Systems with Applications*, 122786. doi: [10.1016/j.eswa.2025.122786](https://doi.org/10.1016/j.eswa.2025.122786).

Wessels, F. et al. (2023). *A self-supervised vision transformer to predict survival from histopathology in renal cell carcinoma*. World Journal of Urology, 41(8), pp.2233–2241. doi: 10.1007/s00345-023-04489-7.

Yan, B., Wen, Z. and Jing, R. (2025) ‘GAN-generated H&E from Raman imaging’, International Journal of Oral Science, 17(1). doi: 10.1038/s41368-025-00346-y.