

AI-based selection of tumor regions for genomic profiling in neuropathology

Narmin Ghaffari Laleh¹, Lukas Friedrich, Fuat Kaan Aras, Katherine J. Hewitt, Leonille Schweizer, Zunamys I. Carrero, Dilan Savran, Daniel Haag², Silvia Barbosa, Felix Sahm², and Jakob Nikolas Kather¹

All author affiliations are listed at the end of the article

Corresponding Authors: Jakob Nikolas Kather, MD, MSc, Else Kröner Fresenius Center for Digital Health, Technische Universität Dresden, Fetscherstrasse 74, 01307 Dresden, Germany (jakob_nikolas.kather@tu-dresden.de); Felix Sahm, PD DR. DR. Med., Department of Neuropathology, Heidelberg University Hospital, Im Neuenheimer Feld 672, 69120 Heidelberg, Germany (felix.sahm@med.uni-heidelberg.de).

Abstract

Automating pathology workflows with deep learning is increasingly feasible and clinically relevant. We present an AI-based method that identifies diagnostically relevant areas directly from H&E-stained slides, trained on 250 glioma cases using sparse, incomplete annotations. First, we show that attention-based multiple instance learning achieves accurate predictions despite noisy labels, easing the annotation burden. Second, the model highlights tumor regions with high cellularity or grade, offering reproducible guidance for tissue selection. In a prospective evaluation, AI-selected regions achieved a mean Dice score of 0.743 [± 0.077], supporting integration into neuropathology workflows as reliable guidance for molecular diagnostics.

Recent advances in deep learning and the availability of large-scale histopathology datasets have made it increasingly feasible to automate pathology workflows.¹ This is particularly relevant in neuropathology, where brain tumors are heterogeneous, infiltrative, and often contain variable tumor cell content ranging from 5% to 95%.² Deep learning models have shown promising performance on several tasks using H&E-stained whole slide images (WSIs), including classification of methylation-based CNS tumor subtypes,³ WHO subtype prediction,⁴ and detection of diagnostic alterations.⁵

However, a major challenge in applying such methods in clinical workflows lies in identifying the diagnostically most informative region of interest (ROI). For downstream molecular analyses, such as methylation arrays or next-generation sequencing, only one tissue region is sampled, typically by 1-3 mm punches or by scratching FFPE sections with or without microdissection, depending on histologically estimated tumor cell content.² Selecting a representative tumor region is critical to avoid under-sampling, which may result in missed diagnoses, failure to detect actionable mutations, or uninterpretable molecular profiles that require testing.

Currently, ROI selection is performed manually by a neuropathologist, a process that is subjective, labor-intensive, and increasingly unsustainable given shortages of trained specialists.⁶ Existing computational approaches for automated ROI

detection in WSIs, including rule-based cellularity tools such as QuPath⁶ and weakly supervised methods,⁷ have shown promise but were not specifically developed or validated for the complex tissue heterogeneity characteristic of brain tumors. Automating this step could improve consistency, reduce labor, and enable cross-center harmonization of results. However, training AI models for this task typically requires dense annotations or pixel-level labels across large cohorts, which are too time-intensive to be practical.

In this study, we present a deep learning approach that achieves 2 main goals: (1) it is trained using sparse, intentionally imperfect annotations ("noisy labels") from trainee neuropathologists and (2) it produces ROI heatmaps that can guide tissue selection for molecular profiling. We hypothesized that even coarse labels would be sufficient to train a model capable of distinguishing diagnostically useful tumor regions from non-informative regions, motivated by the success of weakly supervised learning in histopathology.⁸

Using 250 H&E-stained WSIs from 250 patients (Table S1), we trained a weakly supervised multiple instance learning (MIL) model (Figure 1A). For each case, 2 trainee neuropathologists (FK, LF) marked tumor regions quickly and without attempting exhaustive delineation. For each patient, we extracted tiles within these annotations (Patient_N_ROI) and tiles from the remaining tissue (Patient_N_nonROI, Figure 1B). These were

Received October 7, 2025; accepted June 4, 2026.

© The Author(s) 2026. Published by Oxford University Press, the Society for Neuro-Oncology and the European Association of Neuro-Oncology. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

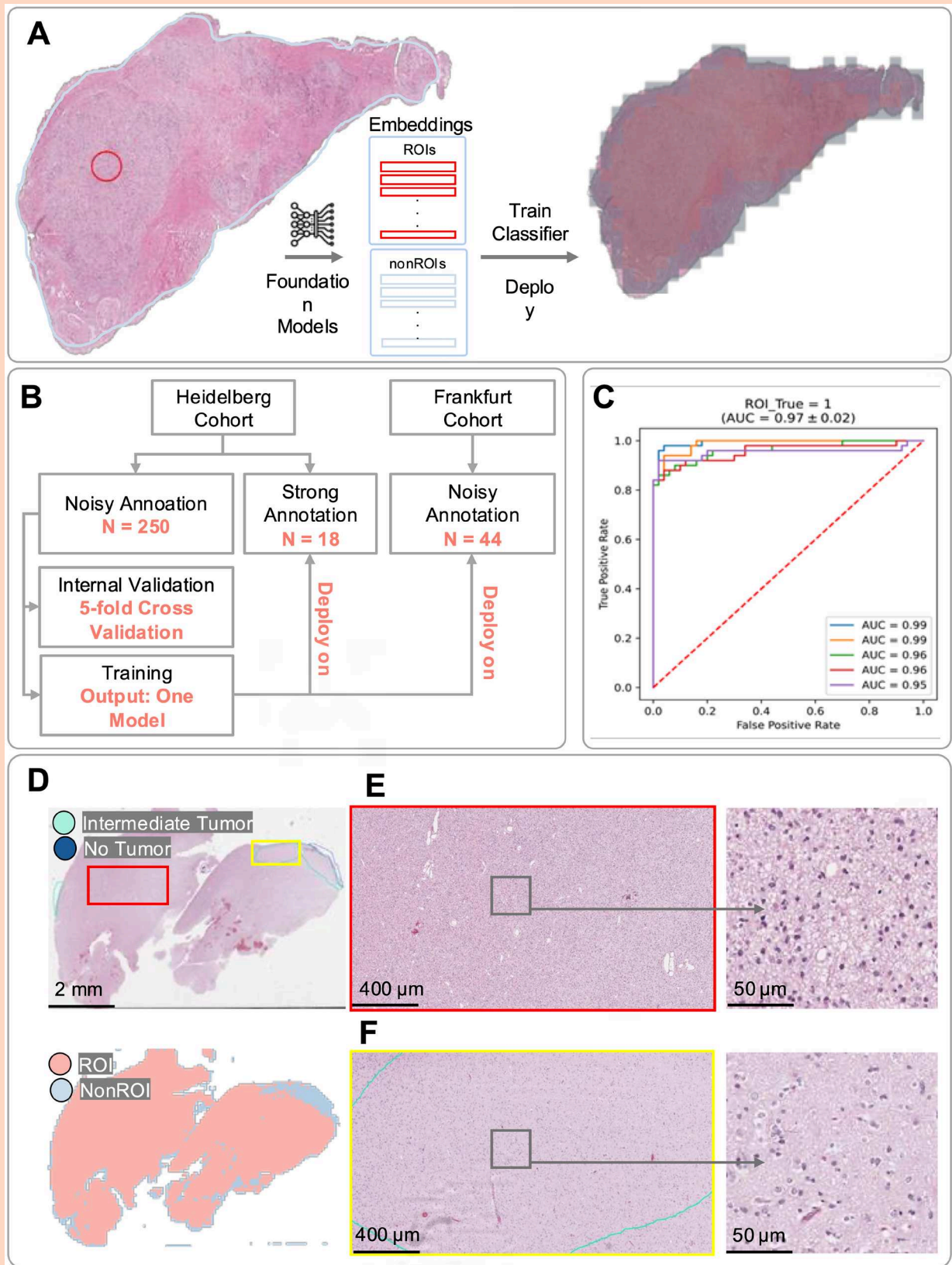


Figure 1. Workflow for automated ROI detection. (A) The pipeline begins with a whole slide image (WSI), from which candidate regions of interest (ROIs) are extracted. ROI and non-ROI tiles are embedded and passed through a neural network model for classification. (B) The model is trained using 250 WSIs and then deployed on 2 external datasets ($n=18$ and $n=44$). (C) The AUROC of 5-fold cross-validation demonstrated high performance of the ROI detection model with an average AUC of 0.97 ± 0.02 . (D) Representative whole slide image with expert annotations indicating intermediate tumor and non-tumor regions, alongside the corresponding AI-generated heatmap showing predicted ROI (red) and non-ROI (blue) regions. (E)

Figure 1. Continued

High-magnification view of the tumor region (red box in D), showing increased cellularity characteristic of solid tumor tissue (400 μm overview with 50 μm inset). (F) High-magnification view of the intermediate tumor region (yellow box in D), showing infiltrative tumor morphology with lower cellularity (400 μm overview with 50 μm inset).

used to create positive and negative bags of tile-level embeddings, generated using a pretrained histology encoder (virchow_2),⁹ which has been shown to outperform alternative foundation models across a wide range of downstream tasks.¹⁰ This setup resulted in 500 embedding bags across all cases (Supplementary Methods).

The MIL model was trained to classify embedding bags using 5-fold cross-validation, ensuring that ROI and non-ROI bags from each case remained in the same data split (<https://github.com/KatherLab/marugoto/tree/main>). The model achieved an average AUROC of 0.97 [± 0.023 (95% CI)] across folds (Figure 1C), confirming strong discriminatory performance. Following cross-validation, we trained a final model on the full training cohort and evaluated it on 2 external validation sets ($n=18$ and $n=44$). The model was used to generate ROI heatmaps by classifying tile embeddings across the WSI (Figure 1D). This yielded segmentation maps that highlight regions most suitable for tissue punching and downstream molecular analysis (Figure S1).

To quantitatively assess the model's localization performance, detailed ROI maps were generated for a subset of the validation cohort ($n=18$). Binary masks were created for tumor, intermediate tumor, and non-tumor regions (Figure 1D). Dice similarity coefficients were calculated for predicted ROI regions versus these masks, yielding a mean Dice score of 0.743 [± 0.162] for tumor-only regions and 0.764 [± 0.154] when including both tumor and intermediate tumor areas. Additionally, a neuropathologist (LF) reviewed the heatmaps for all external cases and scored them on a 0 to 4 scale based on the proportion of tumor highlighted (0: ROI highlighted non-tumor tissue only; 1: ROI predominantly highlighted non-tumor tissue; 2: ROI highlighted approximately 50% tumor and 50% non-tumor tissue; 3: ROI predominantly highlighted tumor tissue; 4: ROI highlighted tumor tissue only) (Figure S2). The model achieved a score of 3 or 4 in 95% of cases, indicating strong alignment with expert judgment (score 0—0%, score 1—0%, score 2—5%, score 3—35%, and score 4—60%). While isolated cases showed reduced discrimination between solid tumor and infiltrative regions, the distinction between infiltration and non-tumor tissue was consistently excellent (Table S2).

Finally, to assess clinical utility, we performed differential methylation analysis in 4 additional cases, comparing 1 ROI and 1 non-ROI punch selected by the model per case (Figure S3A). In 1 case the non-ROI sample had a non-evaluable methylation score. Two cases showed CNV plots with lower quality, whereas 1 case yielded equal results. ROI punches consistently produced correct methylation classifications and high-quality molecular profiles (DNA methylation classification results). Detailed DNA methylation classification results and CNV profiles for all 4 cases are available via the link provided in the Data Availability section.

Together, these findings demonstrate that a simple deep learning model trained in a weakly supervised manner can effectively localize diagnostically relevant tumor regions in brain tumors. This study focuses on feasibility rather than exhaustive benchmarking against alternative methods. Our results show that coarse, “noisy” annotations are sufficient to train useful models, substantially reducing annotation effort while maintaining robust performance. Given that high-quality, pixel-level annotations are rarely available at scale in routine practice, this supports a more realistic and scalable approach. While more precise annotations may further improve quantitative metrics, our findings indicate that fast, coarse annotations can still achieve clinically meaningful ROI selection. Automated ROI selection could serve as a foundation for integrating AI into the broader diagnostic workflow in neuropathology, including tumor classification and molecular testing, tumor grading, spatial transcriptomics, and efficient annotation of large histopathology datasets. In edge cases with low or absent tumor content (Figure S3B), the model is expected to produce predominantly non-ROI heatmaps; however, pathologist review remains essential in such scenarios to confirm whether the available slide is suitable for molecular analysis or whether an alternative section should be selected. This consideration could be further supported in future work by incorporating a confidence-based ROI area metric to objectively flag cases with insufficient tumor tissue. As the model is trained on real neuropathology cases rather than engineered features, it likely captures a broader spectrum of morphological characteristics beyond cellularity, offering potential advantages over rule-based computational approaches in complex infiltrative settings. Furthermore, extending the pipeline to process multiple slides per case simultaneously would reduce the need for initial slide selection by the pathologist. Automated ranking of detected regions based on histological high-grade features, such as mitotic activity, could further streamline the diagnostic workflow and decrease pathologist burden in routine practice. The model is designed to distinguish diagnostically relevant from non-relevant regions and does not explicitly quantify tumor cellularity or rank multiple tumor regions with different histological characteristics.

Supplementary Material

Supplementary material is available online at *Neuro-Oncology Advances* (<https://academic.oup.com/nea>).

Keywords

glioma | molecular diagnostics | multiple instance learning | tumor region segmentation | weakly supervised learning

Author Contributions

N.G.L., F.S., and J.N.K. conceptualized, designed the study and contributed to critical data interpretation. N.G.L., L.F., F.K.A., K.J.H., L.S., and D.H. contributed to the data analysis and/or provided technical support. L.F., F.K.A., and K.J.H. contributed to the pathological revision of attention heatmaps and tiles from WSIs. L.G., F.K.A., K.J.H., L.S., D.S., D.H., S.B., F.S., and J.N.K. provided scientific input. N.G.L. wrote the first draft of the manuscript under the supervision of F.S. and J.N.K. L.F., F.K.A., K.J.H., L.S., D.S., D.H., and S.B. have been involved in critical revision of the manuscript.

Conflict of Interest Statement

J.N.K. declares consulting services for Bioprimus; Panakeia; AstraZeneca; and MultiplexDx. Furthermore, he holds shares in StratifAI, Synagen, Tremont AI and Ignition Labs; has received an institutional research grant by GSK; and has received honoraria by AstraZeneca, Bayer, Daiichi Sankyo, Eisai, Janssen, Merck, MSD, BMS, Roche, Pfizer, and Fresenius. F.S. is co-founder and shareholder of Heidelberg Epignostix GmbH. N.G.L. was employed as a postdoctoral researcher at Heidelberg University and her position was funded by GlaxoSmithKline (GSK).

Funding

JNK is supported by the German Cancer Aid DKH (DECADE, 70115166), the German Federal Ministry of Research, Technology and Space BMFT (PEARL, 01KD2104C; CAMINO, 01E02101; TRANSFORM LIVER, 031L0312A; TANGERINE, 01KT2302 through ERA-NET Transcan; Come2Data, 16DKZ2044A; DEEP-HCC, 031L0315A; DECIPHER-M, 01KD2420A; NextBIG, 01ZU2402A; PROSURV, 01KD2509C), the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) as part of Germany's Excellence Strategy – EXC 2050/2 – Project ID 390696704 – Cluster of Excellence “Centre for Tactile Internet with Human-in-the-Loop” (CeTI) of Technische Universität Dresden, as well as through DFG-funded collaborative research projects (TRR 412/1, 535081457; SFB 1709/1 2025, 533056198), the German Academic Exchange Service DAAD (SECAI, 57616814), the German Federal Joint Committee G-BA (TransplantKI, 01VSF21048), the European Union EU's Horizon Europe research and innovation programme (ODELIA, 101057091; GENIAL, 101096312), the European Research Council ERC (NADIR, 101114631), the Breast Cancer Research Foundation (BELLADONNA, BCRF-25-225) and the National Institute for Health and Care Research NIHR (Leeds Biomedical Research Centre, NIHR203331). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. This work was funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. This publication was supported through state funds approved by the State Parliament of Baden-Württemberg for the Innovation Campus Health + Life Science alliance Heidelberg Mannheim.

Ethics

The sample collection of experimental and clinical data was approved by the Ethics Committee of Heidelberg University Hospital (under IRB vote S-649/2021). The study was conducted in accordance with the Declaration of Helsinki.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request. The DNA methylation classification results are publicly available at: https://drive.google.com/file/d/1gUET3Nz3WY40u8M9SaMbo79rN9CmsKWT/view?usp=drive_link.

Affiliations

Else Kroener Fresenius Center for Digital Health, Faculty of Medicine and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Dresden, Germany (N.G.L., K.J.H., Z.I.C., J.N.K.); Dept. of Neuropathology, University Hospital Heidelberg, Heidelberg Germany and CCU Neuropathology, German Consortium for Translational Cancer Research (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany (L.F., F.K.A., D.S., D.H., F.S.); Frankfurt Cancer Institute (FCI), Frankfurt am Main, Germany (L.S.); German Cancer Consortium (DKTK), Partner Site Frankfurt/Mainz, German Cancer Research Center (DKFZ), Heidelberg, Germany (L.S.); Institute of Neurology (Edinger Institute), University Hospital Frankfurt, Goethe University, Frankfurt am Main, Germany (L.S.); Medical Oncology, National Center for Tumor Diseases (NCT), University Hospital Heidelberg, Heidelberg, Germany (S.B., J.N.K.); Department of Medicine I, Faculty of Medicine and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Dresden, Germany (J.N.K.); Pathology & Data Analytics, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, United Kingdom (J.N.K.).

References

1. Bilal M, Aadam Raza M, Altherwy Y, et al. Foundation models in computational pathology: A review of challenges, opportunities, and impact. arXiv. 2025. <http://arxiv.org/abs/2502.08333>, preprint: not peer reviewed.
2. Capper D, Stichel D, Sahm F, et al. Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. *Acta Neuropathol.* 2018;136:181-210. <https://doi.org/10.1007/s00401-018-1879-y>
3. Jin D, Shmatko A, Patel A, et al. Hetairios is a histology-based artificial intelligence model for predicting central nervous system tumor methylation subtypes. *Nat Cancer.* 2026. <https://doi.org/10.1038/s43018-026-01186-3>
4. Schumann Y, Dottermusch M, Schweizer L, et al. Morphology-based molecular classification of spinal cord ependymomas using deep neural

- networks. *Brain Pathol.* 2024;34:e13239. <https://doi.org/10.1111/bpa.13239>
5. Hewitt KJ, Löffler CML, Muti HS, et al. Direct image to subtype prediction for brain tumors using deep learning. *Neurooncol Adv.* 2023;5:vdad139. <https://doi.org/10.1093/naojnl/vdad139>
 6. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep.* 2017;7:16878. <https://doi.org/10.1038/s41598-017-17204-5>
 7. Afonso M, Bhawar PMS, Saha M, et al. Multiple Instance Learning for WSI: a comparative analysis of attention-based approaches. *J Pathol Inform.* 2024;15:100403. <https://doi.org/10.1016/j.jpi.2024.100403>
 8. Gadermayr M, Tschuchnig M. Multiple instance learning for digital pathology: a review of the state-of-the-art, limitations & future potential. *Comput Med Imaging Graph.* 2024;112:102337.
 9. Zimmermann E, Vorontsov E, Viret J, et al. Virchow2: Scaling self-supervised mixed magnification models in pathology. arXiv. 2024. <http://arxiv.org/abs/2408.00738>, preprint: not peer reviewed.
 10. Neidlinger P, El Nahhas OSM, Muti HS, et al. Benchmarking foundation models as feature extractors for weakly-supervised computational pathology. *Nat Biomed Eng.* 2025. <https://doi.org/10.1038/s41551-025-01516-3>